Over-Expression, Purification and Biochemical Characterization of DOXP Reductoisomerase and the Rational Design of Novel Anti-Malarial Drugs

By

Delia Caroline Tanner

Department of Biochemistry, Microbiology and Biotechnology,
Rhodes University, Grahamstown
December 2003

A dissertation submitted to the Faculty of Science, Rhodes University, in fulfilment of the requirements for the Degree Master of Science in Biochemistry by Research

DECLARATION

I declare that this is my own, unaided work. It is being submitted for the degree of Master of Science at Rhodes University, Grahamstown. It has not been submitted before for any degree or examination in any other University.

Delia C. Tanner

14th day of March 2004.

ABSTRACT

Malaria poses the greatest threat of all parasites to human life. Current vaccines and efficacious drugs are available however their use is limited due to toxicity, emergence of drug resistance, and cost. The discovery of an alternative pathway of isoprenoid biosynthesis, the non-mevalonate pathway, within the malarial parasite has resulted in development of novel anti-malarial drugs. 1-Deoxy-D-xylulose-5-phosphate (DOXP) reductoisomerase, the second enzyme in this pathway, is responsible for the synthesis of 2-*C*-methyl-D-erythritol 4-phosphate (MEP) in an intramolecular rearrangement step followed by a reduction process involving NADPH as a hydrogen donor and divalent cations as co-factors. Fosmidomycin and FR900098 have been identified as inhibitors of DOXP reductoisomerase. However, they lack clinical efficacy. In this investigation recombinant DOXP reductoisomerase from *Escherichia coli* (EcDXR) and *Plasmodium falciparum* (pfDXR) were biochemically characterized as potential targets for inhibition. $(His)_6$-EcDXR was successfully purified using nickel-chelate affinity chromatography with a specific activity of 1.77 μmoles/min/mg and K_m value 282 μM. Utilizing multiple sequence alignment, previous structural data predictions and homology modeling approaches, critical active site amino acid residues were identified and their role in the catalytic activity investigated utilizing site-directed mutagenesis techniques. We have shown evidence that suggests that Trp^{212} and Met^{214} interact to maintain the active site architecture and hydrophobic interactions necessary for substrate binding, cofactor binding and enzyme activity. Replacement of Trp^{212} with Tyr, Phe, and Leu reduced specific activity relative to EcDXR. EcDXR(W212F) and EcDXR(W212Y) had an increased K_m relative to EcDXR indicative of loss in affinity toward DOXP, whereas EcDXR(W212L) had a lower K_m of ~8 μM indicative of increased affinity for DOXP. The W212L substitution possibly removed contacts necessary for full catalytic activity, but could be considered a non-disruptive substitution in that it maintained active site architecture sufficient for DOXP reductoisomerase activity. EcDXR(M214I) had 36-fold reduced enzyme activity relative to EcDXR, while its K_m (~8 μM) was found to be lower than that of EcDXR. This suggested that the M214I substitution had maintained (perhaps improved) substrate and active site architecture, but may have perturbed interactions with NADPH. Rational drug design strategies and docking methods have been utilized in the development of furan derivatives as DOXP reductoisomerase inhibitors, and the synthesis of phosphorylated derivatives (**5**) and (**6**) has been achieved. Future inhibitor studies using these novel potential DOXP reductoisomerase inhibitors may lead to the development of effective anti-malarial drug candidates.

ACKNOWLEDGEMENTS

Foremost I would like to thank my family for their support throughout these past two years. To my parents, David and Marge Tanner, thank-you not only for the opportunities that you have given me in terms of my studies but also for your love and support. To my sister Claire, thank you for being a wonderful "best" - friend!

To Eve, Jim, and Garth Cambray, and Janice Limson thank-you for being wonderful friends and for your support throughout my studies and always encouraging me to succeed!

In addition I would like to acknowledge the following people for their support and assistance during this study:

- My Supervisor Professor Greg Blatch and Co-supervisor Professor Perry Kaye for their dedication, support and unfailing advice
- Jomaa Pharmaka GmbH, the National Research Foundation and the Medical Research Council for funding
- Rhodes University
- To my proof-readers Dr W.S. Nicoll, Prof G.L. Blatch and Prof P.T. Kaye - a very big thank-you
- Fritha Hennessy for her assistance with Molscript and Whatif and her friendship
- Andy Soper for technical assistance
- To all my friends in G-town, thanx for all the good times!
- The members of the Chaperone Research Group
- Kevin Lobb, Ingrid Sabbagh, Andrew Duggan and Mike Wisch from the Asymmetric Synthesis Group for their assistance, support and friendship

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

α	alpha
β	beta
ε	extinction coefficient
λ	lambda
μ	micro
Å	angstrom
ºC	degree Celsius
mm	millimetres
cm	centimetres
ng	nanogram
μg	microgram
g	gram
μl	microlitre
ml	millilitre
L	litre
pmol	picomol
nmol	nanomoles
μmol	micromoles
mmol	millimoles
mol	moles
μM	micromolar
mM	millimolar
M	molar
%	percentage
>	greater than
±	standard deviation
AGE	agarose gel electrophoresis
A_x	Absorbance at x nm
Amp	amperes

B. subtilis	*Bacillus subtilis*
bp	base pairs
BSA	Bovine Serum Albumin
cfu	colony forming unit
CD	circular dichroism
CD-ME	4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol
CSP	circumsporozoite protein
3-D	three dimensional
DDT	dichloro-diphenyl- trichloroethane
DFT	Density Functional Theory
DHFR	dihydrofolate reductase
DHODase	dihydroorotate dehydrogenase
DNA	deoxyribonucleic acid
dNTP	dinucleotide triphosphate
DMAP	4-dimethylaminopyridine
DMF	*N*,*N*-Dimethylformamide
DOXP	1-deoxy-D-xylulose-5-phosphate
EBA	erythrocyte-binding antigen
E. coli	*Escherichia coli*
EcDXR	*E. coli* DOXP reductoisomerase
EGF	epidermal growth factor
EP	electrostatic potential
Et	ethyl
FOSM	fosmidomycin
FV	food vacuole
GST	Glutathione-*S*-transferase
$(His)_6$	6 x histidine tag
HIV	Human Immunodeficiency Virus
HPLC	High Performance Liquid Chromatography
IPTG	isopropyl-β-D-galactopyranoside
IPP	isopentenyl diphosphate units
K	Kelvin
Kb	kilobase
kbp	kilo base pairs

kDa	kilodalton
Mb	megabase
Me	methyl
MECDP	2-*C*-methyl-D-eryhtritol-2,4-cyclodiphosphate
MEP	2-*C*-methyl-D-erythritol 4-phosphate
MHz	megahertz
min	minutes
MM	molecular mechanics
mRNA	messenger ribonucleic acid
MSP	merozoite surface protein
MVA	mevalonate
NADH	nicotinamide adenine dinucleotide (reduced form)
NADPH	nicotinamide adenine dinucleotide phosphate (reduced form)
NCBI	National Center for Biotechnology
N.D.	not determined
nm	nanometres
Ni	Nickel
NMR	Nuclear Magnetic Resonance
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	Protein Data Bank
Ph	phenyl
pH	- log [H]
P. falciparum	*Plasmodium falciparum*
P. gallinaceum	*Plasmodium gallinaceum*
P. malariae	*Plasmodium malariae*
P. ovale	*Plasmodium ovale*
P. vinckei	*Plasmodium vinckei*
P. vivax	*Plasmodium vivax*
PMSF	phenylmethylsulfonylfluoride
ppm	parts per million
QSAR	quantitative structure-activity relationship
RMS	root mean square

RNA	ribonucleic acid
rpm	revolution per minute
S. cerevisiae	*Saccharomyces cerevisiae*
SERA	serine-rich antigen
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
TBV	transmission blocking vaccine
TFA	trifluoroacteic acid
THF	tetrahydrofuran
tRNA	transfer ribonucleic acid
USA	United States of America
V	volts
v/v	volume/volume
w/v	weight/volume

The IUPAC-IUBMB three and one letter codes for amino acids were used, and single letter codes were used for nucleotides

CHAPTER 1 : LITERATURE REVIEW

1 INTRODUCTION

Malaria is one of the worlds most infectious and fatal tropical diseases resulting in approximately 1.5-3 million deaths annually (Shi *et al*., 1999). The majority (greater than 90 %) of these fatalities occur in Sub-Saharan Africa affecting mainly pregnant woman and children (Shi *et al*., 1999). Fatal cases of malaria are often the result of infection from *Plasmodium falciparum* (*P. falciparum).* Currently malaria poses a health risk to over 90 countries of which 40% of the global population inhabit (Beier, 1998). Not only does malaria pose a health risk but it also accounts for the loss of gross domestic product. In Sub-Saharan Africa this loss was estimated between 1-4%, which amounts to roughly 12 billion dollars (Beier, 1998). With no vaccine or cure against the disease, and many of the traditional anti-malarial drugs are failing due to growing drug resistance in *P. falciparum.* With the prevalence of malaria on the increase and the compounding problem of multiple drug resistance there is an urgent demand for new anti-malarial drugs.

With the discovery of the alternative pathway of isoprenoid synthesis within malaria parasites, and the validation of the 1-deoxy-2-xyulose-5-phosphate (DOXP) pathway has come the realization of the importance of this pathway in the fight against malaria, and the development of new anti-malarial drugs. Consequently, the development of inhibitors of the alternative pathway for isoprenoid biosynthesis in the respective pathogen will allow one to target and thus kill the pathogen with no subsequent side effects for the human host, which lacks the DOXP pathway.

In this research, the DOXP reductoisomerase enzyme, of the alternative mevalonate pathway, will be biochemically characterized as a target site for inhibition. Inhibition of this enzyme will hopefully lead to the development of novel anti-malarial drugs for the future fight against malaria.

1.1 Malaria

Sir Patrick Manson first discovered that malaria was being transmitted via the female mosquito in 1876 (http://www. Malaria.freewire.co.uk/malaria.html; date accessed 04/2002). The malaria parasite is a member of the protozoan parasites belonging to the genus *Plasmodium* as well as a

group of one-celled organisms referred to as apicomplexans (Foley and Tilley, 1997; Ridley, 1999). Four species of *Plasmodium* are responsible for human malaria: *P. falciparum, P. vivax, P. ovale,* and *P. malariae*, the most fatal of these being *P. falciparum* (Foley and Tilley, 1999; Beier, 1998; Phillips, 2001).

In the past, many attempts to eradicate malaria have been unsuccessful. This has resulted in malaria still being one of the world's most prominent diseases with approximately 300 to 500 million infections and approximately one million deaths from the disease annually, meaning a death every 30 seconds as a result of malaria (Gardner *et al.*, 1998; Greenwood and Mutabingwa, 2002). Malaria therefore remains a public health threat for more than 90 countries, which comprise 40% of the global population (Beier, 1998). A large percentage of fatalities due to malaria occur within Africa. However malaria is endemic through most of South East Asia, the Indian sub-continent and the South Pacific region (Foley and Tilley, 1997). The disease is endemic to the tropical regions, with extensions into the subtropics. However, due to the ignorance of travelers toward prophylactic drugs there has been an increase in deaths due to malaria in non-malarious regions. Thus, there is a drastically deteriorating situation with the disease returning to areas where the disease had been eradicated and the spread of the disease into new areas such as Eastern Europe and Central Asia (Beier, 1998). Additionally, due to the increase in resistant strains of the parasite there has been a resurgence of malaria including within areas that were previously eradicated of the disease.

1.2 The Biology and Life Cycle of Plasmodium parasites

All the members of the phylum *Apicomplexa*, generally referred to as sporozoans, are parasites of mammals (Campbell, 1996). However, *Apicomplexa* and *Plasmodium* parasites are not exclusive parasites of mammals or man, respectively. The female Anopheles mosquito transmits malaria, although out of the 380 species of this mosquito only 60 are vectors for malaria (Beier, 1998). The reason that not all mosquitoes are vectors of malaria lies in the presence of a toxin to *Plasmodium* within their cells preventing them from being suitable vectors (http:www-micro.msb.le.ac.uk/ 224/ Bradley /Biology.html). In 1998 malariologists celebrated the centenary of the demonstration that the female mosquito transmits malaria. It was the British GP,

Dr Ronald Ross, that demonstrated the biology behind the disease in August 1897 and was awarded the Nobel Prize for his research (reviewed in Phillips, 2001).

The majority of *Apicomplexans* have intricate life cycles in which there is both a sexual and asexual stage with the requirement of at least two hosts for the completion of the cycle (Campbell, 1996). As seen in Figure 1.1 of the life cycle of the *Plasmodium* species in mammals, the malaria parasite disseminates into minute infectious thread-like cells known as sporozoites. These sporozoites are produced within the female mosquito salivary gland and are subsequently injected into the human blood stream as the mosquito must inject anticoagulant saliva to ensure an "even-flowing" meal. Once the sporozoite has gained access to the bloodstream its first destination is the liver whereupon it infiltrates the hepatocytes utilizing a complex group of organelles concentrated within the apex of the sporozoite specialized for the purpose of penetrating host cells and tissues. The sporozoite then remains within the hepatocytes for a period of 6 to 15 days, depending on the species, for the sole purpose of multiplication (Phillips, 2001).

Upon the rupture of the hepatocyte, due to the effects of multiplication of the sporozoite, the parasite is once again released into the blood stream in the form of a merozoite. At this point the pathogen begins to target the hosts red blood cells where they devour haemoglobin and either produce new merozoites through shizogony (asexual division) or the sexual micro and macrogametocytes. Normally several cycles of shizogony occurs prior to the production of any gametocytes (Phillips, 2001).

The new merozoites penetrate additional red blood cells repeatedly rupturing the erythrocytes at intervals of 48 to 72 hours depending on the species of parasite. At this stage of infection the host will begin to take affects of what is occurring within his/hers body, the symptoms being periodic chills, fever and sweating (Campbell, 1996).

The micro and macrogametocytes, which were produced from merozoites, have no further role to play within the human host and are therefore transmitted to the next mosquito that happens to feed upon the infected hosts blood. Upon transmission to the mosquito the gametocytes increase

in volume resulting in the rupture of the erythrocyte within the mosquito. After which, they travel toward the gut where exflagellation of the microgametocytes occurs and fertilization of the macrogametocytes takes place resulting in an ookinete. The ookinete will then infiltrate the midgut wall of the mosquito where it will develop into an oocyst. The oocyst is then responsible for the production of sporoblasts from which sporozoites will bud off. The sporozoites travel to and penetrate the salivary gland cells where they lie in vacuoles for 59 days. During this period of time the sporozoites develop and become a 1000 fold more infectious and more antigenic. A biting mosquito will then transfer at least 10% of the infectious sporozoite load into the capillaries or perivascular tissue of the human host. At this point the sporozoites will begin to evade the hosts immune response on their passage towards the liver. All sporozoites will have left the peripheral circulation within 45 minutes of being transmitted (as reviewed in Phillips, 2001; Bannister and Mitchell, 2003).

Throughout this complicated life cycle the parasite is able to take advantage of several different biological environments (Bannister and Mitchell, 2003). Within the human host the parasite masks its invasion by remaining within host cells and avoiding contact with the hosts circulating antibodies. Additionally, immune attack on vulnerable parasites is avoided through the alteration in the expression of parasitic genes (Bannister and Mitchell, 2003). Another key feature to the survival of the parasite is its ability to adhere to the linings of the blood vessel walls thus hindering tissue perfusion (Bannister and Mitchell, 2003; Miller *et al*., 2002).

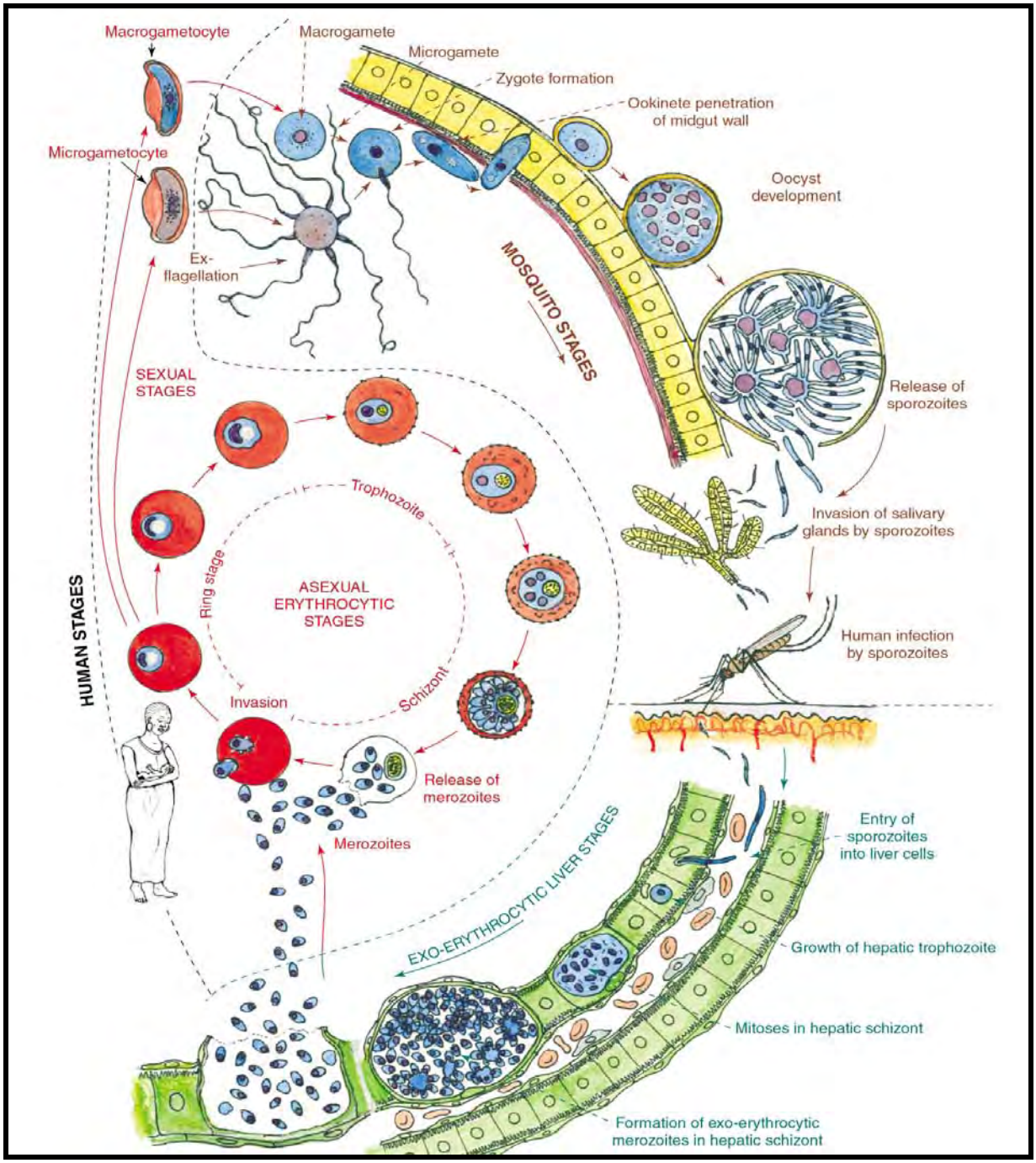


Figure 1.1: The life cycle of the malaria parasite in man and mosquito

The phases of the life cycle in the liver and erythrocytes of the human host, and in the mid-gut and salivary glands of the *Anopheles* mosquito are shown (taken from Bannister and Mitchell, 2003).

1.3 Prevention and Treatment of Malaria

Chemotherapy and prophylaxis are the principal means of combating malaria in the human host. In the past 75 years, during which there was an introduction of synthetic anti-malarials, only a small minority of these compounds have been proven useful in clinical trials and this limited protection against malaria is diminishing with the emergence of drug resistance parasite strains (Hyde, 2002).

Prevention may be implemented through the interruption of the transmission of the disease from the vector to the host. This involves the use of insecticides, bed nets etc. to contain the mass populations of mosquitoes in any particular area. This in itself presents several problems and has declined as a preventative method over the last few years due to the consequences of the dichloro-diphenyl- trichloroethane (DDT) catastrophe, lack of personnel, inefficient insecticide usage and mass population dynamics (Phillips, 2001). However, since many of the techniques for genetically manipulating the mosquito are readily available, another approach has been to genetically manipulate the vector itself to render it unable to transmit the parasite (Ito *et al*., 2002). Ito *et al*. (2002) demonstrated hampered malarial transmission by utilizing a transgenic approach, which allowed for the expression of anti-parasitic genes within the midgut epithelium of culicine and anopheline mosquitoes. Though their preliminary results were successful, one of the main disadvantages to implementing transgenic mosquitoes as a means for combating malaria is the dispersion of foreign genes amongst other mosquito populations. Another option is the most common preventative measure, and involves the use of prophylactic drugs. The current anti-malarial drugs target the various stages of the parasite life cycle with the majority being specific for the intra-erythrocytic stages (Olliaro, 2001).

Most people suffering from malaria in the tropics are treated as outpatients with treatment mainly involving antipyretics (paracetamol and aspirin) and oral anti-parasitics which target the asexual blood stages of the diseases. However, there are a limited number of anti-malarial drugs for widespread use and a lack of affordable new ones. Some common anti-malarial drugs are shown in Figure 1.2 (Winstanley, 2000).

Treatment of malaria is becoming an ever-increasing problem due to the rise in drug resistant species of malaria; this is especially the case with *P. falciparum*. With the exception of artemisinin and its derivatives, all other anti-malarial drugs have been shown to elicit drug resistance within the parasite. This has consequently spurred much research into the mechanism of drug resistance (Phillips, 2001). The knowledge of the mode of action of anti-malarial drugs, and the mechanism of drug resistance is vital information that will lead to the discovery of novel therapeutics and anti-malarial drugs for the future.

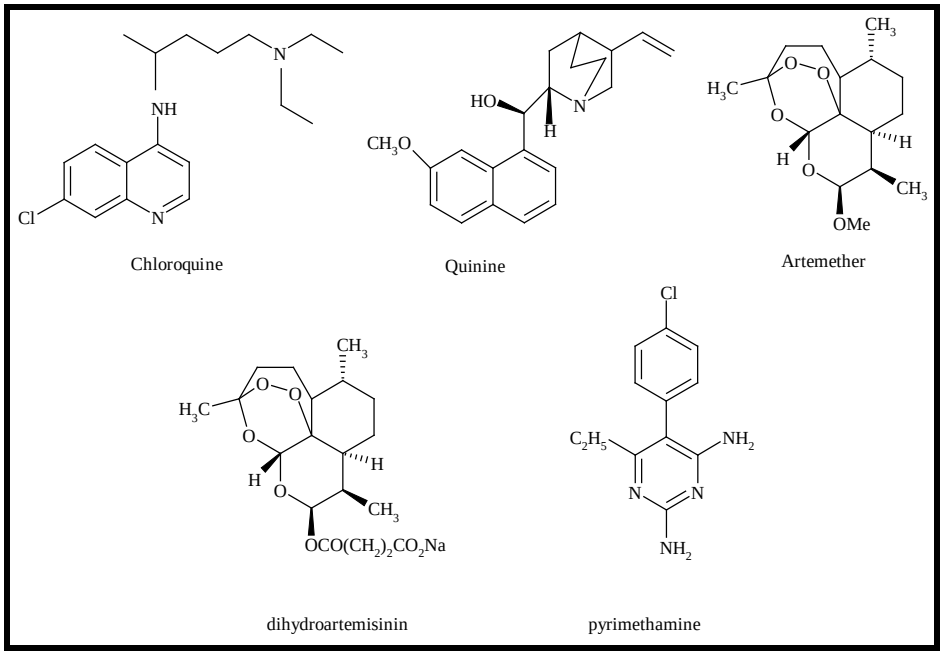


Figure 1.2: The chemical structures of common drugs utilized for malaria treatment

Resistance to several anti-malarial drugs emerged due to their widespread use. This problem was compounded due to multiple drug resistance also being reported due to the use of several closely related anti-malarial compounds (Phillips, 2001). Furthermore, the rapid development of resistance, as well as cross-resistance, also resulted due to the drugs not being administered in high enough doses to ensure a cure to the infection. Selective pressure under such conditions results in the development of increased drug resistance. For this reason, drugs with longer half-lives will exert a larger selection pressure for resistance due to the drug being able to linger in the body at levels which are not suitable to ensure eradication of the infection (Phillips, 2001). Additionally, a further aspect for consideration is whether the parasite was able to undergo

biological modifications due to the effects of acquiring drug resistance. This may allow for certain advantages, or disadvantages, for the parasite regarding survival within the host as well as upon transmission from one host to the other. Studies involving these aspects have already indicated that the resistant parasites are not as effectively transmitted as the non-resistant (reviewed within Phillips, 2001). Table 1 illustrates the mode of action and resistance for some common anti-malarial drugs (Olliaro, 2001).

Various *in vivo* and *in vitro* techniques can be employed to assess drug efficacy but mainly *in vitro* techniques are currently used to detect drug sensitivity of malarial parasites due to the low risk to patients. Several novel *in vitro* techniques have been developed namely the WHO schizont maturation assay, isotopic and colourimetric assays based on the parasitic protein lactate dehydrogenase (pLDH) and the histidine-rich protein 2 (HRP-2) and are reviewed within Noedl *et al*. (2003).

1.3.1 The Antifolates

A large proportion of anti-malarial drugs fall into this class of nucleic acid inhibitors, even though their role in combating malaria is now diminished due to the emergence of drug resistance under drug pressure (Olliaro, 2001). In contrast to other anti-malarials the primary targets for the antifolates has been established with a large degree of research in understanding the mechanism of drug resistance towards these drugs being emphasized upon (Hyde, 2002).

Malarial parasites attain their folate requirements via *de novo* synthesis. The purpose of antifolate drugs is to inhibit two main enzymes within this pathway namely dihydrofolate reductase (DHFR) and dihydropteroate synthase. Pyrimethamine and proguanil target plasmodial DHFR while the sulfur drugs target dihydropteroate synthase, hence their synergistic activity when used in combination therapies (Phillips, 2001). Consequent inhibition of the enzymes of the folate pathway results in decreased pyrimidine synthesis, as well as reduced DNA, serine and methionine formation (Olliaro, 2001).

Table 1: Mechanism of action and resistance for the main classes of anti-malarial drugs

Drug Class	Stage of life Cycle affected	Pathway	Target and mode of Action	Resistance Status
Antifolates	Pre-erythrocytic cycle and the asexual intra-erythrocytic cycle	Folate synthesis		
Type-1			Inhibit DHPS through mimicking the substrate PABA	Mutations in the active site of DHPS reduce binding of the drug
Type-2			Inhibit DHFR by mimicking the substrate dihydrofolate	Mutations at the active site reduce binding of drug. Introduction of bulky amino acids cause steric hindrance for rigid drug molecules.
Atovaquone			Affects mitochondrial electron transport and inhibits DHODase	Mutations in cytochrome *b* include T431C, G399A, and G850T.
Quinolines	Asexual and intra-erythrocytic cycle	Haem detoxification and disposal	Haem crystallization and food vacuole functioning	Point mutations in the *pfmdr1* gene as well as other genes result in proteins that confer resistance
Artemisinin	Asexual and intra-erythrocytic cycle	Questionable but may be linked to oxidative stress	Reductive cleavage of endoperoxides	No known cases of resistance

(adapted from Rastelli *et al*., 2000; Olliaro, 2001; Hyde , 2002)

The DHFR inhibitors mode of action is to mimic the pteridine ring of the natural substrate DHF and as a result compete with it for the active site upon the enzyme (Olliaro, 2001). The enzyme DHFR plays a vital role within the folate pathway maintaining a constant supply of folate cofactors that are essential in allowing for the single carbon transfer reactions of which only the fully reduced (tetrahydro) forms are functional (Hyde, 2002). These reactions are vital for the provision of nucleotides for DNA synthesis and in the metabolism of certain amino acids (Hyde, 2002). Pyrimethamine (PYR), proguanil (PG), the sulfonamides, such as sulfadoxine (SDX) and dapsone are the most commonly used antifolate drugs.

Resistance towards the DHFR and DHPS inhibitors is conferred via single mutations of the gene encoding for the respective enzyme, and consequently resulting in the substitutions in the respective amino acid chain. Isolates that have failed to respond to such drugs have resulted in the identification of the corresponding mutations within these genes that confer resistance (reviewed in Olliaro, 2001). Currently seven point mutations, occurring within the DHFR gene, have been identified as being linked with reduced drug binding capacity in resistant DHFR strains of *P. falciparum* (reviewed in Phillips, 2001; Sirawaraporn *et al*., 1997). In addition, DHFR and DHPS mutations seem to occur in a progressive stepwise fashion i.e. resistance slowly increases by the accumulation of mutations under drug pressure (Olliaro, 2001). Pyrimethamine resistance, for example, has been associated with a point mutation of Serine 108 to Asparagine 108. This mutation has been considered as the essential precursor to additional mutations that confer resistance and in due course lead to treatment failure (as reviewed in Olliaro, 2001). It has been proposed that by introducing bulkier amino acids at the site of action, especially the serine 108 position, steric hindrance results and as a consequence specific drugs, in particular those with a characteristically rigid structure such as pyrimethamine, are hampered from fitting within the active site (Olliaro, 2001). Rastelli *et al*. (2000) have demonstrated that this may indeed be the case using a homology model of the *P. falciparum* dihydrofolate reductase (pfDHFR) enzyme. The homology models of the wild-type and various mutant proteins associated with cycloguanil resistance were used to assess binding of pyrimethamine, cycloguanil, and WR99210 to the pfDHFR active site. Their data supports the steric constraint predictions previously mentioned and sheds light upon the basis of cycloguanil-resistant and pyrimethamine-sensitive strains. In addition the binding of WR99210, a drug structurally similar to cycloguanil, is not affected by steric mutations within the active site that confer cycloguanil resistance and therefore WR99210 is still effective against both resistant and pyrimethamine sensitive strains. This data may aid in the development of novel anti-malarial drugs.

1.3.1.1 Atovaquone

Atovaquone is also regarded as a hydroxynaphthoquinone and is utilised as a drug in both the prevention and treatment of malaria. It is vital to consider that it is only administered as a combination drug together with proguanil, an antifolate classed anti-malarial drug. Even though

it is known to target the mitochondrial functionality its mode of action, especially in combination with proguanil, is not fully elucidated. The general dogma underlies the ability of atovaquone to affect the mitochondrial electron transfer mechanism, although in recent studies and reviews it has been reported that its activity and synergy with proguanil has an impact on the mitochondrial membrane potential. These two processes are in fact related as the mitochondrial electron transfer chain serves to generate and maintain the membrane potential (reviewed within Olliaro, 2001).

Dihydroorotate dehydrogenase (DHODase), a critical enzyme within the mitochondrial electron transport chain, has been proposed as a potential site for atovaquones activity. This enzyme catalyzes the reaction where-upon dihydroorotate is oxidised to orotate. Additionally, it is considered the link between pyrimidine *de novo* synthesis and the mitochondrial electron transport system as it serves as the vital source of electrons. The consequent inhibition of DHODase results not only in the interference of the mitochondrial electron transport chain but additionally it blocks pyrimidine synthesis (Olliaro, 2001).

Fry and Pudney (1992) indicated that atovaquone indeed inhibits the respiratory mechanisms of the malarial mitochondria at the cytochrome bc_1 complex through its ability to mimic the natural substrate, ubiquinone. The result is the consequent depolarisation of the malarial mitochondria and the hampering of electron flow from ubiquinone to cytochrome c.

Despite the novel mode of action that atovaquone possesses, resistance is still apparent upon the utilization of atovaquone independently. For this reason it is no longer administered alone but in fixed-ratio combination therapies along with proguanil, as mentioned previously. The pharmacokinetics surrounding this drug may explain why monotherapy failed due to parasite resistance. Due to atovaquones slow uptake and high lipophilicity the parasite again is exposed to relatively low levels of the drug over extended periods of time, consequently leading to resistance. The mechanism of resistance was associated with single point mutations within the cytochrome b gene itself. This concurrently resulted in amino acid changes in the binding sites of ubiquinone (coenzyme Q). Three such mutations have been identified within the genome of *P.*

berghei, namely T431C, G399A and G850T, with the resulting amino acid mutations in the putative cytochrome b products at the residues 133, 144 and 284 (as reviewed in Olliaro, 2001).

1.3.2 The Quinolines

The first treatment to reach the Western world came with the discovery of the medicinal properties of quinine. Quinine, a quinolinemethanol, is an alkaloid type compound isolated from the bark of the cinchona tree, which is indigenous to specific areas of South America (Raynes, 1999). This anti-malarial is no longer used for prophylaxis or routine treatment as it is considered far too toxic however it still remains an option to treat severe cases of malaria (Foley and Tilley, 1997). The quinoline anti-malarials include the cinchona alkaloids, quinine and quinidine and the aminoalcohol quinine analogues commonly referred to as mefloquine and halofantrine (Phillips, 2001). In addition to these are the 8-aminoquinoline primaquine and the 4-aminoquinolines, chloroquine and amodiaquine (Phillips, 2001). These anti-malarials are categorised as blood schizontocides meaning that they act on the intra-erythrocytic (asexual and to an extent also sexual) parasites (Olliaro, 2001).

Despite the utilization of these anti-malarials for many years, as well as their widespread usage, fairly little is known regarding their mode of action. Some of the proposed mechanisms include inhibition of protein synthesis, inhibition of the food vacuole lipase, aspartic proteinase and inhibition of DNA and RNA synthesis. These mechanisms have been refuted, as they would require higher drug concentrations than those that can be achieved *in vivo* (Olliaro, 2001). Studies regarding the effects of quinoline anti-malarials upon *P. falciparum* indicated the swelling of the food vacuole (FV) or digestive organelle, vesiculation and the accumulation of undigested haemoglobin. From these reported morphological changes within the parasite, it was suggested that the action of these drugs may involve the inhibition of the functioning of the FV, specifically targeting haem disposal, a process whereby intra-erythrocytic-stage malaria parasites detoxify heamoglobin in the FV (as reviewed in Raynes, 1999; Olliaro, 2001). Most theories regarding this statement concur only on two points. Firstly, that chloroquine selectively affects stages of the parasite life cycle after haemoglobin is digested and secondly that chloroquine selectively accumulates within the FV (Olliaro, 2001).

1.3.2.1 Mechanism of Action of Chloroquine

Chloroquine, a diprotic weak base, is only effective against the blood stages of the parasite and only against those stages in which the parasite is actively degrading haemoglobin (Foley and Tilley, 1997). In its unprotonated form, choloroquine is able to pass through the membranes of the parasitised erythrocyte, and by means of the pH gradient, accumulates within the acidic food vacuole (pH5) (Foley and Tiley, 1997). At this point, chloroquine becomes protonated and is now membrane-impermeable and therefore trapped within the FV (Raynes, 1999). The FV of the parasite is responsible for the degradation of host haemoglobin, which in turn provides a valuable source of amino acids vital for growth (Raynes, 1999). From this observation it is thought that chloroquine affects the metabolic processes involved in the endocytosis and digestion of haemoglobin. Chloroquine is considered capable of inhibiting endocytosis through disrupting the recycling of receptors, which are potentially necessary for the uptake of lysosomal enzymes, leading to the inhibition of lysosomal function and haemoglobin degradation ultimately resulting in parasite starvation (Raynes, 1999).

Additionally, during this metabolic pathway there are several enzymes, which have been considered as potential drug targets. However, recent evidence supports the hypothesis that detoxification of haem is the most probable site of action (Foley and Tilley, 1997; Raynes, 1999). During catabolism of host haemoglobin, a toxic by-product namely ferriprotoporphyrin IX (FP) is produced (as reviewed in Raynes, 1999; Foley and Tilley, 1997). The plasmodia cannot degrade these potentially lytic and extremely toxic haem molecules and therefore polymerize the haem into harmless crystals referred to as haemozoin (also known as malarial pigment). Regardless of the nature of catalysis that occurs, chloroquine is thought to interrupt the polymerization process and as a result there is a toxic accumulation of haem molecules within the FV leading to the ultimate poisoning of the parasite (as reviewed in Foley and Tilley, 1997). A key point to keep in mind however is that it is still not clear whether this mechanism of chloroquine action can apply to the other forms of quinoline anti-malarials.

Resistance towards the quinoline anti-malarials is thought to result as the direct consequence of decreased drug accumulation within the FV and is probably conferred by multiple gene mutations

(Raynes, 1999). Theories surrounding the possible mechanisms of drug resistance generally fall into two main categories. The first being that there is an altered accumulation/exclusion mechanism ultimately resulting in the diminished drug levels within the digestive organelle, and secondly that transport mechanisms at the cytoplamsic and/or FV membrane have been affected directly impacting on the drug uptake/efflux processes (Raynes, 1999).

Initially the evolution of parasites, which have a considerably higher pH in their FV than chloroquine sensitive strains were analysed (Ginsburg and Stein, 1991). It was found that the chloroquine resistant isolates did indeed have a diminished or weakened proton pump on the vacuolar membrane, which resulted in the decrease in the concentration of the drug within the FV. From this initial point of discovery it was found that this condition of resistance could be 'reversed' *in vitro* utilizing drugs that reverse drug resistance in tumor cells (reviewed in Winstanley, 2000). Within tumor cells, p-glycoprotein pumps are encoded via multi-drug resistant genes and are responsible for the expulsion of drugs from the cell. Similar genes, namely *pfmdr1* and *pfmdr2*, were identified in *P. falciparum* however it proved unlikely that such genes were indeed responsible for conferring full resistance toward chloroquine. Alternatively, some studies revealed that a point mutation of the *pfmdr1* gene (Tyr 86) could be connected with chloroquine resistance of field isolates *in vitro.* Su *et al*. (1997) initiated a genetic cross experiment whereby neither gene was connected to resistance but alternatively the chloroquine phenotype was mapped to a 36-kb segment of chromosome 7. Additionally, fine-scale analysis of this segment resulted in the identification of multiple potential genes that (together with *pfmdr1*) might encode for proteins that play a role within chloroquine resistance.

1.3.3 Artemisinin-type Compounds

This class of anti-malarials include either the natural extract artemisinin itself or alternatively the semi-synthetic derivatives which include: dihydroartemisinin, artesunate, artemether and arteether (Olliaro, 2001). Artemisinin compounds are active throughout the phases of the asexual and intra-erythrocytic schizogonic cycle of the parasite as well as having the ability to affect the young gametocytes. In addition, these drugs are particularly favoured as they are capable of

reaching higher reduction rates than any of the other anti-malarials known to date, and are therefore utilized within the treatment of uncomplicated as well as severe malaria (Olliaro, 2001).

The mode of action of these anti-malarials is also not completely understood. The current hypothesis is that reductive cleavage of peroxide via ferroheme ferrous-protoporphyrin IX (Fe(II)PPIX) produces carbon-centered radicals, which ultimately would alkylate biomolecules and result in the death of the parasite (reviewed in Olliaro, 2001). This hypothesis is commonly referred to as the "carbon-centered radical theory" and several arguments regarding its validity exist. It is doubted whether the formation of the carbon centered radicals is indeed common in the reductive cleavage of endoperoxides at all, and even if this is the case it is still questionable whether this process can be linked appreciably with anti-malarial activity. In addition, there is the concern as to whether the putative adducts of this process are likely to survive long enough to even have any bearing on biomolecules (Haynes *et al.*, 1999; reviewed in Olliaro, 2001). Upon consideration of these arguments a further suggestion is that it is the 1,2,4-trioxane pharmacophore within artemisinin that confers anti-malarial activity, and in so doing provides the source of hydroperoxide through the generation of an oxo-stabilised cation upon heterolysis of the C_3-O_2 bond. From the reductive cleavage with exogenous iron (II) or other reducing agents the electrophilic oxygenating species/hydroxyl/alkoxyl radicals which are responsible for hydroxylating biomolecules or removing hydrogen atoms are thought to be produced (Olliaro *et al*., 2001). Additionally Ecktein-Ludwig *et al*. (2003) have demonstrated that artemisinin could also specifically and selectively inhibit the sarco/endoplasmic reticulum Ca^{2+} -ATPase (SERCA) of *P. falciparum*. Furthermore, their results confirm the Fe^{2+} dependent activation and consequent antimalarial activity of artemisinins being independent of haem binding.

An additional benefit for these anti-malarials so far is that there are no known cases of resistance, nor have there been any reported therapeutic failures as a result of decreased parasitic sensitivity (Olliaro, 2001).

1.4 Vaccines Against Malaria

Acquired immunity plays a vital role in malaria involving both the humural and cellular immune response (Phillips, 2001). The mechanism of naturally acquired immunity relies on the antibody response toward the parasitic proteins, which are expressed on the surfaces of infected red blood cells, merozoites and within the apical organelle of the merozoite (reviewed in Doolan and Hoffman, 2001). It was from the observation of acquired immunity that the initial idea of developing a vaccination against malaria originated.

The complex life cycle of the parasite and its intricate relationship with the vector and host as well as allelic diversity and antigenic variations makes the development and implementation of effective malaria control problematic (Phillips, 2001). However vaccination towards malaria, if successful, would be one of the most efficient and vital prophylactic methods against malarial infection. In comparison to several viral vaccines, malarial vaccines would have to be subunit vaccines due to the impracticalities of growing any stage of the malarial parasite life cycle on a mass scale to produce enough antigenic material to produce a whole-parasite vaccine (Phillips, 2001; de Roode and Read, 2003). Subunit vaccines are defined as those containing one or more pure or semi-pure antigens. In order to develop such a vaccine, it is vital to identify the individual components out of a myriad of proteins and glycoproteins of the pathogen that play a role in inducing protection. This allows for an additional difficulty when developing the malarial vaccine.

Effective vaccines elicit immune responses that destroy infectious agent or the host cells in which they reside, or inhibit a function that is vital for the parasites survival. The complex life cycle of the malarial parasite allows for several areas as potential targets for vaccination. There are currently three main categories in which malarial vaccine development lies. These include the pre-erythrocytic stage, the asexual stage, and the sexual as well as the transmission stages of the parasites life cycle (Doolan and Hoffman, 2001; Carter, 2001). The best candidates for malaria vaccines are those proteins required for parasitic survival, those with a low mutation rate and conserved epitopes. Due to these proteins playing essential roles within the multiple or alternative steps during the invasion process, they should be the targets against which a protective

immune response should be elicited (refer to Table 2; Phillips, 2001; Doolan and Hoffman, 2001). In order to increase the efficiency of any malarial vaccines under investigation, a combination of specific antigens, and/or epitopes from various stages in the life cycle are required and both a humoral as well as cellular immunity stimulation are required for optimal results.

1.4.1 Pre-erythrocytic Vaccines

Sporozoites are the first stage of the parasites life cycle that comes into contact with the human host. The pre-erythrocytic vaccine is engineered to either prevent sporozoites from invading the hepatocytes or alternatively to prevent liver-stage parasites from developing to maturity and subsequently producing and releasing infective merozoites, or both (Doolan and Hoffman, 1997). Ultimately, by targeting the sporozoites circulating in the blood stream, prior to their invasion into the hepatocytes, a preventative vaccine against malaria would be obtained (Phillips, 2001).

The anti-sporozoite vaccine, designed to prevent infection, was first effective upon using radiation-attenuated sporozoites (as reviewed in Philips, 2001). The immunity gained from such a vaccination was shown to be species specific and independent of parasite strain. However, it carried the severe disadvantage of requiring vast quantities of infected mosquitoes for each individual immunized (reviewed in Phillips, 2001). The main target in recent anti-sporozoite vaccine development has been revolving around the CSP or circumsporozoite protein.

CSP is the dominant protein located on the surface of the sporozoite and was shown to have substantial antigenic properties and was thus considered as a possible candidate for vaccine-induced immunity (Nussenweig and Nussenweig, 1989). Initial clinical studies were performed using both the recombinant protein as well as synthetic peptides based upon the repetitive immunodominant sequence NANP (Asn-Ala-Asn-Pro) within the CSP protein amino acid sequence (reviewed in Anders and Saul, 2000 and Phillips, 2001). Unfortunately, from the vaccination trials of both the synthetic and recombinant vaccine, it was found that only one out of three vaccinees were protective against exposure to *P. falciparum* sporozoites.

However, promising results have been reported recently utilizing a recombinant hepatitis B surface antigen vaccine, known as the RTS,S vaccine developed by GlaxoSmithKline, incorporating the repeat and non-repeat CSP sequences in adjuvant where-upon six out of seven vaccinees were protected in introductory trials (reviewed in Anders and Saul, 2000, Doolan and Hoffman, 1997 and de Roode and Read, 2003).

Table 2: Malaria vaccines and their antigenic targets

Stage of *P. falciparum* life cycle	Features	Antigenic Target and Vaccine Description	Evaluation
LIVER STAGE			
Pre-erythrocytic	Stage or species specific; antibody inhibits infection of liver; large dosage required; can terminate an infection	Irradiated sporozoites; Circumsporozoite protein (CSP) or peptides CS protein RTS,S (incorporating Hepatitus B antigen) CS protein (NANP) (synthetic sporozoite conjugated vaccine) PfCSP(DNA) (VCL-2510 DNA plasmid containing genes for *P. falciparum* CSP expression) Liver Stage Antigens –1 (LSA-1)	Protection Protection Short-term protection No protection Cytotoxic and IFN-γ-producing Tcells present
BLOOD STAGE			
Merozoite and Erythrocytes	Species as well as stage specific; cannot terminate an infection; inhibits invasion of erythrocytes, therefore reducing severity of infection	Erythrocyte Binding Antigen (EBA-175), Merozoite Surface Protein MSP19kd (19kDa C-terminal fragment of MSP1) Merozoite Surface Antigen 1&2 (MSA-1&2); Ring Infected Erythrocytic Surface Antigen (RESA); Serine Repeat Antigen (SERA); Rhoptry Associated Protein (RAP); Histidine Rich Protein (HRP); Apical Membrane Antigen-1(APM-1)	Atibody and Tcell production
SEXUAL STAGE			
Gametocytes & gametes	Prevents infection of mosquitoes; antibody to this and antigen inhibits fertilization and maturation of gametocytes, zygotes or ookinetes; available in endemic areas but not suited for travelers; antibody blocks transmission cycle	Pfs 25, 48/45k, Pfs 230	
MULTI-STAGE			
Combined vaccine (cocktail)	Incorporation of expressed antigens from various life cycle stages into one vaccine to produce an immune response, blocking all stages of the parasite development.	SPf 66 (incorporates pre-erythrocytic and asexual blood stage proteins of *P. falciparum* conjugated to alum)	No protection
	Multi-stage vaccine with multiantigenic variants, highly attenuated vaccinia virus	NYVAC-Pf7	Tcells and antibodies

(adapted form Doolan and Hoffman, 2001; Plebanski *et al*., 2002)

1.4.2 Asexual Blood-Stage Vaccines

Animal models for crude-blood stage vaccines have successfully resulted in a strong induced resistance against the malarial parasite. Research now lies in the identification of those antigens, present within the crude extracts, which are responsible for the induction of protection (Phillips, 2001). From monoclonal antibody studies against the rodent and human malarial parasites several antigenic molecules have been identified (Holder and Freeman, 1981). The two most promising antigenic candidates are the merozoite surface protein 1 (MSP-1) and apical membrane antigen 1 (AMA-1) (as reviewed in Phillips, 2001 and Anders and Saul, 2000). Both of these antigens are present within all four of the malarial species and due to their homologues in both the rodent and simian parasites the vaccine potentials of these two candidates has been tested utilising numerous animal models (Anders and Saul, 2000).

The MSP-1 protein is the major surface protein of the merozoite, with its 19 kDa carboxy terminal portion remaining on the surface of the merozoite even after infection of the hosts red blood cells has occurred (Holder and Freeman, 1981). As discussed in Phillips (2001) the 19-kDa fragment of the protein is conserved amongst all MSP-1 proteins and consists of two cysteine-rich epidermal growth factor domains (EGFD), which are possibly essential for the parasite to invade the erythrocytes of the host. The first EGF domain has been shown to be the target of monoclonal antibodies directed against *P. falciparum* proteins, which are responsible for hindering the development of the parasite, *in vitro* (Chapel and Holder, 1993). From immunization studies the recombinant proteins from the MSP-1 of *P. yoelii* and *P. falciparum* were shown to be effective due to the induction of protective antibody (Phillips, 2001).

The AMA-1 antigen is located in the rhoptry organelles, which are responsible for the invasion of the erythrocytes. Other antigens, which are currently being studied as potential vaccine candidates, involved in this organelles task include the erythrocyte-binding antigen (EBA-175) and the serine-rich antigen (SERA) (as reviewed within Phillips, 2001). Immunization studies involving the recombinant forms of AMA-1 have been successful against both *P. fragile* and *P. chabaudi* with current studies involving the AMA-1 of *P. falciparum* (Phillips, 2001).

1.4.3 Transmission – Blocking Vaccines

Several efforts have been attempted in the identification of existing, as well as potential, antigenic targets for the development of a Transmission-blocking vaccine (TBV). Currently, these studies have led to the identification of proteins such as chitinase, which occurs on the surface of the malarial ookinete and are vital for the penetration of the mosquito mid-gut, as well as other surface proteins of the extracellular male and female gametes, zygote and ookinete as potential antigenic targets (as reviewed in Carter, 2001).

TBV`s are targeted against the sexual stages of the life cycle of the parasite, including the stages within the mosquito mid-gut, and essentially prevent the parasite from infecting the mosquito vector. Due to the vaccines role in reducing or abolishing the transmission of the disease they were named transmission blocking vaccines or TBV`s (Carter, 2001). There are three main categories in which transmission blocking vaccines currently lie. The first involves the pre-fertilization antigens - which are expressed either entirely or predominantly in the gametocytes, the second involves the post-fertilization antigens which are expressed completely or predominantly on the zygotes or ookinetes, and the third category includes the late-mid-gut antigens which are essential for the ookinete to penetrate the peritrophic membrane (Phillips, 2001).

Numerous efforts in development of pre-fertilization target antigen identification resulted in the production of monoclonal antibodies namely: Pf230, Pf48/45, and Pf11.1. These antibodies have shown considerable transmission-blocking activity *in vitro*, however validation of whether these antibodies occur naturally has not yet been manageable (reviewed within Phillips, 2001).

Post-fertilization antigens are antigens that are not expressed by the parasite in the host and thus they have the undesirability of serving as potential vaccine targets due to the antibody response failing to be stimulated after natural infection. However, these antigens are still considered possible targets due to the lack of exposure they have had toward the immune response and thus their lack of exposure toward selection of immune evasion mechanisms (Phillips, 2001). Currently two potential post-fertilization antigenic targets are being investigated, namely P25 (or

Pfs25 in *P. falciparum*) and P28 (or Pfs28 in *P. falciparum*) (Barr *et al.*, 1991; Duffy *et al.*, 1993). Pfs25 was the first sexual-stage antigen to be cloned and is minimally expressed in maturing gametocytes, and mainly expressed after the dispersal of the gametes and formation of the zygote. It is the main surface protein on the zygote and ookinete, whereupon it is always expressed (Barr *et al*., 1991). Pfs25 has been shown to have antigenic diversity and the recombinant form has been expressed within mammalian expression vectors, as well as yeast expression systems, and has been reported to have transmission-blocking activity and induce the respective antibodies in both mice and monkeys (Barr *et al.*, 1991; as reviewed in Phillips, 2001). From this initial success, Pfs25 is currently in Phase I trials. Pfs28, in comparison, has only been shown to have transmission-blocking activity in *P. gallinaceum* (Duffy *et al.*, 1993).

Shortcomings surrounding vaccine development against malaria begin with the inability to grow the parasite in large quantities and range from difficulties of evaluation to the fact that there are multiple antigens, which are species as well as stage specific. Further problems include antigenic diversity and immune evasion. This is especially the case regarding the blood-stage vaccines as they are polymorphic and the antigenic diversity resulting form these polymorphisms is thought to limit the effectiveness of the malarial vaccine. Additionally there are multiple problems within conducting clinical and field trials. Several researchers have hypothesized that a vaccine for a specific spectrum of antigens is unlikely to be protective due to the variation in the target for the immune response at different stages of the parasitic life cycle. An additional aspect to consider is that immunizing against a specific group of antigens may result in the parasite changing its "identity" due to the chances of mutation amongst the parasite itself also being a possibility.

1.5 Genomics, Proteomics and Malaria

The release of the genome sequence of the anopheles mosquito has allowed for the development of new tools for the control of malaria with regard to vector-related strategies. In addition, several organizations collaborated their efforts in order to sequence the *P. falciparum* genome, some of which included The Institute for Genomic Research and the Naval Medical Research Center (TIGR/NMRC), the Sanger Institute and Stanford University. The whole-chromosome shotgun strategy was chosen for the task of sequencing the 22.8-Mb genome, which is comprised

of 14 chromosomes, encoding 5300 genes, and the mitochondrial and apicoplast genome sequences (Gardner *et al.*, 2002). The publication of the *P. falciparum* genome sequence in *Nature* during October 2002, has allowed for new hope in the battle against malaria. The identification of potential leads for anti-malarial drugs and targets for vaccine development have already been assisted through the availability of genome sequence data in the past (Roberts *et al.*, 1998; Surolia and Surolia, 2001; and Jomaa *et al*., 1999). Now, with the genome sequence in hand researchers are looking for novel vaccine and drug targets amongst the 5 279 predicted proteins however it will still be several years before any success is envisaged. In the case of vaccine development this is not due to a lack of vaccine candidates but due to their effectiveness to stimulate an immune response as even though irradiated parasites can illicit a protective immune response several proteins on their own cannot (reviewed in de Roode and Read, 2003).

1.5.1 Target and Lead Identification for drugs based on genome sequence data

Through homology modeling strategies between *P. falciparum* chromosome 2 sequence data, and known genes of plants and algae, researchers have already identified several lipid biosynthesis genes within *P. falciparum* and have demonstrated the characteristic targeting toward the apicolast of these gene products (Waller *et al*., 1998). The proteins of these genes are critical drug targets as they are vital for the survival and viability of the parasite and also due to the absence of the type II fatty acid biosynthesis pathway in humans. An enzyme of this pathway, namely enoyl-acyl-carrier protein (enoyl-ACP) reductase (FabI), was identified from sequence data and subsequent inhibition of this enzyme was shown to have an *in vitro* affect against *P. falciparum* activity (Surolia and Surolia, 2001).

The additional identification of the enzymes 1-deoxy-D-xylulose 5-phophate (DOXP) synthase and reductoisomerase, from the sequence data of chromosome 14, has supplied the evidence that can now link *P. falciparum* with the alternative DOXP pathway for isoprenoid biosynthesis, which is essential for parasitic survival (Jomaa *et al*., 1999). Furthermore, there was no indication within the genome of the mevalonate pathway for isoprenoid synthesis common within eukaryotes thus making this pathway an ideal target for anti-malarial drugs (Gardner *et al*., 2002). The two DOXP reductoisomerase inhibitory drugs, fosmidomycin and its derivative FR-900089,

have been demonstrated as being successful against *P. falciparum*, as well as multi-drug resistant *P. falciparum*, activity *in vitro* and *in vivo* against *P. vinckei* in mice (Jomaa *et al*., 1999).

More recently however, researchers have managed to identify a metal-dependent RNA triphosphate protein family as a novel target for drug design within the parasite *P. falciparum* (Ho and Shuman, 2001). This family of proteins are vital for mRNA cap formation and therefore eukaryotic gene expression. The structure of the active site and catalytic mechanisms of these proteins are similar amongst fungi and *P. falciparum* however differ from the RNA triphosphate domain found in metazoan capping enzymes. For this reason genomic and structural similarities between these enzymes amongst fungi and *P. falciparum* offer the opportunity to generate antifungal as well as anti-malarial activity amongst a single class of mechanism-based inhibitors (Ho and Shuman, 2001; as reviewed in Hoffman *et al*., 2002). Additional protein families within *P. falciparum*, which are considered as potential targets for drug design, include the cysteine and aspartyl proteases, with special interest focusing on the metacaspase of the family of cysteine proteases.

1.5.2 Target identification for vaccines based on genome sequence data

Potential vaccine targets are those that are often present on the surface of the infected erythrocyte or merozoites and parasitic proteins essential for erythrocytic invasion. Several such protein targets have been identified from genomic sequence data and their potential as vaccine targets are being investigated (Duffy *et al*., 2001; reviewed in Hoffman, 2002; Gardner *et al*., 2002). The availability of genomic sequence data has not only allowed the identification of such targets but also revealed clues as to how these genes are regulated. One such example is with the *P. falciparum* erythrocytic membrane protein (PfEMP) gene family (Deitsch, *et al*., 2001). Many proteins, which can be considered as potential vaccine targets, are those responsible for initiating infection within the host and stimulating the immune response. However, for there to be any success in the incorporation of genome sequence data into the development of a malaria vaccine, development of high-throughput immunological assays as well as techniques for ensuring the stability of a protective immune response is needed (Gardner *et al*., 2002).

1.6 Novel Approaches and Strategies in Combating Malaria

1.6.1 Targeting Metabolic Pathways of the Parasite

Several novel strategies for targeting malaria have involved the possibilities of inhibiting the metabolic pathways of the parasite. Utilizing a combination of inhibitors, which target metabolic pathways that converge to produce a final product, has been proposed as being far more effective and beneficial in the struggle against malaria than merely the utilization of a single inhibitor (Macreadie *et al.*, 2000). These studies are hopeful in deterring the evolution of drug resistance however, such studies are promising but still at preliminary stages and examples of such approaches are limited.

The discovery of the apicoplast or apicoplexan plastid within the malarial parasite is an exciting prospect for novel drug development against malaria. The apicoplast contains a circular, 35 kb, extrachromosomal strand of DNA that comprises genes of only 30 proteins. Gardner *et al*. (2002) also identified 551 nuclear-encoded proteins that are destined to the apicoplast, and together with the apicoplast encoded proteins may be promising drug target candidates (Vial, 2000; Gardner *et al*., 2002). This non-photosynthetic plastid is protected by four membranes and possesses a genome that is homologous amongst algae and plants. Due to the evolutionary status of this organelle in all parasites it is essential for the survival of the parasite. Exactly how the organelle ensures the survival of the parasite however, remains a mystery (Vial, 2000). The apicoplast is known to house several metabolic pathways namely those of fatty acid (Gardner *et al*., 1998b; Waller *et al*., 1998), isoprenoid (Jomaa *et* al., 1999) and haem synthesis (Sato and Wilson, 2002). Jomaa *et al.*, (1999) have identified the mevalonate-independent pathway of isoprenoid biosynthesis in *P. falciparum* as a potential source for drug inhibition. This pathway is thought to occur in the apicoplast and operates via a different mechanism to that in humans. Through the identification of the enzymes within this pathway a new target for anti-malarial chemotherapy has originated (Jomaa *et al.*, 1999). The alternative pathway for isoprenoid synthesis will also allow for immunotherapeautic drug development, as many pathogens rely on the alternative pathway of isoprenoid synthesis to produce toxic compounds that interfere with the hosts immune system (www.jomaa-pharmaka.com; dates accessed 01/2002-12/2003). By

utilizing the inhibitor compounds therapeutically, diverse immune reactions common in disorders such as multiple sclerosis and several allergies may be corrected and pathology prevented (www.jomaa-pharmaka.com).

The shikimate pathway (also known as the aromatic biosynthesis pathway), responsible for the synthesis of aromatic compounds comprises a sequence of seven enzymatic steps in which the starting materials (generated via carbohydrate metabolism) phosphoenolpyruvate and erythrose-4-phosphate are converted to chorismate. This pathway exists in prokaryotes, fungi and the plastids of plants but is absent in mammals, which rely solely upon exogenous folates (Hermann and Weaver, 1999). In plants the shikimate pathway enzymes are targeted to the plastid via an amino-terminal leader sequence (Waller *et al*., 1998). However, the apicomplexan pathway is thought to occur within the cytosol due to the lack of the leader sequence within any of the shikimate pathway proteins (Keeling *et al*., 1999). Apart from this, the shikimate pathway is still a valid target for novel anti-parasitic drugs especially due to the absence of the pathway within mammals. Roberts *et al*. (1998) demonstrated the presence of this pathway within *P. falciparum* together with other parasites and in addition have shown its inhibition by the herbicide glyphosate, a known shikimate pathway inhibitor. Glyophosate inhibits the enzyme 5-enolpyruvyl shikimate 3-phosphate synthase and administering p-aminobenzoate can reverse its inhibitory affects. The discovery of the apicomplexan shikimate pathway is therefore a promising metabolic pathway for drug targeting and harbours possibilities for a broad-spectrum drug in the future (Roberts *et al*., 1998).

1.6.2 Rational Drug Design

Another leading field that has not yet been fully explored is that of rational drug design supported by the existing structural information of malarial protein targets (Macreadie *et al.*, 2000). Great hopes have been pinned upon the computational methods for structure based ligand design and some success has been reported within other areas of research, especially in cases regarding water soluble targets, with 3-D structures having been accurately determined via the technique of X-ray crystallography and NMR studies. The main requirement of these methods is knowledge pertaining to the 3D topology of the targets ligand-binding pocket, with the ligand docked in

place (Farber, 1999). The identification of lead compounds from library compounds has also been explored. Wiesner *et al*. (2001) have shown the identification of a novel lead compound, namely 2,5-bis-acylaminobenzophenone, from a library of 61 compounds through inhibitory assays against a resistant strain of malaria.

In order for rational drug design to be a success, over-expression of the gene of interest is essential to provide sufficient levels/quantities of recombinant protein for biophysical studies; these include techniques such as X-Ray crystallography and NMR. The over-expression of parasitic genes, especially *P. falciparum*, in *Escherichia coli* (*E. coli*) has been difficult due to the codon-bias of these organisms. The *P. falciparum* genome in particular is extremely AT-rich which has added to the problem of heterologous protein expression and production within bacteria (Saul and Battistutta, 1988). With the construction of the RIG plasmid researchers are now able to overcome the potential problems of codon-bias apparent in the *P. falciparum* and other parasitic genes (Baca and Hol, 2000). This plasmid is engineered to encode for the three tRNA`s, namely Arg, Ile and Gly, which recognize codons that are found frequently within parasitic genes but are rare within highly expressed genes in *E. coli*.

Another interesting field is that of serial analysis of gene expression (SAGE) which in tandem with the malaria genome project will allow for the identification of novel drug targets. The technology available will allow one to identify genes, or gene clusters, which have characteristically high expression levels under normal conditions, however, when treated with a particular drug have affected expression levels (i.e. considerably lower). Additionally this technology will further allow for researchers to monitor the effects of any particular drug on the parasite (Macreadie *et al.*, 2000).

Inherent to all approaches in combating malaria, discussed in the sections above, is the ultimate need to achieve a reasonable level of progress for malaria control. Intensive measures have been made in malaria control in the dispensation and administration of several chemotherapeutic therapies, as well as physical and chemical strategies to control the vector i.e. the mosquito. Despite these efforts the disease is still on the increase, with *P. falciparum* now ranking amongst the top four infectious disease killers worldwide (de Koning-Ward *et al.*, 2000).

The parasites complex relationship with the vector, as well as the host, poses a challenge towards medical research in the field of novel drug development. However, through the utilization of genetic tools, such as transfection, a deeper understanding of the parasites biology will hopefully serve as a catalyst for a more rational approach toward both vaccine and anti-malarial drug development (de Koning-Ward *et al.*, 2000). Advances in proteomics will additionally aid the battle against malaria, allowing for the identification of differentially expressed proteins, which along with comparative genomic analysis and the utilization of protein interaction maps and the knowledge of the mechanisms of metabolic pathways, will aid the development of therapeutic intervention.

1.7 The Mevalonate/Non-Mevalonate Pathways for Isoprenoid Biosynthesis

Isoprenoids are a heterogenous collection of biomolecules that are paramount for the survival of all organisms, with representatives in all taxonomic groups. Of the 23 000 known isoprenoids some common examples include cholesterol, steroid hormones, and ubiquinones in animals, carotenoids, plastoquinones, and the phytyl side chains of chlorophyll in plants, and hopanoids and undecaprenylpyrophosphate in bacteria. Isoprenoids are involved in several cellular processes such as electron transport within respiration and photosynthesis, hormone-based signaling, control of transcription, post-translational processes in lipid biosynthesis, meiosis, apoptosis, protein cleavage and degradation.

In all organisms isoprenoid biosynthesis relies on the condensation of varying numbers of isopentenyl diphosphate units (IPP). Initially it was accepted that all organisms produced IPP via the mevalonate pathway. Within this pathway IPP is synthesized from three molecules of acetyl-CoA with the intermediate being mevalonate (MVA). However, it was discovered after 1993 that several bacteria, green algae and the chloroplasts of higher plants generate IPP via the non-mevalonate or 1-deoxy-D-xylulose 5-phosphate/2-*C*-methyl-D-erythritol 4-phosphate (DOXP/MEP) pathway (reviewed in Lichtenhaler, 1999 and Rohmer, 1999).

This DOXP/MEP pathway or non-mevalonate isoprenoid pathway has also been shown to exist in the malaria parasite as well as several other potentially pathogenic bacteria however, it is absent in humans and other mammals (Jomaa *et al.*, 1999; Lichtenhaler *et al.*, 2000). The discovery that the alternative pathway is also present in the malaria parasite has allowed for the discovery of a completely novel therapeutic concept for malaria treatment. For this reason it promises the development of novel drug candidates, which would have no adverse affects upon humans and other mammals. Additionally, the lack of the mevalonate pathway within the parasite makes it impossible for this pathway to rescue any inhibition of the alternative pathway.

The discovery of the apicoplast in the phylum apicomplexa, and its indispensability within the *Toxoplasma* and *Plasmodium* parasites, resulted in much curiosity regarding the origin and metabolic importance of this plastidic organelle (Waller *et al.*, 1998). In comparison to other organisms the production of IPP in plants was found to occur via both pathways, with the mevalonate pathway residing in the cytosol and the alternative pathway in the plastid. For this reason, it has been thought that the plastids developed evolutionary from prokaryotes due to a secondary endosymbiotic event (Palmer and Delwiche, 1996). Additionally, two genes of the alternative non-mevalonate pathway of isoprenoid biosynthesis have been identified in the apicoplast of *P. falciparum*, thus directly linking the DOXP/MEP pathway with this organelle (Jomaa *et al*., 1999).

Evidence for the existence of an alternative isoprenoid biosynthesis pathway emerged from the independent incorporation studies in the research groups of Rohmer and Arigoni who found that the isotopic labeling patterns observed within their studies could not be explained via the mevalonate pathway (reviewed within Eisenreich *et al.*, 1998; Lichtenhaler, 1999 and Rohmer, 1999). Not much information regarding the reaction mechanisms of this novel pathway has yet been confirmed, especially regarding the last steps leading to release of IPP. However, it is currently expected that the following outline of the reactions steps is common amongst all species harbouring the DOXP/MEP pathway.

The initial step of this novel pathway involves the condensation of pyruvate and glyceraldehyde-3-phosphate. 1-deoxy-D-xyulose-5-phosphate (DOXP) synthase catalyses this initial step, which

results in the production of 1-deoxy-D-xyulose-5-phosphate (refer to Figure 1.3; Part A). DOXP reductoisomerase then synthesizes 2-C-methyl-D-erythritol 4-phosphate (MEP) in an intramolecular rearrangement step followed by a reduction process using NADPH as a hydrogen donor (Vial, 2000; Proteau *et al.*, 1999; Reuter *et al.*, 2002)(Figure 1.3 Part B). Divalent cations such as Mn^{2+}, Co^{2+} and Mg^{2+} are also required as co-factors for this reaction, with Mn^{2+} being the most effective (Reuter *et al.*, 2002). From here, MEP is cytidylated via 4-(cytidine 5´-diphospho)-2-*C*-methyl-D-erythritol (CDP-ME) synthetase with the release of pyrophosphate and subsequently phosphorylated by the CDP-ME kinase to result in the production of 4-diphosphocytidyl-2-C-methyl-D-erythritol-2-phosphate (CDP-ME-2-phosphate)(Figure 1.3; Part C and D). CDP-ME-2-phosphate is then cyclised by the enzyme MECDP synthase resulting in the release of CMP and the production of 2-C-methyl-D-erythritol-2,4-cyclodiphosphate (MECDP)(Figure 1.3; Part D). The terminal reaction steps are, however, still not fully elucidated (reviewed within Rohmer, 1999; Reuter *et al.*, 2002)(Figure 1.3; Part E).

Additionally, this pathway has been reported as synthesizing other intermediates other than those of IPP in bacteria, for example thiamine (vitamin B1) and pyridoxol (vitamin B6) (Takahashi *et al.*, 1998). Importantly, the novel pathway for IPP biosynthesis is presently not fully understood especially regarding the terminal reaction steps, which lead to the production of IPP itself.

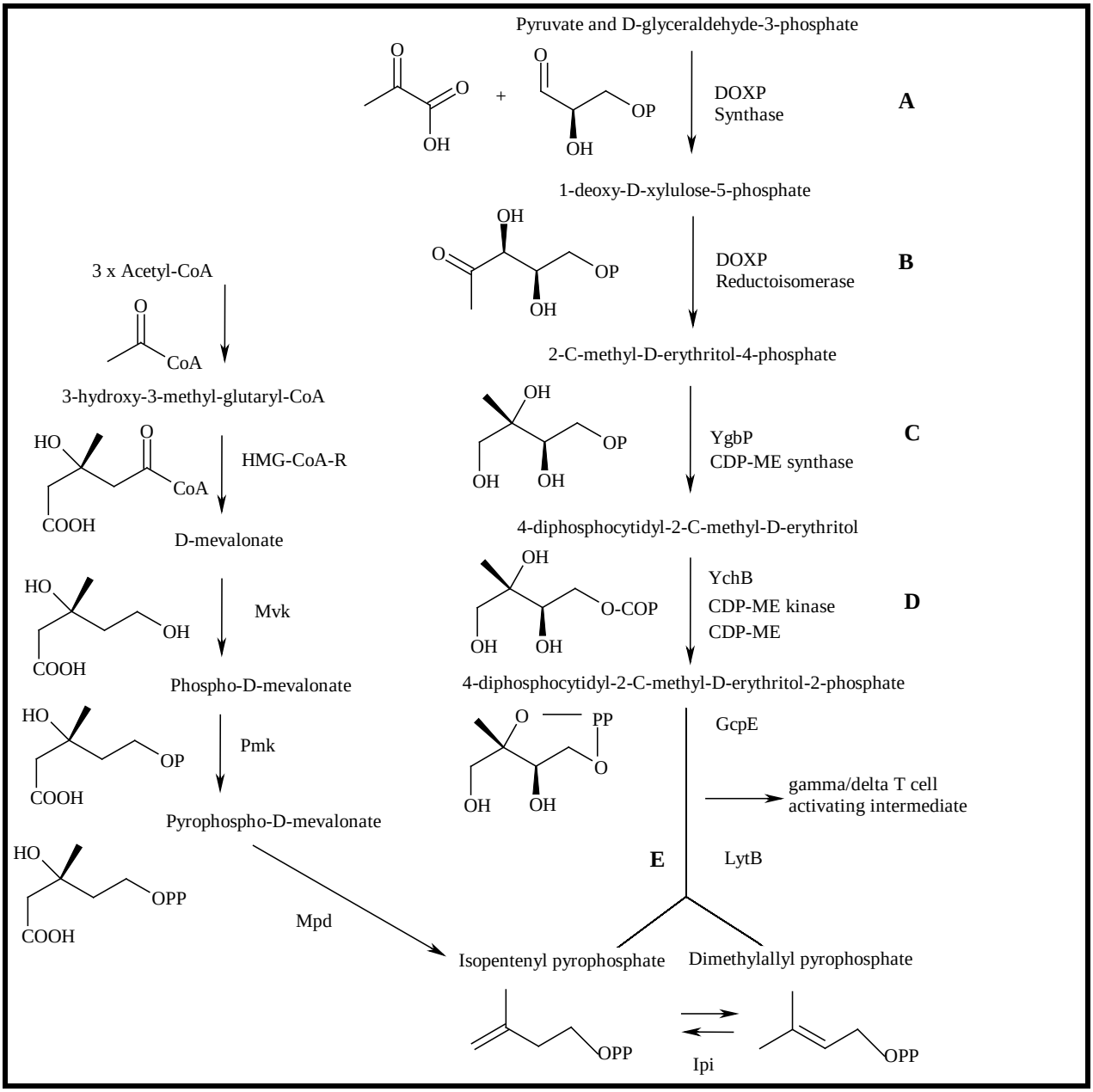


Figure 1.3: The mevalonate & alternative pathway for the biosynthesis of isoprenoid units

The mevalonate (left) and the alternative (right) pathway of isoprenoid synthesis. The terminal steps are still not fully elucidated (adapted from www.jomaa-pharmaka.com).

1.8 1-Deoxy-D-xylulose-5-phosphate Reductoisomerase

Enzymes within the alternative pathway of isoprenoid synthesis have become promising targets for novel drug development against several bacteria and pathogens including malaria. This lies within the importance of the production of isoprenoids to ensure survival of the respective pathogen. The enzyme, which is responsible for the second step of the alternative pathway, namely 1-deoxy-D-xyulose-5-phosphate reductoisomerase (DOXP reductoisomerase or DXP reductoisomerase) has attracted much interest within this field as a potential drug target. Additionally, both the DOXP synthase and DOXP reductoisomerase genes have been cloned, and the expression of these genes in the erythrocytic stages of *P. falciparum* has been confirmed by Jomaa *et al*, 1999 utilizing reverse transcriptase- polymerase chain reaction (Jomaa *et al.*, 1999; Reuter *et al.*, 2002; Sprenger *et al.*, 1997; Takahashi *et al.*, 1998). The manipulation of the non-mevalonate pathway that exists in *E. coli* has shown that the over-expression of DOXP synthase can be correlated with the increase in levels of carotenoid and ubiquinone synthesis (Harker and Bramley, 1999). Additional studies regarding the expression levels of these enzymes, was demonstrated within *Arabidopsis*, however the roles of these enzymes and the reductase isoforms in regulating the non-mevalonate pathway are still unclear (Estèvez *et al*., 2000). Recently, DOXP reductoisomerase proteins were also cloned from *Mentha x piperita* (Lange and Croteau, 1999), *Arabidopsis thaliana* (Schwender *et al*., 1999), *Synechocytis sp.* (Proteau *et al*., 1999), *Synechococcus leopoliensis* (Miller *et al*., 2000), *Zymomonas mobilis* (Charon *et al*., 1999, Grolle *et al*., 2000) and *Pseudomonas aeruginosa* (Altincicek *et al*., 2000).

Reuter *et al.* (2002) were the first group to have determined the crystal structure of the *E. coli* DOXP reductoisomerase to within 2.5 Å. DOXP reductoisomerase was previously reported as being a homotetramer via gel filtration and native polyacrylamide gel electrophoresis (PAGE). However, from the crystal structure the enzyme was shown to be present as a homodimer with each monomer forming a V-shaped molecule, where DOXP reductoisomerase was concluded as having a molecular mass of 86 kDa (each monomer being 43.5 kDa) (Reuter *et al.*, 2002).

The three dimensional structure of the enzyme was shown to consist of three domains, which ultimately form the overall V-shape. Thus each monomer is comprised of an amino-terminal

dinucleotide binding domain, a connective domain, and a carboxy-terminal domain. The amino-terminal domain stretches from residues 1 to 150 and forms one arm of the V. It represents the dinucleotide binding domain responsible for binding NADP(H). The connective/catalytic domain (residues 150-285) is accountable for the dimerization of the enzyme as well as containing the majority of the active site of the enzyme. In addition, the DOXP reductoisomerase enzyme has been shown to have high intrinsic flexibility within the connective domain. The carboxy-terminal domain spans residues 312 to 398 and is present as a four-helix bundle and consequently forms the second arm of the V (Reuter *et al.*, 2002).

The apical region of the DOXP reductoisomerase enzyme is mostly comprised of the connective α/β domain. This region forms a deep cleft in which the substrate binding and catalytic site are situated. Within the connective domain several highly conserved acidic residues are clustered at the active site and are thought to be responsible for the binding of the divalent cations essential for enzyme activity. These residues include Asp^{150}, Glu^{152}, Glu^{231}, and Glu^{234} (Reuter *et al.*, 2002).

In addition, Yajima *et al*. (2002) have obtained the crystal structure of *E. coli* DOXP reductoisomerase in complex with the co-factors NADPH and a sulphate ion from *E. coli* at 2.2 Å. The structure obtained was also reported as being in a homodimer conformation with the monomer consisting of the three domains reported by Reuter *et al.* (2002). However, they initially reported and identified the third domain present within DOXP reductoisomerase that seems to be absent amongst other NADPH-dependent oxidoreductases. This flexible loop region covers the substrate-binding domain and is therefore thought to be critical in the enzymatic reaction and the determination of substrate specificity. Additionally, they propose that this loop may act as a "hatch" that seals the active site once the substrate has entered the catalytic pocket. From the analysis of the crystal structure and subsequent mutational analysis of DOXP reductoisomerase, Yajima *et al*. (2002) demonstrated that the four amino acid residues His^{153}, His^{209}, Glu^{231}, and His^{253}, present within the central domain with their side chains protruding into the binding pocket, are essential for enzyme catalysis.

Steinbacher *et al*. (2003) have released the most recent crystal structure, which is in complex with NADPH and the metal ion cofactor as well as the inhibitor fosmidomycin (see Figure 1.4). They report the protein structure as a homodimer both within the crystal and in the solution state with the presence of a groove within each monomer that is protected via a flexible hatch region that was previously reported by Yajima *et al*. (2002). The monomer was reported as containing the 3 domains reported previously by Reuter *et al*. (2002) and Yajima *et al*. (2002). Additionally, they describe the presence of the binding sites for the divalent cation, the phosphate moiety of the substrate DOXP, and the catalytic loop within the connective/catalytic domain (residues 150-285) of the protein structure with the carboxy terminal maintaining the structural integrity of the active site pocket. Binding of the metal cation was found to be in an octahedral co-ordination sphere with three water molecules and the residues Asp^{150}, Glu^{152} and Glu^{231}.

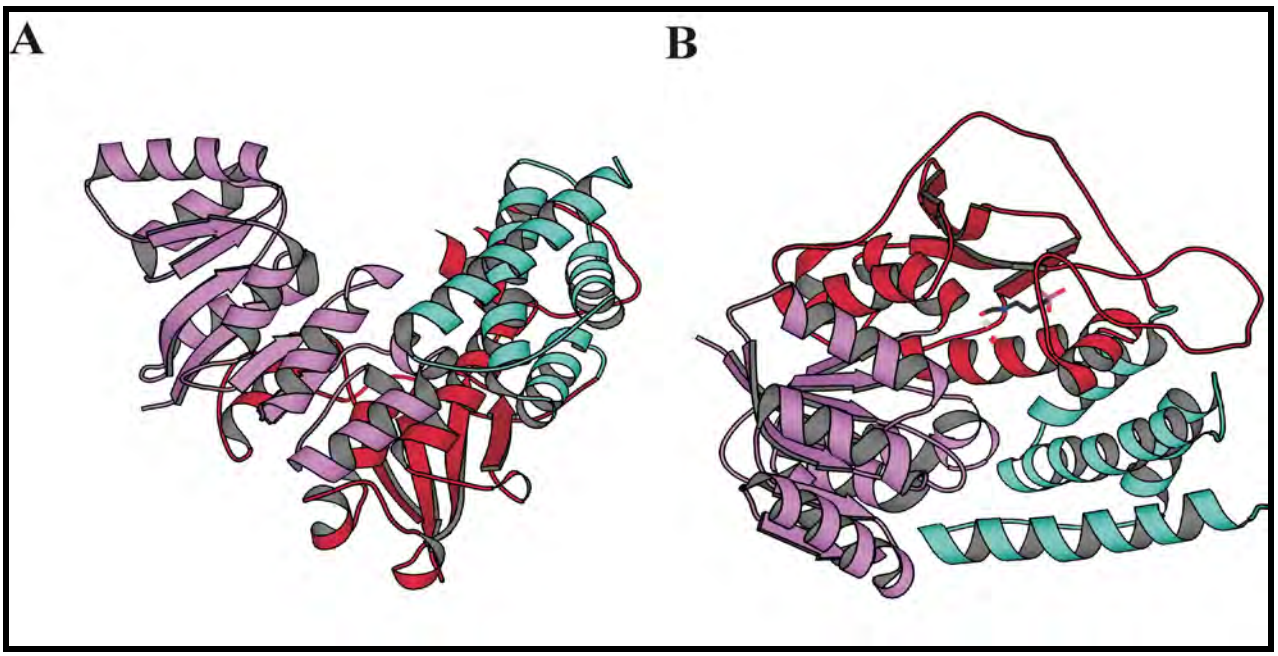


Figure 1.4: Ribbon representation of the DOXP reductoisomerase monomer in complex with the inhibitor fosmidomycin

This figure was generated using Molscript (Kraulis, 1991) using the crystal structure data generated via Steinbacher *et al.* (2003) (Protein Data Bank accession codes 1ONN and 1ONP). (**A**) The enzyme comprises three domains, which form the overall V-shape of the molecule. The N-terminal NADPH binding domain is shown in violet. The central connective/catalytic domain shown in crimson binds the divalent cation (magnesium, manganese or cobalt) and the phosphate moiety of the substrate. The carboxy terminal provides structural integrity for the active site pocket and is shown in turqouise. (**B**) Fosmidomycin is shown bound to the active site in ball and stick configuration.

1.9 Stereochemistry of the Reduction Step Mediated via 1-Deoxy-D-xylulose 5-Phosphate Reductoisomerase

DOXP reductoisomerase is the second enzyme of the methylerythritol (MEP) pathway and is involved in the conversion of DOXP to MEP by catalyzing the rearrangement of the carbon backbone as well as the nucleotide adenosine diphosphate hydrogen (NADPH)-dependent reduction of the carbon 1 position of the putative intermediate 2-*C*-methylerythrose 4-phosphate (Proteau *et al*., 1999). Proteau *et al*., (1999) originally elucidated the stereochemistry of this reductive step of the MEP pathway (see Figure 1.5). Through the utilization of the *Synechocystis* sp. protein, which has a 42% identity to the *E. coli* DOXP reductoisomerase sequence, they addressed the following stereochemical issues surrounding the reaction this enzyme catalyses: "1) which hydrogen at the carbon 1 position of the 2-*C*-methyl-erythritol phosphate originates from the carbon 3 of DOXP and which from NADPH? 2) delivery of the hydride from NADPH occurs on which face of the intermediate aldehyde – *re* or *si*? and 3) which hydride from NADPH is delivered in the reduction (class A or class B dehydrogenase)".

Utilizing [3-^{2}H]DOXP the H' NMR spectrum indicated that the *pro-S* proton/hydrogen of MEP resonates at 4.156 ppm and is derived from the carbon 3 position of deoxyxylulose whereas the *pro-R* proton/hydrogen at the carbon 1 position of MEP, is derived from NADPH and resonates at 3.90 ppm. In addition this result further established that the hydride delivery occurred at the *re* face of the proposed intermediate (Proteau *et al*., 1999). Additionally, from incubation experiment studies involving [1-^{2}H] glucose in cyanobacterium *Synechocystis* sp. Proteau *et al*., (1999) have proposed that the 4 S-hydride of NADPH is delivered, this therefore classifies DOXP reductoisomerase as a class B dehydrogenase.

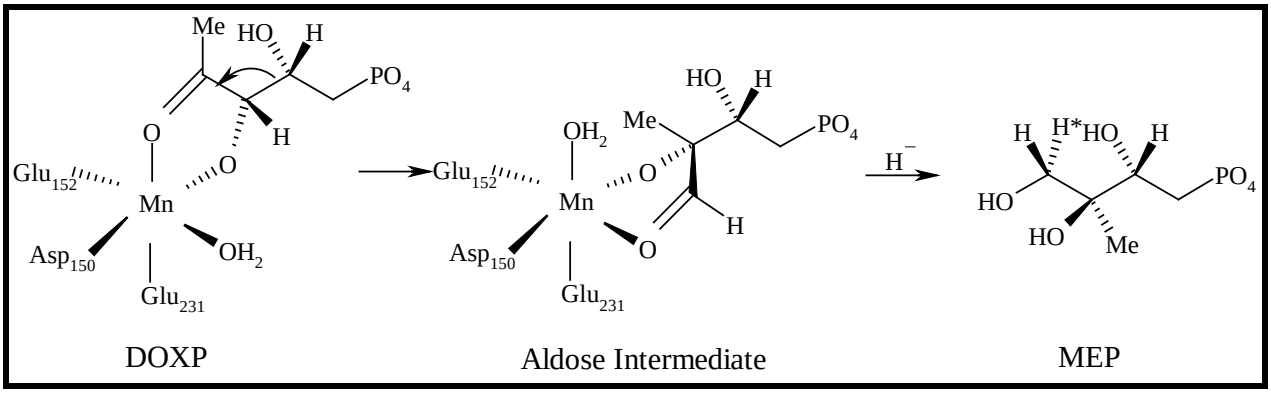


Figure 1.5: Stereochemical view of the reaction catalysed by DOXP reductoisomerase

The manganese metal cation is anchored by the catalytic residues Asp^{150}, Glu^{152} and Glu^{231}. 1-deoxy-D-xylulose-5-phosphate is converted into 2-C-methyl-D-erythritol 4-phosphate via the aldose intermediate 2-C-methyl-erythrose 4-phosphate. The hydride of NADPH and C_4 group of DOXP attack opposite sides of the C_3 and C_2 respectively. The asterisk represents the hydride that is transferred via the cofactor NADPH (adapted from Steinbacher *et al*., 2003).

1.10 Inhibitors of the DOXP synthase and DOXP reductoisomerase enzymes of the Non-Mevalonate Pathway

Tools for the discovery of potential inhibitors of DOXP reductoisomerase as well as DOXP synthase have been investigated as potential drug candidates for malaria. In order to identify possible inhibitors towards the DOXP pathway, initial database searches to identify antibiotics that were active against *E. coli* and *Bacillus subtilis*, harboring the non-mevalonate pathway and inactive against *Staphylococcus aureus* possessing the mevalonate pathway were undertaken in the hope that such inhibitors may already have been reported as antibacterial drugs (Kuzuyama *et al*., 1998). It was from such searches that the inhibitor fosmidomycin (FR-31564) was identified as being a potential inhibitor of the DOXP pathway. Fosmidomycin (3-(*N*-formyl-*N*-hydroxyamino)propylphosphonic acid) (FOSM) has been reported as having a powerful antibacterial activity against the majority of Gram-negative bacteria and some Gram-positive bacteria. From the identification that this inhibitor had a direct impact upon the biosynthesis of menaquinone and carotenoids in *Micrococcus luteus* it was proposed that this inhibitors mechanism of action could be linked with the biosynthetic pathway of isoprenoid and terpenoid synthesis and thus a useful inhibitor towards the nonmevalonate pathway (reviewed within Kuzuyama *et al*., 1998).

Altincicek *et al.* (2000) investigated this approach utilizing enzymes from the pathogenic bacterium *Pseudomonas aeruginosa*. From their research it was found that the recombinant enzymes, of both DOXP synthase and DOXP reductoisomerase, expressed within *Escherichia coli* were inhibited by submicromolar concentrations of fosmidomycin and its derivative FR-900098 (Figure 1.; Jomaa *et al.*, 1999; Reichenberg, 2001). In addition, it has been demonstrated that these antibiotics are efficient inhibitors toward DOXP reductoisomerase from both *P. falciparum* as well as *E. coli* (Jomaa *et al.*, 1999). These results were additionally confirmed through the demonstration of the anti-malarial activity of both these antibiotics *in vitro* and *in vivo*.

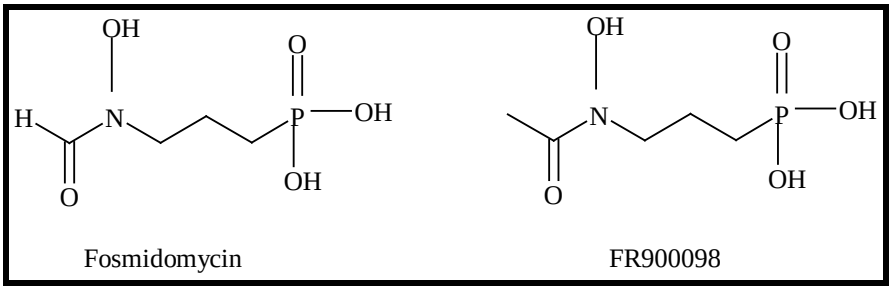


Figure 1.6: Structures of the antibiotic fosmidomycin and its derivative FR-900098

Binding of the inhibitors has been demonstrated within the crystal structure generated via Steinbacher *et al*. (2003). Fosmidomycin binds to the active site through its (N-formyl-N-hydroxy) amino head group by displacing two water molecules present in the octahedral sphere of the metal cation. The phosphate moiety is then secured within the active site through hydrogen bonding with residues Ser^{186}, Ser^{222}, Asn^{227} and Lys^{228} with the N-formyl oxygen placed *trans* to Glu^{231} and the N-hydroxyl oxygen placed *trans* to Asp^{150}. Additionally, Trp^{212} and His^{257} are proposed to form a stacking role, which is maintained by Pro^{274} within the catalytic hatch region only in the presence of NADPH. NADPH is essential for this role to be fulfilled, as it seems to orchestrate the conformation of the active site.

Due to the low oral efficacy of fosmidomycin and its derivatives, prodrugs of the fosmidomycin derivative FR-900098 were then investigated for increased anti-malarial activity (Reichenberg *et al*, 2001). Utilizing FR-900098 as the lead molecule, because of its characteristic 2-fold greater anti-malarial activity, the charged phosphonate moiety was altered into phosphodiaryl esters.

This transformation was chosen as it was expected, and demonstrated, that bioavailabilty would be enhanced through the hydrolysis of the phosphodiaryl esters via non-specific plasma esterases (Reichenberg *et al.*, 2001).

Unfortunately, no other inhibitors against additional enzymes of the nonmevalonate pathway have been reported and due to the short half-life of fosmidomycin and its derivatives there has been a diminished impact upon the clinical application of these drugs. The advantage of fosmidomycin and its derivatives is their low toxicity and therefore they may be potential starting points from which lead molecules may be designed in order to identify new anti-malarial drugs (Zubrzycki and Blatch, 2001).

Zubrzycki and Blatch (2001) have investigated the electrostatic properties of the substrate DOXP and the inhibitors FOSM and FR-900098 through Density Functional Theory (DFT) studies. To assess the possible regions of hard and soft nucleophilic and electrophilic attack, analysis of the charge distribution as well as highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) analyses were performed on both the substrate and inhibitor molecules. From the stick conformations of FOSM and FR-900098 it was determined that the positive charges were concentrated at the head while in the natural substrate, DOXP, there was a negative charge, with the similarity amongst the substrate and inhibitors being the negative charges upon the tail groups. It was concluded that the inhibiting properties of FOSM and FR-900098 were due to the physical properties of these tail groups. From the electrostatic analysis the nitrogen atom, within FOSM and FR-900098, was found to possess a high electronegativity, as well as an increased charge density, with its introduction being directly linked towards the inhibitory properties of the molecule. It was additionally confirmed upon comparison of the charge upon the nitrogen atom within FOSM and FR-900098, with the charge upon the carbon at position 3 in DOXP. The nitrogen atoms within the inhibitory molecules FOSM and FR-900098 were reported as having the charges –0.149 and 0.195 respectively, whereas the carbon atom within DOXP was found to be almost neutral with the charge –0.011 (Zubrzycki and Blatch, 2001).

From the DFT studies of Zubrzycki and Blatch, (2001) it was confirmed that the introduction of the nitrogen atom, carrying a significant negative charge as compared to the carbon atom within DOXP, influences the molecular environment of the active residues in the active centre of the enzyme. This charge was reported to possibly alter the three dimensional microstructure of the active pocket of the enzyme and competitively bind to the active site.

1.11 Hypothesis Statement:

DOXP reductoisomerase is a potential target for inhibition within the mevalonate-independent pathway and serves as a potential site for novel drug design against malaria.

1.12 Approach

Within this investigation *E. coli* and *P. falciparum* DOXP reductosiomerase will be biochemically characterized as a potential target for inhibition. In particular, this will involve the biochemical characterization of the active site (in specific the catalytic hatch region) of *E. coli* and *P. falciparum* DOXP reductoisomerase by rational amino acid analysis. Rational drug design and *in silico* docking strategies will be used to identify novel DOXP reductoisomerase inhibitors for organic synthesis and inhibitor assays.

1.13 Aims and Objectives

For the biochemical characterization of both *E. coli* and *P. falciparum* DOXP reductoisomerase the respective recombinant proteins were to be produced as $(His)_6$ fusion proteins within the heterologous expression system *E. coli* and purified to homogeneity, such that the protein can be utilized in further functional and inhibitor analysis, via nickel-chelate affinity chromatography.

Homology modeling of the *P. falciparum* DOXP reductoisomerase, using the *E. coli* DOXP reductoisomerase structure as a template, were performed to generate a model of the 3-D predicted structure for *P. falciparum* DOXP reductoisomerase. From this potentially critical

amino acids that are involved in the catalytic mechanism of the *P. falciparum* DOXP reductoisomerase activity were identified.

Protein engineering studies upon *E. coli* and *P. falciparum* DOXP reductoisomerase proteins were then performed to assess their involvement in DOXP reductoisomerase activity. Since there is only ~56.5 % amino acid variation between the *E. coli* and *P. falciparum* DOXP reductoisomerase amino acid sequences, the *E. coli* DOXP reductoisomerase could be used as a model system on which functional predictions could be made and tested that should reflect the parasitic system as well. Ultimately, all functional predictions should be verified through analysis with the recombinant *P. falciparum* DOXP reductoisomerase protein.

Optimization of the NADPH DOXP reductoisomerase coupled enzyme assay for the determination of DOXP reductoisomerase activity was performed to allow for the biochemical characterization of the *E. coli* and *P. falciparum* DOXP reductoisomerase proteins and their derivatives.

For the identification of novel DOXP reductoisomerase inhibitors a rational drug design approach utilizing the natural substrate (DOXP) and the inhibitors fosmidomycin and FR900098 as lead molecules were performed. Ligand docking methods, with the lead compounds and novel drugs, allowed for the analysis of potential inhibitors and their identification for organic synthesis. The synthesized drugs were to be assessed for their inhibitory affect *in vitro* using the NADPH DOXP reductoisomerase coupled enzyme assay with the purified recombinant DOXP reductoisomerase protein.

CHAPTER 2: HETEROLOGOUS PROTEIN PRODUCTION, PURIFICATION AND CHARACTERIZATION OF DOXP REDUCTOISOMERASE

2 HETEROLOGOUS PROTEIN PRODUCTION AND PURIFICATION OF DOXP REDUCTOISOMERASE

2.1 Introduction

A major requirement of biochemical research is to purify large quantities of pure protein for further analysis and research. The production of proteins from their natural source often does not fulfill these requirements as well as such cases often being problematic due to scarcity or pathogenicity. The production of proteins in recombinant expression systems has therefore become an established field in both biochemical and biotechnological research. Typical expression systems utilize the protein production machinery of well-characterized organisms such as *Saccharomyces cerevisiae, E. coli* and *Baculovirus* infected insect cells (Higgins and Hames, 1999).

2.1.1 Heterologous Production of Recombinant Proteins

E. coli specifically has often been the host system of choice to many researchers due to its relatively simple genetics and rapid growth rate. Considerable genetic and biochemical knowledge has accumulated on *E. coli* making it a desirable host for the production of large quantities of heterologous proteins and as a consequence many *E. coli* host systems for direct expression, fusion protein expression, and secretion have been developed and optimized. Fusion protein strategies ensure favourable translation initiation and avoid instability problems regarding small peptides. Within this strategy the fusion partner encodes an affinity tag to facilitate purification. The coding region of interest is introduced into an expression vector 3' to a sequence encoding for the amino terminus or whole length of a highly expressed protein, referred to as a carrier protein (Studier *et al*., 1990). Within this approach two separate concepts are addressed, the carrier protein facilitates solubility and high-level expression whereas the affinity tag allows for purification.

However, there are several problems in expressing proteins of eukaryotic origin in prokaryotic host expression systems. One such disadvantage in utilizing *E. coli* to heterologously over-express proteins is that the proteins are often expressed as insoluble aggregated folding

intermediates, known as “inclusion bodies”. Several techniques are employed to overcome such difficulties in protein folding. Aggregation can often be prevented by altering the growth parameters such as lowering the growth temperature, thus decreasing the rate of transcription and translation resulting in higher levels of native recombinant protein (Lilie *et al.*, 1998; Schein and Noteborn, 1988). In addition, one can also recover an active protein expressed within inclusion bodies by initially solubilizing the ‘insoluble’ protein using denaturants such as urea and guanidine hydrochloride and then exposing the solubilized protein to refolding strategies. However, the level of recombinant protein obtained in such cases is often low and hampered by the optimization of the refolding conditions to that specific target protein.

Another emerging technique for increasing heterologous protein production in the native form has been to co-express molecular chaperones, which aid in protein folding. This often requires several experiments to match the target protein to its specific molecular chaperone (Weickert *et* al., 1996).

Codon usage is a further difficulty that can be encountered when utilizing *E. coli* as the host expression system. The problem arises due to the protein coding region containing rare codons that are not common in the host expression system. In such cases the translation product is often truncated and the resulting protein unstable. Such expression problems can be combated by altering growth conditions of the host cell to lower transcription levels prior to induction of the protein of interest or “codon optimizing” the gene of interest to that of the host (Kurland and Gallant, 1996).

Bacillus subtilis, a gram-positive bacteria, is another bacterial expression system that is often considered. This organism, although its cloning and gene characterization are not as well known as *E. coli*, has the advantages of being non-pathogenic, lacks endotoxin production, and secretes the protein into the surrounding medium (Madigan *et al*., 1997). The *B. subtilis* system is utilized in the pharmaceutical production of hormones such as insulin and human growth hormone. This system is suited to industry due to protein yields being maintained after scale-up, however the system does have limitations due to plasmid instability within *B. subtilis*.

An additional downfall when using *E.* coli as a host expression system is the lack of enzymes involved in post-translational modifications, which can often affect the recombinant proteins function. *S. cerevisiae*, *Schizosaccaromyces pombe*, *Pichia pastoris*, and *Pichia angusta* are yeast strains often used protein expression systems, offer an alternative system allowing for sufficient expression levels along with most eukaryotic post-translational modifications (Bradley, 1990). Proteins that were insoluble in the bacterial system are often soluble in this system due to the protein processing within yeast being more efficient and intricate. Another advantage when using yeast expression systems is that proteins can be secreted into the growth media upon induction. This feature is often engineered to prevent toxicity problems and to simplify protein purification (Bradley, 1990). However, one severe drawback in yeast expression systems is the reduction of recombinant protein levels due to the presence of active proteases, however this is often immaterial when utilizing yeast strains that have had protease-encoding genes eliminated. Additionally, when utilizing the His-tag system for protein purification (discussed in section 2.1.3) yeast cells may pose an additional problem due to their ability to acidify the culture media. This results in compounds being present that may interfere with the binding of the His-tag to the nickel charged resin, thus interfering with protein purification.

In addition to yeast and bacterial expression systems *Baculovirus* vectors are useful in the expression of recombinant proteins in insect cells. Recombinant proteins in this system are expressed with the majority of mammalian post-translational modifications and can either be produced within the insect cell itself or secreted into the media. This system does have the limitation of being rather difficult to set-up and scale-up and that insect cells have a far slower growth rate than bacteria and yeast cells (Madigan *et al*., 1997).

2.1.2 Heterologous Production Problems Encountered for *P. falciparum* Recombinant Proteins

Malarial proteins present protein production difficulties especially when producing large quantities necessary for biochemical and biophysical analysis. Expression within *E. coli* is hampered due to the codon bias with *P. falciparum* having an AT rich genome. Codons common within parasitic protein coding regions are rare within *E. coli* and thus protein synthesis is

hampered due to translational problems. Strategies to codon optimise parasitic genes to suite the codon usage of *E. coli* are often costly and lengthy. Baca and Hol, (2000) have described the successful use of the RIG plasmid to overcome translational limitations. This plasmid is designed to encode the genes for tRNA's that are present at low levels in *E. coli,* i.e. tRNA`s that recognise rare codons for Arg, Ile, and Gly common in parasitic genes.

Parasitic proteins are often toxic to *E. coli* even with the RIG plasmid present. This is due to the inducible promoter present within the expression vector allowing for the basal transcription of the protein to occur prior to induction, resulting in hampered growth of the host expression system. Cinquin *et al.* (2001) have developed a hybrid plasmid, which allows for the expression of toxic proteins on an expression vector containing a T7 inducible promoter within *E. coli*. The hybrid plasmid was constructed from the RIG plasmid and the pLysS plasmid, which allows for low levels of T7 lysozyme, a potent natural T7 RNA polymerase inhibitor, to be expressed. This allows for the control of levels of transcription allowing toxic proteins to be accommodated within the cell. Thus the hybrid plasmid contains the properties of both the pLysS and RIG plasmids allowing for toxicity and codon bias to be overcome when expressing toxic parasitic proteins.

Overall, the expression and subsequent purification of recombinant proteins facilitates the production and biochemical characterization of almost any protein. There are many expression systems that have been developed for protein production in *E. coli* each with their own advantages and disadvantages. Within this research the application of fusion proteins and metal-chelate affinity chromatography was utilized and will be discussed in detail. Several methods for protein purification are reviewed within Sheibani (1999).

2.1.3 Metal Chelate Affinity Chromatography

Affinity-based techniques for the isolation and purification of proteins have become a common practice in protein biochemistry. Affinity chromatography is the only technique available whereby a biomolecule may be purified due to its biological function or individual chemical structure. It allows for the rapid separation of proteins, based on reversible interactions between

the protein and a specific ligand, which is attached to a chromatographic matrix. This technique has the advantages of selectivity, thus high resolution, and high capacity (Porath *et al*., 1975; Chaga, 2001).

The coupling of affinity chromatography techniques with the development of epitope tagging of recombinant proteins has allowed for rapid affinity-based purification methodology, such as immunoprecipitation. Such techniques have aided studies involving protein production, expression, purification and modification and interactions in various biological systems.

The first successful description of the purification of recombinant proteins via a chromatographic technique was by Porath *et al*. (1975). This study combined Metal-Chelate Affinity Chromatography (MCAC) in the purification of human serum proteins. The basis of MCAC involves the binding of metal ions to solid chromatographic matrices, which in turn interact and bind to the exposed imidazole and thiol groups of proteins. The exploitation of this technique has been to use tagging systems wherein the protein is tagged, in this case with a poly histidine tag, allowing the purification of the tagged protein on a nickel charged column. The advantage to the poly histidine tag system is that, with sufficient levels of protein, one-step purification is possible. In addition to this it is poorly immunogenic, and at a pH of 8.0 the tag is small as well as uncharged ($(His)_6$–0.84 kDa vs.Glutathione-*S*-transferase GST–26 kDa). These characteristics make it unlikely that the tag will interfere with secretion, compartmentalization, or folding of the protein nor will it affect the structure or function of the purified protein. Unfortunately, the size of the tag bears its own problems as it can often be 'buried' as a consequence of protein folding, particularly if it is situated at the C-terminus of the protein (Sheibani, 1999). Sometimes the $(His)_6$ tag itself can interfere with protein solubility and a range of histidine tags are now available that are still applicable for protein purification yet confer better solubility characteristics to the expressed protein, these include the 6xHN and HAT tags (www.clontech.com).

The high level production of $(His)_6$ DOXP reductoisomerse was carried out in *E. coli* XL1 Blue using the QIAexpress pQE vector system. This system is based on the T5 promotor transcription-translation system. The pQE vector plasmids belong to the PDS family of plasmids described by Bujard *et al.* (1987) and were derived from plasmids pDS56/RBSII and

PDS781/RBSII-DHFRS (Stüber *et al.,* 1990). The *E. coli* and *P. falciparum* DOXP reductoisomerase proteins were purified for use in functional and biochemical characterization.

2.2 Aim

To heterologously over-produce, purify and biochemically characterize *E. coli* and *P. falciparum* DOXP reductoisomerase.

2.3 Experimental Procedures

2.3.1 Reagents and Chemicals

All reagents and chemicals utilized in the expression and purification of *P. falciparum* and *E. coli* DOXP reductoisomerase are described in the Appendices (Chapter 6; section 6.1).

2.3.2 Heterologous Protein Production and Purification of DOXP Reductoisomerase

2.3.2.1 Constructs and Host Expression Systems for the Over-Expression and Purification of *E. coli* and *P. falciparum* DOXP Reductoisomerase

Both plasmid contructs containing the coding regions for *P. falciparum* and *E. coli* DOXP reductoisomerase were kindly provided by Jomaa Pharmaka GmbH (Germany; pQE9EcDXR and pQE31PfDXR). DOXP reductoisomerase coding sequences were provided within pQE (Qiagen) expression constructs incorporating N-terminal $(His)_6$ fusion tags for nickel affinity purification. The *E. coli* DOXP reductoisomerase coding sequence was inserted into the pQE9 vector in frame with an N-terminal $(His)_6$ tag. An ampicillin resistance cassette provides a selectable marker (Jomaa *et al*., 1999). The malarial parasite *P. falciparum* has different codon preferences to that of the host expression system *E. coli* and often in such cases codons, which are preferential in *P. falciparum* genes are rare in *E. coli* genes. In order to combat this problem the *P. falciparum* DOXP reductoisomerase coding sequence was optimized to the preferential codon usage of *E. coli*. This coding sequence was then inserted within pQE31 (Jomaa *et al*., 1999). The genotype

of the host *E. coli* cells utilized for the expression of the pQE expression vector constructs was *E. coli* XL1Blue *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac[F' proAB lacI^{q}ZΔM15 Tn10 (Tet^{r})]^{c}*.

2.3.2.2 Transformation of Competent Host Cells with Expression Constructs

E. coli XL1 Blue cells were made competent using a derivation of the calcium chloride method described by Dagert and Ehrlich, (1979). 50 ml of sterile 2xYT broth (1.6% tryptone, 1% yeast extract, 0.5% NaCl) was inoculated with 0.25 ml of an *E. coli* XL1 Blue overnight culture. The culture was then incubated at 37°C with shaking until the culture reached early log phase with an absorbance (A_{600nm}) of between 0.3 and 0.6. The cells were harvested via centrifugation (5000x *g* for 15 minutes at 4°C) and the cell pellet resuspended in one culture volume (50 ml) of sterile ice cold 0.1 M $MgCl_2$. After incubation on ice for 20 minutes the cells were collected as previously described via centrifugation and the cell pellet resuspended in half the culture volume (25 ml) of sterile ice cold 0.1 M $CaCl_2$. The cell suspension was incubated for 2 hours on ice and the cells subsequently collected following the same centrifugation procedure. The resulting cell pellet was resuspended in one-tenth the culture volume (5 ml) ice cold 0.1 M $CaCl_2$. Aliquots of competent cells were stored as 200 µl fractions (100 µl cell suspension and 100 µl of sterile ice cold 30 % glycerol) at –80°C.

Aliquots (100 µl) of competent *E. coli* XL1-Blue cells were transformed with aliquots (~200 ng) of double stranded plasmid DNA, pUC18 plasmid DNA (1 µl; 0.1 ng/µl) as a transformation positive control and sterile deionized milli-Q water as a no DNA control. Transformation reactions were incubated on ice for 30 minutes after which the reactions were heat pulsed for 2 minutes at 42°C, followed by incubation on ice for a further 2 minutes. Sterile 2xYT Broth (900 µl), preheated to 42°C was supplemented to the transformation reactions followed by incubation at 37°C with shaking at 225-250 rpm for 1 hour to allow for the expression of the antibiotic resistance gene specific to the plasmid construct being transformed. All reactions, excluding pUC18, were micro centrifuged (12 000x *g*) for 1 minute and 800 µl of the supernatant removed from each reaction, and the cell pellet resuspended in the remaining supernatant (100 µl). All transformation reactions were spread plated onto 2xYT broth agar plates containing 1.5 % agar

and 0.1 mg/ml ampicillin, except for pUC18. pUC18 (5 µl) was plated onto 5 % MacConkey agar plates (1 % Lactose) containing 0.1 mg/ml ampicillin. Plates were inverted and incubated at 37˚C for >16 hours. The efficiency of transformation was reflected through the number of colony forming units obtained from the pUC18 transformation reaction, and the water transformation control ensured sterility. Transformants, colonies containing the plasmid construct, were then picked and streaked onto 2xYT broth agar plates containing 0.1 mg/ml ampicillin and analyzed to confirm the presence of the plasmid DNA.

2.3.2.3 Isolation of Plasmid DNA from Transformants for Confirmation of Plasmid Identity via Diagnostic Restriction Enzyme Analysis

E. coli XL1Blue transformants for each putative plasmid were incubated in sterile 2xYT broth (5 ml) containing 0.1 mg/ml ampicillin at 37˚C overnight with shaking at 225 -250 rpm. Cells were collected via was centrifugation at 12 000x *g* for 1 minute and preparation of purified plasmid DNA was performed using either the High Pure Plasmid Isolation Kit (Roche) or the QIAprep Spin Miniprep Kit (Qiagen).

The High Pure Isolation Kit procedure involved the resuspension of the bacterial cell pellet in Suspension Buffer containing RNase (100 µg/ml)(250 µl; 50 mM Tris-HCl, 10 mM EDTA, pH 8.0, 25˚C). To the resuspended solution 250 µl of Lysis Buffer (0.2 M NaOH; 1% SDS) was then added and the suspension inverted a few times and incubated for 5 minutes at room temperature. Chilled Binding Buffer (350 µl; 4 M guanidine hydrochloride; 0.5 M potassium acetate; pH 4.2) was added and the tube inverted, followed by incubation on ice for 5 minutes. The cloudy precipitate was collected via centrifugation (12 000x *g* for 1 minute) and the supernatant transferred to a High Pure filter tube reservoir and centrifuged (12 000x *g* for 1 minute). The flow-through solution was discarded and Wash Buffer II (700 µl; 20 mM NaCl; 2 mM Tris-HCl; pH 7.5) added to the upper reservoir and the filter tube centrifuged (12 000x *g* for 1 minute). The flow-through solution was again discarded and residual wash buffer removed by centrifugation (under previous conditions) once more. The plasmid DNA was eluted into a sterile 1.5 ml eppendorf tube with Elution Buffer (100 µl; 10 mM Tris-HCl; pH 8.5) via centrifugation (12 000x *g* for 1 minute).

The QIAprep Spin Miniprep Kit Protocol involved the resuspension of the bacterial cell pellet in Buffer P1 containing RNase (100 µg/ml)(250 µl; 50 mM Tris-HCl, 10 mM EDTA, pH 8.0, 25˚C) after which Buffer P2 (250 µl; 200 mM NaOH; 1% SDS) was added and the tube inverted to allow for alkaline lysis. Lysis was followed by the addition of Buffer N3 (350 µl)(buffer constituents were not revealed by the manufacturere) and additional inversion of the tube. The resultative cloudy solution was centrifuged (12 000x *g* for 10 minutes) and the supernatant transferred to a QIAprep column and centrifuged for a further minute. The flow-through was discarded and the column washed by adding Buffer PE (750 µl)(buffer constituents were not revealed by the manufacturere) and centrifuging for 30-60 seconds. The resultative flow-through was discarded and the plasmid DNA eluted with Buffer EB (50 µl)(buffer constituents were not revealed by the manufacturere) into a sterile 1.5 ml eppendorf via centrifugation (12 000x *g* for 1 minute). For sequencing purposes sterile distilled milli-Q water was utilized for elution purposes instead of Buffer EB. Each of the plasmid DNA preparations was then utilized in diagnostic restriction enzyme analysis procedures to confirm the presence of the correct plasmid construct.

Restriction Analysis of ~100-500 ng of double stranded plasmid DNA was performed with ~10 Units of the relevant diagnostic restriction enzyme at the relevant temperature for ~2 hours (Amersham; MBI Fermentas; Roche; Promega). Diagnostic Restriction Enzyme Digests were analyzed via agarose gel electrophoresis (AGE). Agarose gel loading buffer (6x buffer; 30% glycerol, 0.25% bromophenol blue) was added to each restriction endonuclease reaction (20 µl) and the samples resolved on the required percentage agarose gel containing 0.5 µg/ml ethidium bromide in 0.5x TBE Buffer (45 mM Tris, 45 mM borate, 1 mM EDTA, pH 8.3) at 100V for ~1.5 hours and visualized under ultraviolet (UV) light. *Pst*I – digested lambda (λ) DNA was resolved as a molecular weight marker.

2.3.2.4 Heterologous Production of DOXP Reductoisomerase

Cultures of *E. coli* [pQE9EcDXR] and *E. coli* [pQE31PfDXR] were incubated in sterile 2xYT broth (25 ml) containing 0.1 mg/ml ampicillin at 37˚C overnight with shaking at 225 -250 rpm. Overnight cultures were diluted 1:10 into sterile 2x YT broth (225 ml) containing 0.1 mg/ml

ampicillin and incubated at 37°C with shaking until the mid-logarithmic stage was reached (A_{600} of between 0.6 and 1). For the production of recombinant protein, cultures were induced with the addition of isopropyl-β-D-galactopyranoside (IPTG)(final concentration of 1 mM). Samples (2 ml) were taken hourly for a period of 5 hours after induction and overnight: A_{600} of a 1 ml sample was determined and the remainder (1 ml) collected by centrifugation at 12 000x *g* for 1 minute. The cell pellet was resuspended in sterile ice cold Phosphate Buffered Saline (PBS)(137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, 1.4 mM KH_2PO_4) to a ratio of 50 µl PBS per 0.5 A_{600} unit. PBS cell suspensions (20 µl) were added to sodium dodecyl sulphate (SDS) polyacrylamide gel loading buffer (4 µl of a 5x loading buffer)(62.5 mM Tris-HCl pH 6.8, 10 % (v/v) glycerol, 2 % (w/v) sodium dodecyl sulphate (SDS), 5 % (v/v) β-mercaptoethanol, 0.05 % (w/v) bromophenol blue) and boiled at 95°C for 5 minutes for analysis via discontinuous 0.1 % SDS 12 % Polyacrylamide Gel Electrophoresis (SDS-PAGE)(12 % acrylamide resolving gel; 4 % acrylamide stacking gel)(BioRad Minigel Electrophoresis Set) at 200 V for 1 hour (Running Buffer: 25 mM Tris-HCl; 250 mM Glycine; pH 8.0; 0.1 % SDS)(Laemmli, 1970). The gel was removed and stained with Coomassie Brilliant Blue stain (0.1 % (w/v) Coomassie Brilliant Blue R-250, 40 % (v/v) methanol, 10 % (v/v) glacial acetic acid) for protein visualization, for 1 hour with shaking and destained for 3 hours in destain (40 % (v/v) methanol, 10 % (v/v) glacial acetic acid). Gels were dried by equilibrating in 20 % ethanol, 10 % glycerol for at least 30 minutes and placing in-between cellophane sheets using the Gel Drying frames supplied by Diversified Biotech (Sigma).

For the solubility analysis cultures of *E. coli* [pQE9EcDXR] and *E. coli* [pQE31PfDXR] were incubated and harvested as described above. Cell pellets were resuspended in PBS (5 ml) and phenylmethylsulfonylfluoride (PMSF)(100 mM), and aliquoted into 5 x 1 ml fractions in sterile eppendorf tubes. The aliquoted cells were lysed on ice by sonication (60 Amp), for 30 seconds pulses to allow for 30 seconds, 60 seconds, 90 seconds, 120 seconds and 150 seconds sonicated samples. The cell debri was removed from the soluble cell extract via centrifugation at 12 000x *g* for 20 minutes at 4°C. Samples (80 µl) of soluble cell extract were added to SDS loading buffer (20 µl) for analysis on 0.1 % SDS 12 % PAGE. The lysed cell pellet was resuspended in PBS (1 ml) and samples (80 µl) were added to SDS loading buffer (20 µl) for analysis on 0.1 % SDS 12% PAGE.

2.3.2.5 Western Blot Analysis and Chemiluminescence-Based Immunodetection of 6x His-tagged DOXP Reductoisomerase

The protein samples separated on the SDS-PAGE gel (described in section 2.3.2.4) were transferred to a nitrocellulose membrane (Hybond ECL) in chilled transfer buffer (25 mM Tris, 192 mM glycine, 20 % (v/v) methanol) at 100V for 1 hour utilizing the Biorad Western Transfer Apparatus (Towbin *et al*., 1979). The transfer was visualized by incubating the membrane in Ponceau S stain (0.5 % (w/v) Ponceau S, 1 % (v/v) glacial acetic acid) for 1 minute and rinsing with milli-Q water. The stained membrane was then scanned and rinsed to remove excess Ponceau stain and subjected to chemiluminescent immunodetection using either the Boeringer Mannheim (BM) Chemiluminescence Western Blotting Kit (Boeringer Mannheim; Germany) or the ECL Advance Western Blotting Kit (Amersham Biosciences). In both cases the nitrocellulose membrane was initially washed with Tris-Buffered-Saline (TBS; 50 mM Tris-HCl (pH 7.5), 150 mM NaCl). The membrane was then incubated å€4overnight in blocking solution (5 % (w/v) non-fat milk powder in TBS) to prevent non-specific binding of the antibody. Detection of $(His)_6$ proteins was performed by incubating the membrane with anti-His monoclonal mouse antibody (1:5 000 dilution of the antibody within blocking solution) for 1 hour at room temperature with shaking. The primary antibody was removed and the membrane was washed four times for 20 minutes with TBS-Tween (0.1 % (v/v) Tween-20 in TBS) at room temperature with shaking. At this point either the BM Chemiluminescence Western Blotting Kit secondary antibody (anti-mouse IgG-peroxidase labeled [POD]/anti-rabbit IgG-POD) or the ECL kit secondary antibody (anti-mouse IgG, peroxidase-linked species-specific whole antibody) was diluted 1:10 000 in blocking solution and added to the membrane. The membrane and secondary antibody was then incubated for 2 hours at room temperature with shaking after which the antibody was removed and the nitrocellulose membrane washed with TBS-Tween four times for 20 minutes with shaking at room temperature. The membrane was incubated with the chemiluminescent substrate luminol (BM Chemiluminescence Western Blotting Kit solution A (5ml) and solution B (50 µl) or ECL Advance Western Blotting Kit solution A and solution B (in a 1:1 ratio) in the dark for 1minute and placed in a HyperFilm Developing Cassette (Agfa, USA) and exposed to high performance autoradiography film (Agfa, USA) for various lengths of time

(30 seconds – 1 hour) depending on the intensity of the signal emitted due to the oxidation of the chemiluminescence substrate by the peroxidase-conjugated secondary antibody used. The X-ray film was developed in developer solution (Agfa, USA) for 1-5 minutes, rinsed in stop solution to stop the reaction (2 % glacial acetic acid), and fixed in fixer solution (Agfa, USA), and rinsed in water before drying.

2.3.3 Nickel Chelate Affinity Purification of *E. coli* DOXP Reductoisomerase

2.3.3.1 Preparation of the cleared lysate of *E. coli* DOXP Reductoisomerase

Cultures of *E. coli* [pQE9EcDXR] were incubated in sterile 2xYT broth (4 x 25 ml) containing 0.1 mg/ml ampicillin at 37°C overnight with shaking at 225 -250 rpm. Overnight cultures were diluted 1:10 into sterile 2x YT broth (225ml) containing 0.1 mg/ml ampicillin and incubated at 37˚C with shaking until the mid-logarithmic stage was reached (A_{600} of between 0.6 and 1). For the production of recombinant protein, cultures were induced with the addition of IPTG (final concentration of 1 mM). After 4 hours of recombinant protein induction, the cells were harvested via centrifugation (5 000x *g*; 15 minutes; 4°C). The cell pellet was resuspended in Native Lysis Buffer (10 ml per litre culture volume)(20 mM Tris-HCl; 300 mM NaCl; 10 mM Imidazole; pH 8.0; 4°C) and Phenylmethylsulfonyl fluoride (PMSF)(final concentration of 1 mM) and stored overnight at –80˚C. Cell lysis was achieved by adding the enzyme lysozyme (100µg/ml) for ~2 hours on ice and sonication for a single 30 second pulse at 4C. The cell debris was removed via centrifugation (12 000x *g*; 25 min; 4C) and the lysate containing soluble protein (supernatant) utilized in native batch purification via nickel chelate affinity chromatography.

2.3.3.2 Native Batch Nickel Purification of *E. coli* DOXP Reductoisomerase

The cleared lysate obtained was added to nickel charged sepharose resin (Pharmacia Biotech)(500 µl) and agitated gently overnight at 4°C. The resin was collected via centrifugation (4 000x *g*; 5 minutes; 4C) and washed, with gentle agitation at 4˚C, three times with Nati ve Wash Buffer (5 ml)(20 mM Tris-HCl; 300 mM NaCl; 100 mM Imidazole; pH 8.0; 4°C). The $(His)_6$ protein was then eluted from the beads using Native Elution Buffer (20 mM Tris-HCl; 300

mM NaCl; 1 M Imidazole; pH 8.0; 4°C) in 4 times the bead volume. The resin was collected between each wash and elution via centrifugation at 4 000x g; 5 minutes; 4°C and 12 000x *g*; 1 minute, 4°C respectively. The protein purification profile was monitored by discontinuous SDS-PAGE (outlined in section 2.3.2.4) of the fractions collected throughout the purification procedure. The eluted protein volume was then buffer exchanged into 100 mM Tris-HCl (pH 7.5) via size exclusion/gel filtration chromatography at 4°C to remove imidazole from the protein fraction. This procedure involved subjecting the pooled eluates to a buffer equilibrated Sephadex G-25 (Amersham Biosciences) column (length 10 cm; width 1 cm) and eluted using 100 mM Tris-HCl (pH 7.5) to allow for buffer exchange. Protein and imidazole elution from the column was monitored at 280nm (100 µl)(using a flat-bottomed 96-well microtitre UV plate; Corning Costar) and via Bradford's Assay (Bradford, 1976). A sample/standard (10 µl) was incubated with Bradford's Reagent (200 µl) at room temperature for 5 minutes and the absorbance measured at 595 nm using a Powerwave Microtitre Plate Reader (Bio-Tek Instruments). A standard curve of protein concentration (0 – 250 µg/ml) was determined using a standard protein, Bovine Serum Albumin (BSA)(Sigma) and was used to monitor elution of protein from the column. The protein fractions eluted from the column were pooled and concentrated on Polyethylene glycol 20 000 (PEG) in dialysis tubing at 4°C. The protein concentration was then determined using the Bradford's Assay described above and the protein stored as a concentrated solution at –80°C in aliquots.

2.3.3.3 The *in vitro* Analysis of *E. coli* DOXP Reductoisomerase Activity

The substrate DOXP was a kind gift from Jomaa Pharmaka GmbH, Germany and the DOXP reductoisomerase assay was adapted from Takahashi *et al*. (1998). The assay mixture consisted of 100 mM Tris-HCl (pH 7.5), 1 mM $MnCl_2$; freshly made 0.3 mM NADPH (Sigma) and 0.3 mM of DOXP substrate. The assay was initiated by addition of 20 µl of purified enzyme (8.73 µM, final concentration in assay 0.873 µM) to the complete assay mixture (final volume 200 µl). The oxidation of NADPH was monitored at 340 nm at 37°C for 5 minutes using a Microtitre Plate Reader and a flat-bottomed 96-well microtitre UV plate. The controls utilized included a no DOXP (i.e. no substrate therefore no oxidation of NADPH should be monitored), no NADPH (i.e. no cofactor therefore no A_{340} reading), no enzyme (i.e. no oxidation of NADPH at 340 nm),

boiled enzyme (i.e. no oxidation of NADPH should occur), and no DOXP and no NADPH (i.e. no DOXP present and no cofactor present therefore no A_{340nm} reading). This allowed for the correction of any basal level of activity to be accounted for as well as validating that any oxidation of NADPH monitored at 340 nm was due to the active soluble enzyme within the assay. The extinction coefficient (ε) for NADPH at 340 nm is 6.3 x 10^3 L/mol/cm. The pathlength for the microtitre plate reader was determined as 0.52 cm.

2.3.4 Nickel Chelate Affinity Purification of *P. falciparum* DOXP Reductoisomerase

2.3.4.1 Preparation of the cleared lysate of *P. falciparum* DOXP Reductoisomerase

Cultures of *P. falciparum* [pQE31PfDXR] were incubated, as stipulated for the production of *E. coli* [pQE9EcDXR], until the mid-logarithmic stage was reached (A_{600} of between 0.6 and 1). For the production of recombinant protein, cultures were induced with the addition of IPTG (final concentration of 1 mM). After 4 hours of recombinant protein induction, the cells were harvested via centrifugation (5 000x *g*; 15 minutes; 4°C). Several methods were utilized during the optimization of the purification protocol for *P. falciparum* DOXP reductoisomerase. In each case the resultant cell pellet was resuspendend in the appropriate Lysis Buffer (detailed in the purification protocols below)(10 ml per litre culture volume) and PMSF (final concentration of 1 mM). The resuspended pellet was then stored overnight at –80°C after which lysis was achieved through the addition of lysozyme and sonication as for *E. coli* DOXP reductoisomerase cleared lysate preparation. The cell debris was removed via centrifugation (12 000x *g*; 25 min; 4°C) and the resultant cleared lysate (supernatant), containing soluble protein, utilized in the batch purification procedures outlined below involving nickel affinity chromatographic techniques.

2.3.4.2 Denaturing Lysis, Native Wash and Elution Purification of *P. falciparum* DOXP Reductoisomerase

During the preparation of the cleared lysate for *P. falciparum* DOXP reductoisomerase, the cell pellet was resuspended within Denaturing Lysis Buffer (8 M Urea; 20 mM Tris-HCl; 300 mM NaCl; 10 mM Imidazole; pH 8.0; 4°C). The cleared lysate obtained from the procedure outlined

above was then applied to charged nickel sepharose beads (500 µl per litre culture volume) and allowed to agitate gently overnight at 4°C. The beads were then collected via centrifugation (4 000x *g*; 5 minutes; 4°C) and washed four times successively with Native Wash Buffer I (20 mM Tris-HCl; 300 mM NaCl; 4 mM Imidazole; pH 8.0; 4°C) and Native Wash Buffer II (20 mM Tris-HCl; 300 mM NaCl; 50 mM Imidazole; pH 8.0; 4°C) and collected between each wash step as described. The elution of the $(His)_6$ recombinant protein was achieved using Native Elution Buffer (20 mM Tris-HCl; 300 mM NaCl; 1 M Imidazole; pH 8.0; 4°C). Collection of the beads during elution was via centrifugation (12 000x *g*; 1 minute; 4°C). The protein purification profile was monitored via discontinuous SDS-PAGE and protein concentration determined via the Bradford Assay (as described in section 2.3.2.4 and 2.3.3.2).

2.4 Results and Discussion

2.4.1 Heterologous Production and Purification of *E. coli* DOXP Reductoisomerase

2.4.1.1 Confirmation of the Identity of pQE9EcDXR Plasmid

Previous expression of the recombinant *E. coli* DOXP reductoisomerase utilizing this plasmid has been demonstrated by Jomaa *et al*. (1999) for inhibitor analysis. Within this work we were interested in producing the recombinant protein as a $(His)_6$ protein from this construct for use in biochemical characterization of the enzymes function as well as inhibitor analysis. Additionally, due to the sequence similarity of the *E. coli* DOXP reductoisomerase protein to its malarial homologue, the *E. coli* protein could allow for a model system for initial inhibitor assay screening and biochemical characterization. This was especially the case due to the problems experienced in the purification of the *P. falciparum* DOXP reductoisomerase in comparison to the high levels of pure native protein obtained in the *E. coli* DOXP reductoisomerase purification. In order to confirm the identity of the plasmid construct a series of diagnostic restriction enzyme digests were used to confirm the plasmid identity (Figure 2.1). The *E. coli* DOXP reductoisomerase coding sequence was inserted into the pQE9 expression vector at the *Bam*HI and *Hind*III sites and upon restriction enzyme digestion with these enzymes the coding sequence was released yielding two DNA fragments of 3419 bp and 1203 bp (Figure 2.1; lane 5). A partial

digest of 4622 bp representative of linearized plasmid DNA is also present (Figure 2.1; lane 5). Digestion with either of these enzymes individually resulted in the linearization of the plasmid construct and a single DNA fragment of 4.622 kbp (Figure 2.1; lanes 3 and 4). Within lane 4 partial digestion occurred and uncut conformations of plasmid DNA are seen at ~12 kbp. Additionally, to ensure the correct coding sequence for both expression vector as well as recombinant *E. coli* DOXP reductoisomerase protein within the construct, restriction with the enzyme *Nco*I was performed. This enzyme has single cut sites within both vector and the protein coding sequence releasing two DNA fragments of 3599 bp and 1023 bp (Figure 2.1; lane 6).

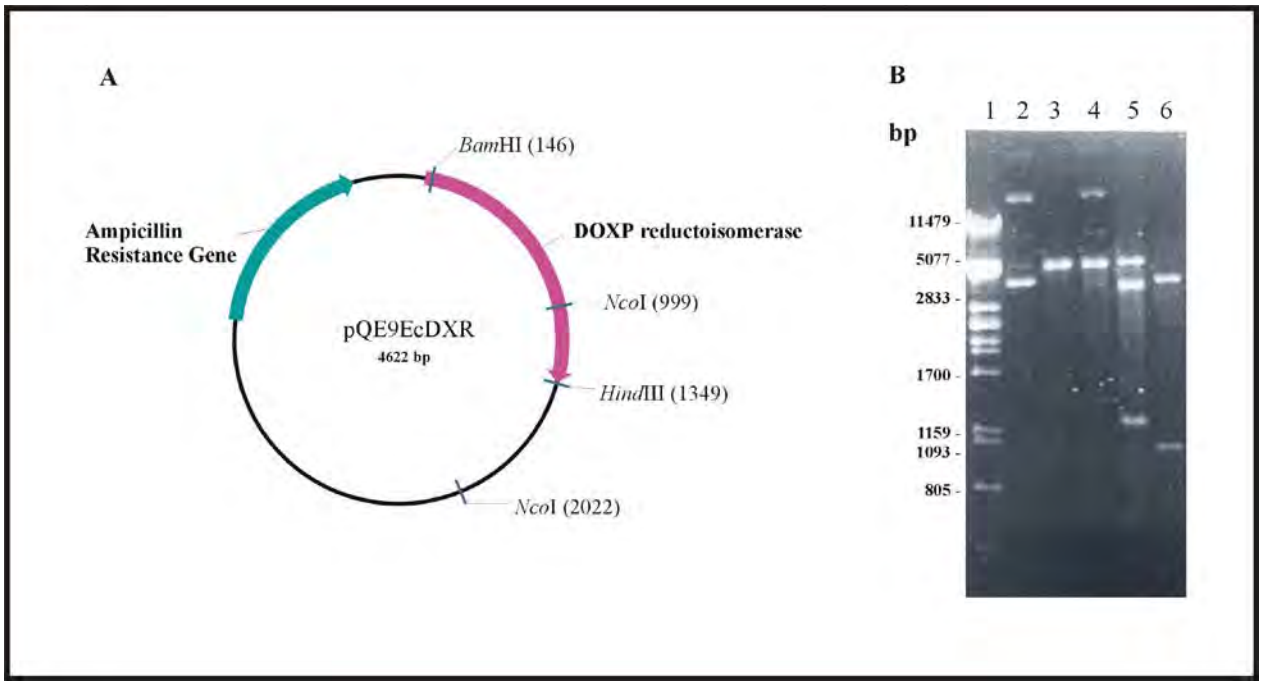


Figure 2.1: Confirmation of the pQE9EcDXR plasmid

(A) Plasmid Map of pQE9EcDXR showing the diagnostic restriction enzyme sites. The *E. coli* DOXP reductoisomerase coding region is shown in pink and the ampicillin resistance marker in blue. (B) 0.8 % Agarose gel showing confirmation of the pQE9EcPfDXR plasmid by diagnostic restriction enzyme analysis. Lane 1 – Lambda DNA restricted with *Pst*I as a molecular marker in base pairs (bp); Lane 2 – Uncut pQE9EcPfDXR; Lane 3 – Linearized plasmid from *Bam*HI restriction resulting in a fragment of 4622 bp; Lane 4 – Linearized plasmid from restriction with *Hind*III resulting in a fragment of 4622 bp and partial digest with an uncut fragment at ~12000 bp; Lane 5 – *Bam*HI and *Hind*III double digest resulting in release of the coding region with two fragments of 34199 bp and 1203 bp. A partial digest fragment of 4622 bp is also present; Lane 6 – *Nco*I restricted pQE9EcDXR resulting in two fragments of 3599 bp and 1023 bp.

2.4.1.2 Over-Production and Purification of *E. coli* DOXP reductoisomerase

The over-production of $(His)_6$- *E. coli* DOXP reductoisomerase in *E. coli* XL1 Blue [pQE9EcDXR] is depicted within Figure 2.2(A). Whole cells were monitored at hourly intervals after the induction of the *E. coli* DOXP reductoisomerase protein with IPTG (1 mM) and the protein induction profile analyzed via discontinuous SDS-PAGE and Coomassie Blue Staining. The production of the recombinant protein reached its optimal level 4-6 hours after induction with IPTG and did not degrade with overnight induction (Figure 2.2(A); lanes 6-8/9). The increase in intensity of the protein fragment of ~ 46 kDa is representative of the $(His)_6$-DXR subunit molecular mass. Low levels of the recombinant protein were present within un-induced whole cells (Figure 2.2(A); lane 2) most likely due to basal levels of transcription occurring. Solubility analysis of the inducible *E. coli* DOXP reductoisomerase was performed to determine protein stability and to guide protein purification set-up. A protein corresponding to the subunit molecular mass of $(His)_6$-EcDXR was obtained within the soluble fractions after sonication and centrifugation of IPTG induced *E. coli* XL1 Blue[pQE9EcDXR] (Figure 2.2(B). The soluble recombinant protein reaches a maximal level after a 30 second sonication pulse (Figure 2.2(B); lane 2). The presence of the recombinant protein within the insoluble fractions most likely represents protein derived from whole unlysed cells that pelleted with the insoluble fraction during centrifugation. Recombinant *E. coli* DOXP reductoisomerase is present within all insoluble fractions due to high levels of over-production and pelleted whole cells during centrifugation.

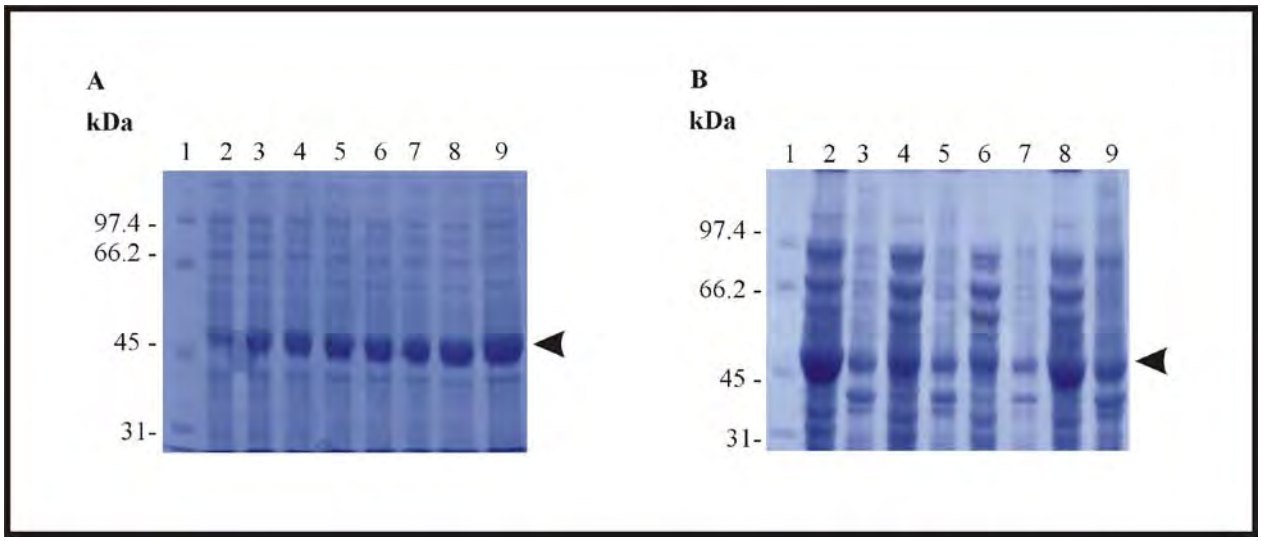


Figure 2.2: Heterologous over-production of soluble *E. coli* DOXP reductoisomerase

(A) Whole cell extracts of pre-induction and IPTG induced logarithmic phase *E. coli* XL1Blue[pQE9EcDXR] cells were analysed by SDS-PAGE. Biorad Broad Range SDS Molecular Weight Marker (kDa) was separated in lane 1. Whole cell samples were removed before induction (lane 2), hourly after addition of 1 mM IPTG (lanes 3-8) and overnight (lane 9). Over-production of *E. coli* DOXP reductoisomerase (~46 kDa) is indicated (arrow). (B) Fractionation study of soluble *E. coli* DOXP reductoisomerase. Biorad Broad Range SDS Molecular Weight Marker (kDa) was separated in lane 1. Soluble fractions were obtained after sonication, for 30 seconds, 60 seconds, 90 seconds, 120 seconds and 150 seconds, and centrifugation (lanes 2, 4, 6 and 8 respectively). A protein of molecular size ~46 kDa represents soluble $(His)_6$-EcDXR (indicated with an arrow). Insoluble fractions containing cell debri were separated within lanes 3, 5, 7 and 9.

From solubility analysis, the protein was regarded both stable and soluble and was successfully purified using native batch nickel-chelate affinity chromatography (Figure 2.3(A)). Nitrilotriacetic acid (NTA; Qiagen) interacts with four of the six binding sites in the coordination sphere of the nickel ion leaving two sites free to bind to the $(His)_6$ protein. The cleared lysate was exposed to the resin overnight at 4˚C to allow for maximal binding. The resin was saturated with $(His)_6$-EcDXR representative of the high levels of unbound protein within the flow through fraction (Figure 2.3(A); lane 7). The DOXP reductoisomerase protein was eluted competitively in the presence of 1 M imidazole (Figure 2.3(A); lane 11-13) with >95 % purity, in sufficient levels for biological assays (~3 mg per litre of culture), however it was also present within the wash steps (Figure 2.3(A); lane 8-10). This is most likely due to weak binding of the recombinant protein to the resin due to the effects of over-saturating the resin. Furthermore, there are low levels of DOXP reductoisomerase still bound to the column after elution, possibly as a result of non-specific interactions as a consequence of protein misfolding (Figure 2.3(A); lane 14). The pooled eluates were buffer exchanged, concentrated and stored in aliquots at -80˚C.

Figure 2.3 (B) represents the western analysis and chemiluminescence-based immunodetection of the purified recombinant $(His)_6$-EcDXR protein. The high molecular mass band detected within lane 4 possibly represents aggregates that were recalcitrant to the SDS/heat treatment for SDS-PAGE analysis.

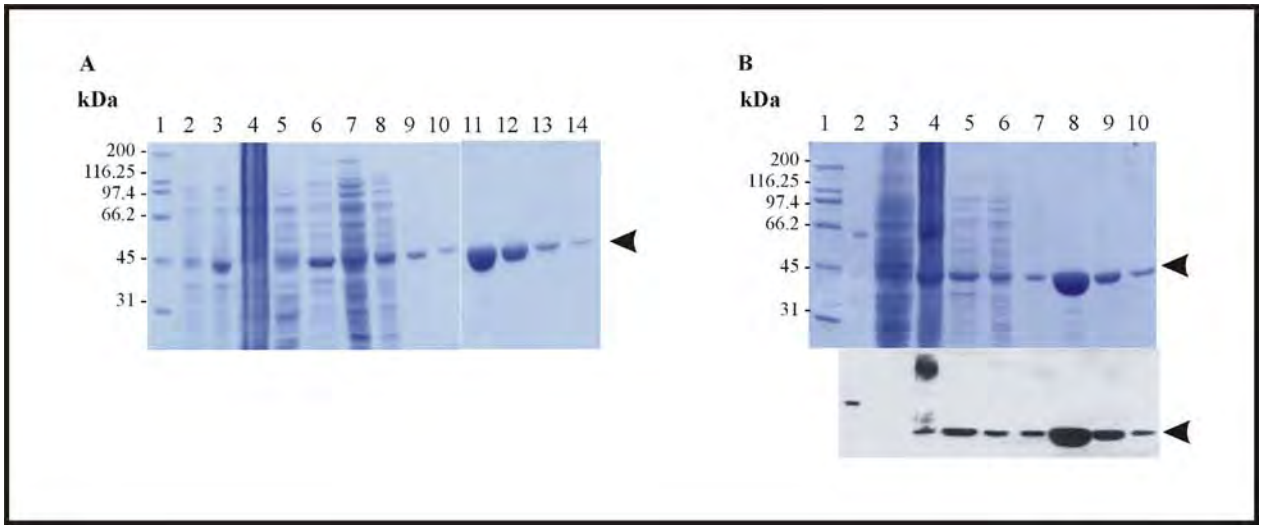


Figure 2.3: Purification of $(His)_6$-EcDXR by nickel affinity chromatography

(A) Purification of $(His)_6$-EcDXR by nickel affinity chromatography was analysed after each purification step using SDS-PAGE. The Biorad Broad Range Molecular Weight Marker (kDa) was separated in lane 1. The purification samples included un-induced whole cell extract (lane 2), induced whole cell extract (lane 3), whole cell lysate (lane 4), insoluble fraction (lane 5), soluble fraction (lane 6), flow-through (lane 7), native wash (lanes 8-10), native elution (lanes 11-13) and Ni-affinity sepharose beads (lane 14). The EcDXR was purified successfully and is indicated with an arrow. (B) SDS-PAGE and Detection of $(His)_6$-EcDXR by Western Analysis and chemiluminescence-based immunodetection. Biorad Markers are separated within Lane 1. Lane 2 and 3 represent a 66 kDa His-tagged protein positive control and *E. coli* XL1Blue whole cell extract as a negative control respectively. The purification samples for the soluble (lane 4), insoluble (lane 5), wash (lane 6-7) and elution fractions (lane 8-10) are detected by the anti-His-antibody within lanes 4 – 10 indicative of the $(His)_6$-EcDXR being present within these samples (indicated with an arrow).

2.4.1.3 The *in vitro* Analysis of *E. coli* DOXP Reductoisomerase Activity

The kinetic parameters for *E. coli* DOXP reductoisomerase were determined using the NADPH coupled enzyme assay. A plot of activity (μmol/min) versus enzyme concentration (μM) indicated that an enzyme concentration of 8.73 μM was satisfactory to ensure linearity (data not shown). The Michaelis-Menten and Lineweaver-Burk plots for *E. coli* DOXP reductoisomerase are shown in Figure 2.4. The reciprocal Lineweaver-Burk (Figure 2.4(B)) plot exhibited K_m and V_{max} values of 282 μM and 0.74 μmol/min/mg respectively. A specific activity of 1.77

μmol/min/mg ± 0.34 was obtained. The analyses for the determination of the kinetic parameters K_m and V_{max} were determined using protein with compromised activity and therefore resulted in a lower V_{max} value calculated than expected. A single independent purification of EcDXR was used to assay for DOXP activity in duplicate. The Michaelis-Menten and Lineweaver-Burk plots were generated using SigmaPlot 8.0 with each data point representing the average of 2 individual experiments on the same purified protein. Kinetic studies were limited due to the substrate not being commercially available and thus a limiting factor.

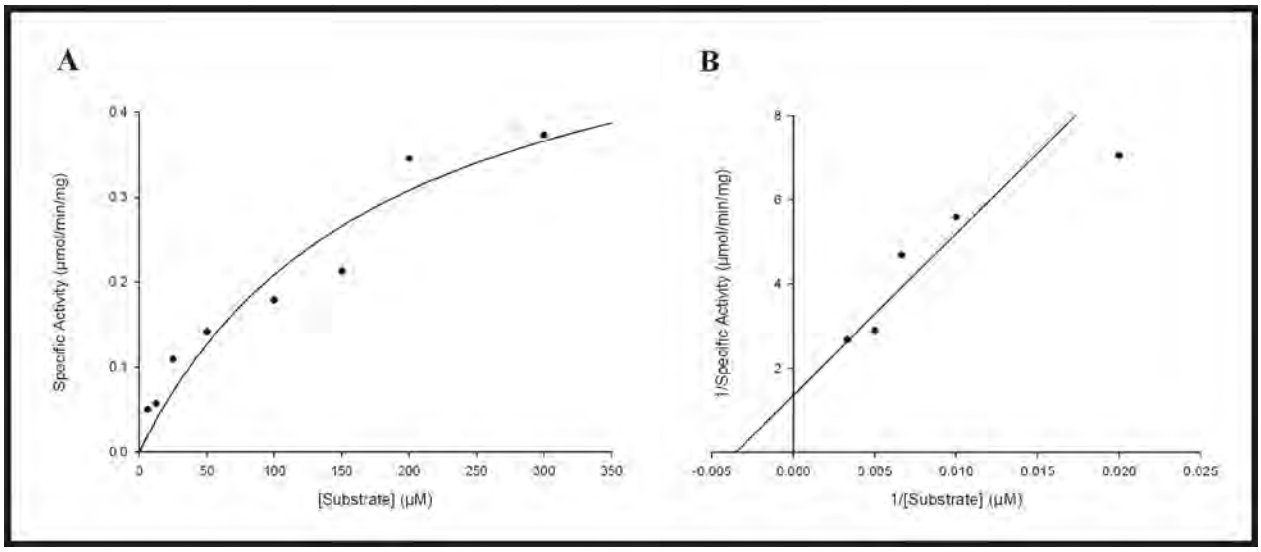


Figure 2.4: Michaelis-Menten and Lineweaver-Burk Plots for the DOXP reductoisomerase activity of EcDXR

The NADPH DOXP reductoisomerase coupled enzyme assay was monitored by the oxidation of NADPH at 340nm and the kinetic parameters were determined using a final enzyme concentration of 0.873 μM and a substrate concentration range of 0 – 300 μM. (A) The Michaelis-Menten plot for *E. coli* DOXP reductoisomerase (B) The reciprocal Lineweaver-Burk plot. Kinetic parameters, K_m of 282 μM and V_{max} of 0.74 μmol/min/mg, were calculated. The Michaelis-Menten and Lineweaver-Burk plots were generated using SigmaPlot 8.0 with each data point representing the average of 2 individual experiments on the same purified protein.

2.4.2 Heterologous Production and Purification of *P. falciparum* DOXP Reductoisomerase

2.4.2.1 Confirmation of the Identity of pQE31PfDXR Plasmid

Work involving the expression of the *P. falciparum* DOXP reductoisomerase protein has been demonstrated in the preparation of purified parasitic DOXP reductoisomerase for inhibitor

analyses, in particular the purification from the plasmid pQE31PfDXR has been reported by Jomaa *et al*. (1999). Within our work we were interested not only in the inhibitor analysis of the parasitic protein in comparison to the *E. coli* protein, but also the biochemical characterization of the parasitic protein as well as the identification of subtle sequence differences between the parasitic DOXP reductoisomerase and its homologues in other organisms that could further elucidate any unique functions of the parasitic protein. The *P. falciparum* coding sequence for DOXP reductoisomerase has been codon-optimized for the expression within the *E. coli* expression system and inserted into the pQE31 expression vector Jomaa *et al*. (1999). This allowed for the production of the $(His)_6$ tagged protein for affinity purification via nickel chelate affinity chromatography. In order to confirm the identity of the pQE31PfDXR plasmid a series of restriction enzyme digests were performed (Figure 2.5(B)). Restriction with *Bam*HI resulted in two fragments of 3436 bp and 1279 bp (Figure 2.5; lane 3) due to the release of a fragment containing the coding sequence of *P. falciparum* DOXP reductoisomerase. To confirm the coding sequence of the recombinant protein, restriction with *Pst*I was performed as it has two restriction sites within the coding sequence and therefore produces two DNA fragments of 4067 bp and 648 bp (Figure 2.5; lane 4). Additionally, the vector sequence was also confirmed through *Sac*I and *Nde*I double restriction digestion yielding two fragments of 2490 bp and 2225 bp (Figure 2.5; lane 7). Restriction with either of these enzymes individually resulted in linearized pQE31PfDXR of size 4.715 kbp (Figure 2.5; lanes 5 and 6).

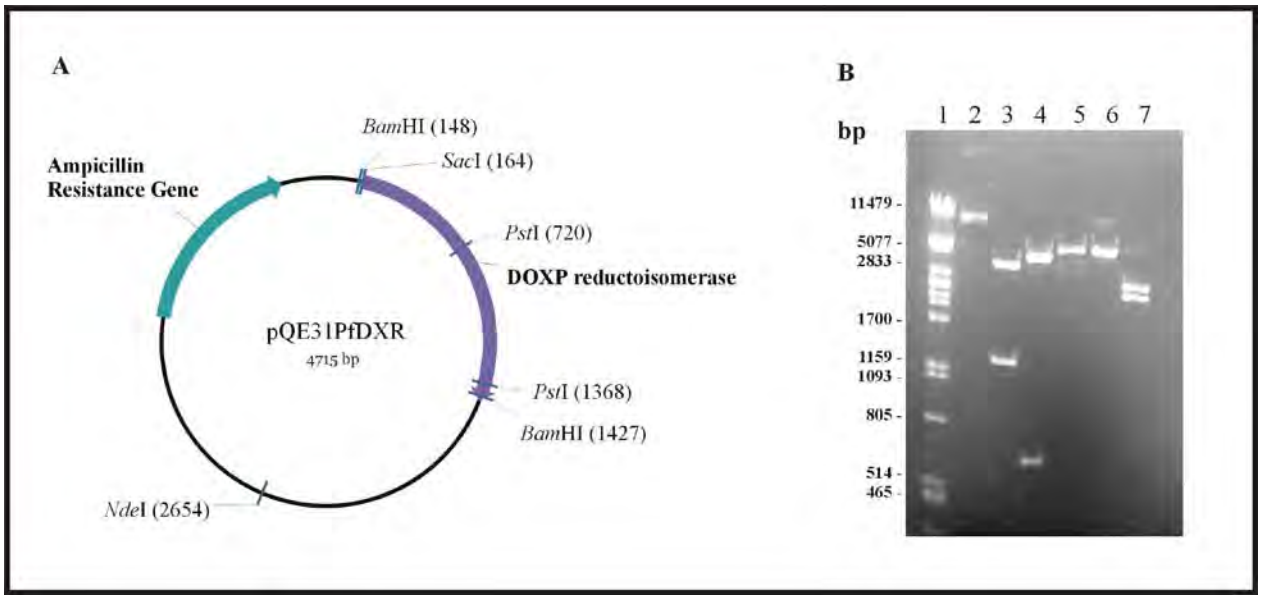


Figure 2.5: Confirmation of the pQE31PfDXR plasmid

(A) Plasmid map of pQE31PfDXR showing the diagnostic restriction enzyme sites. The *P. falciparum* DOXP reductoisomerase coding region is shown in purple and the ampicillin resistance marker in blue. (B) 0.8 % Agarose gel showing confirmation of the pQE31PfDXR plasmid by diagnostic restriction enzyme analysis. Lane 1 – Lambda DNA restricted with *Pst*I as a molecular marker in base pairs (bp); Lane 2 – Uncut pQE31PfDXR; Lane 3 – *Bam*HI restricted pQE31PfDXR resulting in two fragments of 3436 bp and 1279 bp; Lane 4 – *Pst*I restricted pQE31PfDXR resulting in 2 fragments of 4067 bp and 648 bp; Lane 5 – Linearized pQE31PfDXR by digestion with *Sac*I (4715 bp); Lane 6 – Linearized pQE31PfDXR by digestion with *Nde*I; Lane 7 – *Sac*I and *Nde*I double digest resulting in two fragments of 2490 bp and 2225 bp.

2.4.2.2 Over-Production and Purification of *P. falciparum* DOXP Reductoisomerase

The over-production of $(His)_6$-PfDXR in *E. coli* XL1 Blue [pQE31PfDXR] is depicted within Figure 2.6(A). Whole cells were monitored at hourly intervals after the induction of the recombinant DOXP reductoisomerase protein with IPTG (1 mM) and the protein induction profile analyzed via discontinuous SDS-PAGE and Coomassie Blue Staining. The over-production of the recombinant protein was problematic with very low levels of induction occurring with an optimal level being reached at 2 hours after induction with IPTG (Figure 2.6(A); lane 5). Degradation did not seem to occur with overnight induction (Figure 2.6(A); lane 10). The increase in intensity of the protein fragment of ~ 45 kDa is representative of the $(His)_6$-tagged *P. falciparum* DOXP reductoisomerase subunit molecular mass. Extremely low levels of the recombinant protein were present within un-induced and induced whole cells (Figure 2.6(A); lane 4) are likely due to basal levels of transcription occurring and the protein being toxic toward

the host expression system. Additionally, the low growth rate of the *E. coli* host cells was indicative of a toxic protein being produced that inhibited cell growth (data not shown). Figure 2.6 (B) represents the detection of the recombinant $(His)_6$-tagged protein by western analysis and chemiluminescence-based immunodetection. Solubility analysis of the inducible $(His)_6$-PfDXR was performed to determine protein stability and to guide protein purification set-up. A protein corresponding to the subunit molecular mass of $(His)_6$-PfDXR was obtained within the soluble fractions after sonication and centrifugation of IPTG induced *E. coli* XL1 Blue[pQE31PfDXR] (Figure 2.6(C); lane 7-8). The soluble recombinant protein reaches a maximal level after 120 seconds of sonication with 30 second bursts (Figure 2.6(C); lane 6). Unlike the *E. coli* DOXP reductoisomerase protein, the *P. falciparum* DOXP reductoisomerase doesn't seem to be appearing in the pellet fraction in large amounts. This is most likely accountable due to its low level of induction and over-production.

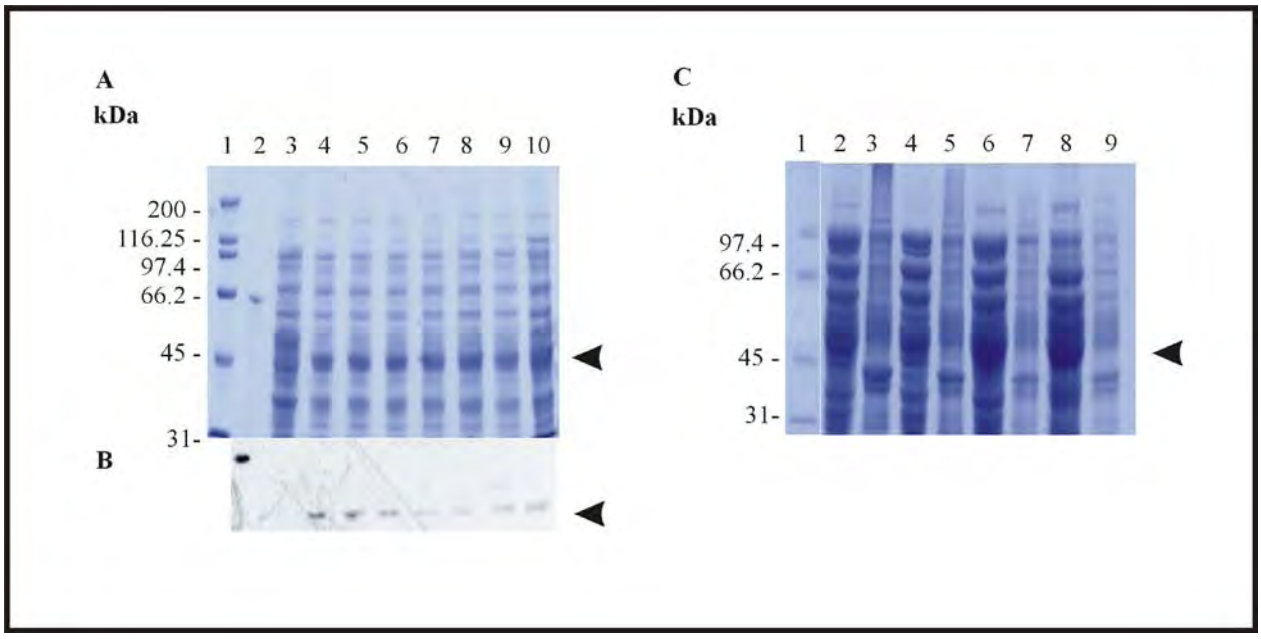


Figure 2.6: Heterologous production of soluble *P. falciparum* DOXP reductoisomerase

(A) Whole cell extracts of pre-induction and IPTG induced logarithmic phase *E. coli* XL1Blue[pQE31PfDXR] were analysed by SDS-PAGE. Biorad Broad Range SDS Molecular Weight Marker (kDa) was separated in lane 1. Whole cell samples were removed before induction (lane 4), hourly after addition of 1 mM IPTG (lanes 5-9) and overnight (lane 10). Over-production of *P. falciparum* DOXP reductoisomerase (~44 kDa) is indicated (arrow). Lane 2 and 3 represent a positive control $(His)_6$ protein of ~66 kDa and a negative control for western analysis. (B) Detection of $(His)_6$ tag on $(His)_6$-PfDXR by western analysis and chemiluminescence-based immunodetection. Lane 2 and 3 represent a 66 kDa His-tagged protein positive control and *E. coli* XL1Blue whole cell extract as a negative control respectively. Un-induced pQE31PfDXR and induced pQE31PfDXR fractions are detected by the His-antibody within lanes 4 – 10 indicative of the $(His)_6$-PfDXR being present within these samples. (C) Heterologous over-production of soluble *P. falciparum* DOXP reductoisomerase. Biorad Broad Range SDS Molecular Weight Marker (kDa) was separated in lane 1. Soluble fractions were obtained after sonication, for 30 seconds, 60 seconds, 90 seconds, 120 seconds and 150 seconds, and centrifugation (lanes 2, 4, 6 and 8 respectively). A protein of molecular size ~44 kDa represents soluble $(His)_6$-PfDXR (indicated with an arrow). Insoluble fractions containing cell debris were separated within lanes 3, 5, 7 and 9. Recombinant *P. falciparum* DOXP reductoisomerase is present within the insoluble fractions to a larger extent than in the soluble fractions however, solubility increased relative to sonication period.

From the solubility analysis the recombinant protein was regarded as stable and soluble which has also been reported by Jomaa *et al*. (1999). However, successful purification of this protein proved to be challenging due to the low levels of induction of the recombinant protein. Initially attempts were to purify the protein through native batch purification that proved successful in the *E. coli* DOXP reductoisomerase purification. This purification strategy failed, however small levels of recombinant protein were appearing within the wash steps of the purification (data not shown). From several attempts to purify *P. falciparum* DOXP reductoisomerase via this strategy

it was decided to attempt native purification using sodium phosphate purification buffers. This attempt at native purification resulted in low levels of pure *P. falciparum* (data not shown).

Buffer composition is important when considering whether the buffer will interact with the protein being purified negatively or with other buffering components. This may in turn impact the stability or even solubility of the protein/enzyme being purified. Many components can be added to buffers in order to improve protein stability depending on the purification procedure and recombinant protein. Thiol compounds are often incorporated within nickel affinity purification procedures to protect protein thiols from oxidation. Dithiothreitol is the choice over 2-mercaptoethanol, due to its higher stability and lower redox potential. Jomaa *et al*. (1999) reported the purification of *P. falciparum* using 2-mercaptoethanol within the purification buffers as a stabilizing agent. This purification procedure was attempted however yielded small amounts of protein within the wash steps of the purification along with several contaminating proteins. No *P. falciparum* DOXP reductoisomerase protein could be detected in the elution samples upon SDS-PAGE analysis (data not shown).

A combination of denaturing and native purification was then attempted (Figure 2.7). The cell lysis was achieved using denaturing conditions in the presence of 8 M Urea followed by native wash to remove the urea, allowing the proteins to renaturate. Elution of *P. falciparum* DOXP reductoisomerase was achieved by competitive elution with imidazole. This was only partially successful with a large majority of protein still bound to the Ni-affinity sepharose beads (Figure 2.7; lane 14). A wide range of proteins other than *P. falciparum* DOXP remained bound to the beads as well (Figure 2.7; lane14). This was only noted to occur under the denaturing purification conditions where there were lower levels of expressed recombinant protein. Under such circumstances the matrix would not have been saturated with *P. falciparum* DOXP reductoisomerase and histidine rich denatured proteins would have attempted to bind to the matrix through specific interactions. Furthermore, denatured proteins would also have bound to the matrix through hydrophobic and electrostatic interactions and would not have been eluted from the Ni-affinity beads easily. Attempts at improving the concentration of protein competitively released from the beads by an increase in the concentration gradient of imidazole was unsuccessful. The difficulties experienced in eluting the *P. falciparum* DOXP

reductoisomerase could be as a result of the protein binding non-specifically to the sepharose matrix of the column, making it difficult to elute. Additional attempts at eluting the recombinant protein using 0.1% SDS and attempting to regain enzymatic activity by refolding the protein could be investigated. A possible reason for the difficulty in obtaining sufficient *P. falciparum* DOXP reductoisomerase could be due to it being toxic toward the host expression system. One solution to this problem could be to express the protein together with the pLysS or pLysE plasmids that are engineered to allow the stable production of toxic proteins within *E. coli*. Furthermore, altering the expression strain may also be an alternative. The *E. coli* M15[pREP4] strain allows for the suppression of toxic protein expression prior to induction thus allowing for an improvement in the expression of recombinant toxic proteins. Utilising western analysis the $(His)_6$-PfDXR protein could only be detected after overnight exposure using chemiluminescence (data not shown). Insufficient amounts of active enzyme were purified for accurate enzyme kinetics analysis to be conducted.

Figure 2.7: Purification of $(His)_6$PfDXR by nickel affinity chromatography

(A) Purification of $(His)_6$-PfDXR by nickel affinity chromatography was analysed after each purification step using SDS-PAGE. The Biorad Broad Range Molecular Weight Marker (kDa) was separated in lane 1. The purification samples included un-induced whole cell extract (lane 2), induced whole cell extract (lane 3), whole cell lysate (lane 4), insoluble fraction (lane 5), soluble fraction (lane 6), flow-through (lane 7), native wash (lanes 8-10), native elution (lanes 11-13) and Ni-affinity sepharose beads (lane 14).

2.5 Conclusion

Heterologous production and nickel chelate affinity purification of $(His)_6$ *E. coli* DOXP reductoisomerase was successful. The native purified EcDXR was shown to be active using NADPH coupled enzyme assays with a K_m of 282 μM, V_{max} of 0.74 μmol/min/mg, and specific activity of 1.77 μmol/min/mg ± 0.340 was obtained.

The purification of *P. falciparum* DOXP reductoisomerase was problematic with the combinatory denaturing and native purification strategy being the only successful purification yielding low amounts of pure protein. The *P. falciparum* DOXP reductoisomerase was found to remain bound to the Ni-affinity sepharose beads even after treatment with increased imidazole concentrations. Due to low levels of pure *P. falciparum* DOXP reductoisomerase being purified, credible enzyme assays were not possible. It was regarded that the recombinant protein had poor expression levels due to it being toxic toward the host cells, which resulted in poor yields within the purification strategies attempted. The low growth rate of the *E. coli* host cells upon heterologous over-production of *P. falciparum* DOXP reductoisomerase further suggested that the *P. falciparum* DOXP reductoisomerase recombinant protein was toxic toward the host cell. A possible solution would be to express the protein in *E. coli* containing either the pLysS or pLysE plasmids suited for expression of toxic proteins within *E. coli.* Furthermore, expression within differing strains of *E. coli* may improve the expression level. This is especially the case for the *E. coli* M15[pREP4] strain, which contains the pREP4 repressor plasmid for suppression of the recombinant protein expression prior to induction. A further consideration may involve the lack of correct folding of the *P. falciparum* DOXP reductoisomerase protein within its "unnatural" cellular environment, *E. coli*. Co-expression of parasitic partner chaperone proteins may improve the folding mechanisms and thus expression of recombinant *P. falciparum* DOXP reductoisomerase. An additional consideration is that the pQE line of plasmids may be flawed for the expression of malarial genes in *E. coli* due to the absence of the *lac* repressor gene in the plasmid backbone, thus leading to inefficient suppression of the plasmid promoter and high background levels of protein expression in the absence of IPTG. A range of different expression plasmids that contain the repressor gene in addition to the Amp resistance gene and expression cassette could be used to improve expression levels. Furthermore, the possibility of using

alternative expression and purification strategies could be explored. This may involve using the GST-fusion approach along with attempts at expression within other heterologous systems (e.g. insect cells).

CHAPTER 3: CHARACTERIZATION OF THE DOXP REDUCTOISOMERASE CATALYTIC ACTIVE SITE BY RATIONAL AMINO ACID SUBSTITUTION ANALYSIS

3 CHARACTERIZATION OF THE DOXP REDUCTOISOMERASE CATALYTIC ACTIVE SITE BY RATIONAL AMINO ACID SUBSTITUTION ANALYSIS

3.1 Introduction

The function of any protein relies strongly on its primary amino acid sequence. Engineered changes/mutations in the genetic code of the cloned gene can create a method for generating structural/functional changes to the protein itself. Advances within recombinant DNA technology have allowed for the ability to manipulate DNA, thereby allowing researchers to manipulate regulatory elements and gene products (reviewed in Ling and Robinson, 1998). The modification of the genetic code often encompasses research in the fields of protein structure-function relationships, gene expression and vector modifications. Additionally, modifications of a specific protein are often beneficial in medical/industrial applications in the improvement of specific biochemical properties.

Bioinformatics is at the forefront of modern gene and protein research. With whole genome sequencing a reality, bioinformatics now holds the key to the identification of genes, proteins and their secondary structure and functional roles. Proteins are vital for all biological processes with the primary amino acid sequence holding all the necessary structural information for its functional conformation. Several computational methods for nucleic acid and amino acid sequence analysis are available from which structural and functional predictions can be made. Genomes are being published rapidly and the genomes of several model organisms have been sequenced and include the yeast *Saccharomyces cerevisiae*, the bacterium *E. coli*, the worm *Caenorhabditis elegans*, the fruit fly *Drosophila melanogaster* as well as the human genome. From these and other sequenced genomes there is an enormous amount of information available for biological analysis to uncover ancestry, gene function, and protein identification, secondary structure prediction and protein function.

Bioinformatic analysis of DOXP reductoisomerase has been performed by several research groups with the most significant being Jomaa *et al*. (1999), Reuter *et al*. (2002), Yajima *et al*. (2002) and Steinbacher *et al*. (2003). Jomaa *et al*. (1999) initially analysed the primary protein

amino acid sequence to identify the gene encoding DOXP reductoisomerase within *P. falciparum*. From the multiple sequence alignment they identified an extended N-terminal sequence that was present only in the bacterial and blue algae sequences. This leader sequence was found to be 74 amino acids in length with the first 30 amino acids representing an endoplasmic reticulum signal peptide and the remainder a plastidal target sequence. Additionally, the N-terminal leader sequence was shown to direct localization to the extranuclear apicoplast, verifying DOXP reductoisomerase as an apicoplast destined protein. Reuter *et al*. (2002), Yajima *et al*. (2002) and Steinbacher *et al*. (2003) have crystalized the *E. coli* DOXP reductoisomerase. The protein has been shown to be a homodimer in solution with a predicted monomer size of ~43.5 kDa. In addition to secondary structure predictions, several catalytically and structurally significant residues within the protein sequence have been identified, some of which have been experimentally tested, and are summarized in Table 3 and will be referred to/discussed in further sections.

Residues that occur within the substrate binding domain, and are conserved amongst DOXP reductoisomerase homologues, were analysed via site-directed mutagenesis techniques by Reuter *et al*. 2002 and Yajima *et al*., 2002. The Gly^{14} to Asp mutation resulted in insoluble protein and was reported as essential in the binding of the NADPH cofactor (Reuter *et al*., 2002). The residues Glu^{231}, His^{153}, His^{209}, and His^{257} were found to be strictly conserved and cluster within the substrate binding region, within the catalytic domain, and were analysed for their functional roles in substrate binding. It was reported that the Glu^{231} to Lys mutation resulted in unaffected affinity of the protein toward substrate and NADPH. However, the k_{cat} was significantly decreased indicative of this residue mainly involved in catalysis rather than substrate or cofactor binding (Reuter *et al*., 2002). The k_{cat} decreased substantially on the His^{153} to Gln mutation suggestive of this residue being vital in the catalysis, while substrate binding was shown to be mediated by the His^{209} and His^{257} residues (Reuter *et al*., 2002). Yajima *et al.* (2002) have also previously conducted mutagenesis analyses upon His^{209}, His^{153}, His^{257} and Glu^{231} residues and reported that these residues were all involved in catalysis. These residues are located with their side chains projecting within the pocket formed by the N-terminal, central domain and C-terminal of the protein allowing it to be identified as the catalytic pocket of DOXP reductoisomerase (Yajima *et al*., 2002).

Table 3: Predicted critical residues within DOXP reductoisomerase

Identified Amino Acid/s	Predicted Role	Experimentally Proven	Reference
Thr^{10}	Anchors 2` phosphate moiety of NADPH through hydrogen bonding		Yajima *et al*. (2002)
Gly^{11} and Ser^{12}	Co-ordinate the pyrophosphonate moiety of NADPH		Reuter *et al*. (2002)
Gly^{14}	Essential in construction of NADPH binding motif	Gly^{14} to Asp resulted in an insoluble protein	Reuter *et al*. (2002)
Ala^{35} and Gly^{36}	Position to accommodate 2` phosphate moiety of NADPH		Reuter *et al*. (2002)
Asp^{150}, Glu^{152}, Glu^{231}, and Glu^{234}	Divalent Cation Binding and form salt bridges with acidic residues Lys^{125}, His^{153}, and Lys^{228}		Reuter *et al*. (2002); Yajima *et al*. (2002); Steinbacher *et al*. (2003)
Glu^{231}	Involved in enzyme catalysis and *Trans* to N-formyl oxygen on Fosmidomycin binding	Glu^{231} to Lys resulted in a change for k_{cat} but no change in K_m	Yajima *et al*. (2002); Steinbacher *et al*. (2003)
Asp^{150}	*Trans* to N-hydroxyl oxygen on Fosmidomycin binding		Steinbacher *et al*. (2003)
Lys^{125}, His^{153}, and Lys^{228}	Acidic residues forming salt bridges with basic divalent cation binding residues listed above		Reuter *et al*. (2002)
His^{153}, His^{209}, and His^{257}	Involved in enzyme catalysis	His^{209} and His^{257} to Gln resulted in decreased affinity toward substrate. His^{153} to Gln results in 10 fold decrease in k_{cat}	Reuter *et al*. (2002); Yajima *et al*. (2002)
His^{209}	Essential residue in catalysis and forms a hydrogen bond between itself and the phosphonate function of the enzyme and in so doing closes the catalytic hatch	His^{209} to Gln resulted in k_{cat}/K_m decreasing substantially	Yajima *et al*. (2002)
His^{257}	Phosphonate moiety binding mediated by catalytic hatch	His^{257} to Gln resulted in K_m for NADPH increasing and k_{cat}/K_m decreasing	Steinbacher *et al.* (2003)
His^{257}, Trp^{212} and Pro^{274}	His^{257} and Trp^{212} from a stacking role, which is supported by Pro^{274} thus ordering the active site structure only in the presence of NADPH. Also interact with carbon backbone of substrate		Yajima *et al*. (2002); Steinbacher *et al*. (2003)
Ser^{186}, Ser^{222}, Asn^{227}, and Lys^{228}	Substrate phosphonate moiety anchoring via hydrogen bonding interactions		Steinbacher *et al*. (2003)
Trp^{212} and Met^{214}	Stacking functions essential in forming hydrophobic contacts with nicotinamide moiety of NADPH		Steinbacher *et al*. (2003)

3.1.1 Homology Modelling of Protein Sequences

Knowledge of the three dimensional (3-D) structure of a protein is essential for rational amino acid analysis via site-directed mutagenesis techniques and can be vital for the design of protein-specific drugs. Structural information can enhance our understanding of protein function. However, many proteins 3-D structures are still not available. Many proteins fail to crystallize and often cannot be obtained or dissolved in large enough quantities for NMR analysis. In such cases model building with known 3-D structural data of a homologous protein is at present the only reliable method to obtain structural information. Homology modelling involves the prediction of a protein's 3-D structure from the primary amino acid sequence of a homologous protein for which there is either X-ray or NMR structural information available. The Protein Data Bank (PDB) (Westbrook *et al*., 2003) is a database of 3-dimensional structural information from which the structural information for the template/homologous protein can be obtained. Creating a homology model of a target protein is useful in that well-defined secondary structure predictions individually do not illustrate the 3-D structure of a protein. The homology model fulfils this role and has several applications within catalytic mechanism analysis, ligand docking analysis, the identification of antibody epitopes, rational amino acid site-directed mutagenesis, and identification of functional sites (www.cmbi.kun.nl/gvteach/hommod/index.shtml).

Homology modelling proceeds via a multi-step process. Initially, the template protein must be recognized and consequently the primary amino acid sequence for the template and protein to be modelled are aligned and optimized through a multiple sequence alignment. After which, the backbone for the model is generated and *ab initio* building of canonical loops is performed. Template side chains are then replaced with model side chains and optimized. Modelling of insertions and deletions is then performed and the model detected for errors and optimized. Iterating of the steps to remove errors is conducted after which the model undergoes energy minimization and dynamic simulation methods (www.expasy.org/swissmod/SWISS-MODEL and www.cmbi.kun.nl/gvteach/hommod/index.shtml). Homology modelling is dependent on the quality of the 3-D structural data of the template protein as well as sequence identity between the template and the protein to be modelled. The choice of the modelling template is crucial and high deviations in the model can be the cause of the modelling template. Additionally the

presence or absence of substrate, cofactors or subunits from the template structural information can impact the experimental protein model (www.cmbi.kun.nl/gvteach/hommod/index.shtml). As differences within the 3-D structure increase with the decrease in sequence identity the homology modelling accuracy deteriorates. With >90 % sequence identity such errors are low however, with 50 % sequence identity root mean square (RMS) errors in the modelled co-ordinates can be as large as 1.5 Å with large local errors. 25 % Sequence identity will introduce problems at the alignment level resulting in large errors and homology is undetected with sequence identity below 25 % (www.cmbi.kun.nl/gvteach/hommod/index.shtml).

Several assumptions regarding the overall three dimensional structure and residue topology are made in the process of creating a homology model. The main assumption is that the overall 3-D structure of the target protein is similar to that of the template with regions of homology having similar structure. Furthermore, when constructing the predicted 3-D model, residues that are homologous within a family of proteins are regarded as structurally conserved and those crucial in the biological function are thought to be topologically equivalent. Loop regions are often not constructed as conserved regions and are regarded as flexible allowing for deletions and insertions to be made within the model. In so doing one assumes that such deletions and insertions do not impact the global structure of the protein and more importantly that the loop region is not performing a biological role within the protein (www.cmbi.kun.nl/ gvteach/hommod/index.shtml).

3.1.2 Approaches to Rational Protein Engineering/Mutagenesis

Several methods have been employed and optimized to mutate DNA. Originally, random techniques were used for manipulating DNA, often involving the generation of random mutations in chromosomal DNA through exposure to X-rays or various chemicals (Ling and Robinson, 1998). These techniques were beneficial in past genetic studies, but were limited in their inability to target the mutation toward a specific gene. An additional drawback was the intensive screening procedures. Recombinant DNA technology now allows for the ability to target DNA *in vitro* and make precise changes within a specific gene.

3.1.3 Site-Directed Mutagenesis

Site-directed mutagenesis is a versatile technique in rational amino acid substitution and allows for a mutation to be engineered at a specific site within a desired gene (Madigan *et al*., 1997). Such methodologies are divided into i) techniques that restructure fragments of DNA, such as cassette mutagenesis; ii) localized random mutagenesis; and iii) oligonucleotide-based mutagenesis (Cosby and Lesley, 1997). Oligonucleotide-based mutagenesis is the most common for introduction of small specific mutations and includes PCR and Non-PCR based approaches. Unlike PCR based approaches in which the product is able to act as a template resulting in exponential amplification, in Non-PCR only the parental plasmid can act as template resulting in a linear amplification and lower error rates. This method is preferred over PCR-based methods for mutagenesis as incorporation of errors is avoided and was used within this work.

Oligonucleotide mutagenesis can be used to generate point mutations, change amino acids, and delete or insert single or several amino acids. The success of the technique relies on the construction of the synthetic oligonucleotide, which is designed to incorporate the desired mutation(s). Parameters such as annealing efficiency and temperature, primer length, and GC content must be considered in ensuring that the oligonucleotide anneals to the plasmid DNA and initiates DNA synthesis of the mutant strand (Cosby and Lesley, 1997).

Site-directed mutagenesis within this work was achieved using two synthetic complementary oligonucleotide primers containing the desired mutation(s) (Jung *et al*., 1992). The oligonucleotides anneal to the opposite strands of the double-stranded template vector DNA and are extended such that successive cycles of annealing and extension results in linear amplification of the whole plasmid. The parental plasmid DNA is removed by restriction with *Dpn*I, which digests the template parental DNA as it is specific for methylated DNA. The resultant mutant plasmid product can then be propagated within a suitable host organism and the DNA screened for the desired mutation.

3.2 Aim

Several amino acid residues have been considered essential in the function and structure of DOXP reductoisomerase. Specifically, the catalytic hatch region within the protein has been regarded to enclose the substrate within the catalytic pocket and therefore not only play a role within substrate binding but active site architectural orientation. Reuter *et al.* (2002), Yajima *et al*. (2002) and Steinbacher *et al.* (2003) have noted the importance of the conserved Trp^{212}, Met^{214} and Pro^{274} residues within the catalytic hatch, however they have not been experimentally determined. Upon 3-D inspection of these residues within the *E. coli* DOXP reductoisomerase structure, amino acid substitution experiments were designed to test the functional predictions for these amino acids.

The Met^{214} residue was replaced with an isoleucine residue to ensure that overall global disruption to the protein was avoided along with the removal of the sulphur group and length of the Met side chain. This would remove possible stacking interactions previously mediated by Met^{214} with Pro^{274} and Trp^{212} and interaction with NADPH. The Pro^{274} was mutated to an isoleucine in order to maintain bulk (to an extent), but interrupt any stacking role the cyclic Pro residue may perform with Trp^{212} and Met^{214}. Trp^{212} substitution with tyrosine, leucine and phenylalanine will allow for the stacking and substrate binding interactions of Trp^{212} to be interrupted. The tyrosine residue introduced a polar hydroxyl group with potential for hydrogen bonding interactions while simultaneously removing the 5-membered ring previously present within the Trp side chain. This would likely affect the stacking role of the residue in the active site as well as any substrate binding interactions. The leucine and phenylalanine residues are both hydrophobic with phenylalanine maintaining a degree of aromaticity. These substitutions would be less detrimental to the hydrophobic nature of the active site and would hopefully elucidate the precise role of the Trp^{212} residue in stacking with Met^{214} and Pro^{274} and substrate binding.

A multiple sequence alignment was used to determine the differences within the *P. falciparum* DOXP reductoisomerase in comparison to *E. coli* and other DOXP reductoisomerase homologues possibly representative of structural/functional differences in the parasitic protein. Homology

modeling of the *P. falciparum* DOXP reductoisomerase, using the *E. coli* DOXP reductoisomerase homologue's primary amino acid sequence and crystal structure data as a template, was performed to obtain the 3-D predicted structure for *P. falciparum* DOXP reductoisomerase. The 3-D predicted structure of *P. falciparum* DOXP reductoisomerase allowed for future biochemical characterization through rational amino acid substitution analysis.

3.3 Experimental Procedures

3.3.1 Reagents and Chemicals

All reagents and chemicals utilized are described in the Appendices (Chapter 6, section 6.1).

3.3.2 Multiple Sequence Alignment Analysis of the Primary Amino Acid Sequences of *E. coli* and *P. falciparum* DOXP Reductoisomerase

Homologous protein sequences were identified using the Basic Local Alignment Search Tool (BLAST)(Altschul *et al*., 1997) and downloaded from the National Center for Biotechnology (NCBI) server (http://ncbi.nlm.nih.gov). Multiple sequence alignment was performed using the software programme ClustalX (Jeanmougin *et al*., 1998).

3.3.3 Homology Modelling of *P. falciparum* DOXP Reductoisomerase

Homology modelling was used to generate a predicted 3-D structure for the *P. falciparum* DOXP reductoisomerase and was performed using SWISS-MODEL (Peitsch, 1996) and WHATIF (Vriend, 1990). Energy minimization and dynamic simulation were performed using the GROMOS 96 method within SWISS-MODEL and via the programme WHATCHECK within WHATIF. The template protein, i.e. the *E. coli* DOXP reductoisomerase homologue, primary amino acid sequence and structural information, was downloaded from the PDB (Westbrook *et al*., 2003). Protein models/structures were visualized using Molscript (Kraulis, 1991). Two homology models were created for *P. falciparum* DOXP reductoisomerase. The first predicted structure was generated with the crystal structure for the *E. coli* DOXP reductoisomerase

obtained by Yajima *et al*. (2002)(PDB accession code 1JVS) as a template, in the absence of the cofactor metal cation and substrate. The second predicted 3-D structure of *P. falciparum* DOXP reductoisomerase was generated using structural data for the *E. coli* DOXP reductoisomerase with the substrate, cofactor metal cation, and inhibitor bound obtained by Steinbacher *et al*. (2003)(PDB accession codes 1ONO, 1ONP and 1ONN).

3.3.4 Identification of Potentially Critical Residues for the Catalytic Functioning of *P. falciparum* and *E. coli* DOXP Reductoisomerase

Residues essential in the catalytic functioning of *E. coli* DOXP reductoisomerase identified from crystal structure information and analysed via site-directed mutagenesis techniques by Reuter *et al*. (2002) and Yajima *et al*. (2002) were discussed in section 3.1. From Reuter *et al*. (2002), Yajima *et al*. (2002) and Steinbacher *et al*. (2003), amino acids that were predicted as crucial in the functioning of DOXP reductoisomerase, but not tested experimentally, were noted and chosen for functional analysis via site-directed mutagenesis. A multiple sequence alignment was performed amongst DOXP reductoisomerase homologues to identify the level of conservation of these residues and to show any differences within the *P. falciparum* DOXP reductoisomerase amino acid sequence that could be of functional/structural importance. Inspection of the 3-D orientation of the residues within the *E. coli* DOXP reductoisomerase and *P. falciparum* DOXP reductoisomerase was performed to validate their predicted functional/structural roles through interaction with the substrate, cofactor or other contacts made within 4 Å. Protein structure visualization was performed using Molscript (Kraulis, 1991).

3.3.5 Site-Directed Mutagenesis of DOXP Reductoisomerase

Plasmids encoding $(His)_6$-EcDXR(W212Y), $(His)_6$-EcDXR(W212L), $(His)_6$-EcDXR(W212F), $(His)_6$-EcDXR(M214I), $(His)_6$-EcDXR(P274I), $(His)_6$-EcDXR(N211K, S213K) and $(His)_6$-PfDXR(K295N, K297S) were prepared via site-directed mutagenesis using a double stranded whole plasmid amplification technique as described in Jung *et al*. (1992). Synthetic oligonucleotide primers were designed using Generunner Software (Hastings Software Inc.) and are tabulated in the appendices (Chapter 6, refer to section 6.2, Table 8).

All mutations were generated using the standard procedure outlined in the QuikChange Site-Directed Mutagenesis Kit supplied by Stratagene (USA) with the respective complementary synthetic oligonucleotide primer pairs. The linear amplification products representative of successful mutagenesis were transformed into supercompetent *E. coli* XL1-Blue cells (Stratagene, USA) and the putative mutant plasmid DNA isolated. Diagnostic restriction enzyme analysis was performed on the plasmids and samples resolved upon agarose gel electropheresis, as described in section 2.3.2.3. Table 4 illustrates the respective mutant construct and the diagnostic restriction enzyme to validate successful incorporation of the desired codon substitutions.

Table 4: Diagnostic restriction enzyme analysis of rational amino acid substitutions within DOXP reductoisomerase

Plasmid Construct	Diagnostic Restriction Enzyme	Expected DNA Fragments (bp)
pQE9EcDXR	*Nco*I	3599 & 1023
pQE9EcDXR(W212Y) pQE9EcDXR(W212L) pQE9EcDXR(W212F)	*Nco*I	3389, 1023, & 210
pQE9EcDXR	*Xma*III	Uncut DNA conformations
pQE9EcDXR(M214I)	*Xma*III	4622
pQE9EcDXR	*Cla*I	Uncut DNA conformations
pQE9EcDXR(P274I)	*Cla*I	4622
pQE9EcDXR(N211K)	*Nco*I	3389, 1023, & 210
pQE9EcDXR(N211K, S213K)	*Nco*I	3599 & 1023
pQE31PfDXR	*Nco*I	2771 & 1944
pQE31PfDXR(K297S)	*Nco*I	2771, 1263 &681
pQE31PfDXR(K295N, K297S)	*Nco*I	2771 & 1944

For the preparation of $(His)_6$-EcDXR(N211K, S213K) plasmid DNA, a two-stage mutagenesis reaction was used to introduce the N211K mutation followed by the S213K mutation. For the generation of the pQE9EcDXR(N211K) plasmid DNA a complementary 27-mer primer pair (refer to the appendices, Chapter 6, section 6.2, Table 8) was synthesised to contain the codon for the lysine amino acid change as well as a silent mutation to introduce the restriction site for *Nco*I. The parental plasmid DNA restricted with *Nco*I results in fragments of 3599 bp and 1023 bp. Through the introduction of a third *Nco*I site the pQE9EcDXR(N211K) plasmid DNA results in

three fragments of 3389 bp, 1023 bp, and 210 bp upon *Nco*I restriction. Samples were resolved using agarose gel electrophoresis, as described in section 2.3.2.3, on a 1 % agarose gel containing 0.5 μg/ml ethidium bromide in 0.5x TBE buffer at 100V. For the incorporation of the lysine residue at 213 a set of complementary primers were synthesised to the pQE9EcDXR(N211K) plasmid DNA (refer to the appendices, Chapter 6, section 6.2, Table 8). Through the introduction of the codon for this amino acid substitution the *Nco*I site engineered through a silent mutation during the lysine codon 211 mutagenesis is removed. Thus identification of pQE9EcDXR(N211K, S213K) by restriction with *Nco*I yields two fragments of sizes 3599 bp and 1023 bp, indicative of the $(His)_6$-EcDXR(N211K, S213K) plasmid DNA. Samples were resolved using agarose gel electrophoresis, as described in section 2.3.2.3, on a 1 % and 0.8 % agarose gel containing 0.5 μg/ml ethidium bromide.

To prepare the $(His)_6$-PfDXR(K295N, K297S) plasmid DNA a two-stage mutagenesis reaction was used. The introduction of a codon for serine at codon 297 was conducted first, and followed by the introduction of a codon for an asparagine residue at codon 295. For the K297S mutation a complementary primer pair was synthesised (refer to the appendices, Chapter 6, section 6.2, Table 8) to comprise the required amino acid change and simultaneously introduce an *Nco*I restriction site. Thus pQE31PfDXR(K297S) could be identified due to three fragments of 2771 bp, 1263 bp and 681 bp compared to only two fragments appearing (2771 bp and 1944 bp) for the parental plasmid DNA upon restriction with *Nco*I. Samples were resolved using agarose gel electrophoresis, as described in section 2.3.2.3, upon a 0.8 % agarose gel containing 0.5 μg/ml ethidium bromide in 0.5x TBE buffer at 100V. The introduction of the codon for an asparagine residue at codon 295 was achieved through the synthesis of a complementary oligonucleotide primer pair (refer to the appendices, Chapter 6, section 6.2, Table 8). Additionally a silent mutation was incorporated to remove the *Nco*I site engineered through the first round of mutagenesis. Upon restriction with *Nco*I the pQE31PfDXR(K295A, K297S) plasmid DNA results in two DNA fragments of 2771 bp and 1944 bp whereas the pQE31PfDXR(K297S) plasmid DNA results in three DNA fragments of 2771 bp, 1263 bp and 681 bp. Samples were resolved using agarose gel electrophoresis, as described in section 2.3.2.3, upon a 1 % and 0.8 % agarose gel containing 0.5 μg/ml ethidium bromide in 0.5x TBE buffer at 100V.

The QuikChange Site-Directed Mutagenesis Kit (Stratagene, USA) was utilized to generate all mutant plasmid constructs according to the reaction scheme outlined in Table 5. A control reaction using the pWhitescript™ 4.5 kb control plasmid is used as a mutagenesis efficiency control. The pWhitescript™ plasmid contains a stop codon within the β-galactosidase gene where a glutamine codon is usually found. The oligonucleotide control primers contain a point mutation that substitutes the T residue within the TAA stop codon to a C residue. This reverts the codon sequence to the wild-type glutamine codon allowing for colonies to be screened after transformation for the β-galactosidase phenotype (β-gal$^+$, blue).

Table 5: Preparation of plasmid DNA site-directed mutagenesis reactions

Volume (µl)	Reagent	Reagent Details
Control Reaction:		
5	10x Reaction Buffer	100 mM KCl, 100 mM $(NH_2)_2SO_4$, 200 mM Tris-HCl (pH 8.8), 20 mM $MgSO_4$, 1 % Triton X-100, 1 mg/ml nuclease free bovine serum albumin (BSA)
2	pWhitescript™ positive control plasmid	5 µg/µl, 10 ng, 4.5 kb
1.25	Forward Primer	100 ng/µl, 125 ng, 34-mer
1.25	Reverse Primer	100 ng/µl, 125 ng, 34-mer
1	dNTP mix	200 mM
38.5	Sterile deionized milli-Q water	
1	*PfuTurbo*™ DNA polymerase	2.5 U/µl
		Total Volume = 50 µl
Sample Reaction:		
5	10x Reaction Buffer	100 mM KCl, 100 mM $(NH_2)_2SO_4$, 200 mM Tris-HCl (pH 8.8), 20 mM $MgSO_4$, 1 % Triton X-100, 1 mg/ml nuclease free bovine serum albumin (BSA)
2	Plasmid Template DNA	~200 ng/µl
1	Synthesised Forward Primer	200 ng/µl, refer to appendix for sequences
1	Synthesised Reverse Primer	200 ng/µl, refer to appendix for sequences
1	dNTP mix	200 mM
39	Sterile deionized milli-Q water	
1	*PfuTurbo*™ DNA polymerase	2.5 U/µl
		Total Volume = 50 µl

3.3.6 Thermal Cycling Parameters utilized in Site-Directed Mutagenesis

The reaction samples were incubated within the GeneAmp PCR System 9700 (Applied Biosystems) utilizing the cycle parameters tabulated in Table 6.

Table 6: Thermal cylcing parameters utilized for site-directed mutagenesis reactions

Segment	Cycles	Temperature (˚C)	Time
1	1	95	30 seconds
2	18	95	30 seconds
		52	1 minute
		68	9 minutes
3	1	78	7 minutes
4	1	4	Holding

After thermal cycling, *Dpn*I (1 µl; 10 U) was added to the mutagenesis reaction products and incubated at 37˚C for 2 hours. Aliquots (10 µl) of the mutagenesis reactions were taken prior to and post *Dpn*I digestion and resolved on a 0.8 % agarose gel as described in section 2.3.2.3.

3.3.7 Screening of $(His)_6$-EcDXR(W212Y), $(His)_6$-EcDXR(W212L), $(His)_6$-EcDXR(W212F), $(His)_6$-EcDXR(M214I), $(His)_6$-EcDXR(P274I), $(His)_6$-EcDXR(N211K, S213K) and $(His)_6$-PfDXR(K295N, K297S) Plasmid Constructs

Aliquots (50 µl) of competent *E. coli* XL1-Blue cells were transformed with 2 µl of *Dpn*I treated test mutagenesis products, *Dpn*I treated pWhitescript™ mutagenesis product (1 µl) as a mutagenesis positive control, pUC18 plasmid DNA (1 µl; 0.1 ng/µl) as a transformation control and sterile milli-Q water as a no DNA control according to the method described in section 2.3.2.2. pUC18 and pWhitescript™ were plated onto 5 % MacConkey Agar plates (10 g/L lactose) containing ampicillin (0.1 mg/ml). The mutagenesis efficiency (ME) for the pWhitescript™ 4.5 kb plasmid positive mutagenesis control transformation reaction was calculated with accordance to the following formula:

$$ME = \frac{\text{Number of red colony forming units (cfu)}}{\text{Total number of cfu}} \times 100\ \%$$

The mutagenesis efficiency was assumed to correlate with that of pWhitescript™ and the efficiency of transformation was reflected through the pUC18 transformation reaction.

The mutant plasmid DNA constructs were confirmed using diagnostic restriction enzyme analysis (discussed in section 3.3.5) and DNA sequencing. Purified plasmid DNA encoding EcDXR(W212Y), EcDXR(W212L), EcDXR(W212F), EcDXR(M214I), and EcDXR(N211K, S213K) (~400 ng) was added to 3.2 pmol primer EcDXRSPF (refer to the appendices, Chapter 6, section 6.3, Table 8 for details) and purified plasmid DNA encoding PfDXR(K295N, K297S) (~400ng) was added to 3.2 pmol primer PfDXRSPF (refer to the appendices, Chapter 6, section 6.2, Table 8 for details). The cycle sequencing reactions (final volume 20 μl) were performed using the ABI PRISM® BigDye™ Terminator Cycle Sequencing Ready Reaction Kit according to manufacturers conditions. AmpliTaq® DNA polymerase, FS, and nucleotide terminators labelled with fluoroscein donor dyes linked to dRhodomine acceptor dye were used to generate DNA sequences.

3.3.8 Over-Production and Purification of $(His)_6$-EcDXR(W212Y), $(His)_6$-EcDXR(W212L), $(His)_6$-EcDXR(W212F), $(His)_6$-EcDXR(M214I), $(His)_6$-EcDXR(N211K, S213K) and $(His)_6$-PfDXR(K295N, K297S)

The over-production of the mutant proteins was performed according to the methodology described in section 2.3.2.4. Purification was achieved through nickel affinity chromatography according to section 2.3.3 for the *E. coli* DOXP reductoisomerase mutant proteins and section 2.3.4 for the *P. falciparum* DOXP reductoisomerase mutant protein.

3.3.9 The *in vitro* Analysis of EcDXR(W212Y), EcDXR(W212L), EcDXR(W212F), and EcDXR(M214I) DOXP Reductoisomerase Activity

DOXP reductoisomerase activity assays were performed using purified EcDXR(W212Y), EcDXR(W212L), EcDXR(W212F), and EcDXR(M214I) according to the procedure outlined within section 2.3.3.3. In some cases the stability of the purified protein was transient possibly due to the mutations themselves and assays were therefore performed directly after purification.

The EcDXR(N211K, S213K) and PfDXR(K295N, K297S) proteins were not characterized further as insufficient amounts of active protein were obtained after buffer exchange.

3.4 Results and Discussion

3.4.1 Multiple Sequence Alignment Analysis of the Primary Amino Acid Sequences of *E. coli* and *P. falciparum* DOXP Reductoisomerase

DOXP reductoisomerase comprises three domains that define the V-shaped structure of the monomer (refer to Figure 1.4). This includes an NADPH binding N-terminal domain, a catalytic/connective domain and the C-terminal domain. The protein has been confirmed to exist as a homodimer within solution with the monomer being ~43.5 kDa (Steinbacher *et al*., 2003; Reuter *et al*., 2002).

The multiple sequence alignment of the DOXP reductoisomerase homologues are depicted in Figure 3.1. An N-terminal leader sequence is present within the *P. falciparum* (residues 1-77) and *Artemisia annua* (residues 1-77) homologues. Jomaa *et al*. (1999) have previously identified this leader sequence within *P. falciparum* and shown it's functioning in targeting the protein to the apicoplast. It is not surprising that a leader sequence is present within the plant homologue due to isoprenoids being synthesized within the plant plastid via the DOXP/non-mevalonate pathway and sterols being synthesized within the cytosol by the classical mevalonate pathway. A high degree of conservation and sequence similarity amongst the various homologues of DOXP reductoisomerase can be observed (refer to Figure 3.1; total sequence conservation is represented by a '*' in the consensus sequence). For general discussion purposes the amino acid numbering will reflect that of the *E. coli* DOXP reductoisomerase primary amino acid sequence.

The cleft within the protein is as a result of two main stretches of strictly conserved residues or similar residues amongst all homologues, His^{251} to Ser^{258} and Pro^{274} to Arg^{277} Steinbacher *et al*. (2003). The NADPH binding domain spans residues 1-149 within the *E. coli* homologue with residues Thr^{10}, Gly^{11}, Ser^{12}, Gly^{14}, Ala^{35}, Gly^{36}, and Lys^{37} being proposed essential for NADPH binding (refer to Table 3; Reuter *et al*., 2002; Yajima *et al*., 2002; Steinbacher *et al*., 2003). Of

the strictly conserved residues Thr^{10} and Gly^{14} are thought to be vital in the binding of NADPH (refer to Figure 3.1) with Gly^{14} experimentally shown to be essential in NADPH binding upon mutation to an Asp residue (Table 3; Reuter *et al*., 2002). Gly^{11} and Ser^{12} co-ordinate the phosphonate of NADPH and are therefore strictly conserved. The residues Ala^{35} and Gly^{36} are not strictly conserved with Ala^{35} substituted as valine within *P. falciparum* and as a glycine within *H. influenzae*, and Gly^{36} substituted for asparagine within *P. falciparum*. This may be due to these residues being positioned slightly differently between homologues to accommodate the 2` phosphate on NADPH binding.

The catalytic domain spans residues 150-285 within the *E. coli* homologue with the catalytic loop comprising residues His^{209} to Met^{214} (Yajima *et al*., 2002). The acidic residues Asp^{150}, Glu^{152}, Glu^{231}, and Glu^{234} are conserved and play a role in the binding of the divalent cation essential for catalysis (Figure 3.2(A)). They additionally form a network of salt bridges with the basic residues Lys^{125}, His^{153}, and Lys^{228} (Reuter *et al*., 2002). Residues His^{209}, His^{153}, Glu^{231}, and His^{257} have been shown to perform a catalytic function and are conserved amongst all homologues within the sequence alignment (Reuter *et al*., 2003; Yajima. *et al*., 2002)(Figure 3.2), with Glu^{231} the most critical catalytic residue (Yajima *et al*., 2002). Residues that have been experimentally shown to be involved within catalysis include His^{153} and Glu^{231}, whereas His^{257} and His^{209} have been experimentally shown to play a role in substrate binding (refer to section 3.1 and Table 3; Reuter *et al*., 2002, Yajima *et al*., 2002). His^{257} seems to perform a catalytic as well as substrate binding role within DOXP reductoisomerase activity. These residues side chains protrude into the active site (as shown within Figure 3.2) and were defined as those which constitute the catalytic pocket formed by the N-terminal, catalytic and C-terminal domains of the protein by Yajima *et al*. (2002).

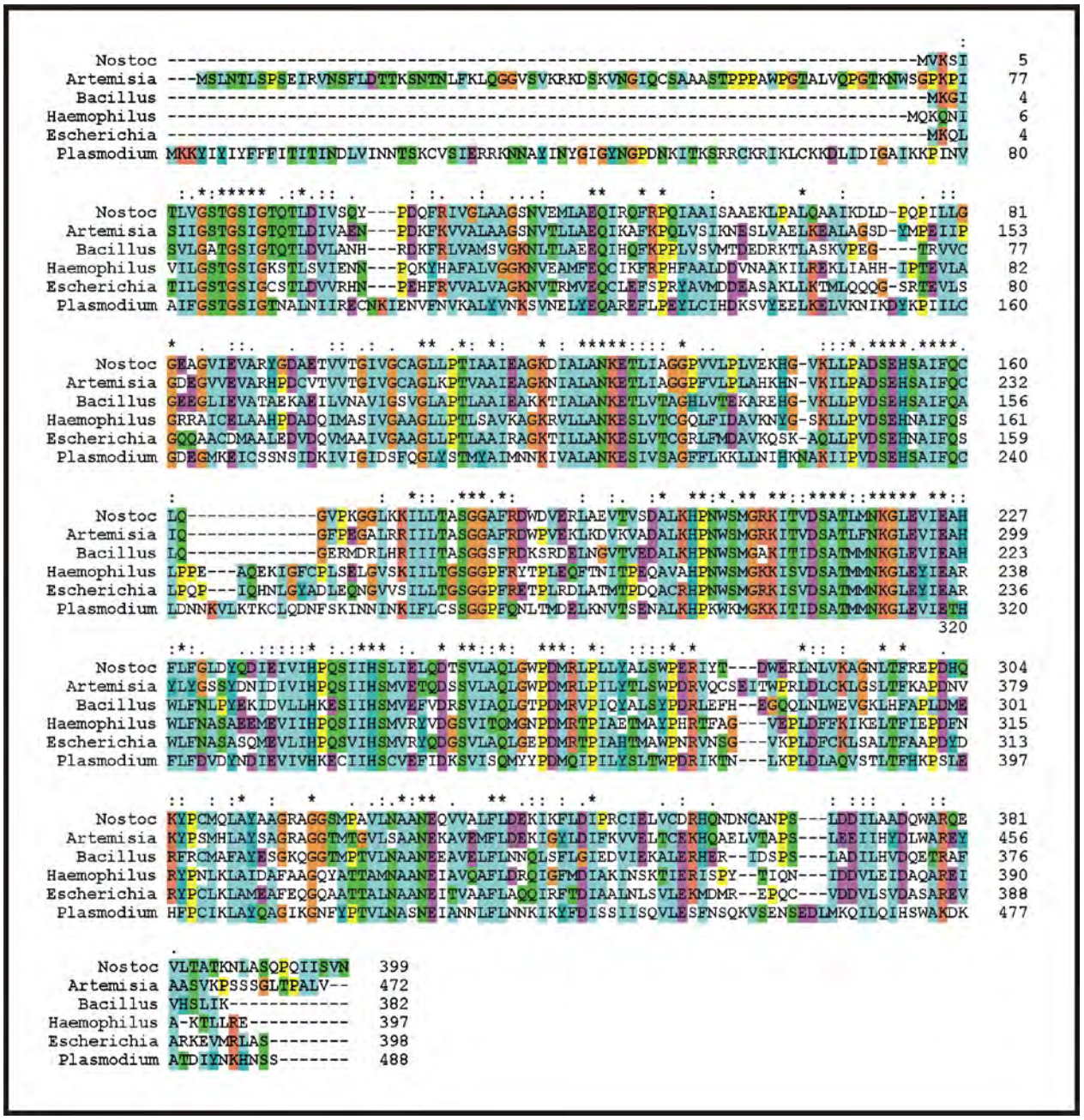


Figure 3.1: Multiple sequence alignment of DOXP reductoisomerase homologues

The multiple sequence alignment of DOXP reductoisomerase homologues was generated via the software programme ClustalX (Jeanmougin *et al*., 1998). The consensus sequence represents identical residues as '*' and similar residues as ':' or '.'. Residues are colour co-ordinated according to their biochemical properties. Both *P. falciparum* and *A. annua* share the amino terminal leader sequence, which harbours the endoplasmic reticulum and apicoplast signal peptides.

Key: Nostoc = *Nostocaceae sp.*PCC 7120 (Genbank accession number AP003596); Artemisia = *Artemisia annua* (Genbank accession number AAD56391.1); Haemophilus = *Haemophilus influenzae Rd* (Genbank accession number NP_438967); Escherichia = *E. coli* (GenBank accession number EDL933); Bacillus = *Bacillus halodurans* (Q9KA69.1) and Plasmodium = *P. falciparum* (GenBank accession number AF111813.1).

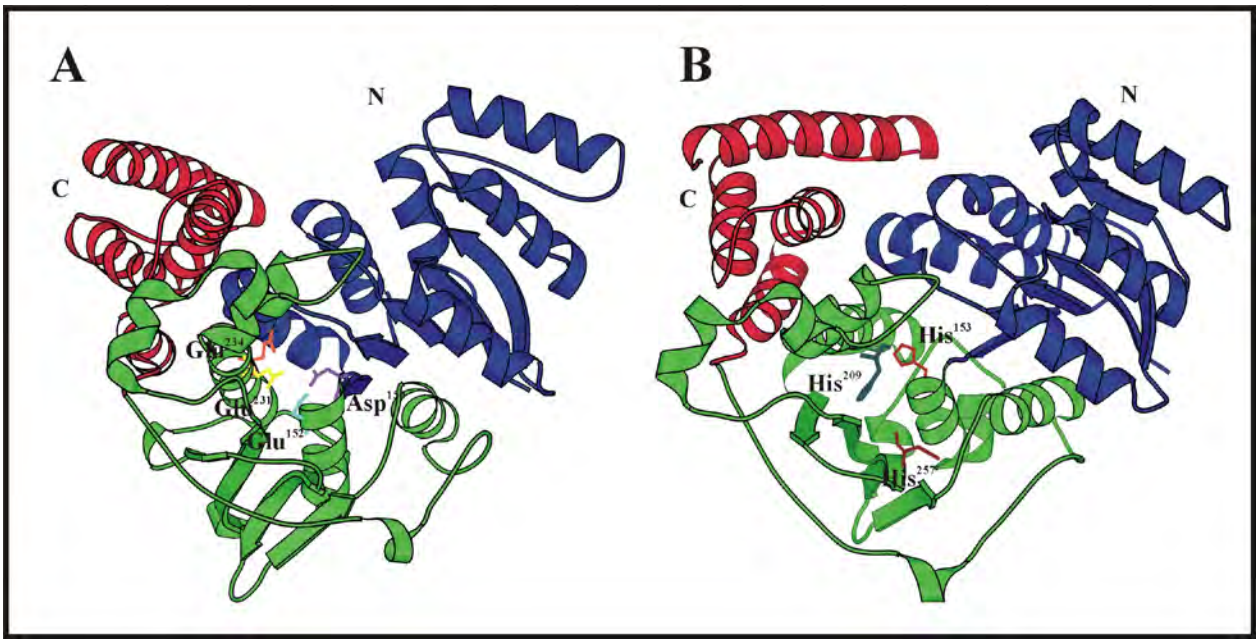


Figure 3.2: Ribbon representation of *E. coli* DOXP reductoisomerase illustrating residues involved in catalysis

The ribbon representation of *E. coli* DOXP reductoisomerase was generated using Molscript (Kraulis, 1991). The N-terminal domain is coloured blue, the catalytic domain green and the C-terminal red. (A) *E. coli* DOXP reductoisomerase with the acidic catalytic residues Asp^{150}, Glu^{152}, Glu^{231}, and Glu^{234} highlighted in stick conformation. (B) *E. coli* DOXP reductoisomerase with the basic catalytic residues His^{153}, His^{209}, and His^{257} highlighted in stick conformation.

The residues Ser^{186}, Ser^{222}, Asn^{227} and Lys^{228} are strictly conserved within the catalytic domain and form hydrogen bonds with the phosphonate moiety of the substrate/inhibitor thus performing an anchoring role (refer to Figure 3.3). Since the catalytic loop performs the essential role of locking the substrate within the active site pocket, and securing it within a hydrophobic environment, residues within this region of the multiple sequence alignment are crucial in protein function and are expected to be strictly conserved to maintain structural and functional integrity within the active site pocket. The residues Met^{214} and Trp^{212} have been predicted from structural data to interact with the nicotinamide moiety of NADPH allowing for the catalytic hatch to close over the substrate once it has bound (Steinbacher *et al*., 2003). These residues perform an essential hydrophobic function, which allows them not only to interact with the NADPH but also the hydrophobic backbone of the substrate or the inhibitor fosmidomycin (Yajima *et al*., 2002). Figure 3.3 illustrates the essential contacts that are made upon binding of the inhibitor. His^{257} and Trp^{212}, together with Pro^{274}, seem to perform the stacking role vital for the formation of the

structure of the active site and are therefore strictly conserved amongst all homologues (Steinbacher *et al*., 2003). Upon binding of fosmidomycin Steinbacher *et al*. (2003) reported that the catalytic loop became disordered possibly suggesting that it can form more than one conformation. Pro^{210} represents another residue that is conserved amongst all homologues within the catalytic loop region. The residues predicted to perform the stacking roles essential in the orientation of the active site as well as within substrate binding have not been verified through experimental work and functional predictions regarding these residues have relied upon 3-D analysis of the crystal structure (Table 3). The Trp^{212} residue is closely associated with the bound inhibitor possibly making it essential in substrate/inhibitor binding through interactions with the substrate/inhibitor carbon backbone. Additionally, the Met^{214} does not seem to be in close proximity to the bound inhibitor, as in the case for Trp^{212} (Figure 3.3). Rather it is in close proximity to the cofactor NADPH and is possibly vital within NADPH binding and the structural orientation of the active site once it has interacted with the cofactor NADPH.

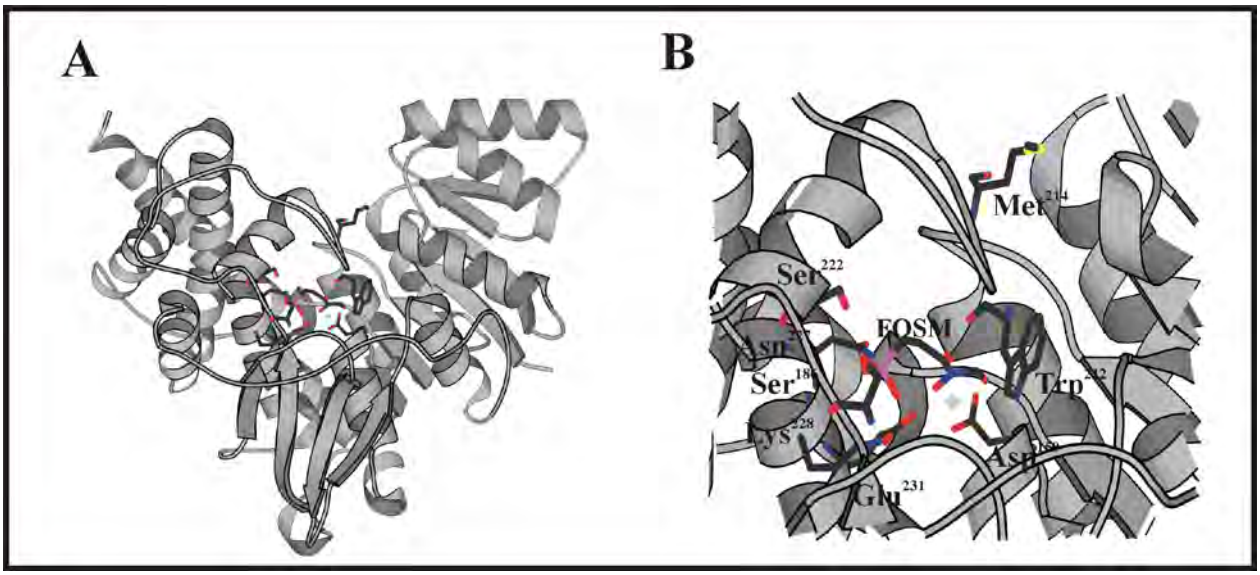


Figure 3.3: Ribbon representation of *E. coli* DOXP reductoisomerase illustrating residues critical for fosmidomycin binding

(A) Fosmidomycin bound to *E. coli* DOXP reductoisomerase with the interacting amino acids highlighted. (B) Close-up view of contacts made between *E. coli* DOXP reductoisomerase and fosmidomycin. The N-formyl oxygen on fosmidomycin is positioned *trans* to Glu^{231} while the N-hydroxyl oxygen is positioned *trans* to Asp^{150}. Pro^{274} orders the active site allowing Trp^{212} and Met^{214} to perform a stacking role and thus securing the cofactor NADPH. The phosphonate moiety of fosmidomycin is anchored via hydrogen bonding with Ser^{186}, Ser^{222}, Asn^{227} and Lys^{228}. The ribbon representations were generated via Molscript (Kraulis, 1991) using the crystal structure data generated via Steinbacher *et al.* (2003)(Protein Data Bank accession codes 1ONO, 1ONN and 1ONP).

3.4.2 Homology Modelling of the *P. falciparum* DOXP Reductoisomerase

P. falciparum DOXP reductoisomerase has an amino acid sequence identity of 43.5 % to that of the *E. coli* homologue sequence. Due to the high sequence similarity the *E. coli* DOXP reductoisomerase homologue it may be used as a template and model for *P. falciparum* DOXP reductoisomerase, from which structural, functional and catalytic predictions regarding the parasitic protein's mechanism can be made. Utilizing the known structural information for *E. coli* DOXP reductoisomerase as a template, homology models for the *P. falciparum* were created (refer to Figure 3.4). Two homology models were created based on crystal structure data available from Reuter *et al*. (2002), Yajima *et al*. (2002) and Steinbacher *et al*. (2003). The first utilized the structure co-ordinates from the unbound protein (Reuter *et al*., 2002; Yajima *et al*., 2002) and the second from structural co-ordinates of the protein with bound substrate/inhibitor and the divalent cation (Steinbacher *et al*., 2003).

Figure 3.4(C) and (D) represent the predicted structure of the *P. falciparum* DOXP reductoisomerase monomer modelled using the ligand-free *E. coli* DOXP reductoisomerase as the template via SWISS-MODEL and the substrate/inhibitor bound and divalent cation-bound *E. coli* DOXP reductoisomerase template via WHATIF respectively. Each monomer forms the V-shaped cleft and contains the 3 domains previously described.

The *P. falciparum* DOXP reductoisomerase homology model closely resembles the 3D structure of the *E. coli* homologue when the substrate/inhibitor bound and divalent cation and NADPH bound template was used and the homology model generated with WHATIF. Variation in the predicted 3-D structure for *P. falciparum* DOXP reductoisomerase is likely due to the differences within the primary amino acid alignment generated in the homology modelling process. Since there is ~56.5 % amino acid difference between *E. coli* and *P. falciparum* DOXP reductoisomerase, with some key differences in the catalytic lid region, one would expect variation within the modelled structure for *P. falciparum* DOXP reductoisomerase to occur. This is especially noted for the NADPH-binding N-terminal domain and catalytic domain in the sequence alignment. In specific, the catalytic hatch region for *P. falciparum* DOXP reductoisomerase contains a Lys^{295} and Lys^{297} amino acid in comparison to the respective

conserved Asn and Ser residues amongst DOXP reductoisomerase homologues. This would suggest a different 3-D orientation of theses residues within the predicted *P. falciparum* DOXP reductoisomerase protein than that of Asn^{211} and Ser^{213} in the 3-D structure of *E. coli* DOXP reductoisomerase. The 3-D predicted structure for *P. falciparum* DOXP reductoisomerase generated by SWISS-MODEL utilizing unbound template differs significantly in the N-terminal and catalytic domains. This is most likely due to the X-ray crystal information for *E. coli* DOXP reductoisomerase being impaired due to the absence of the cofactor NADPH, the metal cation and substrate. In this case, structural conformational changes that occur upon the binding of the cofactor NADPH and the substrate within the active site of the protein are therefore not taken into consideration on generation of the predicted 3-D structure for *P. falciparum* DOXP.

3.4.3 Identification of Potentially Critical Residues for the Catalytic Functioning of *P. falciparum* and *E. coli* DOXP Reductoisomerase

Bioinformatic analysis of the DOXP reductoisomerase protein sequences were performed using the multiple sequence alignment (Figure 3.1) and analyses of the structures of the *E. coli* DOXP reductoisomerase and the predicted structure of the *P. falciparum* DOXP reductoisomerase. Residues that have been reported as critical in the biochemical functioning of DOXP reductoisomerase from crystallographic studies were identified within the multiple sequence alignment and monitored for degree of conservation.

In specific the residues Trp^{212}, Met^{214} and Pro^{274} of the catalytic hatch region of *E. coli* DOXP reductoisomerase were inspected. These residues have been previously identified as critical in the orientation of the active site and binding of NADPH. Trp^{212} has also been considered to interact with the carbon backbone of the substrate/inhibitor (Reuter *et al*., 2002, Yajima *et al*., 2002 and Steinbacher *et al*., 2003). The function of these residues has not been experimentally confirmed and Trp^{212}, Met^{214} and Pro^{274} were therefore considered for site-directed amino acid analysis in this investigation.

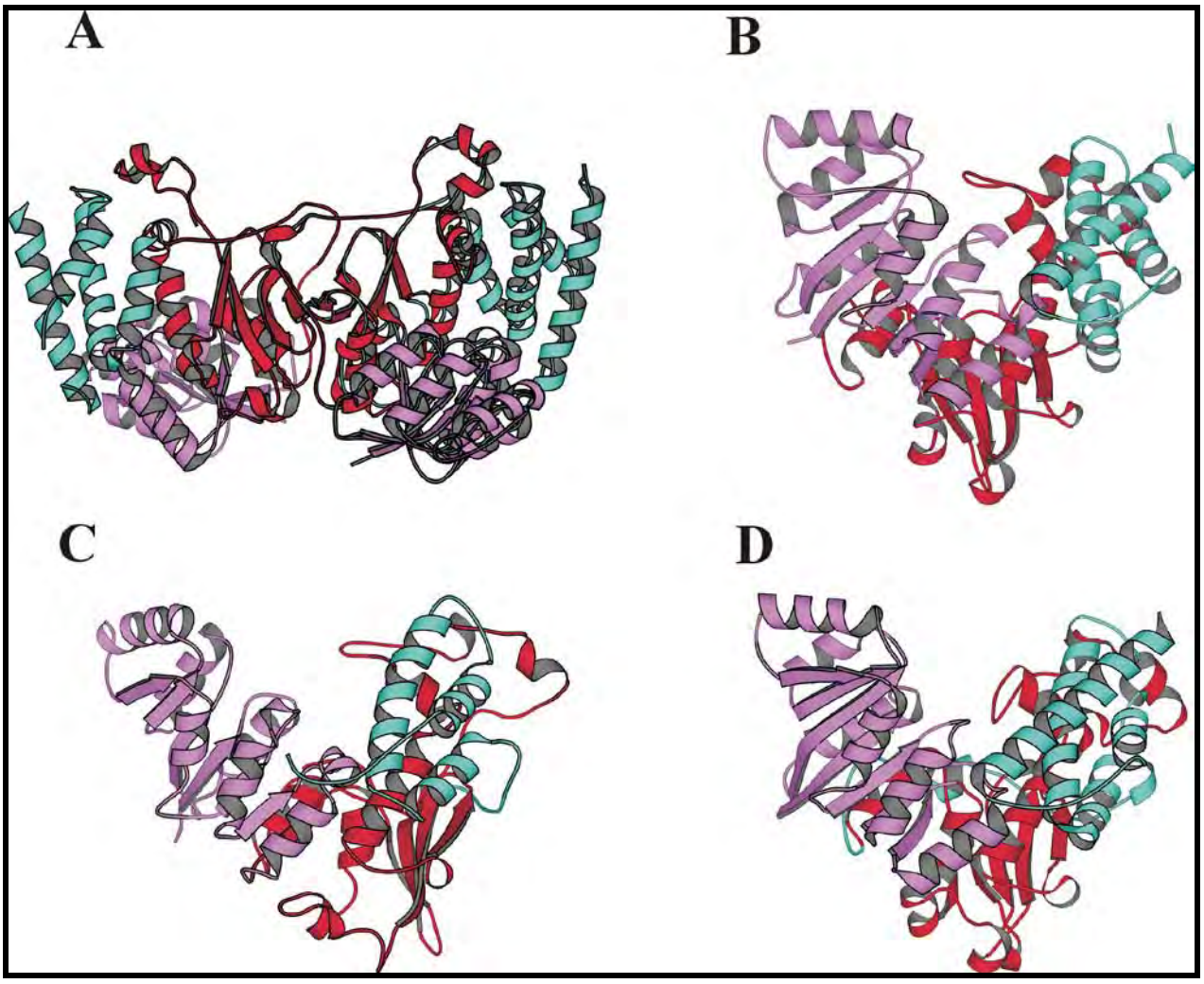


Figure 3.4: Ribbon representation of the structure of *E. coli* DOXP reductoisomerase and the predicted structure of *P. falciparum* DOXP reductoisomerase

(A) *E. coli* DOXP reductoisomerase homodimer. (B) *E. coli* DOXP reductoisomerase monomer. The enzyme comprises three domains, which form the overall V-shape of the molecule. The N-terminal NADPH binding domain is shown in violet. The central connective/catalytic domain shown in crimson binds the divalent cation (magnesium, manganese or cobalt) and the phosphate moiety of the substrate. The carboxy terminal provides structural integrity for the active site pocket and is shown in turquoise. (C) Predicted structure of the *P. falciparum* DOXP reductoisomerase monomer homology modeled using the unbound *E. coli* homologue crystal structure data as a template. (D) Predicted structure of *P. falciparum* DOXP reductoisomerase monomer homology modeled using the substrate/inhibitor and divalent cation-bound *E. coli* homologue crystal structure data as a template. Ribbon representations were generated using Molscript (Kraulis, 1991) using the crystal structure data generated via Reuter *et al.* (2002)(PDB accession code 1K5H), Yajima *et al.* (2002)(PDB accession code 1JVS) and Steinbacher *et al.* (2003)(PDB accession codes 1ONO, 1ONN and 1ONP).

Strictly conserved residues are likely to be essential within the catalytic mechanism or structural integrity within the protein, especially with regards to the catalytic site and hatch of DOXP reductoisomerase. Pro^{274}, Trp^{212}, Met^{214} were found to be conserved within all homologues and due to their potential roles within the binding to the nicotinamide moiety of NADPH and structural integrity of the active site, this is not surprising. The orientation of these residues within the 3-D topology of DOXP reductoisomerase were assessed (Figure 3.5). Trp^{212} is a vital residue for interaction with the substrate/inhibitor as well as NADPH. The formation of the hydrophobic pocket into which the substrate binds in the presence of NADPH is an essential factor in the catalysis process. The residues, which ensure this hydrophobic nature and structural architecture of the active site, are stacked together in a closely associated cluster and form essential contacts with residues on the catalytic hatch, NADPH and the substrate. These residues were chosen for biochemical analysis via site-directed mutagenesis to potentially unravel the biochemical roles essential in the functioning of DOXP reductoisomerase.

Small sequence changes affect the biochemical properties of the protein with regard to structure or function. From the analysis of the multiple sequence alignment of the catalytic hatch region the residues Asn^{211} and Ser^{213} were found not to be conserved within the *P. falciparum* DOXP reductoisomerase sequence in comparison to the *E. coli* and other homologues within the catalytic loop. In *P. falciparum* DOXP reductoisomerase both Asn^{211} and Ser^{213} are replaced by a lysine and represent non-conservative substitutions making it intriguing that both residues are substituted for a lysine in the case of *P. falciparum* DOXP reductoisomerase. This could result in the catalytic hatch of *P. falciparum* DOXP reductoisomerase behaving in a slightly modified manner to that of other species. The orientation of these residues were analysed using the *E. coli* DOXP reductoisomerase model and the *P. falciparum* DOXP reductoisomerase homology models (refer to Figure 3.5). The side chains of the Lys^{295} and Lys^{297} residues protrude into the catalytic pocket of *P. falciparum* DOXP reductoisomerase (Figure 3.5(C)). These residues are polar and charged and are likely to perform essential contacts with the substrate backbone. The topologically equivalent Asn^{211} and Ser^{213} in *E. coli* DOXP reductoisomerase are however orientated differently. Both Asn and Ser are polar yet carry no charge and upon inspection of the 3-D *E. coli* DOXP reductoisomerase Ser^{213} was found to protrude into the active site whereas Asn^{211} seems to protrude outward (Figure 3.5(F)). This is suggestive of Asn^{211} being involved

within other structural contacts. The differences in the orientation of the topologically equivalent residues within *P. falciparum* and *E. coli* DOXP reductoisomerase homologues is suggestive of a variation in the role that the catalytic hatch provides in securing the substrate and the mode in which it performs this function within the parasitic system.

To determine whether the activity of DOXP reductoisomerase and its affinity to substrate are mediated differently between the *P. falciparum* and *E. coli* DOXP reductoisomerase proteins site-directed mutagenesis of the Lys^{295} and Lys^{297} within *P. falciparum* DOXP reductoisomerase to an asparagine and serine respectively and vice versa for the *E. coli* DOXP reductoisomerase were prepared.

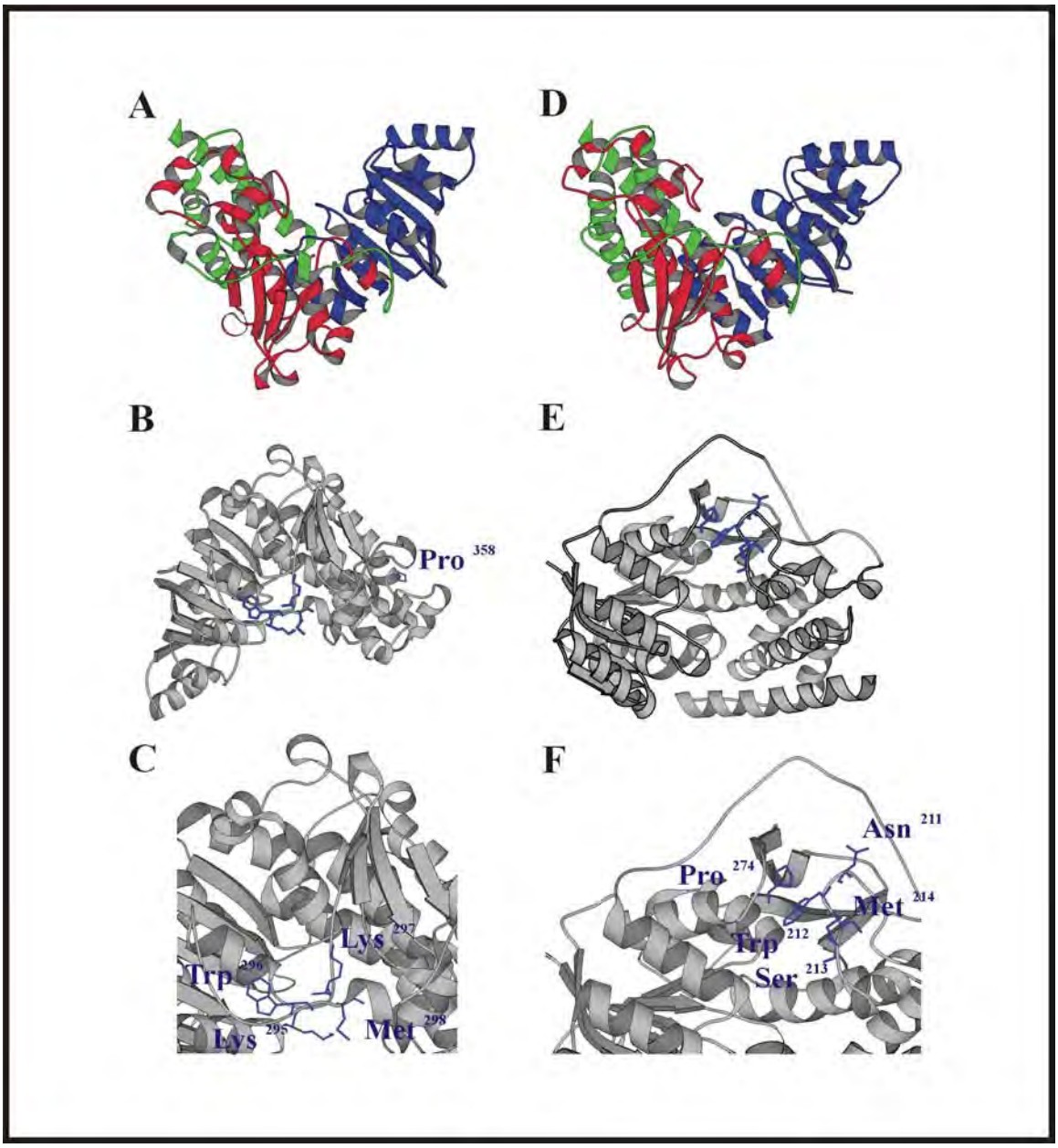


Figure 3.5: Ribbon representation of the *E. coli* DOXP reductoisomerase and the predicted structure of *P. falciparum* DOXP reductoisomerase

(A-C) depict the predicted structure of the *P. falciparum* DOXP reductoisomerase homology model using ligand-bound *E. coli* DOXP reductoisomerase crystal data co-ordinates as a template. (D-F) depict the *E. coli* DOXP reductoisomerase. A and D represent the individual subunit domains with blue representing the N terminal NADPH binding domain, red the catalytic domain and green the C terminal domain. Residues identified as being potentially critical in the function of the enzyme are highlighted in B and E. C and F show these residues magnified. Whatif was used to generate the homology model (Vriend, 1990) and Molscript to generate the ribbon representations (Kraulis, 1991). The *E. coli* DOXP reductoisomerase crystal structure co-ordinates were determined by Steinbacher *et al*. (2003)(PDB accession code 1ONO, 1ONN and 1ONP).

3.4.4 Screening and Purification of DOXP Reductoisomerase Mutant Plasmid Constructs

Apart from the mutagenesis reaction for plasmid construct pQE9EcDXR(P274I), all mutagenesis reactions were successful.

3.4.4.1 Screening of pQE9EcDXR(W212Y), pQE9EcDXR(W212L), and pQE9EcDXR(W212F) and Purification of $(His)_6$-EcDXR(W212Y), $(His)_6$-EcDXR(W212L) and $(His)_6$-EcDXR(W212F)

The mutagenesis reaction products for were resolved on 0.8 % agarose gel according to the method described in 2.3.2.3 (data not shown). Mutagenic amplification products were screened via diagnostic restriction enzyme analysis according to sections 3.3.5 and 3.3.7. Figure 3.6 (A) indicates the successful *Nco*I restriction digests of the pQE9EcDXR(W212Y) and pQE9EcDXR(W212L) plasmid constructs. Upon *Nco*I restriction the putative pQE9EcDXR(W212Y) and pQE9EcDXR(W212L) mutant plasmid constructs resulted in three DNA fragments of 3389 bp, 1023 bp, and 210 bp (Figure 3.6(A); lanes 5 and 7 respectively) thus confirming the presence of mutagenic DNA. The parental/template DNA resulted in only two DNA fragments of 3599 bp and 1023 bp as expected (refer to Figure 3.6(A); lane 3). The presence of the desired mutation on plasmids pQE9EcDXR(W212Y) and pQE9EcDXR(W212L) was further validated through DNA sequencing (data not shown). The desired mutation on plasmid pQE9EcDXR(W212F) could not be confirmed through restriction digest with *Nco*I but was confirmed through DNA sequencing (data not shown). DNA sequencing revealed the incorporation of the required mutation however the silent mutation required for the insertion of the restriction enzyme site for *Nco*I was not achieved possibly due to a fault in the synthesis of the oligonuceotide primer pair.

Over-production and solubility analysis was performed for each mutant construct according to section 2.3.2.4. All mutant DOXP reductoisomerase proteins showed profiles corresponding to that of the wild type protein (*E. coli* DOXP reductoisomerase) with regard to optimal induction period and solubility (data not shown). Purification of the $(His)_6$-tagged EcDXR(W212Y), EcDXR(W212L) and EcDXR(W212F) proteins was via nickel chelate affinity chromatography

under native conditions described in section 2.3.3 (refer to Figure 3.6 B, C, and D respectively). The cleared lysate of the respective mutant DOXP reductoisomerase protein was exposed to the resin overnight at 4 °C to allow for maximal binding. The $(His)_6$-tagged *E. coli* DOXP reductoisomerase mutant proteins saturated the nickel sepharose beads representative of the high levels of unbound protein within the flow through fractions (Figure 3.6 (B, C and D); lane 7). The respective EcDXR(W212Y), EcDXR(W212L) and EcDXR(W212F) protein was eluted competitively in the presence of 1 M imidazole (Figure 3.6 B, C, and D respectively; lane 11-13); however it was also present within the wash steps (Figure 3.6 (B, C and D); lane 8-10) possibly due to weak interaction of the recombinant protein to the resin due to the effects of over-saturating the resin. Buffer exchange to remove the imidazole from the protein fractions was performed and the protein concentrated and utilized in subsequent DOXP reductoisomerase activity assays (refer to Figure 3.6 (B, C, and D); lane 15). All proteins were purified to >95 % purity with yields of 2.5 mg, 2.7 mg and 5.4 mg per litre of culture for EcDXR(W212L), EcDXR(W212Y), and EcDXR(W212F) respectively.

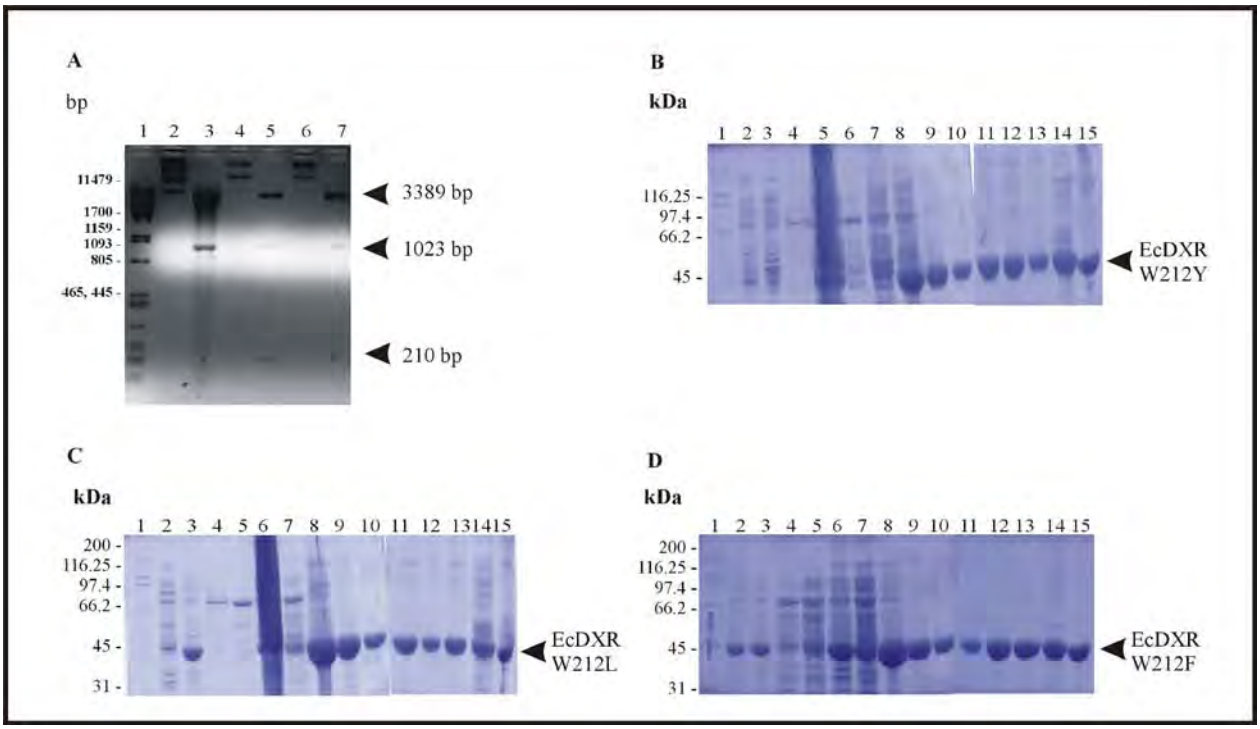


Figure 3.6: Screening of pQE9EcDXR(W212Y) and pQE9EcDXR(W212L) DOXP reductoisomerase mutant plasmid constructs and purification of the $(His)_6$-EcDXR(W212Y), $(His)_6$-EcDXR(W212L) and $(His)_6$-EcDXR(W212F) DOXP reductoisomerase proteins

(A) 1.6 % Agarose gel electrophoresis of the diagnostic restriction enzyme analysis of putative pQE9EcDXR(W212Y) and pQE9EcDXR(W212L) plasmid DNA. Lambda DNA restricted with *Pst*I was separated within lane 1 as a marker (sizes are in bp). The uncut no enzyme pQE9EcDXR parental plasmid DNA products can be seen within lane 2. Within lane 3 the template DNA restricted with *Nco*I results in two DNA fragments of 3599 bp and 1023 bp as expected. Lanes 4 and 6 represent the uncut putative pQE9EcDXR(W212Y) and pQE9EcDXR(W212L) plasmid DNA. pQE9EcDXR(W212Y) restricted with the diagnostic enzyme *Nco*I results in three DNA fragments of sizes 3389 bp, 1023 bp, and 210 bp as expected within lane 5. pQE9EcDXR(W212L) DNA is represented within lane 7 as upon restriction with *Nco*I the resultative DNA fragments of 3389 bp, 1023 bp, and 210 bp can be seen. (B, C and D) Native batch nickel affinity purification of $(His)_6$-EcDXR(W212Y), $(His)_6$-EcDXR(W212L) and $(His)_6$-EcDXR(W212F) respectively. Each purification step was analysed using SDS-PAGE. The Biorad Broad Range Molecular Weight Marker (kDa) was separated in lane 1. The purification samples included un-induced whole cell extract (lane 2), induced whole cell extract (lane 3), whole cell lysate (lane 4), soluble fraction (lane 5), insoluble fraction (lane 6), flow-through (lane 7), native wash (lanes 8-10), native elution (lanes 11-13), Ni-affinity sepharose beads (lane 14) and imidazole buffer exchanged concentrated protein (lane 15). The mutant EcDXR proteins were purified successfully and are indicated with an arrow.

3.4.4.2 Screening of pQE9EcDXR(M214I) and Purification of $(His)_6$-EcDXR(M214I)

The mutagenesis reaction product for the putative mutant pQE9EcDXR(M214I) plasmid DNA was transformed into *E. coli* XL1Blue and screened via diagnostic restriction enzyme digestion

according to section 3.3.5 and 3.3.7 (Figure 3.7(A)). From the digestion with *Xma*III a linearized fragment was obtained representative of mutant plasmid DNA (Figure 3.7(A); lane 7). A further digest with *Bam*HI and *Xma*III was performed to confirm the mutant plasmid DNA and resulted in two bands of 3975 bp and 647 bp as expected (Figure 3.7(A); lane 8). The parental template DNA restricted with *Xma*III resulted in the conformations of uncut DNA being separated, within Figure 3.7(A) lane 2, as expected. Upon digestion with *Bam*HI individually or together with *Xma*III the parental plasmid was expected to be linearized and resolve as a single DNA fragment of 4622 bp (as seen in Figure 3.7(A); lane 4 and 5 respectively). The desired mutation was identified on plasmid pQE9EcDXR(M214I) and validated through DNA sequencing (data not shown).

Over-production and solubility analysis was performed according to the procedure outlined within section 2.3.2.4. The $(His)_6$-EcDXR(M214I) protein was found to have similar induction and solubility characteristics as that of the wild-type protein (data not shown) and was purified through native batch nickel chelate affinity chromatography (section 2.3.3; Figure 3.7(B)). The cleared lysate was bound to the resin maximally overnight representative of little protein of DOXP reductoisomerase monomer mass (~45 kDa) in the flow-through sample (Figure 3.7(B); lane 7). The protein appears within the wash samples possibly as a cause of weak interactions between the resin and protein due to saturation levels of protein present (Figure 3.7(B); lane 8-10). The EcDXR(M214I) protein was successfully purified to >95 % purity at levels sufficient for biological analysis (2 mg per litre of culture) through competitive release with 1 M imidazole (Figure 3.7(B); lane 11-13, 15). A large proportion of pure recombinant protein remains attached to the Ni^{2+} charged resin possibly due to strong interactions and maximal binding having occurred (Figure 3.7(B); lane 14). Additional protein may be eluted from the column by increasing the imidazole concentration further. The EcDXR(M214I) protein was buffer exchanged and concentrated for activity analysis (Figure 3.7(B); lane 15).

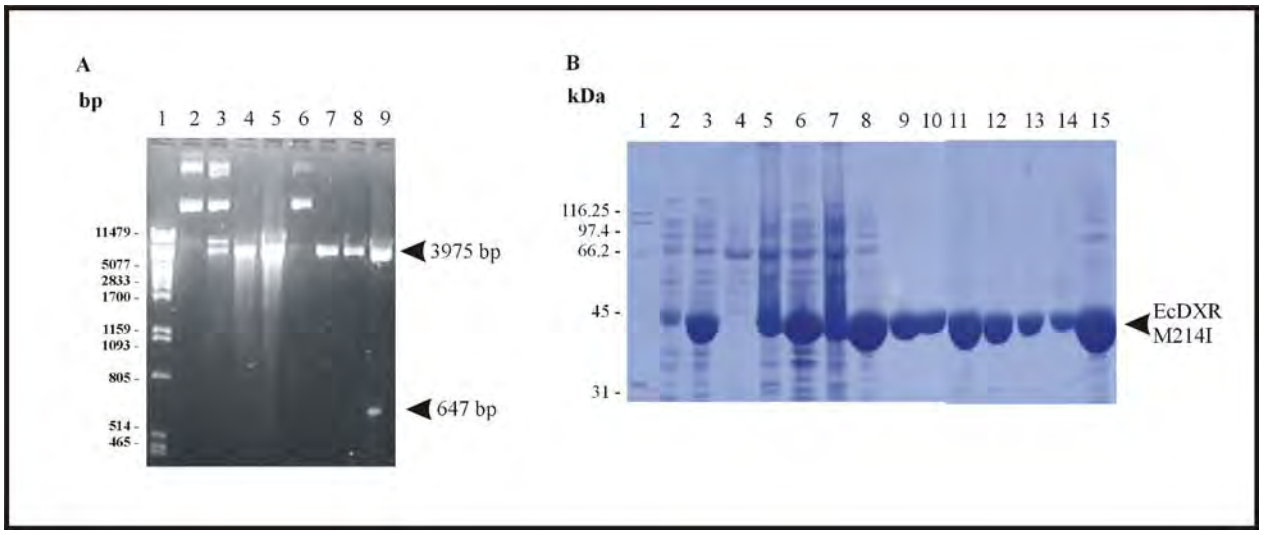


Figure 3.7: Screening of pQE9EcDXR(M214I) and purification of $(His)_6$-EcDXR(M214I) DOXP reductoisomerase

(A) 1.6 % Agarose gel electrophoresis of diagnostic restriction enzyme analysis of putative pQE9EcDXR(M214I). Lambda DNA restricted with *Pst*I was separated within lane 1 as a marker (sizes are in bp). Parental plasmid DNA uncut DNA conformations are separated within lane 2. As no *Xma*III restriction site is present within the template plasmid DNA uncut conformations of plasmid DNA are present (lane 3). Parental plasmid DNA restricted with *Bam*HI results in linearized plasmid DNA and a single fragment of 4622 bp can be seen within lane 4. *Bam*HI and *Xma*III double digest of parental plasmid DNA yields a single DNA fragment representing linearized plasmid DNA (4622 bp) due to *Bam*HI restriction and no *Xma*III nuclease restriction site present in parental plasmid DNA (lane 5). Lane 6 represents the uncut DNA for pQE9EcDXR(M214I). Linearized pQE9EcDXR(M214I) of size 4622 bp is present within lane 7 and 8 due to restriction with *Xma*III and *Bam*HI respectively. Upon the double digest with *Xma*III and *Bam*HI the putative mutant DNA is confirmed through the separation of two DNA fragments of 3975 bp and 647 bp. (B) Native batch nickel chelate affinity purification of EcDXR(M214I). Biorad Broad Range SDS-PAGE Molecular Markers are separated in lane 1 and sizes shown in kDa. Native purification samples were taken after each purification step and analysed on SDS-PAGE these included uninduced and induced *E. coli* XL1Blue[pQE9EcDXR(M214I)] whole cell extracts separated within lane 2 and 3 respectively, cleared lysate (lane 4), insoluble fraction (lane 5), soluble fraction (lane 6), flow-through (lane 7), native wash fractions (lane 8- 10), native elution fractions (lane 11 – 13), nickel sepharose beads (lane 14) and concentrated imidazole buffer-exchanged protein eluate fraction (lane 15). $(His)_6$-EcDXR(M214I) was successfully purified for biological analysis (indicated with an arrow).

3.4.4.3 Screening of pQE9EcDXR(N211K, S213K) and pQE31PfDXR(K295N, K297S) and Purification of $(His)_6$-EcDXR(N211K, S213K) and $(His)_6$-PfDXR(K295N, K297S)

The mutagenesis reaction products for the putative pQE9EcDXR(N211K, S213K) and pQE31PfDXR(K295N, K297S) were transformed into *E. coli* XL1Blue and diagnostic restriction enzyme analysis to confirm the putative mutant DOXP reductoisomerase DNA was performed according to section 3.3.7.

The replacement of the Asn^{211} and Ser^{213} residues within the *E. coli* DOXP reductoisomerase protein with lysine residues was successful. This was achieved through first introducing the codon for lysine at codon 211, confirming this mutation, and then utilizing the pQE9EcDXR(N211K) DNA as a template to incorporate the codon for lysine at codon 213. Verification of the pQE9EcDXR(N211K) DNA is shown within Figure 3.8 (A). Upon the substitution of the Asn^{211} with lysine the restriction endonuclease site for *Nco*I was introduced through a silent mutation. Upon *Nco*I digestion of template DNA only two fragments of 3599 bp and 1023 bp were resolved however are not clear within Figure 3.8 (A), lane 3. For the incorporation of the lysine residue at codon 211 three DNA fragments were obtained (Figure 3.8(A); lane 5 and 7) of 3389 bp, 1023 bp and 210 bp upon *Nco*I digestion as expected indicating the presence of pQE9EcDXR(N213K) DNA. The pQE9EcDXR(N211K) DNA was then used as the template for the introduction of the codon for lysine at codon 213. Confirmation of putative pQE9EcDXR(N211K, S213K) DNA was through *Nco*I restriction analysis. The removal of the previously incorporated *Nco*I site upon substitution of the Ser^{213} to a lysine residue resulted in only two DNA fragments of 3599 bp and 1023 bp as expected (data not shown). Samples were resolved using agarose gel electrophoresis, as described in section 2.3.2.3, upon a 1 % agarose gel, 0.5 μg/ml ethidium bromide in 0.5x TBE buffer at 100V. Additionally the desired mutations in pQE9EcDXR(N211K, S213K) were verified through DNA sequencing (data not shown).

The over-production and solubility analyses for the $(His)_6$-EcDXR(N211K, S213K) protein were performed according to the methodology described in section 2.3.2.4 (data not shown). The over-production for this protein was equivalent with the wild type enzyme upon SDS-PAGE analysis and the protein was found to be soluble. The EcDXR(N211K, S213K) cleared lysate was allowed to bind to the resin overnight for maximal binding. This is apparent due to residual levels of protein present within the flow-through fraction (Figure 3.8(B); lane 7). Elution was achieved through competition with 1 M imidazole (Figure 3.8(B); lane 11-13) and resulted in pure protein being released from the nickel charged resin. Protein was still bound to the resin after purification as can be seen for Figure 3.8 (B) lane 14. Native batch purification of the $(His)_6$-tagged protein was successful (Figure 3.8(B)) however; levels of active protein were not sufficient for activity assays

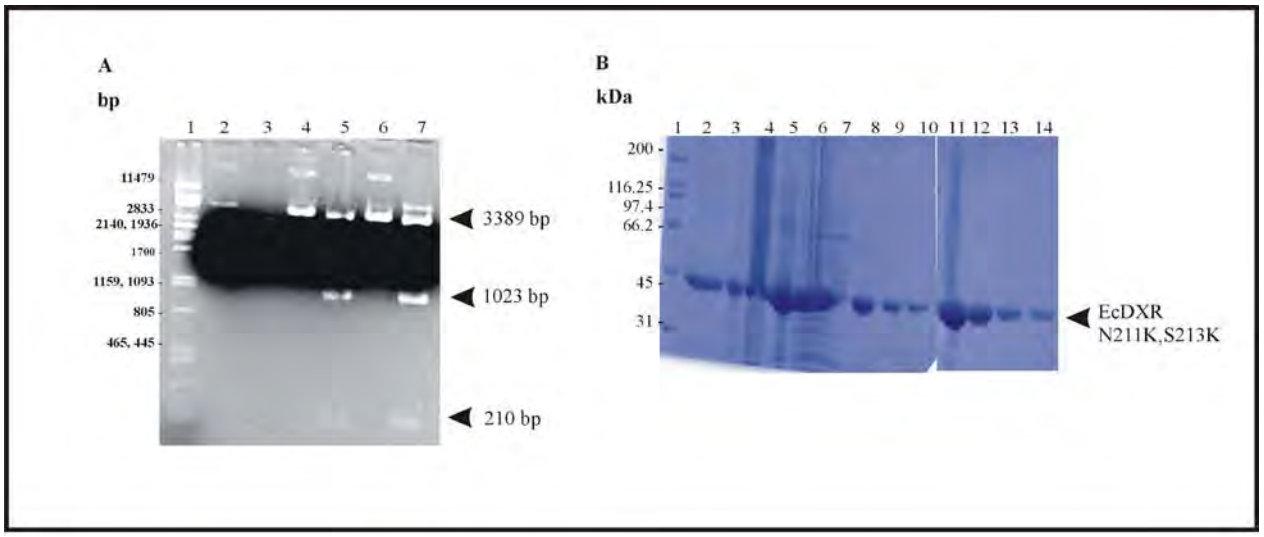


Figure 3.8: Screening of pQE9EcDXR(N211K, S213K) and purification of (His)6-EcDXR(N211K, S213K) DOXP reductoisomerase

(A) 1 % Agarose gel electrophoresis of diagnostic restriction enzyme analysis of putative pQE9EcDXR(N211K, S213K). Lambda DNA restricted with *Pst*I was separated within lane 1 as a marker (sizes are in bp). Parental plasmid DNA uncut DNA conformations are separated within lane 2. Lane 3 represents the restriction of parental plasmid DNA with *Nco*I, which should have resulted in two DNA fragments of 3599 bp and 1023 bp. Lanes 4 and 6 represent uncut pQE9EcDXR(N211K) DNA. Lanes 5 and 7 represent *Nco*I restricted mutant DNA resulting in three DNA fragments of 3389 bp, 1023 bp, and 210 bp as expected. *Nco*I restriction of pQE9EcDXR(N211K, S213K) DNA did not allow for visualization of the mutant DNA clearly and therefore was confirmed by DNA sequencing. (B) Native batch nickel-chelate affinity purification of EcDXR(N211K, S213K). Biorad Broad Range SDS-PAGE Molecular Markers are separated in lane 1 and sizes shown in kDa. Native purification samples were taken after each purification step and analysed on SDS-PAGE these included uninduced and induced *E. coli* XL1Blue[pQE9EcDXR(N211K, S213K)] whole cell extracts separated within lane 2 and 3 respectively, cleared lysate (lane 4), insoluble fraction (lane 5), soluble fraction (lane 6), flow-through (lane 7), native wash fractions (lane 8- 10), native elution fractions (lane 11 – 13), and nickel sepharose beads (lane 14). Purified $(His)_6$-EcDXR(N211K, S213K) is indicated with an arrow.

The mutagenesis reaction product for pQE31PfDXR(K295N, K297S) was transformed into *E. coli* XL1Blue and screened for the putative mutant DNA via restriction enzyme analysis described in section 3.3.7. A similar approach to that of the pQE9EcDXR(N211K, S213K) mutagenesis construction was used. The codon for lysine at codon 297 was replaced by the codon for serine in the first mutagenesis reaction followed by the introduction of the codon for asparagine at codon 295. The confirmation of the pQE31PfDXR(K297S) mutation was achieved through *Nco*I restriction (Figure 3.9(A)). The restriction site was engineered through a silent mutation to allow for three fragments of 2771 bp, 1263 bp and 681 bp to be resolved on agarose gel electrophoresis representing successful mutagenesis (Figure 3.9(A); lane 5). The parental plasmid has two restriction sites for *Nco*I thus resulting in two DNA fragments of 2771 bp and

1944 bp (Figure 3.9(A); lane 3). The pQE31PfDXR(K297S) DNA was then used as the template for the introduction of the codon change for the K295N amino acid substitution. Upon successful replacement of the codon for lysine at codon 295 with a codon for asparagines, the *Nco*I restriction site introduced within the K297S substitution was removed. Thus upon restriction with *Nco*I two fragments of 2771 bp and 1944 bp are present representing pQE31PfDXR(K295N, K297S) DNA (Figure 3.9(B); lane 5).

Over-production and solubility analysis (data not shown) indicated that the same problems were still apparent as experienced for the PfDXR protein, but to a lesser degree. The over-production of the PfDXR(K295N, K297S) protein improved compared to PfDXR (refer to Figure 3.9(C), lane 2 and 3) and the protein seemed to be mainly soluble. Native batch purification was not successful even though the protein is soluble under native conditions (data not shown). For purification purposes the combination of denaturing and native batch purification was then attempted (Figure 3.9(C)). The lysate was prepared under denaturing conditions and the protein allowed to bind maximally to the resin overnight, apparent due to the low levels of *P. falciparum* DOXP reductoisomerase present within the flow-through sample (Figure 3.9(C); lane 7). The protein bound resin was then washed under native conditions allowing for renaturation (Figure 3.9 (C); lane 8-10). The protein was then removed from the column through competitive elution with increasing imidazole concentrations (Figure 3.9(C); lane 11-17). A 31 kDa protein within the protein elution can be seen in Figure 3.9 (C); lane 12 – 17 and is possibly an *E. coli* protein contamination as a result of non-specific proteins that have bound to the nickel charged resin and eluted off in the presence of the imidazole. Additionally, the recombinant protein was found to remain associated with the nickel sepharose beads after purification (Figure 3.9(C); lane 18). Attempted removal of the protein using increasing concentrations of imidazole was attempted but unsuccessful (data not shown). Due to the low concentration and low purity of the PfDXR(K295N, K297S) protein obtained, biological assays were not performed.

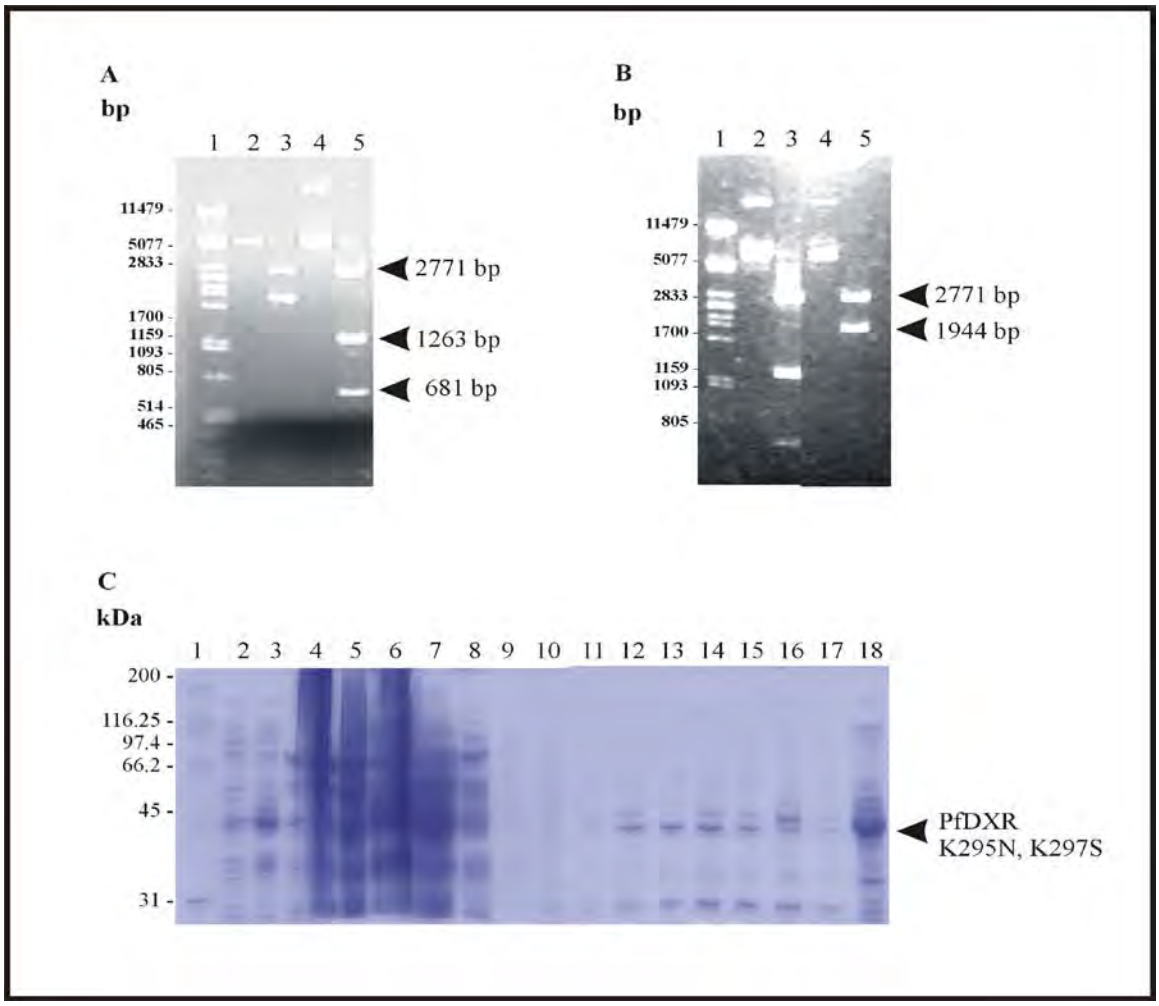


Figure 3.9: Screening of pQE31PfDXR(K295N, K297S) and purification of $(His)_6$-PfDXR(K295N, K297S) DOXP reductoisomerase

(A) 0.8 % Agarose gel electrophoresis of pQE31PfDXR(K297S). Lambda DNA restricted with *Pst*I was separated in lane 1 with sizes depicted in bp. Lane 2 represents the conformations of uncut parental DNA. *Nco*I restricted parental DNA resulted in two fragments of 2771 bp and 1944 bp within lane 3. Lane 4 and 5 constitute the uncut and *Nco*I restricted pQE31PfDXR(K297S) DNA respectively. Three fragments of 2771 bp, 1263 bp and 681 bp are present within lane 5 indicative of the pQE31PfDXR(K297S) DNA. (B) 0.8 % Agarose gel electrophoresis of pQE31PfDXR(K295N, K297S). *Pst*I restricted Lambda DNA in bp was separated in lane 1. Lane 2 and 3 represent the pQE31PfDXR(K297S) DNA as for (A) lane 4 and 5. pQE31PfDXR(K295N, K297S) DNA was present within lane 4 and 5. Lane 4 represents the uncut DNA conformations and lane 5 the *Nco*I restricted putative mutant DNA. Upon *Nco*I restriction the disappearance of the 681 bp fragment was noted and two DNA fragments of 2771 bp and 1944 bp are present indicative of pQE31PfDXR(K295N, K297S) DNA. (C) Denaturing/Native batch purification of $(His)_6$-PfDXR(K295N, K297S) reductoisomerase. Samples were analysed using SDS-PAGE at each step of the purification procedure. Biorad Broad Range Molecular Markers (kDa) are separated in lane 1. Uninduced and induced *E. coli* XL1Blue[pQE31PfDXR(K295N, K297S)] whole cell extracts separated within lane 2 and 3 respectively, cleared lysate (lane 4), insoluble fraction (lane 5), soluble fraction (lane 6), flow-through (lane 7), native wash fractions (lane 8- 10), duplicate native elution fractions with 100mM (lane 11), 150 mM (lane 12-13), 200 mM (lane 14-15), and 250 mM imidazole (lane 16-17), and nickel sepharose beads (lane 18) are shown. Purified $(His)_6$-PfDXR(K295N, K297S) is indicated with an arrow.

3.4.5 The *in vitro* Analysis of the Rational Amino Acid Substituted *E. coli* DOXP Reductoisomerase Proteins

The Trp^{212} and the Met^{214} residues present within the catalytic hatch have been considered essential for the DOXP reductoisomerase catalysis based on predictions from structural data. They are considered to provide a stacking role, together with Pro^{274}, essential in ordering of the active site and association with NADPH. Both residues were shown to be conserved amongst DOXP homologues (Figure 3.1) and are thought to be involved in hydrophobic interactions with the NADPH cofactor as well as with the carbon backbone of the substrate, making them essential in DOXP reductoisomerase activity. These residues were targeted for rational amino acid substitution to further elucidate the role they provide in association with the substrate, cofactor and the active site structural residues in ensuring DOXP reductoisomerase activity.

Proteins that contained rational amino acid substitutions were analysed for the effect of the substitution upon the activity of DOXP reductoisomerase according to the NADPH coupled enzyme assay described in section 2.3.3.3 (Figure 3.10). Specific activities were determined for the modified proteins according to the published assay which, uses substrate concentrations that are saturating (excess relative to enzyme), but which are not necessarily steady state conditions. To further characterize the effects of the mutations, Michaelis-Menten and Lineweaver-Burk plots were generated to determine K_m values. While the specific enzyme activities for EcDXR and its derivatives was determined from three independent batches of purified protein, the K_m values for the EcDXR derivatives were determined for only one batch of purified protein. Therefore, the K_m values for the EcDXR derivatives were considered to be preliminary data. However, the K_m values were used as an indicator of any effect on substrate (DOXP) binding affinity. The enzyme kinetic data and specific enzyme activities are illustrated within Table 7 and discussed in the following section.

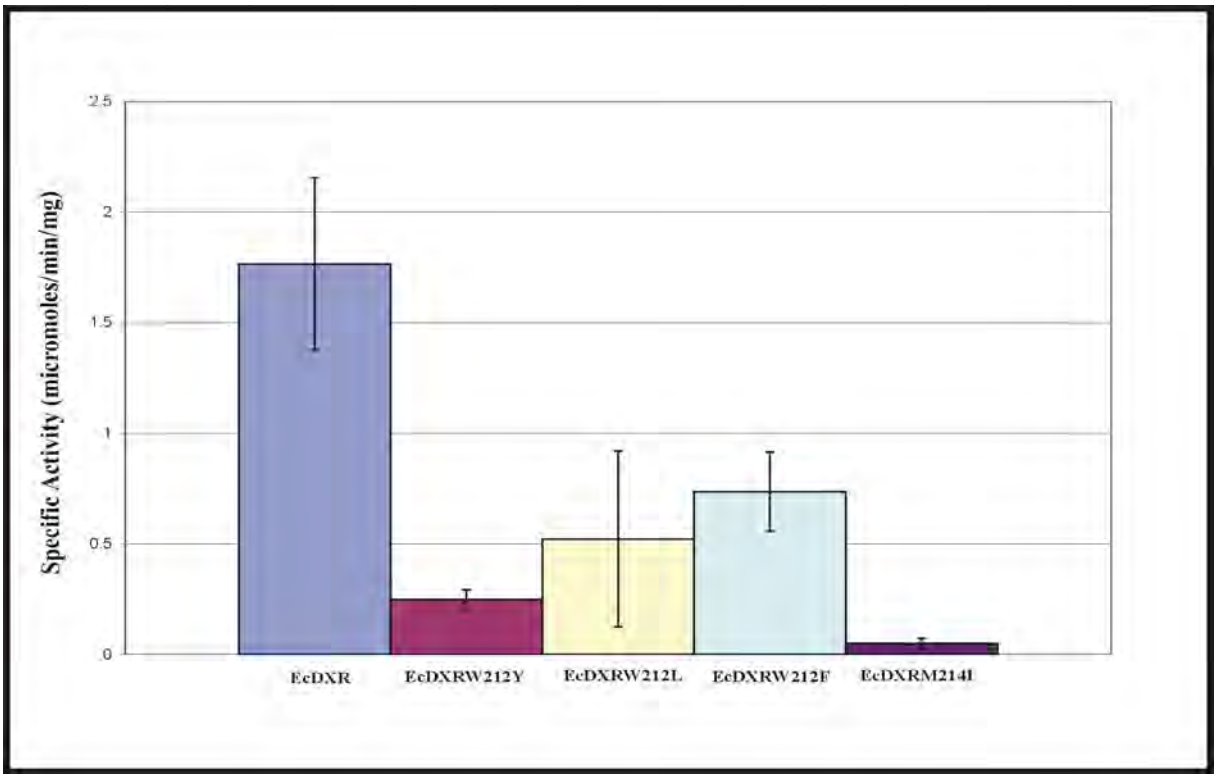


Figure 3.10: DOXP reductoisomerase specific activity for EcDXR and its derivatives

The NADPH coupled DOXP Assay was monitored at 340 nm for the oxidation of NADPH. Final substrate and NADPH concentrations of 0.3 mM were used. Three independent purifications for the EcDXR and its derivatives were used to assay for DOXP activity in triplicate with the bars representative of standard deviation. Bar graphs are representative of the specific activity obtained for EcDXR (blue), EcDXR(W212Y) (dark pink), EcDXR(W212L) (yellow), EcDXR(W212F) (light blue) and EcDXR(M214I) (purple) and are labelled below accordingly. All derivatives of EcDXR resulted in the decline of specific activity with the EcDXRW212Y and EcDXRM214I derivatives causing a significant decrease of specific activity on comparison with the wild-type EcDXR. The specific activities and standard deviations calculated are depicted in Table 7.

Trp^{212} was substituted with three residues; tyrosine, leucine and phenylalanine and Met^{214} with an isoleucine. All purifications of the EcDXR derivatives were successful with the proteins being both soluble and active. There is the assumption that the protein is folded and functional on a gross level and that any perturbations have occurred at a local level at the active site. However, we are aware that thorough characterization of the folded protein is ideally required e.g. circular dichroism (CD) for secondary structure analysis and 3-D crystal structure analysis for tertiary structure analysis. Additionally, validation of the substitutions on substrate affinity, inferred indirectly here from K_m analysis, is still required through the determination of direct binding affinities (K_D) using non-hydrolysable substrates/inhibitors (discussed in Chapter 5, section 5.2).

Upon introduction of a tyrosine residue in EcDXR(W212Y) the specific activity and affinity for the substrate were affected (refer to Figure 3.10 and Table 7). EcDXR(W212Y) had a 7 fold lower specific activity compared to EcDXR (0.25 μmol/min/mg ± 0.04; Figure 3.10), possibly as a result of the tyrosine being unable to engage in contacts that the tryptophan residue was previously involved in. Interaction of the indole ring of His^{257} with Trp^{212} was possibly eliminated, preventing the correct orientation of these residues. It is also possible that the polar hydroxyl group of the tyrosine could not be accommodated within the active site resulting in the loss of structural packing of the active site and the decrease in enzyme activity. In particular, the hydrophobic stacking interactions that Trp^{212} had with the Met^{214} and Pro^{274} were possibly not accommodated by a tyrosine at position 212. The hydroxyl moiety of tyrosine may seek to hydrogen bond with other residues either within the catalytic pocket or elsewhere in the protein. The hydroxyl group could potentially form hydrogen bonds with catalytic residues (e.g. Ser^{222}) thereby competing with the substrate leading to the loss of affinity of enzyme to substrate along with loss of activity. Therefore, the hydrophobic interactions between Trp^{212}, Met^{214} and Pro^{274} which are necessary for the correct active site architecture and binding of substrate, as well as activity, are affected by the W212Y substitution creating a protein no longer as efficient in the catalysis reaction and secure binding of the substrate

When Trp^{212} was substituted with leucine the specific activity of the resultant protein EcDXR(W212L) was lower than that of EcDXR, but higher than that of EcDXR(W212Y)(0.52 μmol/min/mg ± 0.4; Figure 3.10 and Table 7). The EcDXR(W212L) DOXP reductoisomerase activity was determined to have a K_m value of ~8 μM. The leucine may be accommodated within the catalytic pocket more effectively than a tyrosine resulting in a higher specific activity for EcDXR(W212L) compared to EcDXR(W212Y). The affinity for the substrate was found to be increased relative to EcDXR (K_m of ~8 μM). The leucine may be able to maintain the hydrophobic interactions between the residues essential for structural integrity of the active site and DOXP binding. More specifically, the leucine may be able to interact with the Met^{214} and Pro^{274} sufficiently enough to maintain the structural integrity of the active site despite the loss of the tryptophan residue. The interaction between His^{257} and Ser^{151} and interaction between Trp^{212} and His^{257}, through hydrogen bonding would have been removed by introducing the leucine at position 212 and non-polar interactions between the leucine and other residues would be favoured

in EcDXR(W212L). This suggests that while the substrate binding affinity was improved for EcDXR(W212L), possible due to increased hydrophobic interactions within the active site, the orientation of the His^{257}, Ser^{151}, Met^{214} and Pro^{274} may have been disrupted sufficiently to perturb catalytic activity slightly compared to EcDXR.

Trp^{212} was substituted with a phenylalanine residue in EcDXR(W212F) affecting specific activity and substrate affinity (refer to Figure 3.10 and Table 7). EcDXR(W212F) had a 2 fold reduction in specific activity (0.74 μmol/min/mg) compared to EcDXR, but higher than that of EcDXR(W212Y) and EcDXR(W212L). The phenylalanine residue may be accommodated within the active site more effectively than a tyrosine or leucine. Phenylalanine may maintain the hydrophobic interactions within the active site and still interact with Met^{214} and Pro^{274} sufficiently enough to maintain the structural integrity of the active site despite the loss of the tryptophan residue. The interaction between His^{257} and Trp^{212} was possibly interrupted on the phenylalanine substitution and the stacking of the active site slightly altered causing the decrease in enzyme activity in EcDXR(W212F). The affinity for substrate was found to be decreased relative to EcDXR, possibly due to the phenylalanine residue not being able to interact with the carbon backbone of the substrate. Any potential interactions between the Trp^{212} and the substrate backbone are affected by the altered orientation of the active site residues on phenylalanine substitution, thus lowering the affinity for substrate. Therefore, Trp^{212} may make essential contacts with the carbon backbone of the substrate allowing for the secure binding as well as providing a stacking role essential for the correct architecture of the active site.

The Met^{214} residue is closely associated with NADPH and was substituted with an isoleucine residue. This substitution resulted in a specific activity of 0.049 ± 0.02 μmol/min/mg and a K_m value of ~8 μM (Figure 3.10, Table 7). The EcDXR(M214I) specific activity was significantly reduced (36 fold) compared to EcDXR, whereas the K_m value of ~8 μM is suggestive of an increase in affinity of EcDXR(M214I) to the substrate (DOXP). The Met^{214} residue is suggested to interact with Pro^{274} and Trp^{212} and in the presence of NADPH allow for the structural architecture of the active site to be correctly orientated. The isoleucine may be able to maintain the hydrophobic interactions and hydrogen bonding interactions made between the substrate and the structural and catalytic residues. The possible increase in substrate affinity suggests that it is

the interference of the contact made between Met^{214} and NADPH, which has resulted in the loss of enzyme activity. To elucidate if indeed the interaction between Met^{214} and NADPH has been interrupted, calculation of the K_m with regard to NADPH is necessary.

Table 7: DOXP reductoisomerase kinetic parameters for EcDXR and its derivatives

	Specific Activity (μmol/min/mg) ± Std. Dev.	**K_m[a] (μM)**
EcDXR	1.77 ± 0.39	282
EcDXRW212Y	0.25 ± 0.04	N.D.
EcDXRW212L	0.52 ± 0.4	8
EcDXRW212F	0.74 ± 0.18	N.D.
EcDXRM214I	0.05 ±0.02	8

a = The K_m values were determined from Michaelis-Menten and Lineweaver-Burk plots using a substrate concentration range of 0 – 300 μM

N.D. = K_m could not be determined accurately using the standard substrate concentration range most likely due to greatly reduced substrate affinity (K_m in millimolar range).

3.5 Conclusion

The analysis of the sequence identity of *E. coli* and *P. falciparum* DOXP reductoisomerase revealed a relatively high level of identity at the protein level. The construction of a predicted structure for *P. falciparum* DOXP reductoisomerase using the *E. coli* DOXP reductoisomerase as a template was performed. The predicted 3-D structure of *P. falciparum* DOXP reductoisomerase may be used for structural, functional and catalytic predictions regarding the parasitic protein's mechanism. Two homology models were generated based on crystal structure data available (Reuter *et al*., 2002, Yajima *et al*., 2002 and Steinbacher *et al*., 2003). The first predicted 3-D structure for *P. falciparum* DOXP reductoisomerase was generated from crystal structure data from the ligand-free *E. coli* DOXP reductoisomerase (Reuter *et al*., 2002; Yajima *et al*., 2002) whereas the second utilized X-ray crystal structure co-ordinates of the substrate bound, cofactor bound and metal cation bound *E. coli* DOXP reductoisomerase.

The 3-D predicted structure for *P. falciparum* DOXP reductoisomerase, when utilizing the ligand bound template, closely resembled the 3-D structure of *E. coli* DOXP reductoisomerase. The 3-D predicted structure for *P. falciparum* DOXP reductoisomerase based upon unbound template

varied from *E. coli* DOXP reductoisomerase`s 3-D structure greatly. This is possibly as a result of the X-ray crystal structure of *E. coli* DOXP reductoisomerase used being impaired due to the absence of the NADPH, substrate and metal cation.

Bioinformatic analysis of the primary amino acid sequences of *E. coli* and *P. falciparum* DOXP reductoisomerase was via multiple sequence alignment and analyses of their 3-D structures. Residues that have been previously identified as critical in DOXP reductoisomerase activity were identified within the multiple sequence alignment for degree of conservation. In particular, residues within the catalytic hatch of *P. falciparum* and *E. coli* DOXP reductoisomerase were analysed for conservation and 3-D orientation.

From previous analysis of the 3-D structure of *E. coli* DOXP reductoisomerase, amino acids essential in the functioning of the enzyme have been identified. Trp^{212}, Met^{214} and Pro^{274} have been previously identified as critical in the orientation of the architecture of the active site in the presence of NADPH, and interaction with the carbon backbone of the substrate (Reuter *et al*., 2002, Yajima *et al*., 2002 and Steinbacher *et al*., 2003), however this has not been experimentally verified. In this investigation Trp^{212}, Met^{214} and Pro^{274} were analysed by site-directed mutagenesis and the substitution effect on the enzyme monitored via DOXP reductoisomerase activity assays.

Upon introduction of a tyrosine in EcDXR(W212Y) the specific activity decreased 7-fold upon comparison with EcDXR possibly as a result of the tyrosine no longer able to provide the essential contacts that the tryptophan was previously involved in. The hydroxyl group of tyrosine was possibly not accommodated within the active site resulting in loss of interaction with Met^{214} and Pro^{274} and therefore a decrease in enzyme activity. The decrease in affinity in EcDXR(W212Y) to substrate relative to EcDXR could potentially be the result of the hydroxyl group competing with the substrate for hydrogen bonding interactions with the active site residues.

On substitution with leucine in EcDXR(W212L) the leucine may be accommodated within the catalytic pocket more effectively than a tyrosine resulting in a higher specific activity relative to

EcDXR(W212Y). In particular, the leucine may manage to maintain some of the hydrophobic interactions between residues essential for structural integrity of the active site. Additionally, leucine may be able to interact with Met^{214} and Pro^{274} sufficiently enough to maintain the active site structure. However, the Met^{214} and Pro^{274} residues must be disrupted sufficiently to perturb catalytic activity in comparison to EcDXR.

EcDXR(W212F) had a 2-fold lower specific activity compared to EcDXR, but its activity was higher than that of EcDXR(W212Y) and EcDXR(W212L). The W212F substitution may maintain some of the hydrophobic interaction within the active site pocket and still interact to some extent with Met^{214} and Pro^{274}. The affinity of EcDXR(W212F) for DOXP was reduced relative to EcDXR. The phenylalanine residue can likely no longer associate with the carbon backbone of the substrate. Therefore, for EcDXR(W212F) the Phe residue may maintain a degree of structural integrity by interaction with Met^{214} and Pro^{274} but can no longer secure the substrate as effectively within the catalytic pocket.

The Met^{214} substitution with isoleucine resulted in an enzyme with 36-fold lower specific activity and increased affinity for DOXP compared to EcDXR. Since Met^{214} is closely associated with NADPH it is likely that the isoleucine can no longer harbour this interaction causing the decrease in enzyme activity observed. Furthermore, the isoleucine residue may no longer engage in interactions with Trp^{212} and Pro^{274} to allow for the correct architecture of the active site. The increase in substrate affinity observed for EcDXR(M214I) may be due to isoleucine maintaining hydrophobic interactions necessary between active site residues and DOXP. To validate whether the NADPH and Met^{214} interaction has been disrupted, calculation of the K_m with respect to NADPH is required.

The *P. falciparum* and *E. coli* DOXP reductoisomerase primary amino acid sequences were analysed, for any variation in sequence conservation. Variations in the amino acid sequence may affect biochemical properties of the protein and thus its structure and function. Specifically, analysis of the catalytic hatch region revealed residues Asn^{211} and Ser^{213} to be conserved amongst *E. coli* DOXP reductoisomerase and other DOXP reductoisomerase homologues except *P. falciparum* DOXP reductoisomerase where both are replaced with lysine residues. The different

3-D orientation of these topologically equivalent residues within *P. falciparum* and *E. coli* DOXP reductoisomerase is suggestive of a variation in the role of the catalytic hatch in securing the substrate and the mode in which it performs this function in *P. falciparum* DOXP reductoisomerase.

CHAPTER 4: STRUCTURE-BASED DRUG DESIGN: THE RATIONAL DESIGN OF NOVEL ANTI-MALARIAL DRUGS

4 STRUCTURE-BASED DRUG DESIGN: THE RATIONAL DESIGN OF NOVEL ANTI-MALARIAL DRUGS

4.1 Introduction

Computational structure-based drug design first made its appearance in the early 1980`s (Joseph-M^{c}Carthy, 1999). Apart from the development of the HIV-1 protease inhibitors, there have, thus far, been few success stories based on this approach. This apparent lack of progress will probably undergo a drastic change in the near future due to the development of functional genomics, enhanced computer speed and continuing improvements in structural and sequence information (Joseph-M^{c}Carthy, 1999). In addition, the coupling of structure-based drug design, also referred to as rational drug design, with combinatorial chemistry has allowed for a revolution in drug discovery. On its own, combinatorial chemistry had the disadvantage of producing huge numbers of compounds, many of which were irrelevant and ineffective. Now, with structure-based drug design, specific knowledge of the drug target can guide combinatorial chemistry to produce relevant compounds rapidly and in relatively large numbers.

4.1.1 Methods and Tools in Structure-Based Drug Design

Structure-based drug design methods can be grouped into two categories, namely, ligand-based (analogue-based) or target-based (structure-based) drug design; these can be used as separate strategies or together in the process of drug discovery (Bourne and Weissig, 2003). Ligand-based drug design is applicable to all drug discovery strategies as it is dependent upon a group of ligands that are known to bind to the drug target even if the structure of the target has not been established. On the other hand, target-based drug design is dependent on the structure of the drug target or receptor site of a protein being determined.

In cases where no 3-D receptor information is available, analogue-based drug design may permit a 3-D receptor model to be generated from a set of known, active ligands. Analogue-based drug design thus involves the generation of a pharmocophore or quantitative structure-activity relationship (QSAR) from which a lead molecule can be identified based on its ability to interact

efficiently with the generated receptor model relative to the known set of active ligands (Bourne and Weissig, 2003).

Target-based drug design depends on known structural information regarding the target protein and receptor site obtained by X-ray crystallography or NMR spectroscopy. The identification of the active site is therefore based upon known ligand-receptor interactions. *In silico* docking methods then permit an assessment of the binding potential of novel ligands. Docking methods involve the generation of reactivity maps of the ligand docked within the active site of the target protein, and have the advantage of allowing for high-throughput screening. The different binding modes or reactivity maps for each ligand can be used to assess the potential of a ligand in terms of geometry, conformational space and energy. *De novo* synthesis involves the construction of a compound within the active site by joining atoms or chemical fragments together that have been specifically chosen for their ability to interact with the active site. The main disadvantage of this method is that the compounds are often difficult to synthesize (Bourne and Weissig, 2003).

4.2 Aim

The present investigation has aimed to use DOXP, fosmidomycin and FR900098 as lead molecules for the design of novel DOXP reductoisomerase inhibitors. The binding modes and reactivity maps of the lead molecules were assessed and potential DOXP inhibitors designed accordingly. *In silico* docking studies of novel DOXP reductoisomerase inhibitors were performed to permit assessment of the binding potential of the drugs. DOXP reductoisomerase inhibitors, which showed a promising binding affinity, were synthesized and tested using DOXP reductoisomerase inhibition assays.

4.3 Experimental Procedures

4.3.1 Ligand Docking and Rational Drug Design

Computer modelling was effected using the ACCELRYS Cerius^{2TM} Drug Discovery Workbench modelling platform on a Silicon Graphics O^2 computer; the module, LigandFit, was used for the

in silico docking experiments. Structural information regarding the target protein DOXP reductoisomerase was downloaded from the PDB (accession codes 1K5H, 1JVS, 1ONN, 1ONO and 1ONP). All ligands were subjected to energy minimization and dynamic simulation routines prior to docking.

4.3.2 General Organic Synthesis Procedures

NMR spectra were recorded on a Bruker Avance 400 MHz spectrometer at 303K in deuterated chloroform ($CDCl_3$) or deuterated dimethyl sulphoxide (DMSO-d_6). Spectra recorded in $CDCl_3$ were calibrated on the solvent signals at 7.25 ppm and 77.0 ppm in the ^{1}H and ^{13}C spectra, respectively. Spectra recorded in deuterated DMSO-d_6 were calibrated on the solvent signals at 2.5 ppm in the ^{1}H spectrum and 39.4 ppm in the ^{13}C spectrum. For ^{31}P NMR, H_3PO_4 was used as an internal standard at δ 0 ppm.

Normal phase HPLC was performed using a Whatman Partisil 10 Magnum 6 normal phase column, a Spectra-Physics P100 Isocratic pump and a Waters RI210 differential refractometer detector. Flash chromatography was adapted from the procedure described by Pederson and Rosenbohm (2001) using Merck Silica gel 60 (particle size 0.040 – 0.063 mm).

All solvents were dried according to standard procedures outlined by Perrin and Armarego (1988). Diethyl ether (Et_2O) and tetrahydrofuran (THF) were pre-dried over CaH_2 and then distilled from sodium wire and benzophenone under N_2, while *N*,*N*-dimethylformamide (DMF) was pre-dried with $MgSO_4$ and distilled from 3Å molecular sieves under reduced pressure.

4.3.3 Preparation of Phosphorylated Furan Derivatives

The two experimental approaches explored for the synthesis of the phosphorylated furan derivatives were adapted from a method reported by Li and Ganeson (1998).

***Preparation of 3-trityloxyfuran* (2)**

A mixture of triphenylmethyl chloride (8.36 g, 0.03 mol), 3-furanmethanol (**1**) (3 g, 0.03 mol), triethylamine (8 ml; 5.808 g, 0.06 mol) and 4-dimethylaminopyridine (DMAP; 291 mg, 0.002 mol) was stirred under N_2 for 15 hours at 80˚C. Evaporation of the triethylamine *in vacuo* gave the crude *3-trityloxyfuran* (**2**) (7.084 g, 62 %) as a yellow gum [δ_H (400 MHz; $CDCl_3$) 3.41 (2H, s, CH_2) and 7.13 – 7.91 (18H, series of multiplets, Ar-H); δ_C (100MHz; $CDCl_3$) 40.6 (CH_2), 82.0 (Ph_3C), 106.8, 127.2, 127.88, 127.90, 129.6, 139.5 and 146.9 (Ar-C)], which was used without further purification.

***Preparation of 5-trityloxyfuran-2-carbaldehyde* (3) *and 3-trityloxyfuran-2-carbaldehyde* (4)**

In a dry, round-bottomed flask fitted with a septum seal, a solution of crude 3-trityloxyfuran (**2**)(1g, 0.003mol) in dry THF under N_2 was cooled to –30˚C. Butyllithium (1.5 M in hexane; 4 ml, 0.006 mol) was then added drop wise from a syringe and the reaction mixture stirred for 4 hours at -30˚C. DMF (0.46 ml, 6.0 mmol) was added to the resulting red solution. The reaction mixture was allowed to warm to room temperature, and then stirred for a further 2 hours. The mixture was poured into a slurry of 10 % aqueous $NaHCO_3$ and ice and then extracted using Et_2O (2 x 25 ml). Vacuum flash chromatography (elution with hexane-ethyl acetate) afforded a residue (2.04 g), shown by ^1H NMR analysis to comprise a 0.77:1 mixture of the isomeric aldehydes (**3**) and (**4**). Semi-preparative HPLC (elution with MeOH) afforded *5-trityloxyfuran-2-carbaldehyde* (**4**) as a dark yellow oil; δ_H (400 MHz; $CDCl_3$) 4.06 (2H, s, CH_2), 7.1-7.6 (17H, series of signals, Ar-H), and 9.59 (1H, s, CHO); δ_C (100MHz; $CDCl_3$) 57.8 (CH_2), 87.3 (Ph_3C), 126.4, 127.2, 127.9, 128.0, 128.5, 143.6, 145.2 and 153.0 (Ar-C) and 178.0 (CHO).

***Deprotection and phosphorylation of compounds* (3) *and* (4)**

A 0.77:1 mixture of the isomeric aldehydes (**3**) and (**4**) (9.0 g in 10 ml dry THF; 24 mmol) was stirred together with H_3PO_4-THF (1:1 v/v; 10 ml) for 2 days at room temperature. EtOAc (20 ml) and 10 % aqueous $NaHCO_3$ (20 ml) were then added. Effervescence was noted to occur upon addition of the $NaHCO_3$ and the aqueous fraction was collected (pH 9.0). The aqueous fraction was acidified to pH 2.0 with 10 % aqueous HCl and then extracted with EtOAc (3 x 10 ml). The organic extract was dried over anhydrous $MgSO_4$ to afford an oil indicated by NMR analysis to

contain the isomeric phosphorylated aldehydes (**5**) and (**6**). Annotated NMR spectra of the mixture of products are provided in the Appendix (Chapter 6, section 6.3).

Preparation of polymer-bound 3-furanmethanol

A mixture of polymer bound 2-chlorotrityl chloride (~1.9 mmol Cl/g resin; 4 g), 3-furanmethanol (**1**) (0.157 g; 0.0016 mol) and DMAP (0.003g; 0.025 mmol) in dry DMF was stirred under N_2 at 80˚C for 15 hours. Evaporation of the solvent *in vacuo* gave the polymer-bound product (**7**) as pale yellow granules.

***Formylation of the polymer-bound substrate* (7)**

The yellow product (**7**) was suspended in anhydrous THF in a dry, round-bottomed flask fitted with a septum seal under N_2, and the temperature was lowered to -30˚C. Butyllithium (1.5 M in hexane; 5.87 ml, 0.009 mol) was then added drop wise using a syringe and the reaction mixture stirred at -30˚C for 4 hours. The electrophile, DMF (1.17 g; 0.016 mol), was added to the red mixture, which was allowed to warm to room temperature and then stirred for a further 2 hours. Methanol (40 ml) was added and the resin washed with aqueous methanol (20 ml) and THF (20 ml) successively to give the polymer-bond aldehyde (**8**).

Cleavage of 5-(hydroxymethyl)furan-2-carbaldehyde* (9) *from the polymer support

A mixture of the polymer-bound aldehyde (**8**) (3.82 g), trifluoroacetic acid (1 ml), triethylsilane (1 ml), and dichloromethane (18 ml) was stirred for 20 minutes at room temperature under N_2. Evaporation of the filtrate *in vacuo* afforded a solution of *5-(hydroxymethyl)furan-2-carbaldehyde* (**9**) in trifluoroacetic acid.

4.4 Results and Discussion

4.4.1 Ligand Docking and Rational Drug Design

To identify and characterize the binding site within DOXP reductoisomerase, previous X-ray crystal structure data and residues known to be involved in the catalysis were used. The detailed 3-D information gave an indication of where the binding pocket is located. Using a Molecular

Mechanics (MM) approach and the LigandFit module, small molecule binding sites on the protein surface were identified. Overlap of the known catalytic residues with the putative binding pockets on the protein surface permitted the binding site within DOXP reductoisomerase to be identified (Figure 4.1).

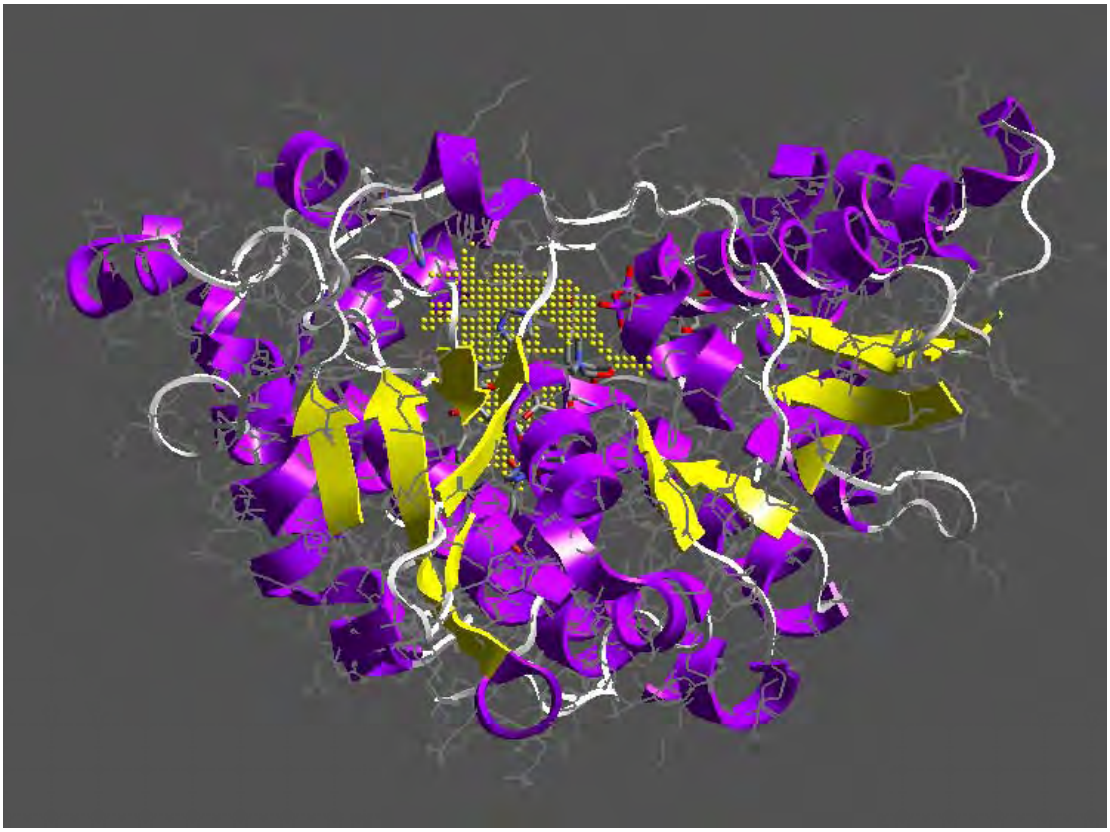


Figure 4.1: The *E. coli* DOXP reductoisomerase monomer with the site model identifying the active site

Ribbon representation of the *E. coli* DOXP reductoisomerase with α-helices shown in purple and the β-sheets shown in yellow. The known catalytic residues (shown in solid stick) are situated within the catalytic domain. The identified binding pocket of DOXP reductoisomerase is indicated by yellow dots and fills the groove within the V-shaped monomer.

For docking purposes, the unbound protein templates were used (1K5H and 1JVS)(Reuter *et al*., 2002, Yajima *et al*., 2002). This was to rule out any fine structural variations within the active site that occur on binding to different ligands, such as fosmidomycin in this case, which can interfere with the docking method. Such details are often not detected within docking strategies (Fradera *et al*., 2002). Initially, the substrate and known inhibitors (fosmidomycin and

FR900098) were flexibly docked into the defined active site of DOXP reductoisomerase to analyse interactions between the respective ligands and the active site residues. Examination of these known ligands was expected to yield valuable information on the capacity of the active site to bind a drug molecule. The binding modes of these ligands were then compared, and Figure 4.2 shows the conformational space occupied by DOXP (A and B), fosmidomycin (B and C) and FR900098 (D and E). The receptor model of the bound ligand was coloured according to the electrostatic potential (EP) [blue (+ve) to red (-ve)]. The change in the electrostatic potential following the introduction of nitrogen within the inhibitor molecule is apparent when compared with the natural substrate DOXP. In previous DFT studies, the nitrogen atoms within the inhibitory molecules FOSM and FR-900098 were reported as having calculated charges of –0.149 and 0.195 respectively, whereas the carbon atom within DOXP was found to be almost neutral with a charge of –0.011 (Zubrzycki and Blatch, 2001). The increase in negative charge permits the drug to bind with increased affinity to the active site.

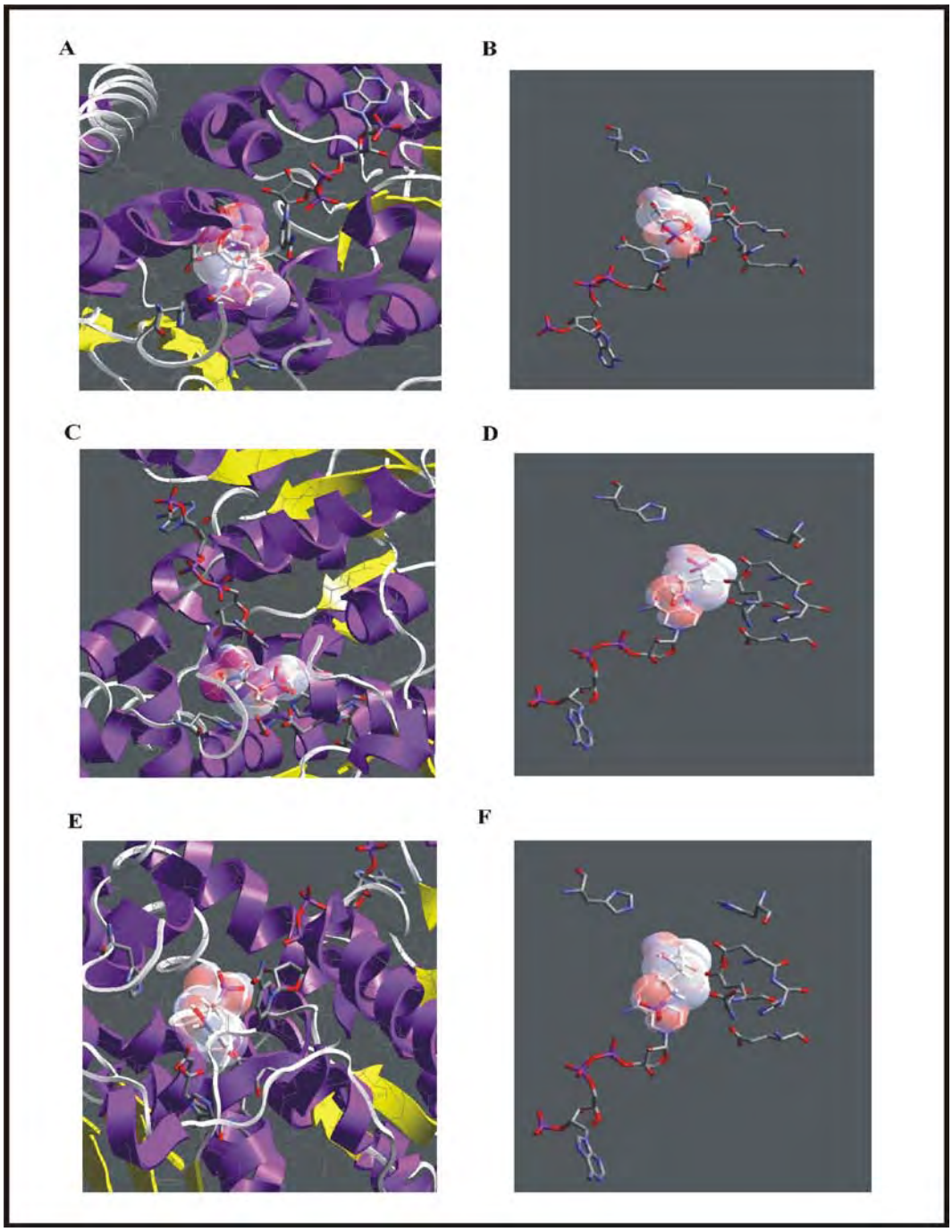


Figure 4.2: Binding of DOXP, fosmidomycin and FR900098 to the identified active site within DOXP reductoisomerase

(A, B) The natural substrate (DOXP) flexibly fitted into the catalytic active site of DOXP reductoisomerase. The electrostatic potential surrounding the ligand is indicated from blue (+ve) to red (-ve). (C, D) The inhibitor fosmidomycin flexibly docked within DOXP reductoisomerase. The electrostatic potential of the receptor model indicates an increase in the negative charge surrounding the N atom. (D, F) The derivative FR900098 docked flexibly within the DOXP reductoisomerase active site. B, D, and F all reflect the same orientation with respect to NADPH in the bottom left corner.

For the rational design of novel DOXP reductoisomerase inhibitors, a combination of activity, specificity and binding stability characteristics were considered. In terms of specificity and activity the conversion of DOXP to MEP mediated by DOXP reductoisomerase was analysed (see Chapter 1; Figure 1.5). DOXP reductoisomerase catalyzes the conversion of DOXP to MEP through an intramolecular rearrangement. Through the introduction of a ring structure within the carbon backbone of a drug mimicking DOXP, it was hypothesised that the conformational flexibility would be decreased and that the resulting entropic advantage would enhance binding to DOXP reductoisomerase. A set of compounds were designed to mimic the carbon backbone of DOXP and the inhibitors FOSM and FR900098, but contain the furan ring (Figure 4.3).

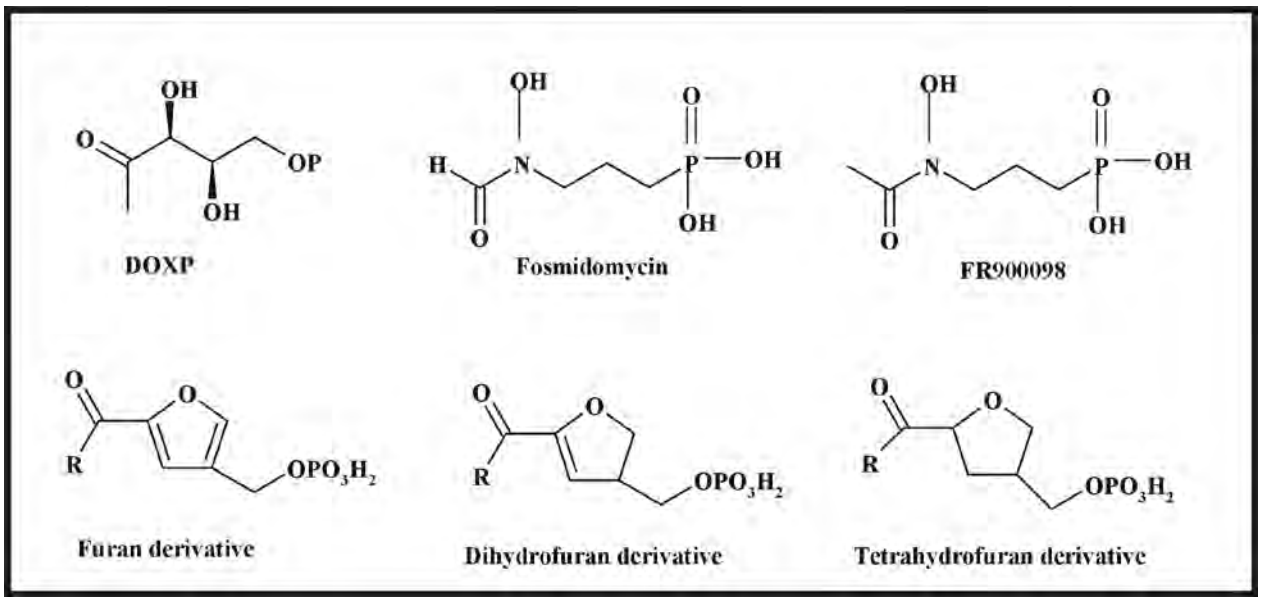


Figure 4.3: Chemical Structures of DOXP, fosmidomycin, FR900098 and the furan derivatives targeted for synthesis (R= H, Me, Et etc.)

The furan, dihydrofuran and tetrahydrofuran derivatives were fitted into the identified active site of DOXP reductoisomerase and scored using LIGSCORE to determine the ranking order of each fitted conformation of the ligand. The ligand that was best conformationally suited to the active site was then aligned with the fitted conformations of the substrate, fosmidomycin and FR900098 (Figure 4.4). Alignment methods allow for the superposition of molecules using least squares fitting, with atom equivalencies specified (in this case manually). The alignment data (Figure 4.4) indicates the structural similarity of the fitted conformations of the different ligands. The

rationally designed furan derivatives align with a high level of confidence suggesting they could adopt the conformational geometry required for binding to the active site of DOXP reductoisomerase. By improving the specificity of the furan derivatives one could generate a novel inhibitor with a higher affinity for the active site. Improving the specificity could involve the introduction of a negative nitrogen within the ring structure or carbon backbone of the furan derivatives to increase the affinity of the drug toward the active site. Furthermore, by altering the R-substituent (R=H, Me, Et *etc*.) one could monitor the effect of different substituents on the binding affinity of the drug.

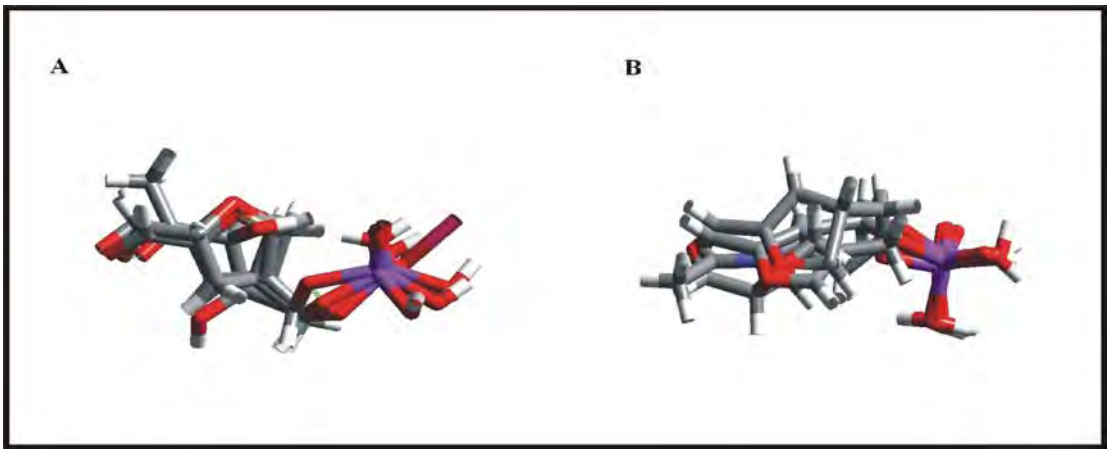


Figure 4.4: Alignment of the fitted conformations of the furan, dihydrofuran and tetrahydrofuran derivatives upon DOXP and FR900098

(A) The alignment of the fitted conformations of DOXP and the furan, dihydrofuran and tetrahydrofuran derivatives (R=H, Me and Et). The main carbon backbone aligns with a high degree of confidence with the DOXP substrate. (B) The fitted conformation of FR900098 aligned with the furan, dihydrofuran and tetrahydrofuran derivatives (R=H, Me and Et). The alignment suggests that the furan ring might orientate in a different conformation within the active site; however, it can still be accommodated by the active site, as the overall conformational space remains largely unaffected.

4.4.2 Synthesis of the Phosphorylated Furan Derivatives

The synthesis of the furan derivatives was based upon the procedure described by Li and Ganeson (1998) for the preparation of 2,4-disubstituted furans and thiophenes. The method involves lithiation of polymer-bound 3-furanmethanol, reaction with a suitable electrophile and subsequent removal from the resin. It was therefore a promising procedure for the production of the furan, dihydrofuran and tetrahydrofuran derivatives outlined in Figure 4.3. The polymer

bound 2-chlorotrityl chloride resin allows for C_5 lithiation to occur selectively due to the steric bulk of the 2-chlorotrityl moiety preventing electrophilic attack at C_2. In addition to using the polymer bound 2-chlorotrityl chloride resin, we decided to "protect" the alcohol function using triphenylmethyl chloride, assuming that the bulky trityl moiety would have the same steric effect as the polymeric system (Figure 4.5).

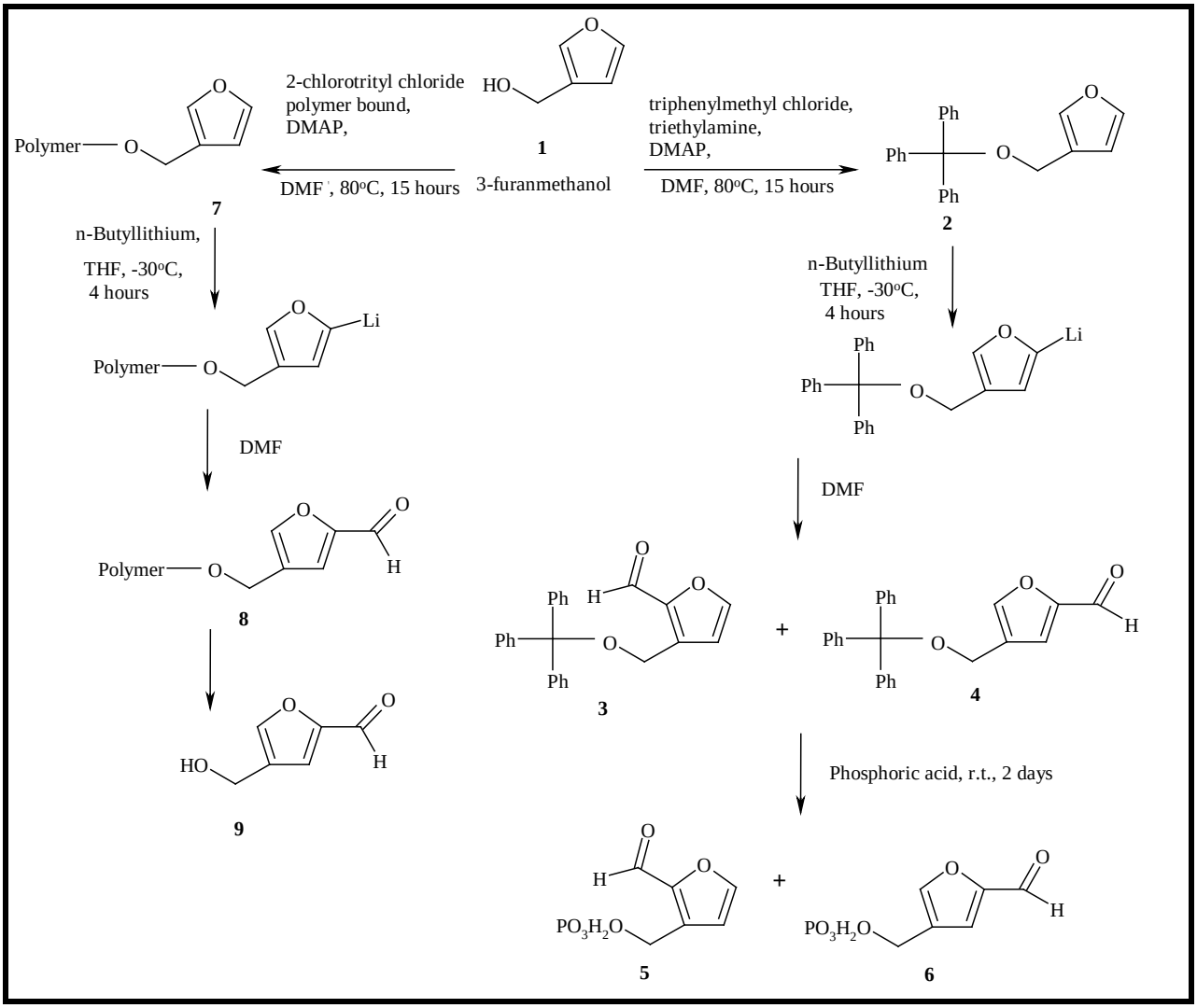


Figure 4.5: Synthesis of the furan derivative

Initially, 3-furanmethanol was reacted with triphenylmethyl chloride and with the 2-chlorotrityl chloride polymer resin. Direct lithiation of the resulting ethers (**2**) and (**7**) generated red

mixtures. Addition of the electrophile, DMF, yielded the formylated furan derivatives (**3**, **4** and **8**). The presence of the isomeric formylated derivatives (**3** and **4**) was indicated by ^{1}H and ^{13}C NMR analysis. Lithiation at both centres, C_2 and C_5, is attributed to the failure of the trityl group to provide as much steric hindrance at C_2 as the polymeric matrix, which only exhibits C_5 lithiation. Separation of the isomers (**3** and **4**) was achieved using normal phase HPLC. Figure 4.6 and Figure 4.7 illustrate the ^{1}H and ^{13}C NMR spectra obtained for compound (**4**).

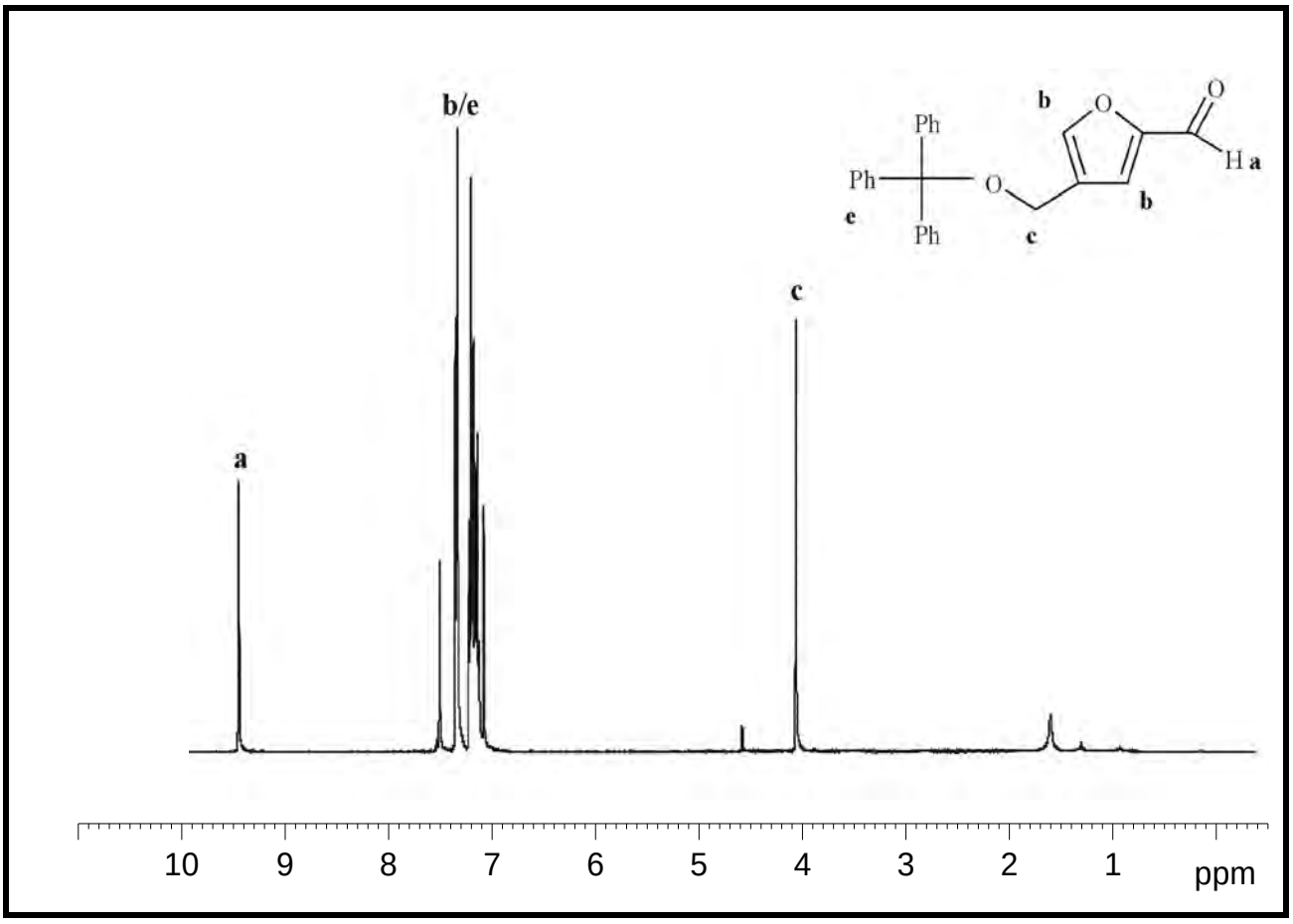


Figure 4.6: δ 400 MHz ^{1}H NMR spectrum for compound 4 (separated by normal phase HPLC) in $CDCl_3$

The aldehyde proton peak (a) can be seen at ~9.5 ppm and the CH_2 signal (c) at 4 ppm. The aromatic phenyl (e) and furan protons (b) are present in the region 7.0 ppm to 7.5 ppm.

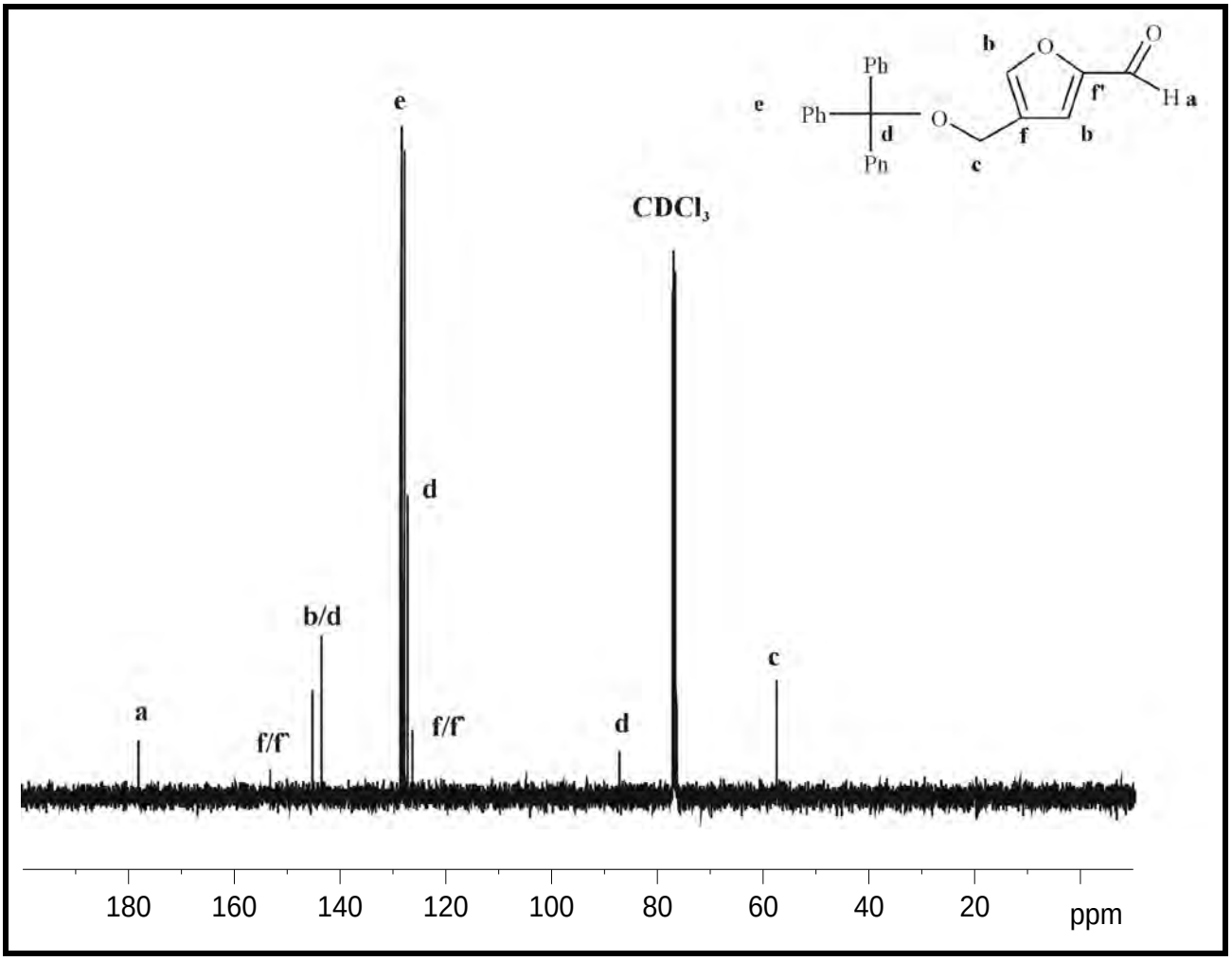


Figure 4.7: 100 MHz ^{13}C NMR spectrum of compound 4 (separated by normal phase HPLC) in $CDCl_3$

Cleavage of trityl ethers has been achieved under mild acid hydrolysis conditions with TFA and through catalytic hydrogenation (Barton and Ollis, 1979). Cleavage of the 5-(hydroxymethyl)furan-2-carbaldehyde (**9**) from the polymer support using TFA was achieved successfully in low yield. Thus, phosphorylation was not attempted within this synthetic route. The deprotection and phosphorylation of the isomeric tritylated ethers (**3**) and (**4**) to afford compounds (**5**) and (**6**) using phosphoric acid appeared to be successful. The presence of the phosphorylated isomeric aldehydes (**5**) and (**6**) was detected using ^{31}P NMR (Figure 4.8). A single signal at 4 ppm is indicative of the phosphorylated products (phosphoric acid was used as an internal standard at δ 0 ppm).

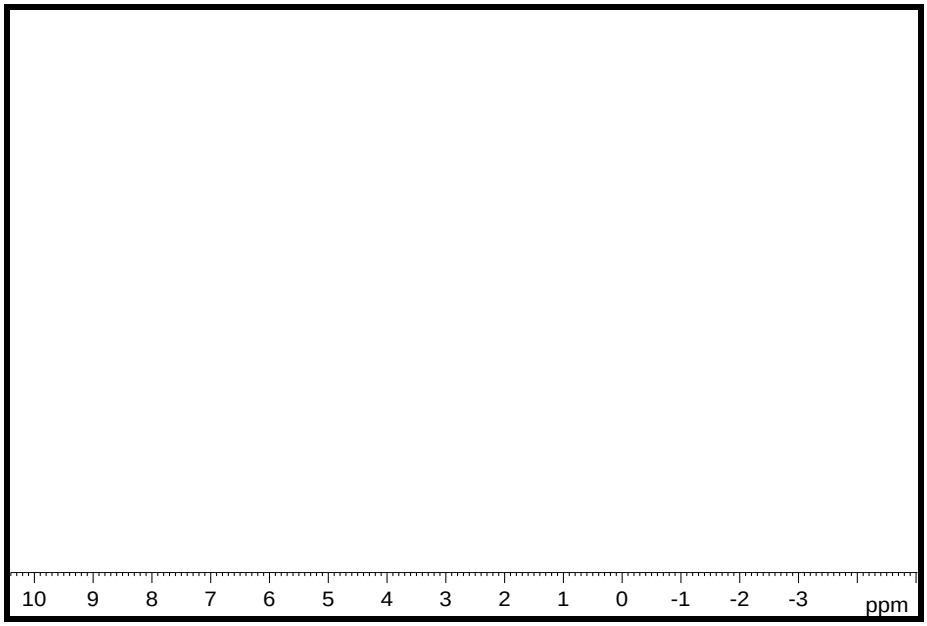


Figure 4.8: δ ^{31}P NMR of compounds 5 and 6 in DMSO-d_6

^{31}P NMR spectrum of the isomeric phosphorylated aldehydes (**5**) and (**6**).

Since the tritylated isomers (**3**) and (**4**) could be separated using normal phase HPLC, it was anticipated that the analogues (**5**) and (**6**) could be separated similarly.

4.5 Conclusion

The aim of the structure-based drug design was twofold. Firstly, detailed knowledge of the biochemically important features of the protein active site would aid in the development of inhibitors that would interact strongly with DOXP reductoisomerase with the added benefit of minimizing secondary interactions with other proteins. Secondly, to rationally design novel inhibitors through ligand docking strategies. The introduction of a ring structure within the carbon backbone of a drug mimicking DOXP, was expected to decrease conformational flexibility and thus enhance binding to DOXP reductoisomerase. A set of furan derivatives, designed to mimic the carbon backbone of DOXP and the inhibitors FOSM and FR900098, demonstrated efficient *in silico* binding modes, and a synthetic route to a phosphorylated target was developed.

CHAPTER 5: CONCLUSIONS AND FUTURE WORK

5 CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

5.1.1 Heterologous Protein Production and Purification of DOXP reductoisomerase

The over-production and nickel-chelate affinity purification of *E. coli* DOXP reductoisomerase in a native state was successfully achieved. The recombinant *E. coli* DOXP reductoisomerase had a specific activity of 1.77 ± 0.34 μmol/min/mg and a K_m and V_{max} of 282 μM and 0.74 μmoles/min respectively. The K_m value determined was in close proximity to the published value of 171 μM (Wungsintaweekul, 2001).

The difficulties in the over-production of *P. falciparum* DOXP reductoisomerase and consequent problems in the purification of the recombinant protein made functional analysis of the protein not possible. Different methods of purification were attempted, as well as alteration of the growth parameters, to attempt to purify sufficient levels of native *P. falciparum* DOXP reductoisomerase. The low levels of the protein prior to and upon IPTG induction suggest that the recombinant protein was toxic toward the host expression system.

5.1.2 Characterization of the DOXP reductoisomerase Catalytic Site by Rational Amino Acid Substitution Analysis

The bioinformatics analysis of the multiple sequence alignment of the DOXP reductoisomerase homologues contained conserved residues known to be critical in the functioning of the DOXP reductoisomerase protein. The Trp^{212}, Met^{214} and Pro^{274} residues have been previously predicted, from crystal structure data, to perform a stacking role essential in the orientation of the active site architecture in the presence of NADPH. These residues have not been tested experimentally and were therefore analysed by site-directed mutagenesis and DOXP reductoisomerase activity assays to further elucidate their roles within DOXP reductoisomerase activity.

The Trp^{212} residue mutations were indicative of this residue performing a vital role in the correct orientation of the active site architecture as previously predicted. Mutations of this residue to a Leu, Phe and Tyr residue respectively, resulted in the decrease of DOXP reductoisomerase

activity. This was possibly as a result of a loss of structural integrity within the active site likely affected due to the interaction of Met^{214} and Pro^{274} being altered on substitution of the Trp^{212} residue. The EcDXR(W212Y) protein had a lower specific activity and reduced affinity for substrate relative to EcDXR, possibly due to the hydrophobic nature of the active site being affected as well as possible competing hydrogen bonding interactions occurring between the hydroxyl of the tyrosine residue and the residues normally involved within substrate binding. The EcDXR(W212F) also had a decreased affinity for substrate possibly due to the Phe residue no longer being able to interact with the carbon backbone of the substrate. The EcDXR(W212L) protein had a K_m value of ~8 μM indicative of an increase in substrate affinity. This is most likely due to the hydrophobic nature of the active site being improved and allowing for improved substrate binding.

The EcDXR(M214I) protein had a reduced enzyme activity however, substrate affinity was improved (K_m of ~8 μM). The M214I substitution resulted in the possible disruption of the binding of NADPH essential for catalysis and the significant decrease in enzyme activity observed relative to EcDXR. The substrate affinity is not likely to be affected negatively due to the interaction with Trp^{212} and Pro^{274} being sufficiently maintained by the Ile residue. The improved substrate affinity is possibly due to the hydrophobic interactions being favoured within the active site pocket. The possible increase in substrate affinity suggests that it is the interference of the interaction between Met^{214} and NADPH that has caused the decrease in enzyme activity.

The bioinformatics analysis of the primary amino acid sequence of *P. falciparum* DOXP reductoisomerase with its homologues revealed a deviation in the consensus sequence within the catalytic hatch of the protein. Two lysine residues occur at position 295 and 297 in the *P. falciparum* sequence compared to conserved Asn and Ser residues respectively within other DOXP reductoisomerase homologues. The orientation of the Lys^{295} and Lys^{297} (Figure 3.5(C)) residues were shown to protrude into the catalytic pocket within the predicted 3-D structure of *P. falciparum* DOXP reductoisomerase generated through homology modelling. The topologically equivalent Asn^{211} and Ser^{213} in *E. coli* DOXP reductoisomerase are however orientated differently. In particular, the Asn^{211} was found to protrude outward suggestive of it not being

involved in direct association with the substrate. These differences in the orientation of the topologically equivalent residues within *P. falciparum* and *E. coli* DOXP reductoisomerase suggests a variation in the role that the catalytic hatch provides in securing the substrate and the mode in which it performs this function within the parasitic system.

To determine whether the activity of DOXP reductoisomerase and its affinity to substrate are mediated differently between the *P. falciparum* and *E. coli* DOXP reductoisomerase proteins site-directed mutagenesis of the Lys^{295} and Lys^{297} within *P. falciparum* DOXP reductoisomerase to an asparagine and serine respectively and vice versa for the *E. coli* DOXP reductoisomerase were prepared. The purification of these recombinant proteins was successful however not in sufficient levels for functional analysis.

Much of this investigation centres on the residues within the catalytic lid region of DOXP reductoisomerase. The Asn^{211}, Trp^{212} and Ser^{213} motif is the most predominant consensus sequence at this position with regard to the bacterial enzyme, and Lys^{295}, Trp^{296} and Lys^{297} present in the *P. falciparum* DOXP reductoisomerase sequence (refer to Figure 3.1). These amino acids may differ in other field isolates and laboratory strains for example mouse *P. yoelii* and *P. berghei* have a Lys^{295}, Trp^{296} and Asn^{297} motif and for plant species an Asn^{274}, Trp^{275} and Ser/Asn^{276} motif occurs. These variations, at the catalytic lid region of the protein, further substantiate the importance of the catalytic hatch within the active site of the enzyme and reinforces the need to determine whether the activity of DOXP reductoisomerase and its affinity for the substrate are mediated differently amongst different species.

5.1.3 Structure-Based Drug Design – The Rational Design of Novel Anti-Malarial Drugs

The rational design of novel DOXP reductoisomerase inhibitors was based upon the lead molecules DOXP, fosmidomycin and FR900098 and *in silico* docking studies. The proposed furan derivatives were designed to mimic the carbon backbone of the natural substrate DOXP upon binding to the active site. Through the introduction of a ring structure within the carbon backbone of the drug mimicking DOXP, it was hypothesised that the conformational flexibility would be reduced and thus enhancing binding to DOXP reductoisomerase relative to the

substrate. This was demonstrated utilizing *in silico* alignment tools to compare the receptor model and binding modes of the furan, dihydrofuran and tetrahydrofuran derivatives to identify possible drugs. A synthetic route for the synthesis of phosphorylated furan derivatives was identified and used to synthesise possible DOXP reductoisomerase inhibitors.

5.2 Future Work

The low levels of over-expression of *P. falciparum* DOXP reductoisomerase suggest that this protein is toxic toward the host expression system, *E. coli*. Further optimization of the over-expression and purification of *P. falciparum* DOXP reductoisomerase is necessary to perform functional and inhibitory studies on the parasitic protein. This could involve the introduction of the pLysS plasmid to improve toxic protein expression in *E. coli*, as well as expression in different strains of *E. coli*, which are suited toward toxic protein production. The co-expression of *P. falciparum* chaperone proteins may additionally improve expression levels of the *P. falciparum* DOXP reductoisomerase recombinant protein as they may aid in protein folding. Additionally, moving into a yeast or *baculovirus* expression system for the expression of *P. falciparum* DOXP reductoisomerase is worth considering due to the benefits of post-translational modification within these expression systems.

Kinetic parameters for *E. coli* DOXP reductoisomerase were determined on purified protein that had a low activity. In subsequent DOXP reductoisomerase enzyme assays from independent purifications of *E. coli* DOXP reductoisomerase the specific activity has been notably consistently higher. This is suggestive of the purified *E. coli* DOXP reductoisomerase protein utilized within the determination of K_m and V_{max} was less active. Therefore, the K_m and V_{max} parameters for *E. coli* DOXP reductoisomerase need to be verified using several batches of independently purified recombinant *E. coli* DOXP reductoisomerase for which the amount of active enzyme is known.

Preliminary kinetic studies of the *E. coli* mutant DOXP reductoisomerase proteins needs to be repeated but suggest a critical role of Met^{214} and Trp^{212} in the functioning of DOXP reductoisomerase. The Met^{214} to Ile mutation specifically indicated a role in the binding of

NADPH mediated by this residue essential in DOXP reductoisomerase activity. Analysis of the K_m and V_{max} of the DOXP reductoisomerase protein and Met^{214} to Ile derivative with respect to NADPH, would further elucidate the Met^{214} NADPH interaction.

All purifications of the EcDXR derivatives were successful with the proteins being soluble and active. This assumes that the protein is folded and functional on a gross level and that any perturbations occur at a local level within the active site architecture rather than gross misfolding. The latter would result in a completely inactive enzyme. Additional characterization of the folded protein is ideally required through either circular dichroism (CD) for secondary structure analysis and X-ray crystal structure analysis for tertiary structure determination.

Determination of the amount of active protein within the DOXP reductoisomerase activity assay system with an inhibitor (e.g. fosmidomycin) would allow for the kinetic parameters k_{cat} and k_{cat}/K_m to be calculated to further analyse the effect of the mutation on the DOXP reductoisomerase functional mechanism. This would be achieved by titrating EcDXR and its derivatives with an inhibitor to determine at what concentration 100 % inhibition is achieved. This in turn is representative of the amount of active enzyme present in an assay.

The K_m values determined for EcDXR and its derivatives are not representative of binding studies and therefore need to be validated. The determination of K_D through binding studies of DOXP reductoisomerase with a non-hydrolysable substrate is necessary to validate the K_m values and calculate K_D. The natural substrate DOXP cannot be utilized in such studies as it is converted to MEP, however the inhibitor fosmidomycin can serve as the non-hydrolysable substrate analogue.

Additional rational amino acid substitutions within the active site of DOXP reductoisomerase would allow for further biochemical characterization of the protein. Ultimately, studies within the *P. falciparum* DOXP reductoisomerase need to be performed to determine if the role of the residues within the *E. coli* DOXP reductoisomerase homologue hold true for the parasitic protein.

The analysis of the catalytic hatch primary amino acid sequence of *P. falciparum* DOXP reductoisomerase revealed a Lys^{295} and Lys^{297} in comparison to asparagine and serine residues in other DOXP reductoisomerase homologues. To examine whether these sequence differences noted on bioinformatics analysis are indeed indicative of different functioning of the catalytic hatch in the *P. falciparum* DOXP reductoisomerase, kinetic analysis of the PfDXR(K295A, K297S) and EcDXR(A211K, S213K) is required.

Information from biochemical characterization and further rational amino acid analysis of the *E. coli* and *P. falciparum* DOXP reductoisomerase active site's will aid in the development of novel DOXP reductoisomerase inhibitors.

In vivo complementation systems could be used to further analyse the ability of the mutant DOXP reductoisomerase proteins to complement the wild-type protein (Takahashi *et al*., 1998). This system would also demonstrate whether the *P. falciparum* DOXP reductoisomerase could complement for the *E. coli*/prokaryotic homologue.

Further optimization of the procedure for the preparation of the furan derivatives is required as well as complete chemical characterization of the compounds synthesised. Enzyme inhibition assays of the phosphorylated aldehydes (**5**) and (**6**) with purified recombinant *E. coli* and *P. falciparum* DOXP reductoisomerase will demonstrate their inhibitory effect. Further improvement of the furan derivatives may also be considered through the introduction of a negative nitrogen within the heterocyclic system and alteration of the R-group substituents. Further structure-based drug design strategies could also involve the screening of structural databases to identify a consortium of possible DOXP reductoisomerase inhibitors, which can be docked, synthesised and tested for inhibitory effects.

CHAPTER 6: APPENDICES

6 APPENDICES

6.1 Chemicals and Reagents

All general buffer components and reagents were from Saarchem, RSA or Sigma-Aldrich, USA. Growth media was supplied by Biolab, UK or Difco, USA. Tabulated below are specialized reagents used within this work.

REAGENT	SUPPLIER
Agar	Oxoid
Agarose	Whitehead Scientific
Ammonium persulphate	Saarchem
Ampicillin	Sigma
Anti-His Antibody (mouse)	Sigma
Biorad Minigel Electrophoresis Set	Biorad
BM Chemiluminescence Western Blotting Kit	Roche
Bromophenol blue	Sigma
Chelating Sepharose	Amersham
2-Chlorotrityl chloride polymer bound	Chemika, Fluka
Coomassie brilliant blue R250	Saarchem
DNA Miniprep Kit	Qiagen
ECL Advance Western Blotting Kit	Amersham Pharmacia Biotech
Ethanol	Merck
Ethidium bromide	Sigma
3-Furanmethanol	Sigma-Aldrich
High Pure Plasmid Isolation Kit	Roche
Hybond^{TM} –c extra Nitrocellulose Membrane	Sigma
HyperfilmTM MP Autoradiography Film	Amersham
Imidazole	Sigma
IPTG	Promega
Magnesium chloride	Merck
β-mercaptoethanol	Merck

Methanol	Merck
Nickel sulphate	Sigma
PMSF	Roche
Polyethylene glycol 12 000	Merck
Ponceau S	Sigma
Restriction endonucleases	Amersham Pharmacia Biotech
Site-directed Mutagenesis Kit	Stratagene
SDS-PAGE Molecular Weight Standards Broad Range	Biorad
SDS	Saarchem
TEMED	Merck
Triphenylmethylchloride	Sigma-Aldrich
Tween 20	Saarchem

6.2 Oligonucleotide Primers

Oligonucleotides were synthesized by Integrated DNA Technologies Inc. (IDT).

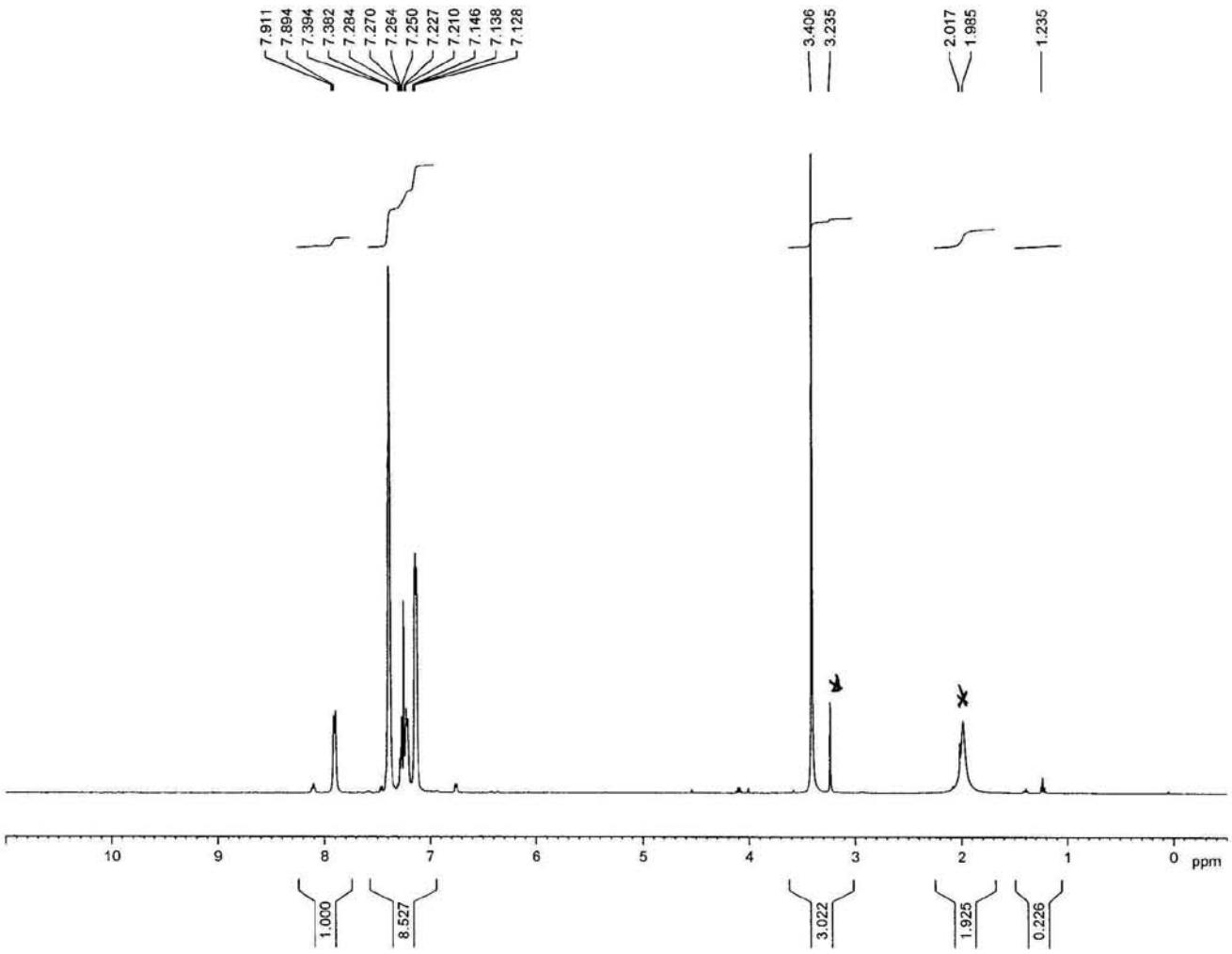
Table 8: DNA Oligonucleotides and their parameters

Name	T_m at 50 nm NaCl (°C)	% GC content	Sequence
pQE9EcDXR sequencing primer	56.5	50	5`-GTT TAC CGC AAC CTA TCC-3`
pQE31PfDXR sequencing primer	42.1	54.5	5`-CTG GAT AAC AAC AAA GTG C-3`
pQE9EcDXRW212YF	75.4	55.6	5`-CGT CAT CCG AAC TAT TCC ATG GGG CGT-3`
pQE9EcDXRW212YR	75.4	55.6	5`-ACG CCC CAT GGA ATA GTT CGG ATG ACG-3`
pQE9EcDXRW212LF	77.4	61.5	5`-CGC CCC ATG GAC AAG TTC GGA TGA CG-3`
pQE9EcDXRW212LR	77.4	61.5	5`-CGT CAT CCG AAC TTG TCC ATG GGG CG-3`
pQE9EcDXRW212FF	77.1	55.6	5`-CGT CAT CCG AAC TTT TCC ATG GGG CGT-3`
pQE9EcDXRW212FR	77.1	55.6	5`-ACG CCC CAT GGA AAA GTT CGG ATG ACG-3`
pQE9EcDXRM214IF	78.6	50	5`-CTG GTC GAT CGG CCG TAA AAT TTC TG-3`
pQE9EcDXRM214IR	78.6	50	5`-CAG AAA TTT TAC GGC CGA TCG ACC AG-3`
pQE9EcDXRP274IF	79	58.6	5`-CAG CTG GGG GAA ATC GAT ATG CGT ACG CC-3`
pQE9EcDXRP274IR	79	58.6	5`-GGC GTA CGC ATA TCG ATT TCC CCC AGC TG-3`
pQE9EcDXRN211KF	76.2	59.3	5`-GTC ATC CGA AGT GGT CCA TGG GGC GTA-3`
pQE9EcDXRN211KR	76.2	59.3	5`-TAC GCC CCA TGG ACC ACT TCG GAT GAC-3`
pQE9EcDXRN211K, S213KF	71.8	56	5`-CAT CCG AAG TGG AAG ATG GGG CGT A-3`
pQE9EcDXRN211K, S213KR	71.8	56	5`-TAC GCC CCA TCT TCC ACT TCG GAT G-3`
pQE31PfDXRK297SF	72.3	48	5`-CAT CCG AAA TGG TCC ATG GGC AAA A-3`
pQE31PfDXRK297SR	72.3	48	5`-TTT TGC CCA TGG ACC ATT TCG GAT G-3`
pQE31PfDXRK295N, K297SF	74.6	41.4	5`-GAA ACA TCC GAA TTG GTC AAT GGG CAA AA-3`
pQE31PfDXRK295N, K297SR	74.6	41.4	5`-TTT TGC CCA TTG ACC AAT TCG GAT GTT TC-3`

6.3 NMR Spectra

The NMR Spectra were recorded on a Bruker AMX400 spectrometer and are given below:

Compound 2 Proton NMR



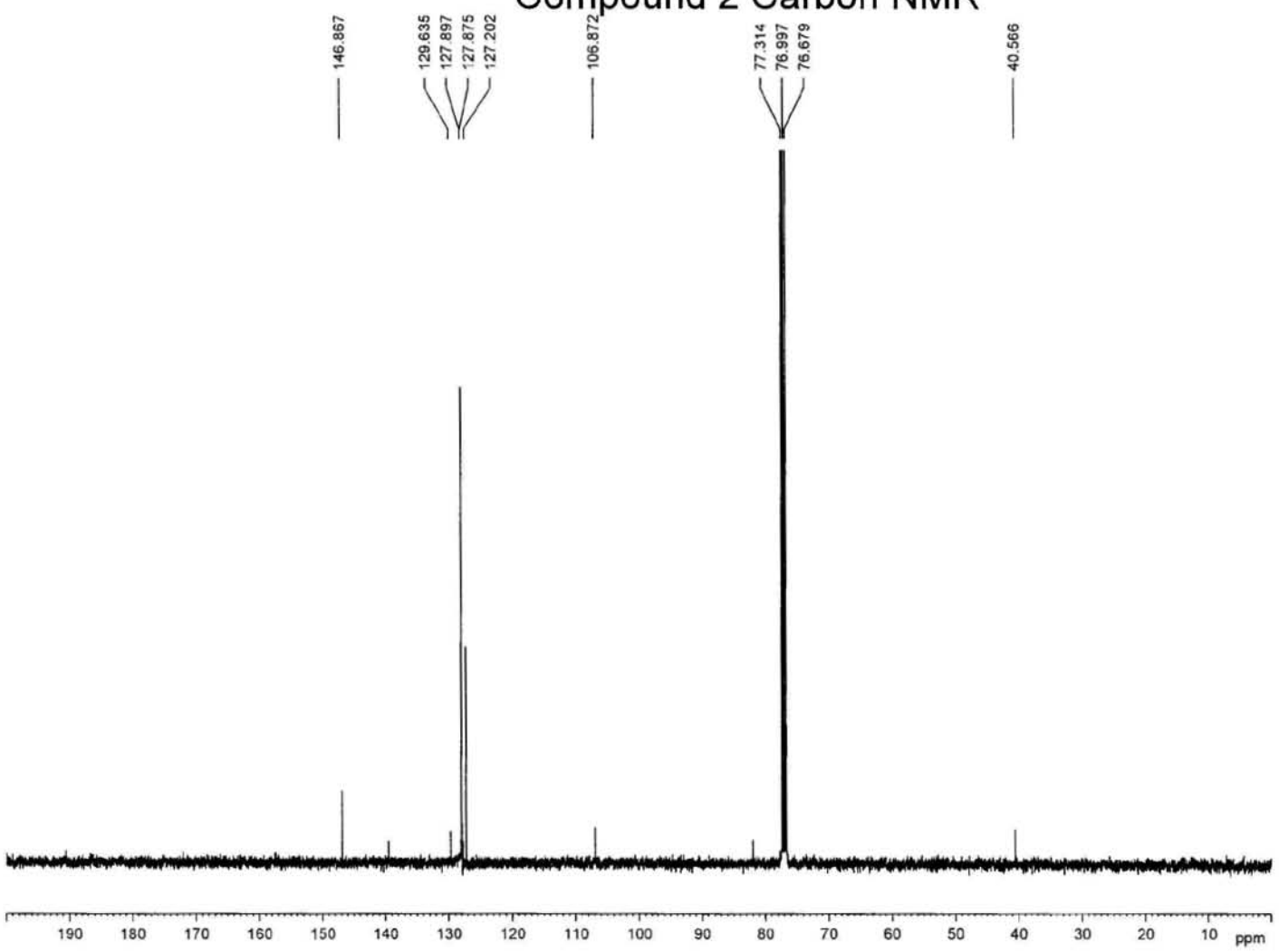
Current Data Parameters
NAME DCT1
EXPNO 100
PROCNO 1

F2 - Acquisition Parameters
Date_ 20030307
Time 15.36
INSTRUM spect
PROBHD 5 mm BBO Mul
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 2
SWH 16233.767 Hz
FIDRES 0.247708 Hz
AQ 2.0185587 sec
RG 256
DW 30.800 usec
DE 6.50 usec
TE 300.0 K
D1 1.00000000 sec

========== CHANNEL f1 ==========
NUC1 1H
P1 14.00 usec
PL1 2.00 dB
SFO1 400.1324710 MHz

F2 - Processing parameters
SI 32768
SF 400.1300134 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

Compound 2 Carbon NMR



Current Data Parameters
NAME DCT1
EXPNO 101
PROCNO 1

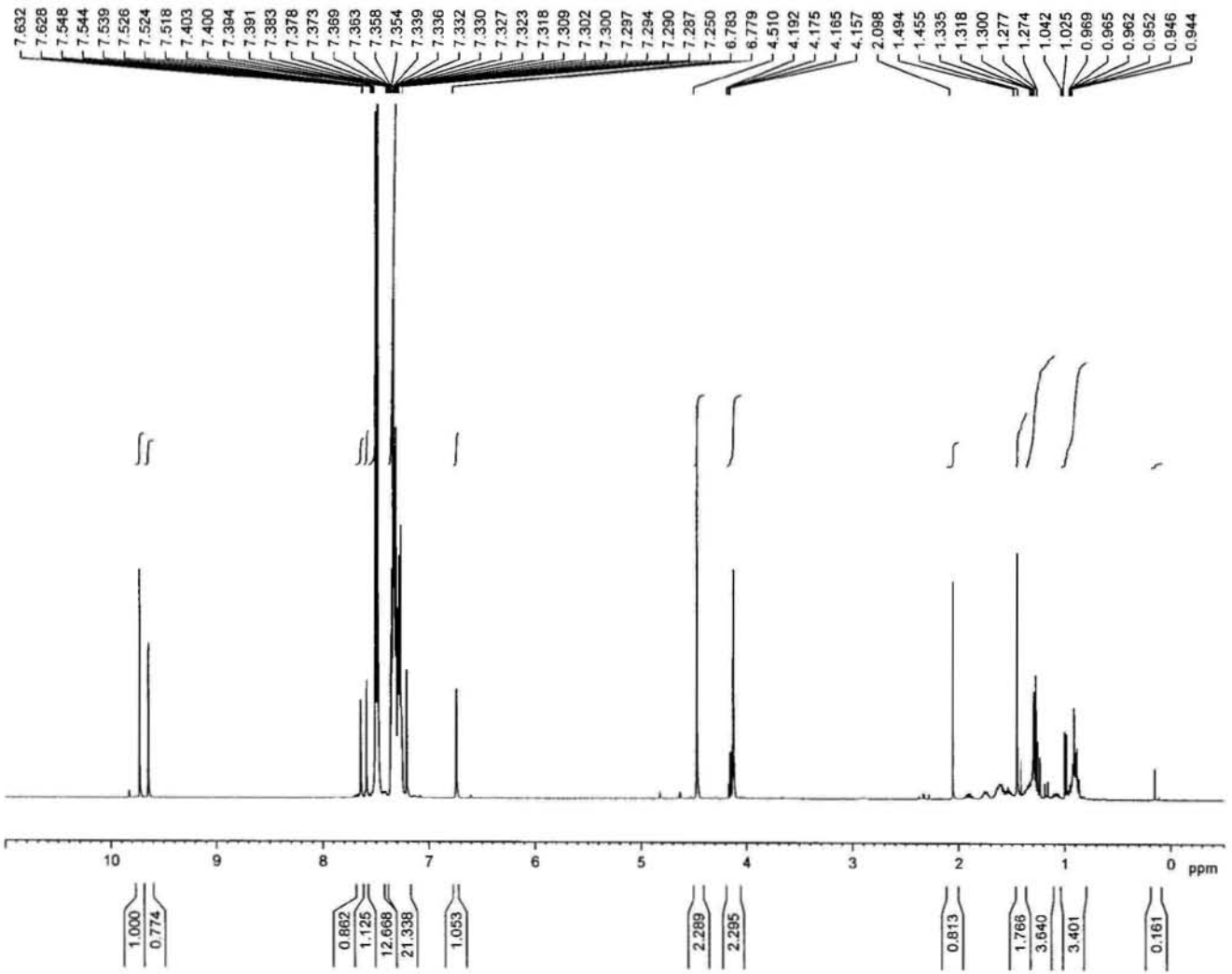
F2 - Acquisition Parameters
Date_ 20030318
Time 9.20
INSTRUM spect
PROBHD 5 mm BBO Mul
PULPROG zgpg30
TD 65536
SOLVENT CDCl3
NS 750
DS 4
SWH 24154.590 Hz
FIDRES 0.368570 Hz
AQ 1.3566452 sec
RG 256
DW 20.700 usec
DE 6.50 usec
TE 300.0 K
D1 1.00000000 sec
D11 0.03000000 sec
D12 0.00002000 sec

========== CHANNEL f1
NUC1 13C
P1 12.80 usec
PL1 2.00 dB
SFO1 100.6233934 MHz

========== CHANNEL f2
CPDPRG2 waltz16
NUC2 1H
PCPD2 90.00 usec
PL2 2.00 dB
PL12 17.20 dB
PL13 17.20 dB
SFO2 400.1316005 MHz

F2 - Processing parameters
SI 32768
SF 100.6127722 MHz
WDW EM
SSB 0
LB 1.00 Hz

Compound 3 and 4 Proton NMR



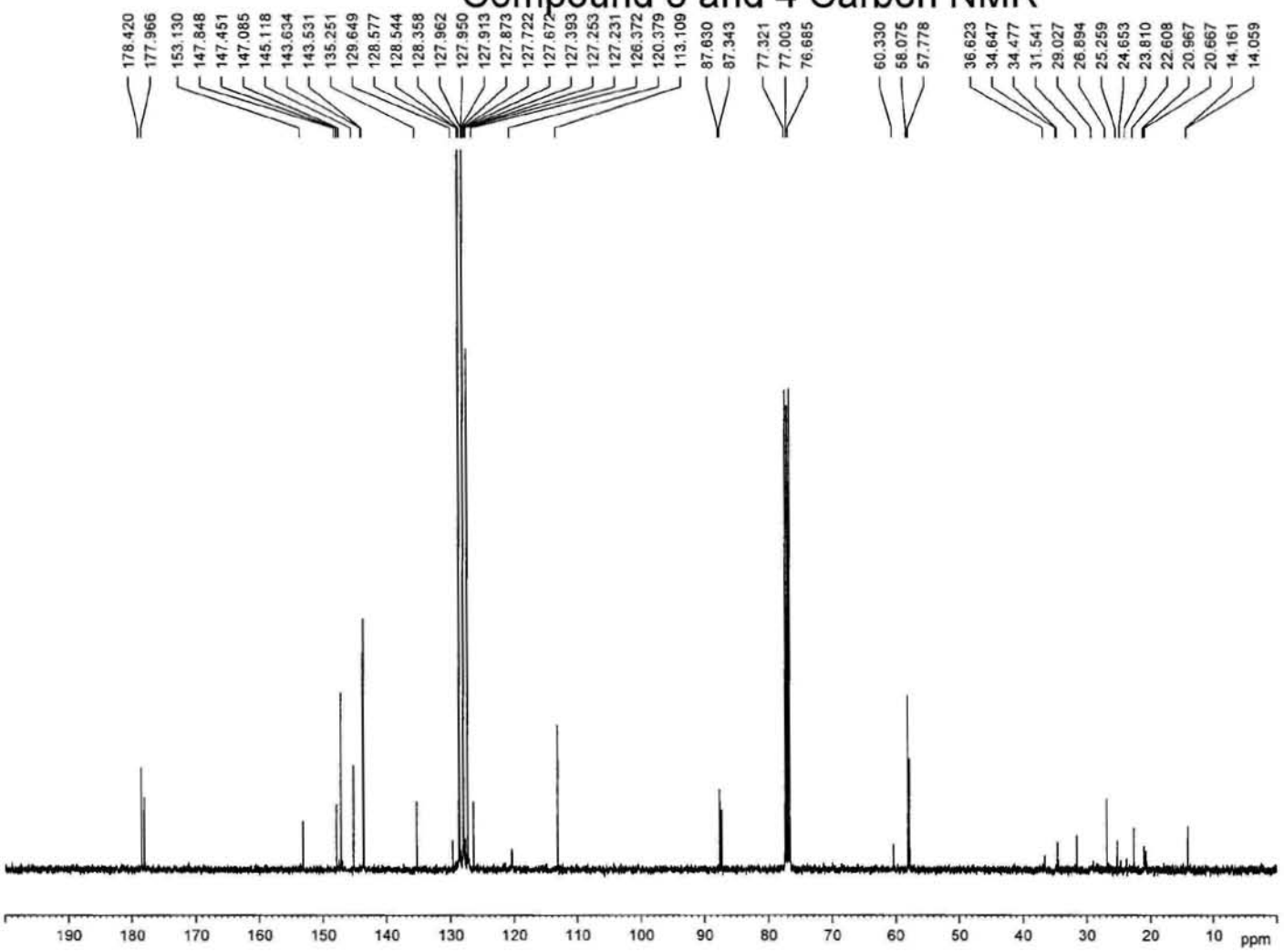
Current Data Parameters
NAME DCT2v
EXPNO 200
PROCNO 1

F2 - Acquisition Parameters
Date_ 20030516
Time 14.43
INSTRUM spect
PROBHD 5 mm BBO Mul
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 2
SWH 8278.146 Hz
FIDRES 0.126314 Hz
AQ 3.9584243 sec
RG 16
DW 60.400 usec
DE 6.50 usec
TE 300.0 K
D1 1.00000000 sec

========== CHANNEL f1 ==========
NUC1 1H
P1 14.00 usec
PL1 2.00 dB
SFO1 400.1324710 MHz

F2 - Processing parameters
SI 32768
SF 400.1300063 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

Compound 3 and 4 Carbon NMR



Current Data Parameters
NAME DCT2v
EXPNO 201
PROCNO 1

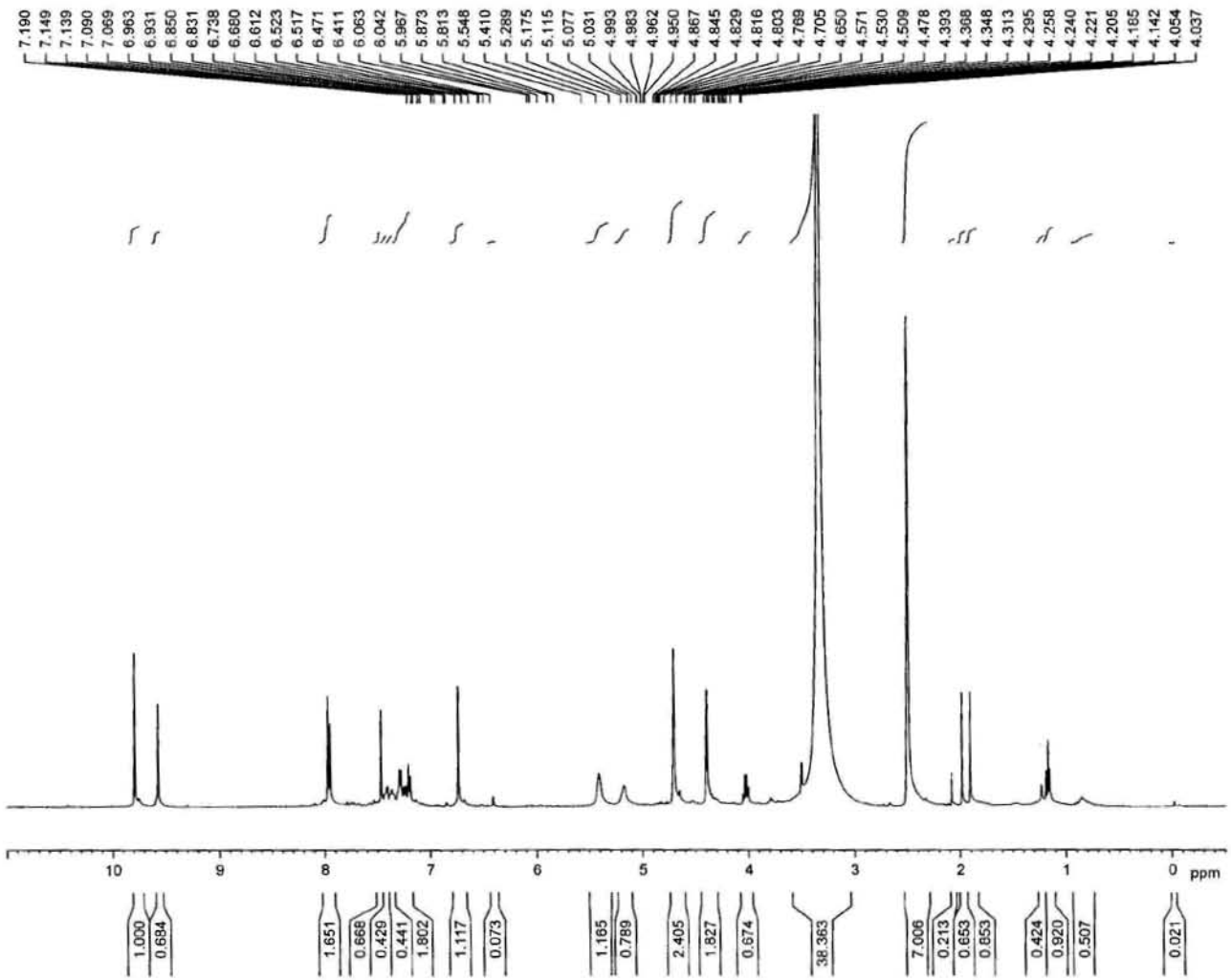
F2 - Acquisition Parameters
Date_ 20030516
Time 15.23
INSTRUM spect
PROBHD 5 mm BBO Mul
PULPROG zgpg30
TD 65536
SOLVENT DMSO
NS 832
DS 4
SWH 24154.590 Hz
FIDRES 0.368570 Hz
AQ 1.3566452 sec
RG 128
DW 20.700 usec
DE 6.50 usec
TE 300.0 K
D1 1.00000000 sec
D11 0.03000000 sec
D12 0.00002000 sec

========== CHANNEL f1
NUC1 13C
P1 12.80 usec
PL1 2.00 dB
SFO1 100.6233934 MHz

========== CHANNEL f2
CPDPRG2 waltz16
NUC2 1H
PCPD2 90.00 usec
PL2 2.00 dB
PL12 17.20 dB
PL13 17.20 dB
SFO2 400.1316005 MHz

F2 - Processing parameters
SI 32768
SF 100.6127736 MHz
WDW EM
SSB 0
LB 1.00 Hz

Compound 5 and 6 Proton NMR



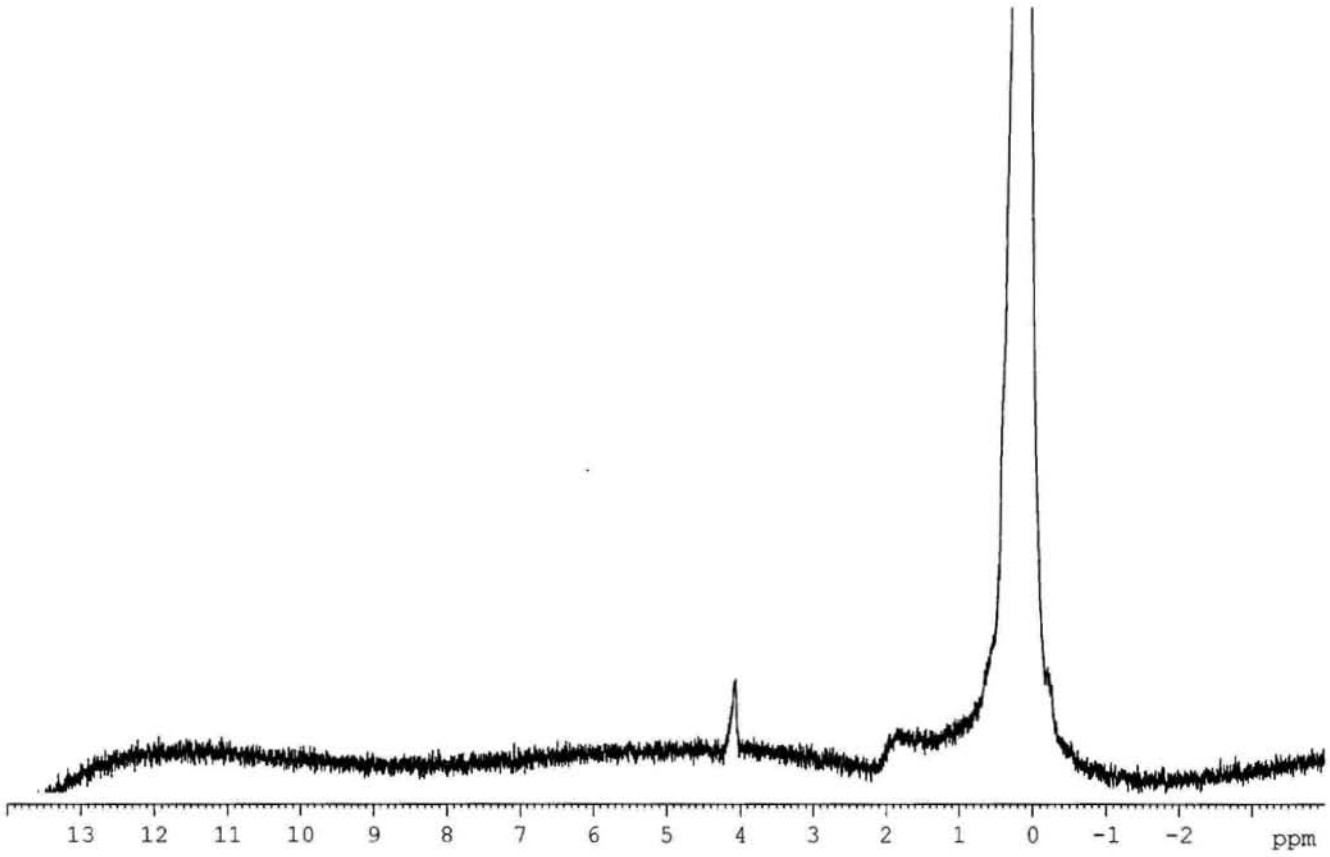
Current Data Parameters
NAME DCT12
EXPNO 100
PROCNO 1

F2 - Acquisition Parameters
Date_ 20031128
Time 12.48
INSTRUM spect
PROBHD 5 mm BBO Mul
PULPROG zg30
TD 65536
SOLVENT DMSO
NS 64
DS 2
SWH 8278.146 Hz
FIDRES 0.126314 Hz
AQ 3.9584243 sec
RG 143.7
DW 60.400 usec
DE 6.50 usec
TE 300.0 K
D1 1.00000000 sec

========== CHANNEL f1 ==========
NUC1 1H
P1 14.20 usec
PL1 2.00 dB
SFO1 400.1324710 MHz

F2 - Processing parameters
SI 32768
SF 400.1300031 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

Compound 5 and 6 + H3PO4 (set as 0 ppm) 257 of 3100 scans



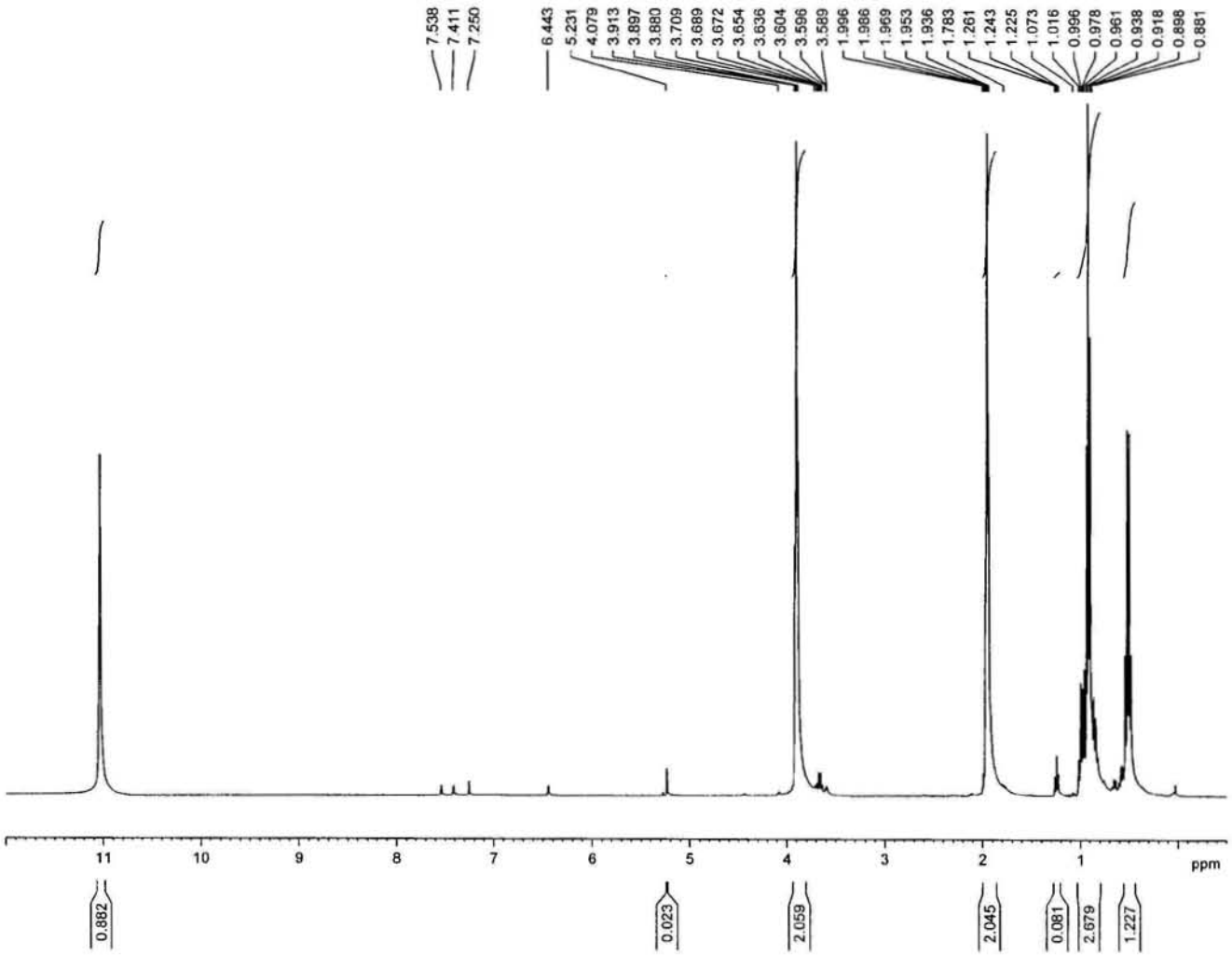
Current Data Parameters
NAME DCT12+H3PO4
EXPNO 102
PROCNO 1

F2 - Acquisition Parameters
Date_ 20031204
Time 21.45
INSTRUM spect
PROBHD 5 mm BBO Mul
PULPROG zg30
TD 65536
SOLVENT DMSO
NS 3100
DS 2
SWH 3255.211 Hz
FIDRES 0.049671 Hz
AQ 10.0663719 sec
RG 2048
DW 153.600 usec
DE 6.50 usec
TE 300.0 K
D1 1.00000000 sec

========== CHANNEL f1 ==========
NUC1 31P
P1 15.50 usec
PL1 2.00 dB
SFO1 161.9761773 MHz

F2 - Processing parameters
SI 32768
SF 161.9755358 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

Comound 8 Proton NMR



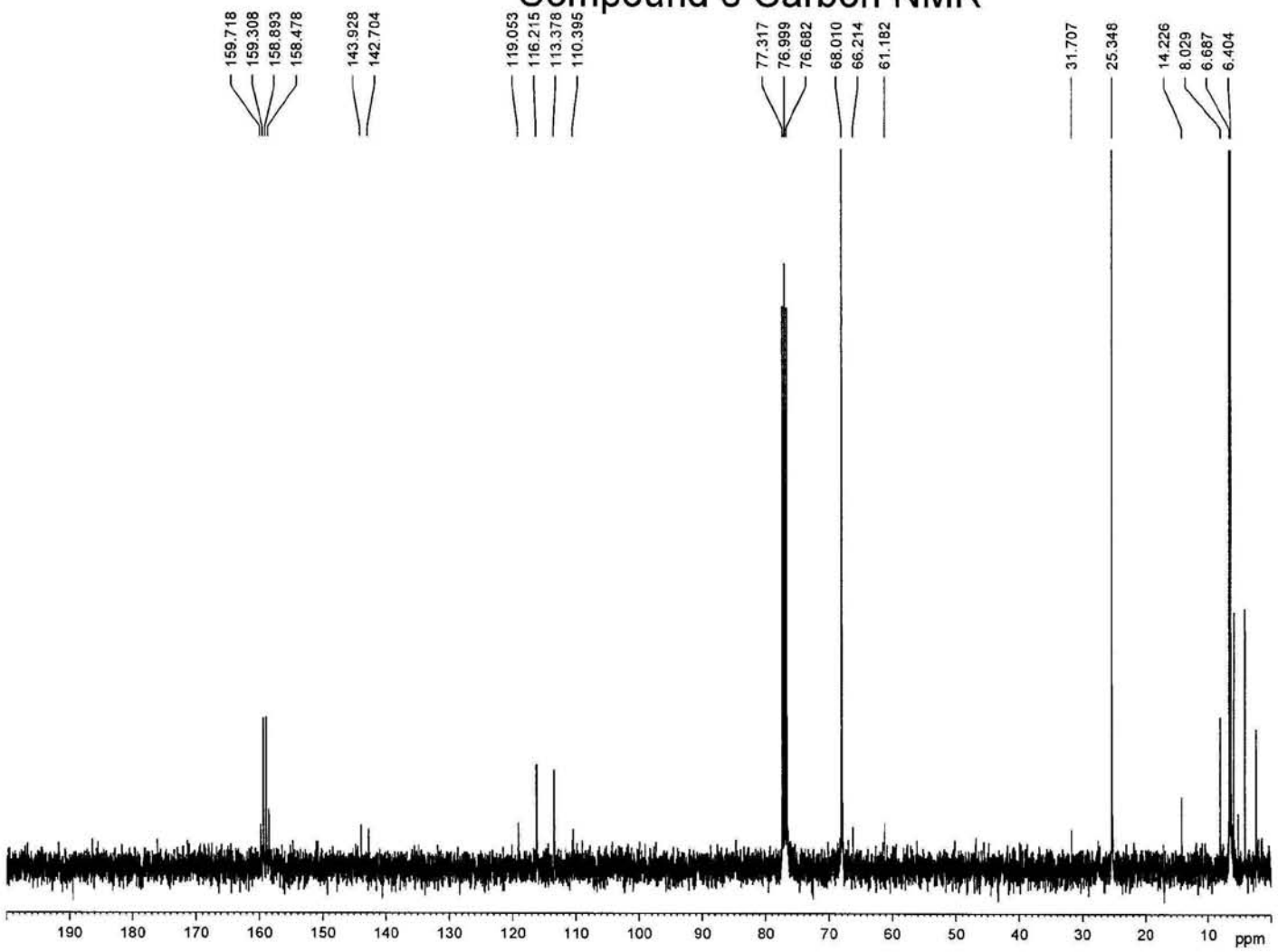
Current Data Parameters
NAME DCT6
EXPNO 100
PROCNO 1

F2 - Acquisition Parameters
Date_ 20031106
Time 10.11
INSTRUM spect
PROBHD 5 mm BBO Mul
PULPROG zg30
TD 65536
SOLVENT CDCl3
NS 16
DS 2
SWH 8278.146 Hz
FIDRES 0.126314 Hz
AQ 3.9584243 sec
RG 8.975
DW 60.400 usec
DE 6.50 usec
TE 300.0 K
D1 1.00000000 sec

========== CHANNEL f1 ==========
NUC1 1H
P1 14.20 usec
PL1 2.00 dB
SFO1 400.1324710 MHz

F2 - Processing parameters
SI 32768
SF 400.1300131 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

Compound 8 Carbon NMR



Current Data Parameters
NAME DCT6
EXPNO 101
PROCNO 1

F2 - Acquisition Parameters
Date_ 20031106
Time 10.13
INSTRUM spect
PROBHD 5 mm BBO Mul
PULPROG zgpg30
TD 65536
SOLVENT CDCl3
NS 245
DS 4
SWH 24154.590 Hz
FIDRES 0.368570 Hz
AQ 1.3566452 sec
RG 128
DW 20.700 usec
DE 6.50 usec
TE 300.0 K
D1 1.00000000 sec
D11 0.03000000 sec
D12 0.00002000 sec

========== CHANNEL f1
NUC1 13C
P1 14.15 usec
PL1 2.00 dB
SFO1 100.6233934 MHz

========== CHANNEL f2
CPDPRG2 waltz16
NUC2 1H
PCPD2 90.00 usec
PL2 2.00 dB
PL12 17.20 dB
PL13 17.20 dB
SFO2 400.1316005 MHz

F2 - Processing parameters
SI 32768
SF 100.6127648 MHz
WDW EM
SSB 0
LB 1.00 Hz

7 REFERENCES

Altincicek, B., Hintz, M., Sanderbrand, S., Wiesner, J., Beck, E. and Jomaa, H. (2000) Tools for the discovery of inhibitors of the 1-deoxy-D-xyulose 5-phosphate (DXP) synthase and DXP reductoisomerase : an approach with enzymes from the pathogenic bacterium *Pseudomonas aeruginosa*. *Federation of European Microbiology Societies Microbiology Letters* **190**, 329-333

Anders, R.F. and Saul, A. (2000) Malaria Vaccines. *Parasitology Today* **16**, 444-447

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acid Research* **25**, 3389-3402

Baca, A.M. and Hol, W.G.J. (2000) Overcoming codon bias: A method for high-level overexpression of *Plasmodium* and other AT-rich parasite genes in *Escherichia coli*. *International Journal for Parasitology* **30**, 113-118

Bannister, L. and Mitchell, G. (2003) The ins, outs and roundabouts of malaria. *TRENDS in Parasitology* **19**, 209-213

Barr, P. J., Green, K.M., Gibson, H.L., Bathurst, I.C., Quakyi, I.A. and Kaslow, D.C. (1991) Recombinant Pfs25 protein of *Plasmodium falciparum* elicits malaria transmission-blocking immunity in experimental animals. *Journal of Experimental Medicine* **174**, 1203–1208

Barton, D. and Ollis, W.D. (1979) Comprehensive Organic Chemistry, Pergamon Press, Volume 1, Oxford

Bradford, M. (1976) A rapid and sensitive method for the quantification of microgram quantities of protein using the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248

Bradley, M. (1990) Overexpression of Proteins in Eukaryotes. *Methods in Enzymology* **182**, 112-132

Beier, J.C. (1998) Malaria parasite development in mosquitoes. *Annual Reviews in Entomology* **43**, 519-543

Bourne, P.E. and Weissig, H. (2003) Stuctural Bioinformatics, John Wiley & Sons Publication, New Jersey.

Bujard, H., Gentz, R., Lanzer, M., Stüber, D., Müller, M., Ibrahimi, I., Häuptle, M.T. and Dobberstein B. (1987) A T5 promotor based transcription-translation system for the analysis of proteins *in vivo* and *in vitro*. *Methods in Enzymology* **155**, 416-433

Campbell, N.A. (1996) Biology. Fourth Edition The Benjamin/Cummings Publishing Company, Inc. California pp 526-527

Carter, R. (2001) Transmission blocking malaria vaccines. *Vaccine* **19**, 2309-2314

Chaga, G.S. (2001) Twenty-five years of immobilized metal ion affinity chromatography: past, present and future. *Journal of Biochemical and Biophysical Methods* **49**, 313-334

Chappel, J. A. and Holder, A.A. (1993). Monoclonal antibodies that inhibit *Plasmodium falciparum* invasion *in vitro* recognise the first growth factor-like domain of merozoite surface protein-1. *Molecular Biochemical Parasitology*. **60**, 303–312

Charon, L., Pale-Grosdemange, C. and Rohmer, M. (1999) On the reduction steps in themevalonate independent 2-*C*-methyl-D-erythritol 4-phosphate (MEP) pathway for isoprenoid biosynthesis in the bacterium *Zymomonas mobilis. Tetrahedron Letters.* **40**, 7231-7234

Cinquin, O., Christopherson, R.I. and Menz, R.I. (2001) A hybrid plasmid for expression of toxic malarial proteins in *Escherichia coli*. *MolecularBiochemistry and Parasitology* **117**, 245-247

Cosby, N. and Lesley, S. (1997) Site-directed mutagenesis. *Promega Notes Magazine* **61**, 12-19

Dagert, M. and Ehrlich, S.D. (1979) Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells. *Gene* **6**, 23-28

De Koning-Ward, T.F., Janse, C.J. and Waters, A.P. (2000) The development of genetic tools for dissecting the biology of malaria parasites. *Annual Reviews Microbiology* **54**, 157-185

De Roode, J.C. and Read, A.F. (2003) Evolution and ecology, after the malaria genomes. *Trends in Ecology and Evolution* **18**, 60-61

Deitsch, K. W., Calderwood, M. S. and Wellems T. E. (2001) Cooperative silencing elements in *var* genes. *Nature* **412**, 875–876

Doolan, D.L. and Hoffman, S.L. (2001) DNA-based vaccines against malaria: status and promise of the multi-stage malaria DNA vaccine operation. *International Journal for Parasitology* **31**, 753-762

Duffy, D.E.,. Pimenta, P. and Kaslow, D.C. (1993). Pgs28 belongs to a family of epidermal growth factor-like antigens that are targets of malaria transmission-blocking antibodies. *Journal of Exploratory Medicine* **177**, 505–510

Duffy, D.E., Craig, A.G. and Baruch, D.I. (2001) Variant proteins on the surface of malaria-infected erythrocytes—developing vaccines. *Trends in Parasitology* **17**, 354–356

Eckstein-Ludwig, U., Webb, R.J., van Goethem, I.D.A., East, J.M., Lee, A.G., Kimura, M., O'Neill, P.M., Bray, P.G., Ward, S.A. and Krishna, S. (2003) Artemisinins target the SERCA of *Plasmodium falciparum*. *Nature* **424**, 957-961

Eisenreich, W., Schwarz, M., Cartayrade, A., Arigoni, D., Zenk, M.H. and Bacher, A. (1998) The deoxyxylulose phosphate pathway of terpenoid biosynthesis in plants and microorganisms. *Chemistry and Biology* **5**, R221-233

Estèvez, J.M., Cantero, A., Romero, C., Kawaide, H., Jimmenez, L.F., Kuzuyama, T., Seto, H., Kamiya, Y. and Leon, P. (2000), Analysis of the expression of CLA1, a gene that encodes the 1-deoxy-D-xylulose 5-phosphate synthase of the 2-*C*-methyl-D-erythritol-4-phosphate pathway in Arabidopsis. *Plant Physiology*, **124**, 95-103

Farber, G.K. (1999) New approaches to rational drug design. *Pharmocology and Therapeutics* **84**, 327-332

Foley, M. and Tilley, L. (1997) Quinoline antimalarials: mechanisms of action and resistance. *International Journal for Parasitology* **27**, 231-240

Fradera, X., de la Cruz, X., Silva, C.H.T.P., Gelpi, J.L., Luque, F.J., and Orozco, M. (2002) Ligand-induced changes in the binding sites of proteins. *Bioinformatics* **18**, 939-948

Fry, M. and Pudney, M. (1992) Site of action of the antimalarial hydroxynaphthoquinone, 2-[trans-4-(4'-chlorophenyl) cyclohexyl]-3-hydroxy-1,4-naphthoquinone (566C80) *Biochemical Pharmacology* **43**, 1545-1553

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., Paulsen, I.T., James, K., Eisen, J.A., Rutherford, K., Salzberg, S.L., Craig, A., Kyes, S., Chan, M., Nene1, V., Shallom, S.J., Suh, B., Peterson, J., Angiuoli, S., Pertea1, M., Allen, J., Selengut, J., Haft, D., Mather, M.W., Vaidya, A.B., Martin, D.M.A., Fairlamb, A.H., Fraunholz, M.J., Roos, D.S., Ralph, S.A., McFadden, G.I.,

Cummings, L.M., Subramanian, G.M., Mungall, C., Venter, J.C., Carucci1, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M. and Barrell, B. (2002) Genome Sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**, 498-511

Gardner, M.J., Tettelin, H., Carucci, D.J., Cummings, L.M., Aravind, L., Koonin, E.V., Shallom, S., Mason, T., Yu, K., Fujii, C., Pederson, J., Shen, K., Jing, J., Aston, C., Lai, Z., Schwartz, D.C., Pertea, M., Salzberg, S., Zhou, L., Sutton, G.G, Clayton, R., White, O., Smith, H.O., Fraser, C.M., Adams, M.D., Venter, J.C. and Hoffman, S.L. (1998) Chromosome 2 Sequence of the Human Malaria Parasite *Plasmodium falciparum*. *Science* **282**, 1126-1132

Garret, R.H. and Grisham, C.M. (1999) Biochemistry. Second Edition, Saunders College Publishing, Harcourt Brace College Publishers, New York, pp 419

Ginsburg, H. and Stein, W.D. (1991) Kinetic modelling of chloroquine uptake by malaria infected erythrocytes. Assessment of the factors that may affect drug resistance. *Biochemical Pharmacology* **41,** 1463-1470

Greenwood, B. and Mutabingwa, T. (2002) Malaria in 2002. *Nature* **415**, 670-672

Grolle, S., Bringer-Meyer, S. and Sahm, H. (2000), Isolation of the *dxr* gene of *Zymomonas mobilis* and characterization of the 1-deoxy-D-xylulose 5-phosphate reductoisomerase. *Federation of Microbiological Societies Microbiology Letters* **191**, 131-137

Harker, M. and Bramley, P.M. (1999) Expression of prokaryotic 1-deoxy-D-xylulose 5-phosphatases in *Escherichia coli* increases carotenoid and ubiquinone biosynthesis. *Federation of European Biochemical Societies LETTERS* **448**, 115-119

Haynes, R.K., Hung-On Pai, H. and Voerste, A. (1999) Ring opening of artemisinin (qinghaosu) and dihydroartemisinin and interception of the open hydroxperoxides with formation of *N*-oxides- a chemical model for antimalarial mode of action. *Tetrahedron Letters* **40**, 4715-4718

Hermann, K.M. and Weaver, L.M. (1999) The Shikimate Pathway. *Annual Reviews in Plant Physiology and Plant Molecular Biology* **50**, 473-503

Higgins, S.J. and Hames, B.D. (1999) Protein Expression. A practical approach. Oxford University Press

Hoffman, S., Subramanian, G.M., Collins, F.H. and Venter, J.C. (2002) *Plasmodium*, human and *Anopheles* genomics and malaria. *Nature* 415,702-709

Ho, C. K. and Shuman, S. (2001) A yeast-like mRNA capping apparatus in *Plasmodium falciparum*. *Proceedings of the National Academy of Sciences, USA* **98**, 3050–3055

Holder, A. A. and Freeman, R.R. (1981). Immunization against blood-stage rodent malaria using purified antigens. *Nature* **294**, 361–364

Hyde, J.E. (2002) Mechanisms of resistance of *Plasmodium falciparum* to antimalarial drugs. *Microbes and Infection* **4**, 165-174

Ito, J., Ghosh, A., Moreira, L.A., Wimmer, E.A. and Jacobs-Lorena, M. (2002) Transgenic anopheline mosquitoes impaired intransmission of a malaria parasite. *Nature* **417**, 452-455

Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G. and Gibson, T.J. (1998) Multiple sequence alignment with ClustalX. *Trends In Biochemical Sciences* **23**, 403-405

Jomaa, H., Wiesner, J., Sanderbrand, S., Altincicek, B., Weidemeyer, C., Hintz, M., Türbachova, I., Eberl, M., Zeidler, J., Lichtenthaler, H.K., Soldati, D. and Beck, E. (1999) Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science* **285**,1573-1576

Joseph-McCarthy, D. (1999) Computational approaches to structure-based ligand design. *Pharmacology & Therapeutics* **84**, 179-191

Jung, R., Scott, M.P., Olivera, L.O. and Neilsen, N.C. (1992) A simple and efficient method for the oligodeoxyribonucleotide-directed mutagenesis of double stranded DNA. *Gene* **121**, 17-24

Keeling, P.J., Palmer, J.D., Donald, R.G.K., Roos, D.S., Waller, R.F. and McFadden, G.I. (1999) Shikimate pathway in apicomplexan parasites. *Nature* **397**, 219-220

Kraulis, P.J. (1991) *MOLSCRIPT:* a program to produce both detailed and schematic plots of protein structures. *Journal of Applied Crystallography* **24**, 946-950

Kurland, C. and Gallant, J. (1996) Errors of heterologous protein expression. *Current Opinions in Biotechnology* **7**, 489-493

Kuzuyama, T., Shimizu, T., Takahashi, S. and Seto, H. (1998) Fosmidomycin, a Specific Inhibitor of 1-Deoxy-D-xylulose 5-Phosphate Reductoisomerase in the Nonmevalonate Pathway for Terpenoid Biosynthesis. *Tetrahedron Letters* **39**, 7913-7916

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. *Nature* **227**, 680-685

Lange, B. M. and Croteau, R. (1999), Isoprenoid biosynthesis via a mevalonateindependent pathway in plants: cloning and heterologous expression of 1-deoxy-D-xylulose-5-phosphate reductoisomerase from peppermint. *Archive of Biochemical Biophysics*. **365**, 170-174

Lichtenthaler, H.K. (1999), The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. *Annual Reviews in Plant Physiology and Plant Molecular Biology*. **50**, 47-65

Lichtenhaler, H.K., Zeidler, J., Schwender, J. and Müller, C. (2000) The non-mevalonate isoprenoid biosynthesis of plants as a test system for new herbicides and drugs against pathogenic bacteria and the malaria parasite *Verlag der Zeitschrift für Naturforschung* **55c**, 305-313

Lilie, H., Schwarz, E. and Rudolph, R. (1998) Advances in refolding of proteins produced in *E. coli*. *Current Opinions in Biotechnology* **9,** 497-501

Li, Z. and Ganeson, A. (1998) Soli-phase functionalization of heterocycles by direct lithiation. *Synthetic Letters* 405-406

Ling, M.M. and Robinson, B.H. (1998) Approaches to DNA Mutagenesis: An Overview. *Analytical Biochemistry* **254**, 157-175

Macreadie, I., Ginsburg, H., Sirawaporn, W. and Tilley, L. (2000) Antimalarial Drug Development and New Targets. *Parasitology Today* **16**, 438-444

Madigan, M.T., Martinko, J.M. and Paker, J. (1997) Brock Biology of Microorganisms, Eighth Edition, Prentice Hall, New York, Chapter10 pp 357-396

Miller, L.H., Baruch, D.I., Marsh, K. and Doumbo, O.K. (2002) The pathogenic basis of malaria. *Nature* **417**, 673-679.

Miller, B., Henser, T. and Zimmer, W. (2000), Functional involvement of a deoxy-D-xylulose 5-phosphate reductoisomerase gene harboring locus of *Synechococcus leopoliensis* in isoprenoid biosynthesis. *Federation of European Biochemical Societies Letters* **481**, 221-226

Noedl, H., Wongsrichanalai, C. and Wernsdorfer, W.H. (2003) Malaria drug-sensitivity testing: new assays, new perspectives. *Trends in Parasitology* **19**, 175-181

Nussenzweig, V. and Nussenzweig, R.S. (1989) Rationale for the development of an engineered sporozoite malaria vaccine. *Advances in Immunology*. **45**, 283–334

Olliaro, P. (2001) Mode of action and mechanisms of resistance for antimalarial drugs. *Pharmocology and Therapeutics* **89**, 207-219

Olliaro, P., Haynes, R.K., Meunier, B. and Yuthavong, Y. (2001) Possible modes of action of the artemisinin-type compounds. *Trends in Parasitology* **17**, 122-126

Palmer, J.D. and Delwiche, C.F. (1996) Second-hand chloroplasts and the case of the disappearing nucleus. *Proceedings of the National Academy of Sciences, USA* **93**, 7432-7435

Peitsch, M.C. (1995) Protein Database Quarterly Newsletter. **72**, 4

Pederson, D.S. and Rosenbohm, C. (2001) Dry Column Vacuum Chromatography *Synthesis* **16**, 2431-2434

Perrin, D.D. and Armarego, W.L.F. (1988) Purification of Laboratory Chemicals, Pergamon Press, Third Edition, Oxford

Phillips, R.S. (2001) Current status of Malaria and Potential for Control. *Clinical Microbiology Reviews* **14**, 208-226

Proteau, P.J., Woo, Y., Williamson, R.T. and Phaosiri, C. (1999) Stereochemistry of the Reduction Step Mediated by Recombinant 1-Deoxy-D-xyulose 5-Phosphate Isomeroreductase. *Organic Letters* **1**, 921-923

Plebanski, M., Proudfoot, O., Pouniotis, D., Coppel, R.L., Apostolopoulos, V. and Flannery, G. (2002) Immunogenetics and the design of *Plasmodium falciparum* vaccines for use in malaria-endemic populations. *The Journal of Clinical Investigations* **110**, 295-301

Porath, J., Carlsson, J., Olsson, I. and Belfrage, G. (1975) Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **258**, 598-599

Rastelli, G., Sirawaraporn, W., Sompornpisut, P., Valaivan, T., Kamchonwongpaisan, S., Quarell, R., Lowe, G., Thebtaranonth, Y. and Yuthavong, Y. (2000) Interaction of Pyrimethamine, Cycloguanil, WR99210 and their analogues with *Plasmodium falciparum* Dihydrofolate Reductase: structural basis of antifolate resistance. *Bioorganic & Medicinal Chemistry* **8**, 1117-1128

Raynes, K. (1999) Bisquinoline antimalarials: their role in malaria chemotherapy. *International Journal for Parasitology* **29**, 367-379

Reichenberg, A., Wiesner, J., Weidemeyer, C., Dreiseidler, E., Sanderbrand, S., Altincicek, B., Beck, E., Schlitzer, M. and Jomaa, H. (2001) Diaryl ester prodrugs of FR900098 with improved *in vivo* antimalarial activity. *Bioorganic & Medicinal Chemistry Letters* **11**, 833-835

Reuter, K., Sanderbrand, S., Jomaa, H., Wiesner, J., Steinbrecher, I., Beck, E., Hintz, M., Klebe, G. and Stubbs, M.T. (2002) Crystal Structure of 1-Deoxy-D-xyulose-5-phosphate Reductoisomerase, a crucial enzyme in the non-mevalonate pathway of isoprenoid biosynthesis. *The Journal of Biological Chemistry* **277**, 5378-5384

Ridley, R.G. (1999) Planting the seeds of new antimalarial drugs. *Science* **285**, 1502-1503

Roberts, F., Roberts, C.W., Johnson, J.J., Kyle, D.E., Krell, T., Coggins, J.R., Coombs, G.H., Milhous, W.K., Tzipori, S., Ferguson, D.J.P., Chakrabarti, D., and McLeod, R. (1998) Evidence for the shikimate pathway in apicomplexan parasites. *Nature* **393**, 801–805

Rohmer, M. (1999) A mevalonate-independent route to isopentenyl diphosphate in"Comprehensive Natural Products Chemistry", ed. Cane, D.E., Elsevier, Oxford, 45-62

Sato, S. and Wilson, R.J. (2002) The genome of *Plasmodium falciparum* encodes an active delta-aminolevulinic acid dehydratase. *Current Genetics* **40**, 391-398

Saul, A. and Battistuttan, D. (1988) Codon usage in *Plasmodium falciparum. Molecular Biochemistry and Parasitology* **27**, 35-42

Scheibani, N. (1999) Prokaryotic gene fusion expression systems and their use in structural and functional studies of proteins. *Preparative Biochemistry &Biotechnology* **29**, 77-90

Schein, C.H. and Noteborn, M.H.M. (1988) Formation of soluble, recombinant proteins in *Escherichia coli* is favoured by lower growth temperature. *Bio/Technology* **6**, 291-294

Shi, Y.P., Hasnain, S.E., Sacci, J.B., Holloway, B.P., Fujioka, H., Kumar, N., Wohlheuter, R., Hoffman, S.L., Collins, W.E. and Lal, A.A. (1999) Immunogenicity and *in vitro* protective efficacy of a recombinant multistage *Plasmodium falciparum* candidate vaccine. *Proceedings of the National Academy of Sciences, USA* **96**, 1615-1620

Sirawaraporn, W., Sathitkul, T., Sirawaraporn, R., Yuthavong, Y. and Santi, D. (1997) Antifolate-resistant mutants of *Plasmodium falciparum* dihydrofolate reductase. *Proceedings of the National Academy of Sciences, USA* **94**, 1124-1129

Sprenger, G.A., Schörken, U., Wiegert, T., Grolle, S., de Graaf, A.A., Taylor, S.V., Begley, T.P., Bringer-Meyer, S. and Sahm, H. (1997) Identification of a thiamin-dependent synthase in *Escherichia coli* required for the formation of the 1-deoxy-D-xylulose 5-phosphate precursor to isoprenoids, thiamin, and pyridoxol. *Proceedings of the National Academy of Sciences, USA* **94**, 12857-12862

Steinbacher, S., Kaiser, J., Eisenreich, W., Huber, R., Bacher, A. and Rohdich, F. (2003) Structural Basis of Fosmidomycin Action Revealed by the Complex with 2-C-Methyl-D-erythritol 4-phosphate Synthase (IspC). Implications for the catalytic mechanism and anti-malaria drug development. *The Journal of Biological Chemistry* **278**, 18401-18407

Stüber, D., Matile, H. and Garotta, G. (1990) System for high-level production in *Escherichia coli* and rapid purification of recombinant proteins: applications to epitope mapping, preparation of antibodies, and structure-function analysis. Immunological Methods, Academic Press, New York, pp 121-152

Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Use of T7 RNA Polymerase to Direct Expression of Cloned Genes. *Methods in Enzymology* **185**, 61-89.

Su, X., Kirkman, L.A., Fujioka, H. and Wellems, T.E. (1997) Complex polymorphisms in an approximately kDa protein are linked to chloroquine-resistant *Plasmodium falciparum* in Southeastern Asia and Africa. *Cell* **91**, 593-603

Surolia, N. and Surolia, A. (2001) Triclosan offers protection against blood stages of malaria by inhibiting enoyl-ACP reductase of *Plasmodium falciparum*. *Nature Medicine* **7**, 167–173

Takahashi, S., Kuzuyama, T., Watanabe, H. and Seto, H. (1998) A 1-deoxy-D-xyulose 5-phosphate reductoisomerase catalyzing the formation of 2-C-methyl-D-erythritol 4-phosphate in an alternative nonmevalonate pathway for terpenoid biosynthesis. *Proceedings of the National Academy of Sciences, USA* **95**, 9879-9884

Towbin, H., Staehelin, T. and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proceedings of the National Academy of Sciences, USA* **76**, 4350-4354

Vial, H.J. (2000) Isoprenoid Biosynthesis and Drug Targeting in the Apicomplexa. *Parasitology Today* **16**, 140-141

Vriend,G. (1990). Whatif: a molecular modelling and drug design programme. *Journal of Molecular Graphics* **8,** 52-56

Waller, R.F., Keeling, P.J., Donald, R.G.K., Striepen, B., Handman, E., Lang-Unnasch, N., Cowman, A.F., Besra, G.S., Roos, D.S. and McFadden, G.I. (1998) Nuclear-encoded proteins target to the plastid in *Toxoplasma gondii* and *Plasmodium falciparum*. *Proceedings of the National Academy of Sciences, USA* **95,** 12352-12357
Wang, R., Doolan, D.L., Le, T.P., Hedstrom, R.C., Coonan, K.M., Charoenvit, Y., Jones, T.R., Hobart, P., Marglith, M., Ng, J., Weiss, W.R., Sedegah, M., de Taisne, C., Norman, J.A. and Hoffman, S.L. (1998). Induction of antigen-specific cytotoxic T lymphocytes in humans by a malarial DNA vaccine. *Science* **282**, 446–480

Weickert, M.J., Doherty, D.H., Best, E.A. and Olins, P.O. (1996) Optimization of heterologous protein production in *Escherichia coli*. *Current Opinion in Biotechnology* **7**, 494-499

Westbrook, J., Feng, Z., Chen, L., Yang, H. and Berman, H.M. (2003). The Protein Data Bank and structural genomics. *Nucleic Acid Research*. **31**, 489-491

Wiesner, J., Wiβner, P., Dahse, H., Jomaa, H. and Schlitzer, M. (2001) Discovery of a novel lead structure for anti-malarials. *Bioorganic and Medicinal Chemistry* **9**, 785-792

Winstanley, P.A. (2000) Chemotherapy for Falciparum Malaria: The armoury, the problems and the prospects. *Parasitology Today* **16**, 146-153

Wungsintaweekal, J. (2001) Enzymes of the alternative terpenoid pathway in *Escherichia coli*. PhD Thesis. Technische Universität, München.

Yajima, S., Nonaka, T., Kuzuyama, T., Seto, H. and Ohsawa, K. (2002) Crystal Structure of 1-Deoxy-D-xylulose 5-phosphate Reductoisomerase Complexed with Cofactors: Implications of a Flexible Loop Movement upon Substrate Binding. *Journal of Biochemistry* **131**, 313-317

Zubrzycki, I.Z. and Blatch, G.L. (2001) DFT study of a substrate and inhibitors of 1-deoxy-2-xylulose-5-phosphate reductoisomerase – the potential novel target molecule for anti-malaria drug development *Journal of Molecular Modelling* **7**, 378-383

Internet references:

www.Malaria.freewire.co.uk/malaria**.**
www.jomaa-pharmaka.com
www.rphwa.gov.au/labs/haem/malaria/index
www.ncbi.nlm.nih.gov
www.clontech.com
www.expasy.org/swissmod/SWISS-MODEL
www.cmbi.kun.nl/gvteach/hommod/index.shtml