Structural modelling of therapeutic targets and inhibitors of the malaria parasite

by

Fourie Joubert

Submitted in partial fulfilment of the degree

Ph.D.

in the Department of Biochemistry

University of Pretoria

July 2000

"Molecular docking is a serious challenge for bio-chemists. There are many ways to fit molecules together but only a few juxtapositions that bring them close enough to bond. On a molecular level success may mean discovering what synthetic structure, what chemical, will form a union with, say, the protein shape on a tumor cell. If you make this high-risk jigsaw work you may have found a cure for carcinoma. But molecules and the human beings they are part of exist in a universe of possibility. We touch one another, bond and break, drift away on force-fields we don't understand."

A quotation from the novel *'Written on the Body'* by the British writer Jeanette Winterson (1992).

Acknowledgements

I would like to express my sincere gratitude to:

- Prof. Braam Louw, my supervisor, who always supported and motivated me, even when things were looking hopeless. Without his ideas and inputs, this project could not have succeeded. He has truly become part of my family.
- Prof. Albert Neitz, co-supervisor and Head of the Department of Biochemistry. He strongly supported all my new ventures into Bioinformatics and Structural modelling with his usual good humour and patience, and has led the whole Department to new heights.
- Prof. Robin Crewe, who was the first to give us financial and moral support to venture into new research directions.
- All my friends and colleagues in the Biochemistry, Genetics and Microbiology Departments for their support, assistance and good humour.
- Ben, Deborah, Karen, Katherine and Lyn-Marie for support, input and enthusiasm through the years.
- My brother Louis, his wife Anita, brother-in-law Michael, sister-in-law Clarel and parents-in-law Michael and Susan for patience, love, support and understanding.
- My wife Annie, who has loved and carried me through the toughest of times. Without her, I would probably not be alive today.
- My Father in Heaven, for giving me the will, srength and good health to complete this thesis.

Contents

List of Figures

List of Tables

List of Abbreviations

3D	Three dimensional
Å	Angstrom
A	Adenine
ADP	Adenosine diphosphate
Ala	Alanine
ALD	Fructose 1,6-bisphosphate aldolase
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
ATP	Adenosine triphosphate
bp	Base pairs
C	Cytidine
cDNA	Complementary DNA
CPK	Corey-Pauling-Koltun
CTP	Cytosine triphosphate
Cys	Cysteine
D-GAP	D-glyceraldehyde phosphate
DHAP	Dihydroxyactone phosphate
DHF	Dihydrofolate
DHFR	Dihydrofolate reductase

DHFR-TS	Dihydrofolate reductase-thymidylate synthase
DHODH	Dihydroorotase dehydrogenase
DNA	deoxy-Ribonucleic acid
dTMP	deoxy-Thymidine monophosphate
dUDP	deoxy-Uridine diphosphate
dUMP	deoxy-Uridine monophosphate
dUTP	deoxy-Uridine triphosphate
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetra-acetate
FDA	Food and Drug Administration
G	Guanine
GAP	Glyceraldehyde phosphate
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GDH	Glucose 3-phosphate dehydrogenase
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
GMP	Guanosine monophosphate
GNDA	Gossyclic nitrile diacetate
GPI	Phosphoglucoisomerase
His	Histidine
HGPRT	Hypoxanthine-guanine phosphoribosyltransferase
Ile	Isoleucine
IMAC	Immobilised metal affinity chromatography
IMP	Inosine monophosphate
IPTG	Isopropyl-β-d-thiogalactopyranoside

k_{cat}	Catalytic constant
K_I	Inhibition constant
K_M	Michaelis-Menten constant
LB-broth	Luria-Bertani broth
L. casei	*Lactobacillus casei*
LDH	Lactate dehydrogenase
Leu	Leucine
Lys	Lysine
MALDI	Matrix assisted laser desorption ionisation
mdr	Multi-drug resistance
$MePhSO_2$-Ph	Phenyl methanethiosulfonate
Met	Methionine
mTIM	Monomeric triosephosphate isomerase
NAD	Nicotinamide adeninedinucleotide
NADH	Nicotinamide adeninedinucleotide (reduced)
NCI	National Cancer Institute
NMR	Nuclear magnetic resonance
OD	Optical density
PABA	Para-amino butyric acid
PBS	Phosphate-buffered saline
P. carinii	*Pneumocystis carinii*
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PGA	Phosphoglycolic acid
PGH	Phosphoglycolohydroxamate
P. falciparum	*Plasmodium falciparum*

PGK	Phosphoglycerate kinase
Phe	Phenylalanine
PMSF	Phenylmethylsulphonylfluoride
Pro	Proline
P. vivax	*Plasmodium vivax*
RFLP	Restriction fragment length polymorphism
R.M.S.	Root mean square
RNA	Ribonucleic acid
S. cerevisiae	*Saccharomyces cerevisiae*
SDS	Sodiumdodecyl sulphate
SDS-PAGE	Sodiumdodecyl sulphate polyacrylamide gel electrophoresis
Ser	Serine
S. pombe	*Saccharomyces pombe*
S-Tag	Ribonuclease-S binding peptide tag
T	Thymine
T. brucei	*Trypanosoma brucei*
TIM	Triosephosphate isomerase
TIM-PGH	Triosephosphate isomerase-phosphoglycolohydroxamate complex
tRNA	Transfer RNA
Trp	Tryptophan
TS	Thymidylate synthase
Tyr	Tyrosine
UDP	Uridine diphosphate
UTP	Uridine triphosphate
Val	Valine
V_{max}	Maximum velocity
XMP	Xanthine monophosphate

Summary

Anti-malarial drug resistance is increasing on a global level. With the progresses made in computerised drug design, developments in the field of therapeutic agents against malaria may be greatly accelerated in the near future. In this study, two malaria metabolic enzymes were identified as potential targets for computer-aided drug design. The enzymes dihydrofolate reductase (DHFR) and triosephosphate isomerase (TIM) were cloned, and expressed in bacterial expression systems. Large amounts of DHFR could be expressed, but not in soluble form. It was shown however, that native malaria genes could be expressed in *E. coli* under certain promoters. TIM was expressed in an active form and purified to homogeneity by means of immobilized metal affinity chromatography (IMAC).

The structures of both DHFR and TIM were modelled by homology based methods and analysed in terms of quality. DHFR was extremely complicated to model due to the presence of several large insertions, and did not yield a high quality model. The model of TIM was of adequate quality to be used for ligand studies.

When the X-ray structure for malaria TIM became available recently, it was compared to the homology model and was found to be virtually the same. The only important difference was the rotation of the *Phe96* sidechain at the active site. The X-ray structure was subsequently used for ligand screening against the NCI-3D small molecule database. Two of the 7 *in vitro* tested compounds yielded inhibition in the <100μM range. These two molecules Direct Red and Direct Violet are derivatives of the anti-trypanosomal Suramin which was originally derived from the dye Trypan Red. According to the literature, Direct Red and Direct Violet inhibit trypanosomal TIM approximately 100x more potently than Suramin.

A common structural region was identified in Direct Red, Direct Violet

and a series of other Suramin-related compounds, and this moiety may be used as a basis for the design of anti-TIM lead drug compounds.

Opsomming

Weerstandbiedenheid teen anti-malaria middels neem toe op 'n globale vlak. Met die vordering in rekenaargebaseerde ontwerp sal die ontwikkeling van terapeutiese agente teen malaria binnekort vinnig toeneem. In hierdie studie is twee metaboliese ensieme van die malaria parasiet geïdentifiseer as potensiële teikens vir rekenaargebaseerde ontwerp van geneesmiddels. Die ensieme dihidrofolaat reduktase (DHFR) en triosefosfaatisomerase (TIM) is gekloneer en uitgedruk in bakteriële uitdrukkingsisteme. Groot hoeveelhede DHFR kon uitgedruk word, maar nie in 'n oplosbare vorm nie. Daar is egter gewys dat natiewe malaria gene in *E. coli* uitgedruk kon word onder die beheer van spesifieke promoters. TIM is uitgedruk in 'n aktiewe vorm en daarna gesuiwer tot homogeniteit deur middel van geïmmobiliseerde metaal affiniteitschromatografie (IMAC).

Die strukture van beide DHFR en TIM is gemodelleer deur middel van homologie metodes en geanaliseer in terme van kwaliteit. DHFR was besonder ingewikkeld om te modelleer as gevolg van die teenwoordigheid van verskeie groot ingevoegde segmente, en het nie 'n model van hoë kwaliteit gelewer nie. Die model van TIM was van voldoende kwaliteit om vir ligandbindingstudies gebruik te word.

Aangesien die X-straalstruktuur van malaria TIM onlangs beskikbaar geword het, is dit vergelyk met die homologie model en min verskille is gevind. Die enigste belangrike verskil was die rotasiehoek van die *Phe96* syketting by die aktiewe sentrum. Die X-straalstruktuur is daarna gebruik vir ligandsoektogte teen die NCI-3D kleinmolekuul databasis. Twee van die 7 *in vitro* getoetste verbindings het inhibisie in die $<100\mu$M konsentrasiegebied gewys. Die twee verbindings, Direct Red en Direct Violet is derivate van die anti-trypanosomale middel Suramin, wat oorspronklik op die kleurstof Trypan Rooi gebaseer is. In die literatuur is gevind dat Direct Red en Direct Violet trypanosomale

TIM ongeveer 100x sterker inhibeer as Suramin.

'n Gedeelde strukturele motief is geïdentifiseer in Direct Red, Direct Violet en 'n reeks ander Suramin-verwante verbindings. Die motief mag gebruik word as basis vir die ontwerp van anti-TIM voorloper geneesmiddels.

Chapter 1

Introduction

1.1 Malaria

Malaria is a disease with severe world-wide implications, with mortality rates approaching 2 million people per year. More than 80% of the world's malaria cases are found in Africa, where at least 800 000 children die from the disease every year (Schapira *et al.*, 1993). A little more than hundred years ago, the use of methylene blue against malaria was described by Guttman and Ehrlich (1891). This was followed by several generations of anti-malarial drugs including: quinine (1820), pamaquine (1924), quinacrine (1930), chloroquine (1934), proguanil (1944), pyrimethamine (1952) and mefloquine (1971). Today the effectivity of nearly every anti-malarial drug is threatened or impaired by resistant forms of the malaria parasite. Management of the disease has become problematic due to the widespread and steady increase of multi-drug resistant parasitic forms in different parts of the world. This has led to the studies of the mechanisms of drug resistance, and the identification of genetic changes involved in these processes.

The World Health Organisation is currently spending approximately 4% of its budget on anti-malarial activities (Kumar, 1999). Despite these efforts, malaria is now present in 102 countries, and is responsible for 300-500 million clinical cases each year.

Malaria is being virtually ignored by drug companies. Recent facts concerning the American drug industry shed some light on the subject. They invested 10.9 billion dollars in research and development in 1992 (more than twice the amount invested 5 years before). However, the

cost of developing a new drug has more than doubled during the last 10 years. In 1995 it took 231 million dollars, and an average of 12 years to take a chemical from the laboratory to the pharmacy shelf. Using data from the US Army's Anti-malarial Drug Develpoment Program (Brueckner *et al.*, 1998), they show that for every 3000 compounds screened only one reached the final testing stage. Only about half of these final compounds were advanced to preclinical studies. According to the Center for the Study of Drug Development, 5 in 4000 compounds undergoing preliminary testing makes it to the clinical stage, with 1 in 5 of those being approved by the Food and Drug Administration (FDA). It is easy to see that developing new drugs is a very long, high-risk and expensive venture, and that the pharmaceutical industry requires a high return to cover their investment. For comparative purposes, in 1993 the USA was testing 91 AIDS medications, but only one anti-malarial drug was undergoing clinical testing. Malaria is a disease of developing countries, which cannot afford to pay for expensive medicines. This places an enormous responsibility on African scientists to contribute to the search for new anti-malarial drugs in order to reduce the initial costs of the development programme.

The traditional mainstay of anti-malarial treatment, chloroquine, now frequently fails against the *Plasmodium falciparum* malaria parasite, with resistance occurring in South America, Africa and Southeast Asia. Resistant strains of *Plasmodium vivax* have also been reported. Chloroquine resistance has been mapped to a 400 kilobase segment of chromosome 7 in *P. falciparum* (Pennisi, 1999). Chloroquine-resistant parasites rid themselves of the drug 40-50-fold faster than sensitive parasites. Although the exact mechanism of resistance is unknown, it seems to be similar in all strains of the parasite, and is always reversible *in vitro* by verapamil.

South African cultures were analyzed in terms of sensitivity to a range of drugs, including amodiaquine, mefloquine, quinine, chloroquine and sulphadoxine/pyrimethamine (Deacon *et al.*, 1994). These strains were collected from Kwazulu/Natal, and reference strains from South Africa, Malawi and Mozambique were also analyzed.

Amodiaquine: All Kwazulu isolates collected from 1984-1986 were sensitive to amodiaquine, but of the 1989-1990 strains, only one was sensitive. Three of the Transvaal isolates from 1987 were amodiaquine sensitive, and two were resistant (Deacon *et al.*, 1994).

Chloroquine: Kwazulu-Natal isolates showed 2/14 to be sensitive to chloroquine, comparable to 1987 results when 2/17 and 1988 when 0/14 were sensitive. Chloroquine resistance has risen dramatically since appearing on the African continent in 1978, but remains effective in some regions of Central America, the Caribbean and the Middle East (Deacon *et al.*, 1994).

Mefloquine: All strains were sensitive to mefloquine. It should be kept in mind, however, that in Zambia mefloquine resistance was observed ranging between 15% and 48% of isolates. It must also be taken into account that these samples were collected during 1989-1990, and that mefloquine resistance has since been reported. Mefloquine may be used to treat multi-drug resistant malaria, but has been shown to cause neuropsychiatric side effects in some patients (Schlagenhauf, 1999) (Deacon *et al.*, 1994).

Quinine: Kwazulu-Natal reference isolates and field strains were shown to have a lower sensitivity to quinine than the four 1990 field isolates. Transvaal and Malawi reference isolates also had reduced sensitivity. Quinine is still used for the treatment of uncomplicated malaria (Raynes, 1999) (Deacon *et al.*, 1994).

Sulphadoxine/pyrimethamine: Southern African reference isolates revealed a wide range of susceptibilities. A sharp increase in the amount of drug necessary for IC_{50} was shown from 1987 to 1989 in Kwazulu-Natal as well as Transvaal and Malawian isolates. The apparent reduced susceptibility of Kwazulu-Natal isolates could have serious implications for treatment of *P. falciparum* malaria, as this drug combination has been used for treatment in this region since 1988. Combinations of sulfadoxine-pyrimethamine are added to chloroquine treatment when the patient contracted malaria in a region such as South America or Asia (Pearson and Hewlett, 1987)(Deacon *et al.*, 1994).

Polymorphisms in pfmdr1, the gene encoding the P-glycoprotein homologue 1 (*Pgh1*) protein of *Plasmodium falciparum*, have been linked to chloroquine resistance; *Pgh1* has also been implicated in resistance to mefloquine and halofantrine. The same mutations influence parasite resistance towards chloroquine in a strain-specific manner and the level of sensitivity to the structurally unrelated compound, artemisinin.

Compounds such as doxycycline have a slow but potent action against

the asexual blood stages of *P. falciparum*, as well as the primary intrahepatic stages (Juckett, 1999). Halofantrine may be used when a poor reponse to quinine is found, but shows poor and variable absorption and may become toxic (Touze *et al.*, 1997). Artemesinin-based drugs are mostly restricted to the treatment of multi-drug resistant malaria from south-east Asia (White and Olliaro, 1998). The World Health Organisation (WHO) has published guidelines regarding artemesinin use to ensure effective treatment aiming to prevent the fast development of resistance. A summarised diagram of geographical malaria drug resistance is provided in Figure 1.1.

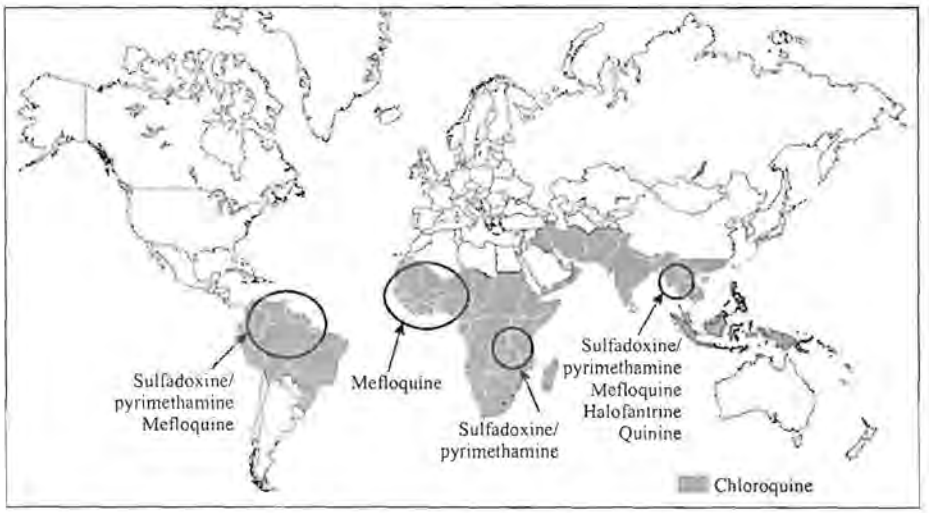


Figure 1.1: Summarised geographical map of recent surveys in malaria drug resistance (Centers for Disease Control and Prevention, Atlanta, 1999).

While various projects are in progress to develop malaria vaccines, realistic estimates put efficient vaccines as far off as 15 years (Holder, 1999). Anti-malarial drugs have contributed heavily towards combatting the disease for the past 100 years, and much potential still exists for developing new drugs against proteins involved in parasite metabolism and invasion. This study will focus on intervention in the major metabolic pathways of *P. falciparum*.

1.2 Metabolic pathways

Antiparasitic drugs have been designed to target a series of metabolic pathways (Subbayya *et al.*, 1997). The glycolytic and nucleic acid biosynthesis pathways have been investigated in detail in many parasites, and enzymes from these two pathways will be addressed experimentally in this study. This is preceded by a literature overview of other work done on malaria metabolism and detailed investigations into some potential drug target enzymes.

1.2.1 Glycolysis

The malaria parasite lacks a functional tricarboxylic acid cycle and is dependent solely on glycolysis for its ATP requirements (Sherman, 1979)(Roth *et al.*, 1988a)(Roth *et al.*, 1988b). Glucose consumption in *Plasmodium*-infected red blood cells is increased by approximately 50-100 fold over that of uninfected cells (Roth *et al.*, 1988a). The parasite glycolytic enzymes are believed to be associated with membrane components (Dobeli *et al.*, 1991). These enzymes come into contact with the immune system, as antibodies have been found to react to triosephosphate isomerase (TIM) in infected patients (Ritter *et al.*, 1993). Glycolytic enzymes from malaria tend to have a relatively high degree of homology with the human forms. Some of these glycolytic enzymes have been investigated as drug targets. An overview of the glycolytic pathway is given in Figure 1.2.

Hexokinase from *P. falciparum* has 26% identity with the human form, and increased hexokinase activity is found on infection (Olafsson and Certa, 1994). A large C-terminal hydrophobic stretch is present which has been proposed to play a role in membrane association (Olafsson *et al.*, 1992).

Phosphoglucoisomerase (GPI) shows a 4-9 times increase in infected erythrocytes (Srivastava *et al.*, 1992). The expression of GPI parallels parasite maturation and reaches a maximum at the trophozoite / schizont stage. The gene contains a 1773-base pair open reading frame, has no introns, and maps to *P. falciparum* chromosome 14. Of the deduced amino acid sequence, 34% is identical to human glucose phosphate isomerase, with highest similarity in regions of the proposed active sites (Kaslow and Hill, 1990). Gene complementation of an *Escherichia coli*

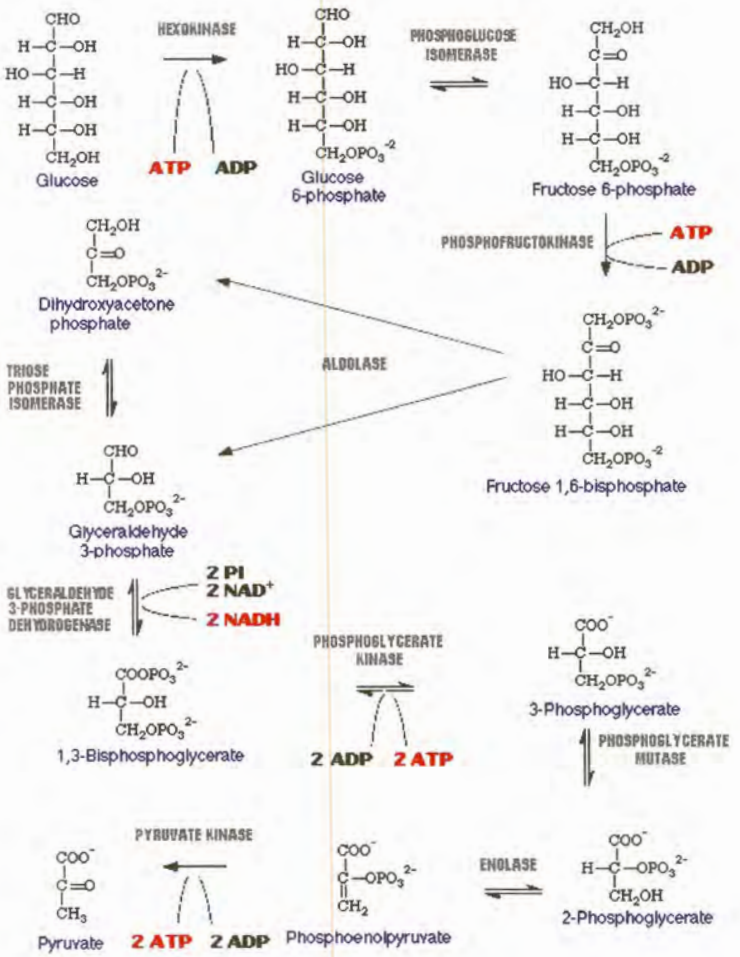


Figure 1.2: The glycolytic pathway

double mutant was used to isolate and express the gene coding for *P. falciparum* glucose phosphate isomerase.

Aldolase is of importance for vaccine development as well as for drug design, since vaccination in monkeys has been shown to offer partial protection (Certa *et al.*, 1988). It shows high sequence homology to human aldolase. Aldolase is known to bind to red blood cell cytoskeletal elements, and the N-terminal end of band 3 is a highly negatively charged inhibitor of aldolase (Dobeli *et al.*, 1990). Inhibitors designed on this basis were unfortunately not selective for *P. falciparum*.

Phosphoglycerate kinase was isolated by Grall *et al.* (1992). The activity of this ATP-generating enzyme in infected red blood cells was found to be 7 times higher than in uninfected cells. Two *P. falciparum* isoen-

zymes with neutral isoelectric points were detected.

Lactate dehydrogenase was recently crystallised by Dunn *et al.* (1996). Some features were shown to differ from the human form such as a displacement of the NADH cofactor. A cleft adjacent to the active site was also identified as a possible drug target. The malaria enzyme was selectively inhibited by gossylic nitrile diacetate (GNDA).

Enolase was isolated and characterised by Read *et al.* (1994). All amino acid residues implicated in substrate/cofactor binding and catalysis are conserved in the malarial enolase molecule. The predicted protein sequence displays approximately 60-70% identity to enolase molecules of other eukaryotes. Of particular significance in this well conserved molecule is a characteristic 5-amino acid insertion sequence that is identical in position and virtually identical in primary structure to that which is otherwise found uniquely in plant enolase proteins.

1.2.2 Purine salvage pathway

P. falciparum like all other parasitic protozoa lacks a *de novo* purine biosynthesis pathway, and obtain purines either by uptake or by a salvage pathway (Sarma *et al.*, 1998).

The major purine source for the parasite is hypoxanthine which is generated from ATP. It is known that the addition of hypoxanthine oxidase to parasite culture medium will lead to the death of parasites (Berman *et al.*, 1991). Since both adenosine and guanosine are formed by the hypoxanthine-guanine phosphoribosyltransferase (HGPRT) pathway via inosine monophosphate, HGPRT would presumably be a good target for antiparasitic drugs. HGPRT was cloned and expressed by Keough *et al.* (1999). Li *et al.* (1999) designed inhibitors corresponding to transition state analogues (the immucillins) which are the most powerful inhibitors yet reported for both human and malarial HGPRTs. Parasite HGPRT has an additional substrate specificity for xanthine, suggesting xanthine analogs as possible inhibitors (Queen *et al.*, 1988). The adenylosuccinate pathway which generates AMP from IMP seems to be unique to parasites, and has been inhibited by the compound hadacidin (Webster *et al.*, 1984). The activity of inosine monophosphate dehydrogenase is inhibited by Bredinin, which affects the formation of GMP from IMP (Scott *et al.*, 1987). Both the aforementioned pathways may be putative targets for selective inhibitors.

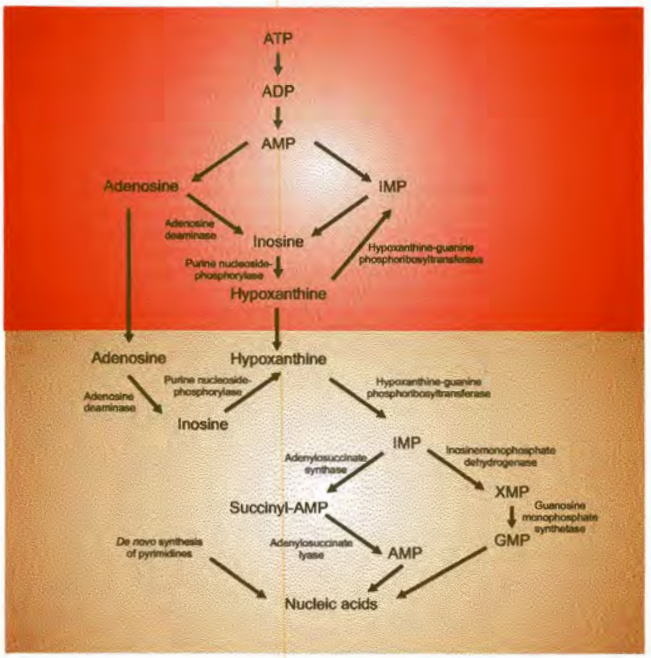


Figure 1.3: The postulated purine salvage pathway in *P. falciparum*. Red indicates the red blood cell pathway and brown that of the parasite.

1.2.3 Pyrimidine biosynthesis

The malaria parasite is solely dependent on the *de novo* pathway for pyrimidine synthesis, as they are unable to use a salvage system and human red blood cells contain low pyrimidine levels. Of the enzymes in the *de novo* pathway, carbamoylphosphate synthetase, dihydrooratate dehydrogenase and thymidylate synthase-dihydrofolate reductase have been cloned.

Derivatives of orotate have been shown to inhibit dihydroorotase activity and dihydroorotase dehydrogenase (DHODH), suggesting these enzymes as targets for antiparasitic drugs (Krungkrai *et al.*, 1992). The sequence of DHODH contains an insertion of 42 amino acids compared to the human form. Such insertion events are found in several malaria metabolic enzymes (LeBlanc and Wilson, 1993).

The investigation of malaria metabolic enzymes as new drug targets will hopefully lead to drugs which do not develop resistance within a short time period.

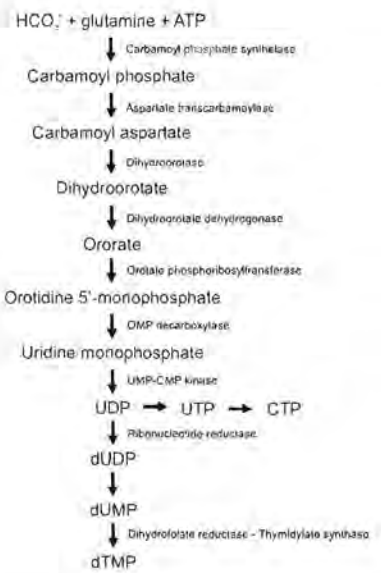


Figure 1.4: *De novo* pyrimidine biosynthesis in *P. falciparum*.

1.3 The target enzymes

Dihydrofolate reductase (DHFR) and triosephosphate isomerase (TIM) were chosen as targets for this study. Although malaria DHFR has not been crystallised, the structure has been studied extensively in other species. DHFR has been characterised in detail regarding point mutations leading to drug resistance, and it has been shown that modifications to existing drugs may lead to the return of therapeutical properties. DHFR has been validated as a drug target, and the impact of inhibition has been characterised in detail. TIM from *P. falciparum* was crystallised during the course of this study by Velanker *et al.* (1997). Although TIM has not been validated as a drug target, its structure has been studied extensively in other species, its mechanism is well known and TIM has been subjected to a wide variety of mutational analyses. As a key enzyme in the parasite glycolytic cycle, TIM holds great potential as a putative drug target.

1.3.1 Dihydrofolate reductase

The vitamin folic acid was discovered in 1930, and found to be abundant in leafy foliage, hence the name folic acid (Mathews and Van Holde, 1990). It consists of three moieties: a pteridine ring (6-methylpterin), *p*-aminobenzoic acid (PABA) and glutamic acid (Figure 1.5). Different

numbers of glutamic acid residues may be conjugated to PABA, but animals can only transport the mono-glutamated form. Poly-glutamated folic acid can thus not escape from the cell following conjugation.

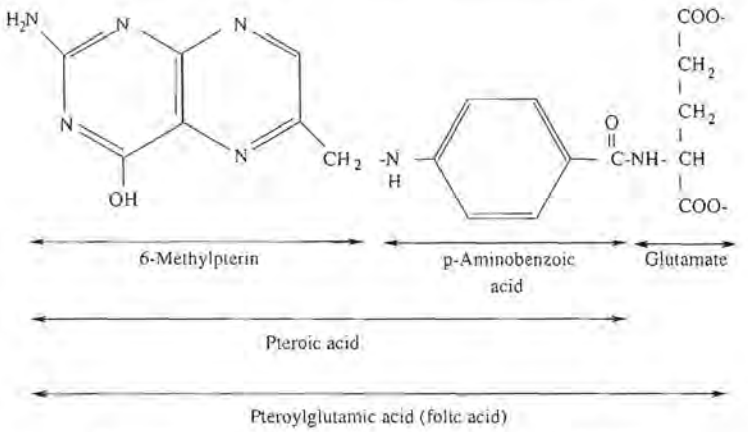


Figure 1.5: Structure of folic acid moieties.

Once inside the cell, folate is converted to active forms by reduction with the enzyme dihydrofolate reductase. The first reduction produces 7,8-dihydrofolate, followed by a second producing 5,6,7,8- tetrahydrofolate (Figure 1.6).

The coenzymatic function of tetrahydrofolate is the mobilisation and utilisation of single-carbon functional groups. These are involved in the amino acid metabolism of serine, glycine, methionine and histidine, as well as purine biosynthesis and methyl group incorporation of thymine.

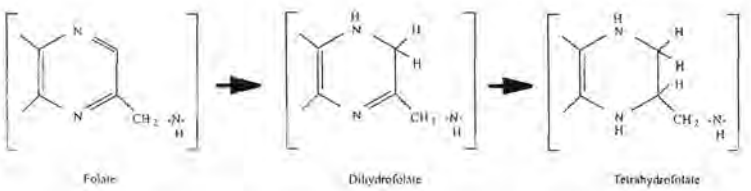


Figure 1.6: Reduction of folic acid.

Dihydrofolate reductase as a therapeutic target

The enzyme dihydrofolate reductase has been widely studied in a variety of species, including human, chicken and bacterial systems. The malaria DHFR enzyme occurs in the cytosol as a bifunctional complex with thymidylate synthase. The key role of DHFR is in the synthesis

of thymidine, reducing dihydrofolate to tetrahydrofolate (THF) in conjunction with NADPH as cofactor. THF is then further metabolised by thymidylate synthase for the methylation of dUMP to dTMP.

Much attention has been given to the human form of DHFR, because of its use as a chemoterapeutic target in diseases such as cancer. Methotrexate is a potent inhibitor of the enzyme, resulting in inhibition of cellular tetrahydrofolate metabolism and ultimately cell death.

Human dihydrofolate reductase

The general structure of human dihydrofolate reductase comprises of a larger, as well as a smaller sub-domain, with the active cleft inbetween these two domains. This cleft contains the substrate as well as the NADPH binding sites (Figure 1.7).

The active site comprises of contacts from different parts of the protein's primary structure in a spatial arrangement to yield catalytic activity. Various residues have been implicated in active site contacts.

Trp24 (red)has been shown to be conserved in all bacterial and vertebrate dihydrofolate reductases. It is part of a loop that connects strand A of the β-sheet and α-helix B (Beard *et al.*, 1991). Vertebrate DHFRs show a conserved segment of this loop *Pro 23-Trp 24-Pro 25-Pro 26* (red). This loop forms one wall of the active site which is mostly hydrophobic except for *Glu30*. *Trp24* shows hydrophobic and van der Waals interaction with the nicotinamide moiety of NADPH, and its indole nitrogen interacts with the C_4 oxygen of bound folate through a bridge provided by a bound water molecule. The synthetic mutation of this *Trp24* to a *Phe* led to a decrease in substrate affinity and hydride transfer, but an increased k_{cat} value.

Phe 31 (yellow) is one of several hydrophobic amino acids lining the active cavity of hDHFR (Chunduru *et al.*, 1994). Previous mutations at this position have shown different and uncorrelated effects on methotrexate binding. The most likely contribution to decreased methotrexate binding is probably due to loss of interaction of the *Phe31* sidechain with methotrexate. The Van der Waals interactions found with the *Phe31* wild type enzyme are not present in mutants containing *Ser*, *Ala* or *Gly*. Similar interactions are found in the wild-type enzyme-folate complex. *Leu*, *Val* and *Gly* mutations may possibly exclude the water molecule. These mutations lead to tight methotrexate binding.

Figure 1.7: Human DHFR (PDB ID: 1DHF) bound to folate. *Trp24* and the conserved loop is indicated in red, *Phe31* in yellow, *Phe34* in green and *Arg70* in cyan. Folate is shown in CPK atom colors.

In each case very little perturbation of the backbone could be shown. However, in each case the presence of an additional molecule of water in the phenyl-occupied cavity could be demonstrated.

One of the components of the catalysis performed by this enzyme is the hydride transfer from C_4 of NADPH to C_6 of H_2folate, which is critically dependent on the spatial arrangement of the carbon atoms and the relative orientation of the nicotinamide and pteridine rings. This step is speculated to be preceded by an isomerisation concerning the protonation of N_5 of H_2folate.

Phe34 (green) is another of the hydrophobic residues lining the active site (Nakano *et al.*, 1994). Its phenyl ring is inclined at about 45° to plane of the pteridine ring of folate, and several atoms of the ring are

in van der Waals contact. A similar structure is found in chicken dihydrofolate reductase. Interaction of the phenyl ring of *Phe34* with methotrexate occurs in a similar fashion as in the case of folate. An investigation was performed where *Phe34* was replaced by smaller residues, and the effect on enzyme activity was studied. Introduction of *Ile*, *Val*, *Thr*, *Ser* or *Ala* at this position led to a marked decrease in affinity for the H_2folate substrate as well as for methotrexate (Cody and Ciszak, 1991).

Arg70 (cyan) has been shown to interact with the α-carboxylate of the terminal L-glutamate of folic acid or methotrexate. X-ray crystallographic data has shown that folate in the active site is directly H-bonded to three conserved hydrophilic residues, of which *Arg70* is one. This residue is invariant in all vertebrate and bacterial DHFR's. Replacement of *Arg 70* by *Lys* led to a significant increase of affinity for methotrexate. Methotrexate is known to bind in the inverse orientation from folate and to be in contact with *Glu 30* via ionic H-bonds. *Ile7* and *Val115* also interact via H-bonds with the 4-amino group of methotrexate (Thompson and Freisheim, 1991).

E. coli and *L. cassei* dihydrofolate reductase

Li and Benkovic (1992) showed that mutations in the α-helix of *E. coli* DHFR resulted in an increase in reaction rate, but decreased Km and hydride transfer rate. Comparison of the *E. coli* and the *L. cassei* structures have shown only 28% amino acid homology, but great similarity in overall structure. The solvent accessible active-site surface of DHFR in *E. coli* is 93% of that of the *L. cassei* form. The *E. coli* enzyme is shown in Figure 1.8.

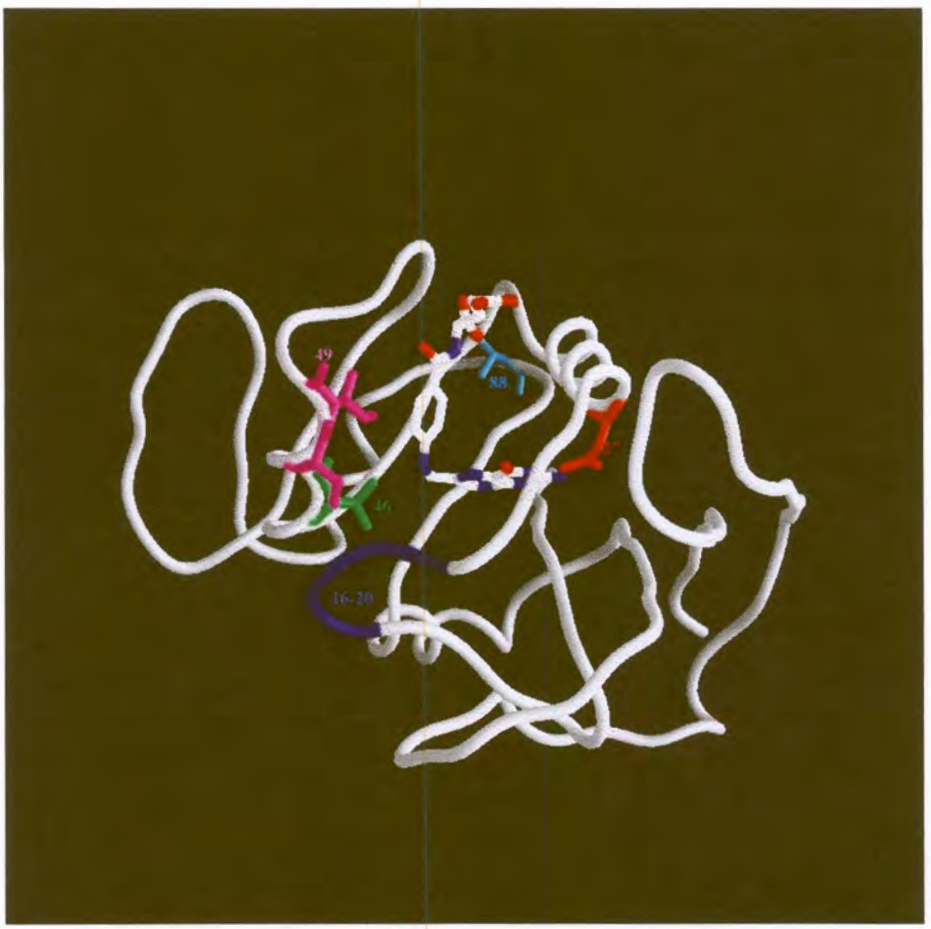


Figure 1.8: *E. coli* DHFR bound to folate (PDB ID: 1DYI). *Asp27* is indicated in red, *Thr46* in green, *Ser49* in purple, *Val88* in cyan and the loop formed by residues 16-20 in blue.

The active site of both enzymes is located in a cavity about 15Å deep, lined by mostly hydrophobic amino acids. NADPH binds with the nicotinamide moiety extended through the entrance of the cavity. Methotrex-

ate binds with the pteridine ring nearly perpendicular to the benzoyl ring. The pyrimidine edge of the pteridine ring penetrates into the center of the active site and interacts with *Asp27* (*E. coli*, red) or *Asp26* (*L. cassei*) through hydrogen bonding. *Thr46* (*E. coli*, blue) or *Thr45* (*L. cassei*) is conserved in all known DHFRs and maintains a hydrophobic interaction with the NADPH nicotinamide moiety. Another highly conserved residue is *Ser49* (*E. coli*, purple) or *Ser48* (*L. cassei*) which interacts with N_{10}-methyl of methotrexate and with the nicotinamide ribose through hydrogen bonding mediated by a water molecule. The side-chain of *Ile50* (*E. coli*) or *Phe49* (*L. cassei*) is in van der Waals contact with the *p*-aminobenzamide moiety of methotrexate. Near the N-terminus of the enzyme, *Arg44* (*E. coli*) or *Arg43* (*L. cassei*) forms an ion pair with the 5'-phosphate of adenine as well as a hydrogen bond to the ribose via a water molecule. The conserved *Leu54* is located in the loop structure following the helix, and interacts hydrophobically with the *p*-aminobenzamide moiety of methotrexate and is critical in the hydride transfer process.

Asp27 (red) has been implicated in the hydride transfer process, being the only ionisable group within the active site. It has been proposed to deliver a proton to N_5 of dihydrofolate through a network of water molecules, or act as proton relay with the solvent as source.

Previous work has shown *Asp27* to be critical for catalysis. Dihydrofolate needs to be protonated at N_5 of the dihydropterin ring in order to facilitate hydride transfer from C_4 of the nicotinamide ring of NADPH to C_6 of dihydrofolate. It is still unclear what the immediate source of the proton is, or how it is transferred to N_5 of dihydrofolate, seeing that the carboxylic side-chain of *Asp27* is about 6Å away from the N_5-C_6 double bond of DHF, and N_5 is buried in a hydrophobic environment composed of *Met20* and *Leu28*. *Asp27* may donate a proton itself, or may catalyze proton donation by a water molecule. The proton may be relayed to N_5 via a fixed water molecule *403*, present in every DHFR structure having a ligand in the pteridine binding site. One of the residues that can form a hydrogen bond with water *403* is an invariant *Trp22*. This residue makes hydrophobic interactions with *Met20* and *Leu4* as well as an indirect hydrogen bond with methotrexate through water *403*. In the DHFR-NADPH-folate complex, an additional Van der Waals interaction with the nicotinamide is found. *Trp22* (yellow) is also a key residue in the loop connecting βA to αB. This loop serves as a lid closing over bound ligands in the ternary complex. Warren *et al.* (1991)

applied site-directed mutagenesis to the conserved *Trp22* residue. Mutation of *Trp22* to *His* or *Phe* did not however yield the expected results. The water *403* remained unperturbed in the expected position.

Ahrweiler and Frieden (1991) investigated the role of a hinge region with three site-directed mutants at *Val 88*. The turn region reverses the direction of the backbone and separates the molecule into two domain-like structures. In the region following this turn, *Gly95-96* overhangs the methotrexate binding pocket. The carbonyl oxygen of *Ile94* forms a parallel β-bulge with *Gly95*, hydrogen-bonding to N_4 of methotrexate and is proposed to hydrogen bond with N_8 of the dihydrofolate pteridine ring. Site specific mutations disturbed the hydrophobic packing in the *85-91* turn region. A *Val 88* deletion, or mutation to *Ile* or *Ala* led to slight increases in the activity of the enzyme, thus the structure is not yet quite optimal in that region.

Li *et al.* (1992) investigated the role of loop I in DHFR. This mobile loop involves residues 9-24. The central section 16-20 (red) is disordered in the X-ray structure. Binding of folate and $NADP^+$ causes the disordered segment to form a hairpin turn defined by hydrogen bonding between the backbone carbonyl oxygen of *Met16* and the amino nitrogen of *Ala19*. The newly formed hairpin folds over the nicotinamide moiety of $NADP^+$ and the pteridine moiety of folate. In the formation of the ternary complex, the loop moves towards the α-C helix and a hydrogen bond is formed between the side-chain of *Asn18* and the backbone carbonyl oxygen of *His45* in the helix. The bending of the N-terminus of the toward the pteridine binding site causes hydrogen bonding between residues 21 and 24 and a supporting β loop to be moved.

Avian dihydrofolate reductase

Volz *et al.* (1982) determined the crystal structure of avian dihydrofolate reductase in complex with NADPH and phenyltriazine. The overall backbone folding is very similar to those in other dihydrofolate reductases. The structure of the DHFR-biopterin complex is shown in Figure 1.9.

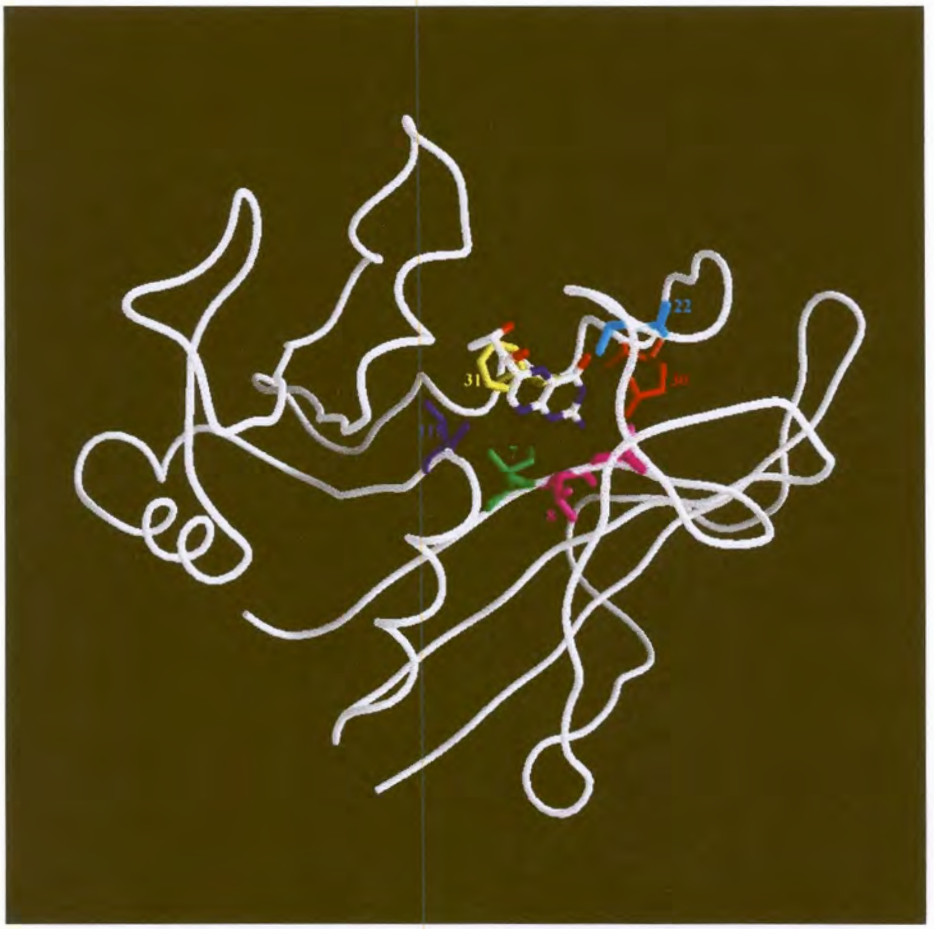


Figure 1.9: Avian DHFR bound to biopterin (PDB ID: 1DR1). *Glu30* is shown in red, *Leu22* in cyan, *Ile7* in green, *Val8* in magenta, *Tyr31* in yellow and *Val115* in blue.

The molecule is folded into an eight-stranded β-sheet consisting of seven parallel strands plus a single anti-parallel strand leading to the carboxy terminus. Four α-helices are packed against this sheet, while the ones remaining are involved in loops. Avian dihydrofolate reductase contains 189 residues, compared to the 159 in *E. coli* and the 161

in *L. cassei* species. About 70% of the extra residues occur in loops far from the active site.

The carboxyl group of *Glu 30* (red) in the chicken enzyme occupies the same position as *Asp27* in the *E. coli* structure, with the negatively charged carboxylate group in both cases hydrogen-bonded to the ring N_1 and 2-amino group of the respective inhibitors. Lining the binding pocket in hydrophobic contact with the triazine groups are side-chains of *Ile7* (green), the peptide bond between *Val8* and *Ala9*, the side-chains of *Ala9*, *Leu22* (yellow), *Tyr31*, *Phe34*, *Val 115* (blue) and the carboxamide of NADPH. The phenyl ring of *Phe34* is nearly parallel to the triazine ring. The hydroxyl of *Thr136* bonds to a carboxylate oxygen of *Glu30* and to the 2-amino group of the inhibitor, probably via a water molecule. Hydrogen bonds are donated by the 4-amino group of triazine to the carbonyls of *Ile7* and *Val115*. The inhibitor's methoxyphenyl group resides at the mouth of the pteridine binding pocket, analogue to pyrazine and C_9-C_{10} portions of methotrexate. It makes Van der Waals contact with the side-chains of *Leu22*, *Phe34*, *Thr56*, *Ser59*, *Ile60*, *Val115* and the nicotinamide ring of NADPH. *Leu22*, *Ser59* and *Ile60* are structurally equivalent to *Leu29* (*Met20*), *Ser48* (*Ser49*) and *Phe49* (*Ile50*) in the *L. cassei* and *E. coli* structures.

P. falciparum dihydrofolate reductase

DHFR in the malaria parasite has been extensively characterised in terms of nucleotide sequences, but little is known of its three-dimensional structure. Sequence characterisation has especially been done to account for parasite resistance to the anti-malarial drugs pyrimethamine and cycloguanil. These drugs were introduced more than 40 years ago, and in combination with sulfa-drugs have been widely used in the treatment of malaria (Peterson *et al.*, 1990). However, the occurrence of resistant strains have been reported. While increased expression of the enzyme may in some cases play a role in drug resistance (Zolg *et al.*, 1989), the major cause is mutations in regions of the enzyme affecting drug binding (Basco *et al.*, 1995).

Proguanil is activated *in vivo* by the mixed function oxidase system of hepatic enzymes to form the active drug cycloguanil. Both cycloguanil and pyrimethamine have been shown to strongly inhibit DHFR (K_i for cycloguanil = 0.78×10^{-9}M) with a selective anti-malarial action. It

is, however, not known if cycloguanil and pyrimethamine share the same binding site on the enzyme. A wide range of resistance has been demonstrated, with simultaneous mutations (positions 16, 59, 108) leading to resistance against either one, or both drugs at once (Bhasin and Nair, 1996).

De Pécoulas *et al.* (1995) developed a rapid and simple method for the detection of mutations at positions 16 and 108 of *P. falciparum* dihydrofolate reductase. The method consists of the PCR amplification of the DHFR gene, followed by restriction enzyme digestion of codons 16 and 108. Three different enzymes are used to cut the wild-type, the *Thr108* mutant and the *Asp108* mutant. Since every natural antifolate-resistant isolate identified until now carries a mutation in codon 108, determination of this codon can predict the sensitivity of any *P. falciparum* isolate. This technique was employed on local isolates as assayed by Birkholtz *et al.* (1998) to greatly increase knowledge regarding the resistance of *P. falciparum* to pyrimethamine in southern Africa.

Bzik *et al.* (1987) cloned and sequenced the gene for *P. falciparum* dihydrofolate reductase-thymidylate synthase. The amino-terminal DHFR domain was joined to the carboxy terminal TS (thymidylate synthase) domain by a 94-amino acid junction sequence. The TS domain was more conserved than the DHFR domain, and both were more homologous to eukaryotic than prokaryotic forms. The primary DHFR-TS protein contained 608 amino acids, with a calculated mass of 70kDa. The coding sequence resembled other *P. falciparum* sequences in that it had a lower A+T content (75%) than the flanking sequences (94% for 5'-UTR and 87% for 3'-UTR). *P. falciparum* DHFR had little homology to other species in amino acid regions 1-10 and 201-228, while region 11-200 has significant homology except for several amino acid insertions.

Basco *et al.* (1995) studied pyrimethamine and cycloguanil resistant strains, showing a total of 5 amino acid changes. These changes were at *Ala16*, *Asn51*, *Cys59*, *Ser108* and *Ile164*. A single mutation at *Asn108* was associated with moderate pyrimethamine and cycloguanil resistance. Additional mutations at *Asn51Ile* and *Cys59Arg* were required for high resistance to cycloguanil and pyrimethamine. A fourth *Ile164Leu* mutation was observed in a highly resistant Asian strain.

Peterson *et al.* (1990) analyzed cycloguanil resistant *P. falciparum* DNA sequences, finding a Ser to Thr mutation at position 108. This position is usually associated with a Ser to Asp mutation in pyrimethamine resistance. A further Ala to Val mutation was found at position 16. Pe-

terson *et al.* (1988) speculated that the *Asn108* mutation occurs at a site analogous to that of the threonine residue in the C α-helix of bacterial, avian and human enzymes. *Thr56* in this helix makes contact with the methoxyphenyl group of a triazine inhibitor. The *Asn108* mutation would be expected to affect these kind of contacts and may therefore prevent pyrimethamine binding.

It is of obvious importance that the wide range of contacts described in the previous paragraphs occurring in the human, bacterial and avian forms of dihydrofolate reductase should be taken into account in all studies performed on the *P. falciparum* enzyme. In summary, the most important residues are *Trp24*, *Asp27*, *Phe34*, *Arg70* and *Val115* which interact with the ligand in most of the species studied.

Homology models of malaria DHFR have been prepared by Lemcke *et al.* (1999) as well as Toyoda *et al.* (1997), as discussed in Chapter 3.

1.3.2 Triosephosphate isomerase

Triosephosphate isomerase (D-glyceraldehyde 3-phosphate ketol- isomerase; EC 5.3.1.1) is responsible for the interconversion of dihydroxyacetone phosphate and D-glyceraldehyde 3-phosphate. TIM is an enzyme critical for energy production, since no organisms without a TIM gene have been described. A deficiency in TIM leads to chronic hemolytic anemia and neuromuscular disorders (Bardosi *et al.*, 1990).

The cysteine residues of TIM were investigated in the enzymes from *Saccharomyces cerevisiae*, *E. coli*, rabbit, chicken and *Schizosaccharomyces pombe* (Garza-Ramos *et al.*, 1996). Methyl disulfides were produced using phenyl methanethiosulfonate. No effect was found on *S. cerevisiae* and *E. coli*, but the chicken and *S. pombe* were inhibited. It was found that the enzymes having a Cys in position 217 were sensitive to $MePhSO_2$-Ph treatment.

Human triosephosphate isomerase

An overview of the structure of human TIM complexed to the active-state analogue, phosphoglycolic acid (PGA) is shown in Figure 1.10, and the TIM active site region is depicted in Figure 1.11.

The mechanism of the TIM reaction has been studied over the years and may be summarised as follows (Aqvist and Fothergill, 1996): After binding of DHAP one of the *C-1* protons is removed by the *Glu165* (red) catalytic base, yielding the corresponding enediolate. It is then protonated at *O-2* by the imidazole ring of *His95* (green), yielding a double protonated enediol and a imidazolate ion. This is followed by the recapture of a proton by *His95* from the *O-1* oxygen creating an enediolate lacking the proton at *O-1*. Protonation of the enediolate occurs at C-2 by the carboxyl group of *Glu165*, thus restoring the general base to release the D-GAP product. The proximity of *Lys12* (blue, forming an ion bond with *Glu165*) to the substrate together with hydrogen bonds from *His95* and *Asn10* (purple) are important in stabilising the enediolate species. The α-helix directed at His95 also contributes to the stability of the imidazolate ion (Lodi and Knowles, 1993). Simulations have indicated that a water molecule situated between *Glu165* and *His95* plays an important role during catalysis, providing stabilisation by hydrogen-binding to *Glu165*, the substrate and negatively charged *His95*. Investigation of a *Ser96Phe* mutant showed that cat-

Figure 1.10: Dimeric human TIM with the active sites of both subunits indicated in green (PDB ID: 1HTI). PGA is shown in red in the active site of one subunit.

alytic activity was dramatically decreased, and that this was probably due to the displacement of a water molecule. *Thr75* (yellow) from the second subunit donates a hydrogen bond to the carboxylate group of Glu*97* (cyan) when *His95* is neutral. This is probably one of the reasons for the low activity of monomeric TIM (Borchert *et al.*, 1994).

Figure 1.11: An overview of the human TIM active site region (PDB ID: 1HTI). PGA is shown in CPK colours. *Glu165* is shown in red, *His95* in green, *Lys12* in blue, *Asn10* in purple and *Glu97* in cyan. *Thr75* and its chain from the second subunit are shown in yellow.

Yeast triosephosphate isomerase

TIM from yeast was first crystallised by Lolis *et al.* (1990). It was shown to be active only in dimeric form (Waley, 1973). The active site structure is shown in Figure 1.12.

Figure 1.12: The active site region of yeast TIM (PDB ID: 2YPI). *His95* is indicated in red, *Glu165* in green, *Lys12* in cyan, *Glu97* in blue and residues 71-77 in magenta. PGA is indicated in CPK colours.

The imidazole Nϵ of *His95* was within hydrogen bond distance (3Å) of a solvent oxygen and a carboxylate oxygen of *Glu165* (green, 3.3Å). Since the carboxylate was unprotonated and both water and the glutamate side-chain may act as hydrogen bond acceptors, is was concluded that the Nϵ2 of *His95* (red) was the hydrogen bond donor and was thus protonated. The Nδ1 atom is within hydrogen bond distance of the backbone amide nitrogen of *Glu97* (blue). Stereochemical restraints

dictate that the amide nitrogen must be the proton donor and the Nδ1 the acceptor. The Nδ1 is thus unprotonated and the imidazole in an uncharged state.The major contacts of the subunit interface were at residues 71-77 (magenta), extending from one subunit into a pocket near the active site of the other subunit. This area is in general more hydrophobic than the rest of the monomer surface. At the interface, 15 water molecules were found of which 10 mediated interaction between the subunits. Little sequence conservation between species occurs at the interface, and the authors speculated about the possibility of designing selective inhibitors preventing TIM dimerisation. A study by Lodi *et al.* (1994) emphasised the importance of a positively charged active site. This was supported by mutation of *Lys12* to *Met* giving rise to a loss in enzyme activity. Substitution with amino acids carrying a positive charge or having the potential to be positively charged, resulted in lower enzymatic activity. The TIM-2-phosphoglycolic acid complex was crystallised (Lolis and Petsko, 1990), and showed that PGA formed hydrogen bonds to the side-chains of *His95* and *Glu165*, confirming that *Glu165* is protonated upon PGA binding. Conformation changes induced by binding were the movement of the *Glu165* side-chain by 2Å and the movement of a 10-residue flexible loop 7Å closer to the active site. This so called "hinged lid motion" closes over the active site upon binding, providing a "hinged lid" structure (Joseph *et al.*, 1990). Crystallisation of the TIM-phosphoglycolohydroxamate complex suggested the mechanism whereby the carboxylate of *Glu165* removes a proton from C1 of DHAP while *His95* donates a proton to the DHAP oxygen to form an enediolate intermediate (Davenport *et al.*, 1991). When *Glu165* was changed to *Asp*, the distance to the substrate increased by 1Å, but the major impact was to decrease catalytic activity due to differences in orientation of the catalytic groups (Joseph-McCarthy *et al.*, 1994).

Avian triosephosphate isomerase

The crystal structure of chicken TIM was determined as early as 1975 by Banner *et al.* (1975). They showed two roughly spherical subunits with a diameter of approximately 35Å each. An inner barrel fold with 8 strands of parallel pleated β-sheets was visible. The TIM-PGH complex was crystallised by Zhang *et al.* (1994), and compared to the uncomplexed structure. The model of TIM-PGH was similar to that obtained for both unliganded avian TIM and TIM from yeast. Upon binding to PGH, the carboxylate group of the catalytic base *Glu165* was shown to

move 2-3Å toward PGH. In the free enzyme, the carboxylate is oriented deeper in the active site pocket, and the carboxylate oxygens form two hydrogen bonds with the backbone amide nitrogen and the side-chain hydroxyl group of *Ser96*. After PGH binding, these hydrogen bonds are broken and the *Glu165* (green) side-chain moves into proximity of PGH (Figure 1.13).

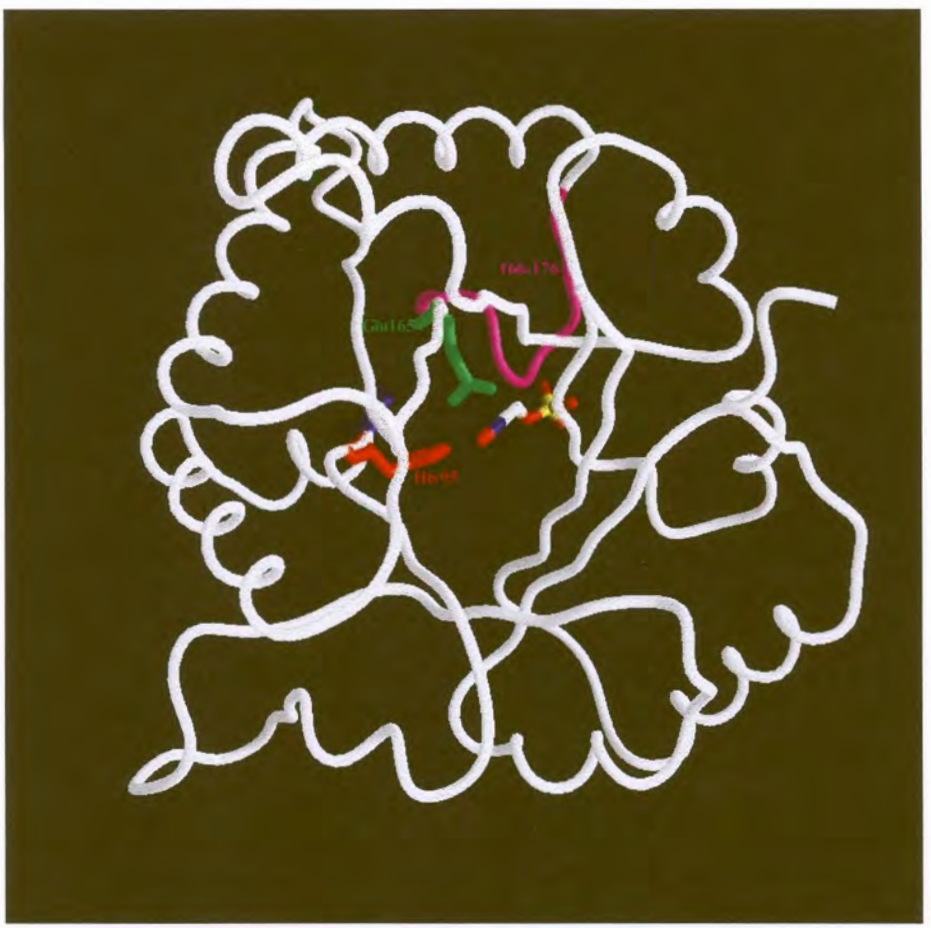


Figure 1.13: The active site of avian TIM (PDB ID: 1TPH). *His95* is indicated in red, *Glu165* in green and residues 166-176 in magenta. PGH is shown in CPK colours.

The loop of residues 166-176 (magenta) in the unbound state is exposed to solvent, leaving the active site accessible to the substrate. In the complex, the loop closes over the active site acting like a rigid lid.

Trypanosoma triosephosphate isomerase

The TIM enzyme from *Trypanosoma brucei* was crystallised by Wierenga et al. (1984). Residues involved in catalysis are conserved between TIM of most species, decreasing the chances of developing selective inhibitors (Kuntz *et al.*, 1992). Three loops are of importance in the formation of dimeric TIM, as indicated in Figure 1.14.

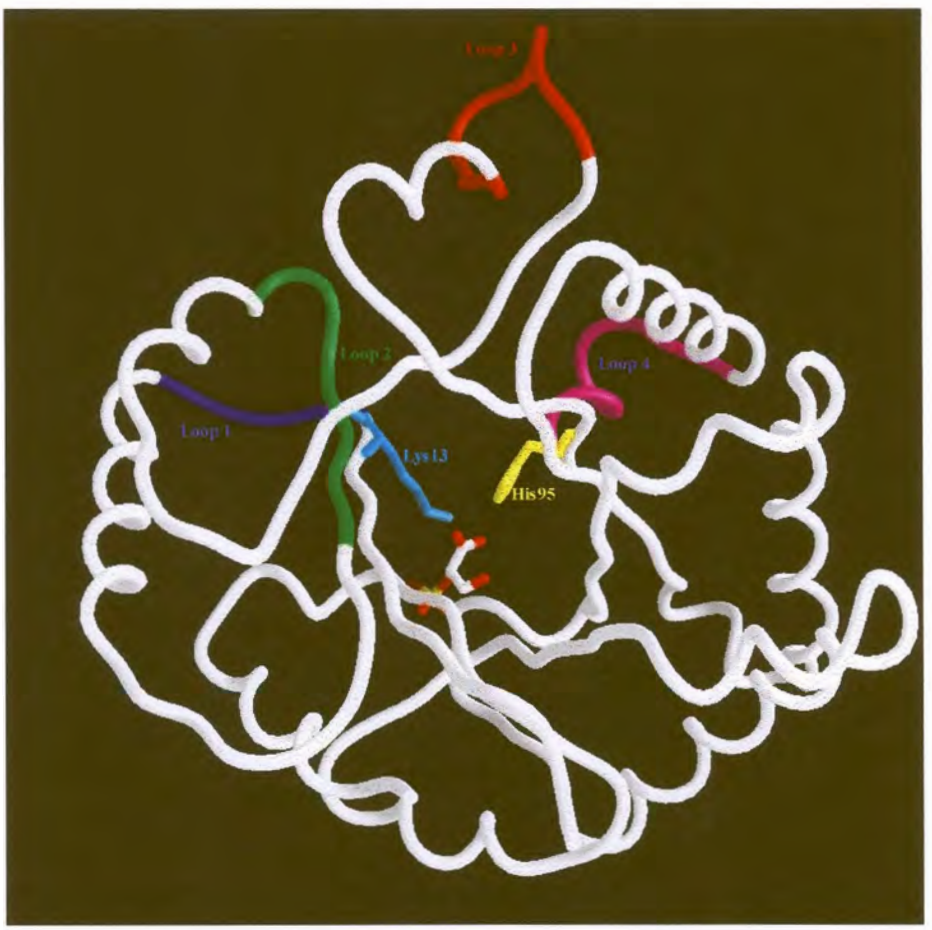


Figure 1.14: Structure of *Trypanosoma* TIM (PDB ID: 4TIM). *Lys13* is indicated in cyan, and *His95* in yellow. Loop 1 is colored blue, loop 2 green, loop 3 red and loop 4 magenta.

Loop 3 (red) is relatively long, with the tip around *Thr75* fitting into a cavity near the classical β turn. A series of cyclic peptides were designed to mimic the loop-2 β-turn (green), but no inhibition of enzyme activity could be shown. A second series of more hydrophobic cyclic peptides were able to inhibit the TIM from *Trypanosoma brucei*, but not

from other species such as rabbit, yeast and *E. coli*. The inhibition was non-competitive and reversible, and it was shown that more than one peptide could bind per active site. The anti-trypanosomal drug Suramin appeared to compete with the enzyme's substrates (K_i = 0.1mM), indicating an influence at the active site. Inhibition was not present at an ionic strength above 0.01M (Lambeir *et al.*, 1987).

In dimeric TIM, loops 1 (blue) and 4 (magenta) are very rigid due to interaction with the other subunit. In monomeric TIM *Lys13* (cyan) and *His95* have been shown to be more flexible. By point mutation, these two residues were shown to be essential for optimal catalysis (Schliebs *et al.*, 1996). The *Lys13Ala* mutant was completely inactive although still being able to bind substrate analogues. The *His95Ala* mutation was 50 times less active than the wild-type enzyme.

In summary, *Lys12*, *His95*, *Glu165* and *Thr75* play a major role in the catalytic process, and should be taken into account in any contacts identified during ligand searching and design.

Plasmodium falciparum triosephosphate isomerase

Triosephosphate isomerase from *Plasmodium falciparum* was first cloned and sequenced by Ranie *et al.* (1993). A cDNA library was used as initial source. TIM cDNA contained a 5'-untranslated region of 172bp, followed by an open reading frame of 744bp which showed an AT-content of 71%. The features surrounding the start codon correlated well with that found for other *P. falciparum* genes, with an adenine preceeding the ATG followed by a purine (Robson and Jennings, 1991). A 42-45% sequence identity was seen with higher and lower eukaryotes, and 40% with bacteria with active site residues being conserved in comparison to yeast, chicken and *Trypanosoma*. Significant variance in comparison to the human TIM active site region was *Ala73Ser*, *Phe75Tyr*, *Ser96Phe*, *His100Lys*, *Val101Tyr*, *Lys130Ser*, *Val167Leu* and *Gly233Asn*. Upon investigation of the *P. falciparum* genomic DNA, an intron of 309 bp was found in the codon for residue 38. Like all introns it begins with GT and ends with AG. The intron was shown to be 86% AT-rich. In comparison, TIM from maize contains 8 introns (Marchionni and Gilbert, 1986), but the location of the *P. falciparum* intron is conserved in chicken (Straus and Gilbert, 1985), human (Brown *et al.*, 1985), rhesus monkey (Old and Mohrenweiser, 1988), chimpanzee (Craig *et al.*, 1991) and maize. The enzyme from

P. falciparum was overexpressed using the pTrc99A plasmid in *E. coli* AA200 which is TIM deficient. Large amounts of a 28kDa protein were visible, and the specific activity in crude lysates was four times higher than in the control *E. coli* which was TIM deficient (Ranie *et al.*, 1993).

1.4 The approach

The identification of new inhibitors is enhanced by structural modelling of the target proteins. Detailed characterisation of the enzyme's active sites may be used in rational approaches to drug screening of chemical software libraries and *de novo* design of new inhibitory compounds. Once ligands have been selected, *in vitro* and *in vivo* screening is needed to evaluate the potential of these compounds as possible lead drugs. In summary, the following approaches were followed in this study as discussed in the successive chapters:

- Cloning, expression, characterisation and purification of the two putative malaria drug target enzymes, dihydrofolate reductase and triosephosphate isomerase.
- Homology modelling of dihydrofolate reductase and triosephosphate isomerase.
- Ligand docking and *in vitro* ligand screening of selected compounds binding to malaria triosephosphate isomerase, and *in vivo* screening of malaria cultures.

Chapter 2

Expression of putative malaria drug target proteins

2.1 Introduction

A large amount of a pure functional protein is required for a study involving inhibitors of an enzyme. Such quantities can usually only be produced by recombinant overexpression of the enzyme. The recombinant expression of *P. falciparum* proteins has met with problems in many different expression systems. The AT-nucleotide ratio of the malaria parasite varies from approximately 86% in non-coding regions to 69% in coding regions (Weber, 1987). This has clear implications on the tRNAs necessary for translation of these proteins (You *et al.*, 1999). The tRNA repertoire of an organism is optimised for the codons most often used for an amino acid, and the introduction of a gene employing different codons may lead to the exhaustion of certain tRNAs. Some promoters seem also to discriminate against the expression of malaria genes. This may be rectified in different ways. A popular approach is the preparation of whole synthetic genes (Sirawaraporn *et al.*, 1993)(Sano *et al.*, 1994). Oligonucleotides in the range of 50-150 bases are assembled either by restriction digestion followed by ligation, or alternatively by PCR-based methods. These synthetic genes were designed to utilise codon preferences of the expression host organism,

and also to take into account any start codon preferences of the promoter used. A recent alternative is to co-transfect the organism with a vector encoding the necessary tRNAs for supplementation purposes (You *et al.*, 1999).

Once translation of the malaria protein has been achieved, the next challenges are those of correct folding and solubility. Especially in bacterial systems, foreign proteins do not always fold correctly and these misfolded proteins tend to form intra-cellular inclusion bodies. The lack of chaperones may play a role in the incorrect folding of some proteins. The presence or absence of chaperones in malaria have not been conclusively proven but chaperone-like proteins have been found (Hassoun *et al.*, 1998)(Foley *et al.*, 1994). The overexpression of large quantities of a recombinant protein may also overload the cellular mechanisms and cause precipitation of a fraction of the protein which can not be accommodated in the cytosol. This may be optimised in several ways. One is to use a tightly controlled promoter, which does not allow any leaky expression in the uninduced stage (Studier *et al.*, 1990). Expression is then induced in a slow manner, so that large amounts of proteins do not suddenly accumulate in the cytosol. The growth of bacteria at lower temperatures has been shown to have a large effect on the solubility of recombinant proteins, probably due to the speed of growth, protein synthesis and also the protein conformation at lower temperatures (Song *et al.*, 1999). Some proteins which were insoluble at 37°C were successfully expressed at 25°C or 30°C.

Various systems have been used for the expression of malaria proteins such as *E. coli* (Sirawaraporn *et al.*, 1993), yeast (Bathurst, 1994), COS cells (Elliott *et al.*, 1990), *Dictyostelium discoideum* (Fasel *et al.*, 1992) and recombinant baculovirus-infected cultures (Matsuoka *et al.*, 1996). Some yeasts have very high AT-ratios similar to *P. falciparum* which may simplify the expression of malaria proteins. Bacterial systems remain the most economic and simple way for expression, as long as the proteins can be expressed in a suitable form for further use. For the purposes of vaccine development the protein may not necessarily need to retain full activity, but for use in drug screenings and mutational studies, the characteristics should be as close as possible to the native parasite protein. Some malaria metabolic enzymes that have been successfully expressed in recombinant systems are listed in Table 2.1.

From this list it can be clearly seen that *E. coli* is by far the most popular expression system for *P. falciparum* genes. Therefore, any global op-

Table 2.1: Some malaria enzymes which have been successfully expressed in recombinant systems.

Enzyme	Organism	Reference
Dihydrofolate reductase (DHFR)	*E. coli*	Sirawaraporn and Yuthavong (1986)
Triosephosphate isomerase	*E. coli*	Ranie *et al.* (1993)
Falcipain	Baculovirus	Salas *et al.* (1995)
Adenine phosphoribosyltransferase	Mouse L cells	Pollack *et al.* (1985)
Antigens	Yeast	Bathurst (1994)
Circumsporozoite protein	*Dictyostelium discoideum*	Fasel *et al.* (1992)
Ookinete surface antigen	Baculovirus	Matsuoka *et al.* (1996)
Lactate dehydrogenase	*E. coli*	Bzik *et al.* (1993)
Glutamate dehydrogenase	*E. coli*	Penny *et al.* (1998)
ADP-ribosylation factor	*E. coli*	Steketee *et al.* (1996)
Plasmepsin I	*E. coli*	Moon *et al.* (1997)
S-adenosylhomocysteine hydrolase	*E. coli*	Creedon *et al.* (1994)
Dihydrofolate reductase - thymidylate synthase	*E. coli*	Prapunwattana *et al.* (1996)
Glucose phosphate isomerase	*E. coli*	Kaslow and Hill (1990)
Hypoxanthine-guanine phosphoribosyltransferase	*E. coli*	Vasanthakumar *et al.* (1990)
Aldolase	*E. coli*	Dobeli *et al.* (1990)

timisations that can be made for *E. coli* expression would be extremely valuable.

The purification of the recombinant protein may present a problem, especially if an expression host possessing a protein very similar to the recombinant one is used. The ideal solution is the addition of some sort of a peptide tag selective for the recombinant product as a fusion protein, such as a poly-histidine tag, a S-tag (Richards *et al.*, 1972) or an antibody epitope tag (Enomoto *et al.*, 1998). These tags do not usually interfere with the correct folding of the protein and can often be removed after purification by protease cleavage.

If the recombinant protein can not be successfully expressed in a soluble form, methods exist for purifying the protein from inclusion bodies in a denatured form (Mukhopadhyay, 1997)(Lin and Cheng, 1991). The denatured protein may then be renatured by dialysis or other slow solvent exchange methods. In the case of malaria DHFR, the protein was denatured with guanidinium chloride, and subsequently renatured by dialysis to remove the strong denaturant. For example with histidine-tagged proteins, methods are available for renaturation during the column purification procedure (Lin and Cheng, 1991).

This chapter addresses the following goals:

- The optimisation of *E. coli* expression systems for malaria metabolic enzymes.
- The expression of *P. falciparum* DHFR in an active soluble form.
- The purification and characterisation of *P. falciparum* DHFR.
- The cloning of *P. falciparum* TIM cDNA.
- The expression of *P. falciparum* TIM.
- The purification and characterisation of *P. falciparum* TIM.

2.2 Materials and Methods

2.2.1 Expression of dihydrofolate reductase

2.2.1.1 The pET17 system

pET17 plasmid containing a synthetic gene for *P. falciparum* DHFR was a kind gift from the group in Thailand (Sirawaraporn *et al.*, 1993). pET17-DHFR was propagated in *E. coli* strain BL21(DE3) (Novagen, Madison) using Luria-Bertani (LB)-broth containing carbenicillin (50μg/ml). Cultures were grown to an OD_{600nm} of 0.8, and expression was induced by the addition of IPTG to a final concentration of 0.4mM. Cultures were grown for 4 hours post-induction, and bacteria harvested by centrifugation. The pellet from 1ml of culture was washed in 50mM Tris pH 7.4 (500μl), and resuspended in H_2O (25μl). An equal volume of 2x SDS loading buffer (100mM Tris pH 6.8, 200mM DTT, 4% SDS, 0.2 bromophenol blue, 20% glycerol) buffer was added, and samples were boiled for 5 min. Expressed products were analyzed by SDS-PAGE on a 12% gel.

Inclusion bodies were purified as follows: pET17-DHFR cultures were grown in LB containing carbenicillin (50μg/ml, Sigma, St. Louis) and thymine (100μg/ml, Sigma, St. Louis) to a OD_{600nm} of 0.8, and induced as above. Cultures were grown for 4 hours post-induction, and harvested by centrifugation. Pellets were resuspended in Buffer A (15ml) comprising of 20mM Tris pH 7.5, 20% sucrose and 1mM EDTA. After incubation on ice for 10 min, cells were centrifuged for 5 min at 4,000xg, and resuspended in 15 ml of ice cold water. Spheroplasts were pelleted by cenrifugation at 8,000xg for 10 min and resuspended in Buffer P (3ml) consisting of PBS, 1μg/ml leupeptin, 20μg/ml aprotinin and 0.5mM PMSF (Boehringer-Mannheim, Mannheim). Samples were sonicated at 50W on a Branson Sonifier (Branson Scientific, Danbury) with 3 cycles of 15s sonication followed by 15s incubation on ice. RNAse A (10μg/ml) and DNAse I (50μg/ml, Boehringer-Mannheim, Mannheim) were added and samples were incubated at room temperature for 10 min. Samples were diluted up to 15ml with buffer P, and centrifuged at 30,000xg for 30 min. Pellets were resuspended in buffer W (15ml) consisting of PBS, 25% sucrose, 5mM EDTA and 1% Triton X-100. After incubation on ice for 10 min, samples were centrifuged at 25,000xg for 10 min. Pellets were resuspended in denaturing solution D (10ml) consisting of 20mM potassium phosphate pH 7.0, 0.1mM

EDTA, 0.2M KCl and 6M guanidine HCl. The suspension was stirred gently overnight at 10°C. The DHFR protein was renatured by dropwise dilution with buffer A containing 0.2M KCl and 20% glycerol (200ml). Precipitates were removed by centrifugation at 10,000xg for 30 minutes.

To assay for enzymatic activity, a NADPH-coupled test was used. When dihydrofolate is reduced to tetrahydrofolate by DHFR, NADPH is converted to $NADP^+$ with a resulting decrease in absorbance at 340nm. Dihydrofolate was prepared as follows: Folic acid (38.2mg) was dissolved in 1M NaOH (1.6ml) and added to 10% ascorbate pH 6.0 (10ml). Sodium dithionite (400mg) was added and the reaction stirred for 5 min in an ice bath. HCl (1M) was added until the pH reached 2.8, and the reaction was stirred for another 10 min. The reaction was centrifuged for 5 min at 1,000xg with a temperature of 0°C, and the precipitate redissolved in 10% ascorbate pH 6.0 (10ml). 1M HCl was added until the pH reached 2.8, and the reaction was centrifuged for 5 min at 1,000xg with a temperature of 0°C. The precipitate was dissolved in 5mM HCl (10ml), and stored in the dark at 4°C. The enzyme assay was done in reaction buffer consisting of 0.5M potassium phosphate pH 7.5, 0.3mM mercaptoethanol, 10mM NADPH (Boehringer-Mannheim, Mannheim) and 2mM dihydrofolic acid. Bovine DHFR (Sigma, St. Louis) was used as a positive control. The decrease of NADPH was monitored at 340nm in a Shimadzu spectrophotometer (Shimadzu, Columbia).

2.2.1.2 The pTrxFus system

DHFR was cloned in the pTrxFus system due to its ability to enhance the solubility of recombinant systems. The synthetic DHFR gene was excised from pET17 by PCR. Two primers were designed incorporating restriction sites: DHFR1 with a KpnI site 5'-CGG **GGT ACC** AAT GAT GGA ACA GGT TTG-3' and DHFR2 with a SalI site 5'-ACG **CGT CGA** CTT AGT TGT TGG TTT TTT TTG-3'. pET17 was digested with EcoRI before use as template. The PCR reaction was conducted with 0.01μg of template using 1 cycle of denaturation at 95°C for 3 min, annealing at 60°C for 30s and extension at 72°C for 1 min. This was followed by 29 cycles with denaturation at 95°C for 30s, annealing at 60°C for 30s and extension at 72°C for 1 min. The PCR product was visualised by ethidium bromide staining of a 1% agarose gel. The band of approximately 800bp was excised, and purified on a Qiagen gel extraction

column according to the manufacturer's instructions (Qiagen, Hilden).

A sticky-blunt ended cloning strategy was used. The pTrxFus plasmid (2μg, Invitrogen, Carlsbad) was digested with SalI (20U) for 2 hours, and the enzyme heat inactivated at 70°C for 10 min. dNTP's were added to 5mM, followed by the addition of Klenow fragment (Boehringer- Mannheim, Mannheim) to the plasmid and the reaction was incubated at room temperature for 30 min. The enzyme was heat inactivated at 70°C for 10 min and the plasmid ethanol precipitated. The plasmid was subsequently digested with KpnI (20U) for 2 hours, followed by heat inactivation and gel purification. PCR product (2μg) was treated with T4 kinase (Boehringer-Mannheim, Mannheim) and ATP for 1 hour at 37°C, followed by heat inactivation. dNTP's were added to 5mM, followed by the addition of Klenow fragment (Boehringer- Mannheim, Mannheim) and the reaction was incubated at room temperature for 30 min at 37°C. The enzyme was heat inactivated and ethanol precipitated. The PCR product was subsequently digested with KpnI (20U) for 2 hours 37°C, and ethanol precipitated. Sticky/blunt-ended ligation was performed with 100ng of plasmid and 50 ng of PCR product using the Epicentre Fast Ligation Kit according to the manufacturer's instructions. Transformation was performed in competent G1724 *E. coli* cells (Invitrogen, Carlsbad). Colonies were screened for the presence of insert by digestion with PstI for 3 hours, and agarose gel analysis. Positive clones were sequenced with the Big Dye Kit (Perkin-Elmer, Foster City) on a ABI 377 sequencer according to the manufacturer's instructions.

Cultures bearing pTrxFus-DHFR were grown overnight in RM-medium consisting of Na_2HPO_4 (6g/l), KH_2PO_4 (3g/l), NaCl (0.5g/l), NH_4Cl (1g/l), casamino acids (20g/l), $MgCl_2$ (0.095g/l), 50% glycerol (20ml/l) and ampicillin (50μg/ml). Subsequently, 0.5ml of the culture was inoculated in 20ml of induction medium consisting of Na_2HPO_4 (6g/l), KH_2PO_4 (3g/l), NaCl (0.5g/l), NH_4Cl (1g/l), casamino acids (Gibco, 2g/l), $MgCl_2$ (0.095g/l), 20% dextrose (25ml/l) and ampicillin (50μg/ml) and grown to a OD_{600nm} of 0.5. Expression was induced by the addition of tryptophan (to 100μg/ml) and the culture was grown for a further 4 hours. Samples were centrifuged to obtain bacterial pellets, and were resuspended in 500μl osmotic shock solution #2 consisting of 20mM Tris pH 8.0 and 2.5mM EDTA. Samples were sonicated with three 15 s bursts as above, and flash frozen in a methanol bath. After thawing two more sonication-freeze-thaw cycles were performed. Soluble and

insoluble fractions were separated by centrifugation at 12,000xg for 10 min at 4°C. Expression was analyzed by SDS-PAGE on a 12% gel.

2.2.1.3 The pET32 system

DHFR was cloned into the pET32 system, as this system is designed to produce soluble products under control of the T7 promoter. The synthetic DHFR gene was excised from pET17-DHFR by PCR. Two primers were designed incorporating restriction sites: DHFR3 with a NcoI site 5'-GCA TG**C CAT GG**G TAT GAT GGA ACA GGT TTG-3' and DHFR4 with a BamHI site 5'-CGC **GGA TCC** TAT TAG TTG TTG GTT TTT TTG-3'. pET17-DHFR was digested with EcoRI before use as template. The PCR reaction was done with 0.01μg of template using 1 cycle of denaturation at 95°C for 3 min, annealing at 60°C for 30s and extension at 72°C for 1 min. This was followed by 29 cycles with denaturation at 95°C for 30s, annealing at 60°C for 30s and extension at 72°C for 1 min. The PCR product was visualised by ethidium bromide staining of a 1% agarose gel. The band of approximately 800bp was excised, and purified by silica-based gel extraction with melting in 6M sodium iodide, washing in Wash Solution consisting of 10mM Tris pH 7.5, 2.5mM EDTA, 50mM NaCl, 50% ethanol and elution in water (Smith *et al.*, 1995). The pET32 plasmid (2μg, Novagen) as well as the PCR product were digested with NcoI (20U) for 20 hours, and heat inactivated at 70°C for 10 min. After ethanol precipitation, digestion was performed with BamHI (20U) for 2 hours at 37°C, followed by heat inactivation at 70°C for 15 min, followed by gel purification. Ligation was performed with 100ng of plasmid and 50 ng of PCR product using the Epicentre Fast Ligation Kit according to the manufacturer's instructions. Transformation was performed in competent BL21(DE3) or AD494 *E. coli* cells (Novagen, Madison). Colonies were screened for the presence of insert by digestion of the plasmids with PstI for 3 hours, and agarose gel analysis. Positive clones were sequenced with the Big Dye Kit on a ABI 377 sequencer according to the manufacturer's instructions.

Cultures bearing pET32-DHFR were grown to a OD_{600nm} of 0.8 and expression was induced by the addition of IPTG to a final concentration of 0.4mM. Cultures were grown for 4 hours post-induction, and bacteria harvested by centrifugation. The pellet from 10ml of culture was suspended in TE buffer (500μl). DNAseI was added to 50μg/ml together with 1/10th volume of 1% Triton X-100. Samples were sonicated as

described above. Soluble and insoluble fractions were separated by centrifugation at 12,000xg for 15 min. Pellets were resuspended in 1ml of 1xSDS buffer, and an equal volume of 2x SDS buffer was added to supernatants. Samples were boiled for 5 min and expression was analyzed by SDS-PAGE on a 12% gel. This procedure was performed at 37°C, 30°C and 25°C.

2.2.2 Expression of triosephosphate isomerase

2.2.2.1 The pET15b system

RNA was isolated from parasite isolate PfUP1 with the Tri-Reagent method (Molecular Research Center, Cincinnati) according to the manufacturer's instructions. Total RNA was used for reverse transcription, and TIM cDNA was amplified by PCR employing TIM-specific primers TIM1 5'-**CAT ATG** GCT AGA AAA TAT TTT GTC G-3' and TIM2 5'-**GGA TCC** TTA CAT AGC ACT TTT TAT TAT ATC-3'. Total RNA (5μg) was incubated with TIM2 (2pmol) at 70°C for 10 min, and chilled on ice. To the 10μl mix was added 5x RT buffer (4μl), 0.1M DTT (2μl) and 2.5mM dNTPs (4μl), followed by incubation at 42°C for 2 min. Superscript II(200U, Gibco, Rockville) was added, and the reaction incubated at 42°C for 50min, followed by heat inactivation at 70°C for 15 min.

A PCR reaction was set up containing cDNA (0.5 μl/1.5 μl), 10x ExTaq buffer (5μl), 5mM dNTPs (4μl), primer TIM1 & 2 (10pmol each), H_2O (38μl) and ExTaq DNA polymerse (Takara, 0.3μl, Takara, Osaka)). Amplification (25 or 35 cycles) was performed with denaturation at 94°C for 45s, annealing at 50°C for1 min and extension at 72°C for 1 min. Samples were analyzed by agarose gel electrophoresis and ethidium bromide staining. The TIM PCR product was recovered from the gel by silica purification, and polished as follows: PCR product (4μl), 5mM dTNPs (1μl), reaction buffer (1.3μl), H_2O (5μl) and Pfu polymerase (Stratagene, 1U) were incubated at 72°C for 30 min.

Blunt-ended ligation was performed with PCR product (250ng), pCR-Script vector (100ng, Stratagene, La Jolla), buffer (1μl), rATP (0.5μl), SrfI (1U), T4 ligase (1U) and H_2O (4μl) in a total volume of 10μl at room temperature for 1 hour according to manufacturer's instructions, followed by heat inactivation at 70°C for 15 min. Transformation was performed in competent DH5α *E. coli* cells. Colonies were screened for the presence of insert by digestion with HindIII for 3 hours at 37°C,

and agarose gel analysis. Positive clones were sequenced with the Big Dye Kit on a ABI 377 sequencer according to the manufacturer's instructions.

The TIM gene was transferred to pET15b for expression. pET15b vector (Novagen, 3μg) as well as pCRScript-TIM (3μg) was digested with NdeI (30U) for 20 hours at 37°C. Following ethanol precipitation samples were digested with BamHI (30U) for 2 hours at 37°C, and purified by gel extraction with silica according to the GeneClean Kit instructions (BIO101, Carlsbad). Ligation was performed with pET15b (250ng), TIM (200ng), 10x ligation buffer (1.5μl), rATP (1.5μl) and Epicentre Rapid Ligase (1U) in a final volume of 15μl for 5 min at room temperature. Transformation was performed in competent BL21(DE3) *E. coli* cells. Colonies were screened for the presence of insert by digestion with HindIII for 3 hours at 37°C, and agarose gel analysis. Positive clones were sequenced with the Big Dye Kit on a ABI 377 sequencer according to the manufacturer's instructions.

Cultures bearing pET15b-TIM were grown to a OD_{600nm} of 0.8 and expression was induced by the addition of IPTG to a final concentration of 0.4mM. The procedure was performed either at 37°C or 30°C. Cultures were grown for 0, 8 and 16 hours post-induction, and bacteria harvested by centrifugation. The pellet from 10ml of culture was suspended in TE buffer (500μl). DNAseI was added to 50μg/ml together with 1/10th volume of 1% Triton X-100. Samples were sonicated as described above. Soluble and insoluble fractions were separated by centrifugation at 12,000xg for 15 minutes. Pellets were resuspended in 1xSDS buffer (1ml), and an equal volume of 2x SDS buffer was added to supernatants. Samples were boiled for 5 min and expression was analyzed by SDS-PAGE on a 12% gel.

2.2.2.2 Analysis of recombinant TIM

Recombinant *P. falciparum* TIM was purified by immobilised metal affinity chromatography (IMAC). Culture (100ml) bearing pET15b-TIM were grown to a OD_{600nm} of 0.8 at 30°C and expression was induced by the addition of IPTG to a final concentration of 0.4mM. Cultures were grown for 4 hours post-induction, and bacteria harvested by centrifugation. Pellets were resuspended in ice-cold binding buffer (4ml) consisting of 5mM imidazole, 0.5M NaCl and 20mM Tris pH 7.9. DNAseI was added to 50μg/ml, and the sample was sonicated as described

above. After centrifugation at 40,000xg for 25 min at 4°C, the supernatant was filtered through a 0.45m filter (Waters, Massachussetts). The IMAC column (bedvolume 2.5ml, Novagen, Madison) was washed with water (3 column volumes) at a flow speed of 10 column volumes per hour and charged with 50mM $NiSO_4$ (5 column volumes) followed by equilibration with binding buffer (3 column volumes). The filtered supernatant was loaded, and the column washed with binding buffer until baseline absorbance was restored. Wash buffer consisting of 60mM imidazole, 0.5M NaCl and 20mM Tris pH 7.9 was added until a contaminant peak had appeared and baseline was restored. Recombinant TIM was eluted with elution buffer consisting of 1M imidazole, 0.5M NaCl and 20mM Tris pH 7.9 and glycerol was added to a volume of 20% for storage at 4°C. Purified TIM was analyzed by SDS-PAGE as described above.

Amino acid analyis for quantitation purposes was performed on a Pico-Tag system (Waters, Bedford) according to the manufacturer's instructions. Mass spectrometry was performed on a DE-PRO MALDI (matrix assisted laser desorption ionisation mass spectrometer, Perseptive Biosystems, Framingham). The assay used is based on the conversion of glyceraldehyde-3-phosphate (GAP) to dihydroxyacetone phosphate (DHAP) coupled to α-glycerolphosphate-dehydrogenase, resulting in the conversion of NADH to NAD^+. Kinetic measurements were performed at 25°C on a Shimadzu UV-5000 spectrophotometer with a constant temperature attachment. GAP was purchased in the diethylacetal monobarium form (Sigma, St. Louis) and deprotected according to the manufacturer's instructions. The concentration of the L-isomer was determined in a GAP to 3-PGA conversion assay according to instructions. All assays were carried out in a total volume of 1ml. The reaction buffer consisted of 20mM triethanolamine pH 7.9 and 1mM EDTA. The GAP substrate was used at concentrations between 0.2 and 5mM. Glyceraldehyde-3-phosphate dehydrogenase (Boe, Mannheim) was added to a concentration of 10μg/ml. The reaction was started by addition of TIM, and absorbance decrease of NADH was followed at 340nm. The pH optimum was measured in the range of 6-10, and temperature stability was measured in the range of 25-70°C by preincubation in a PTC200 thermal cycler (MJ Research, Waltham). MES buffer was used in the pH 6-7 range, 20mM triethanolamine in the 7-8 pH range and Tris in the 8-10 pH range.

2.3 Results

2.3.1 Expression of dihydrofolate reductase

Expression of the synthetic malaria DHFR gene was first done in the pET17 expression system as described by Sirawaraporn *et al.* (1993). This system expresses recombinant proteins under control of the *lac* promoter, with strict control possible by co-transformation with the pLysS plasmid, leading to the expression of small amounts of T7 lysozyme which is an inhibitor of T7 RNA polymerase. After induction, expression was monitored by SDS-PAGE (Figure 2.1). The recombinant DHFR protein was expressed in the induced *E. coli* strain at concentrations of approximately 1mg/ml. The distribution between soluble and insoluble fractions was not investigated at this time, since DHFR had already been described by the original authors to accumulate in the insoluble phase.

Purification of the DHFR inclusion bodies was performed according to the method described by Sirawaraporn (1993). Crude refolded DFHR was assayed for activity above the background level of that of the host *E. coli* strain (Figure 2.2).

The negative control (water) showed no DHFR activity. The control bovine DHFR showed the expected high activity, and crude refolded DHFR showed some activity. Further kinetic analysis was not performed, as it was decided to attempt to find a system which could express DHFR in a soluble form without undergoing a refolding process.

Malaria DHFR was subsequently cloned in the pTrxFus system which is designed to express a recombinant protein as a N-terminal fusion protein with *E. coli* thioredoxin to enhance solubility (LaVallie *et al.*, 1993). PCR rescue of the synthetic DHFR gene from the pET17-DHFR vector yielded a sharp band at the correct molecular mass (not shown). After recloning, transformation of pTrxFus-DHFR in *E. coli* G1724 cells yielded 3 colonies containing the correct insert. The sequence was verified by automated nucleotide sequencing. SDS-PAGE analysis of the control containing native pTrxFus and samples containing pTrxFus-DHFR transformed bacteria could not detect any expression of recombinant DHFR at temperatures of 30°C or 37°C (results not shown).

The pET32 system was tested for the expression of recombinant malaria DHFR. Similar to pTrxFus, the pET32 system expresses recombinant proteins as a N-terminal fusion protein with thioredoxin, but

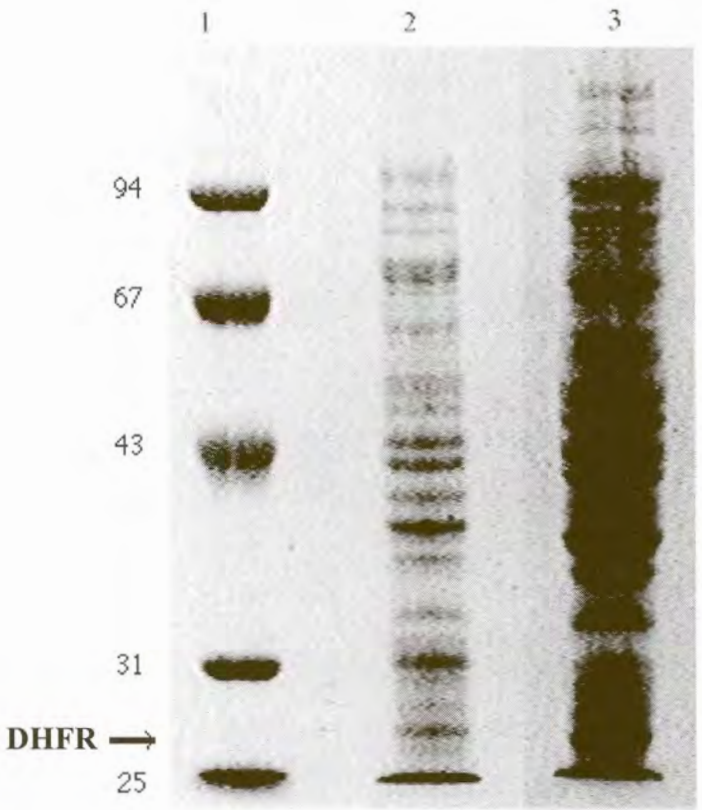


Figure 2.1: SDS-PAGE analysis of pET17-DHFR expression. Lane 1 contained molecular mass markers, lane 2 was a negative control BL21(DE3) sample and lane 3 was an extract of BL21(DE3) expressing malaria DHFR.

employs a different promoter. PCR rescue with pET17-DHFR was done as described above. Following ligation in pET32, transformation in *E. coli* BL21(DE3) cells yielded 4 colonies containing the correct insert. The sequence was verified by automated nucleotide sequencing and shown to be correct. Expression with the pET32-DHFR fusion protein at 37°C was analyzed separately for the insoluble and soluble fractions together with controls containing only pET32. A band was expected at approximately 43kDa, as the fusion proteins includes thioredoxin, two His-tags, a S-tag, a thrombin cleavage site and an enterokinase cleavage site. Very high levels of recombinant DHFR were visible at the correct molecular mass, but almost all protein was once again in the insoluble fraction (Figure 2.3).

This was repeated at a temperature of 30°C, as well as with AD494 *E.*

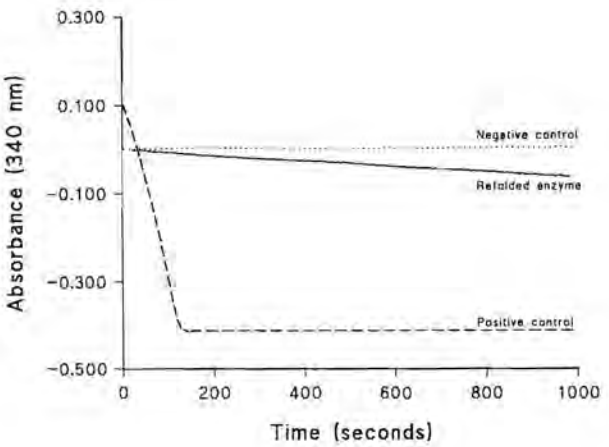


Figure 2.2: Activity assay of crude malaria DHFR. The negative control (water) is indicated by a dotted line, the positive control (bovine DHFR) by a dashed line and malaria DHFR by a solid line.

coli cells which are capable of forming disulfide bonds, but the results were similar to the above with little or no expression in the soluble phase.

2.3.2 Expression of triosephosphate isomerase

Local malaria cultures were used for RNA isolation, and selective TIM cDNA was prepared with TIM primer 2. Total RNA was used for cDNA synthesis rather than purified mRNA. When malaria total RNA is selected against poly-T oligonucleotides for enrichment of mRNA, not only the mRNA poly-A tails, but also various internal poly-A sequences from non-mRNA nucleic acids are bound. This is due to the rich AT-nature of the malaria genome. In this way the poly-T selection system is saturated, and important mRNAs may be lost. PCR with TIM cDNA as template was performed using TIM primers 1 and 2, and a sharp band of the correct size was visible at different numbers of cycles and concentrations of cDNA template (Figure 2.4). If contaminating genomic DNA had been present, it would have yielded an extra band for the intron-containing product.

The PCR product was "polished" to contain blunt ends, and was initially cloned into the pCRScript vector. The ligated construct was trans-

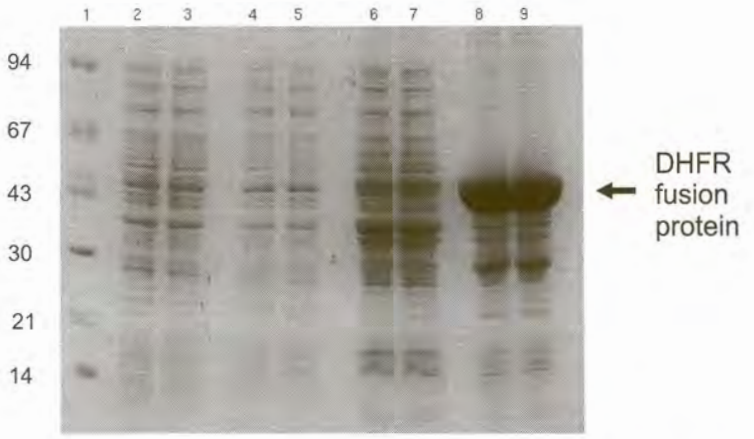


Figure 2.3: Expression of the recombinant malaria DHFR fusion protein with pET32 in BL21(DE3). Lane 1 contained molecular mass markers, lanes 2&3 both contained control soluble phases, lanes 4&5 contained soluble phases from pET32-DHFR, lanes 6&7 contained control insoluble phases and lanes 8&9 contained insoluble phases from pET32-DHFR.

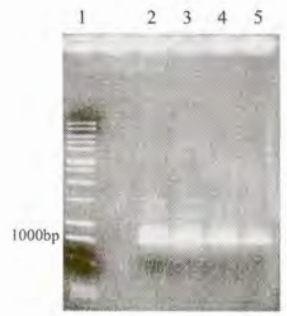


Figure 2.4: Electrophoresis of PCR producs for malaria TIM. Lane 1 contained molecular mass markers, lanes2&4 contained 1.5μl of cDNA template and lanes 3&5 contained 0.5μl of cDNA template. The PCR of samples in lanes 4&5 was performed for 35 cycles and lanes 2&3 for 25 cycles.

formed into SURE *E. coli* cells yielding 5 clones containing the correct insert. Automated sequencing showed the sequence to be identical to that reported in the literature (Ranie *et al.*, 1993) except for one silent mutation. The TIM cDNA was subsequently transferred by NdeI / BamHI restriction digestion followed by ligation to pET15b vector, yielding 2 colonies with correct inserts. Once again automated sequencing was used to verify the correct insertion and sequence of the insert. A part of the electrophoretogram for the TIM sequence is shown as an example in Figure 2.5.

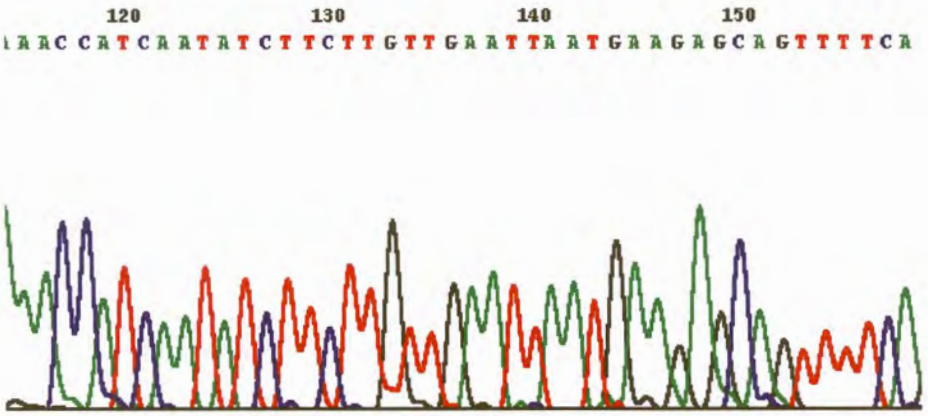


Figure 2.5: A part of the electrophoretogram for the TIM sequence.

IPTG-induced expression was monitored by SDS-PAGE at post-induction times of 0, 4, 8 and 16 hours (Figure 2.6, not all results shown). A strong band of the correct molecular mass was visible in the soluble phase with optimum expression at 8 hours post-induction. An even stronger band was also visible in the insoluble fraction indicating over-expression.

Figure 2.6: Expression of recombinant TIM at 30°C. Lane 1 contained molecular mass markers, lanes 2-7 contained soluble fractions and lanes 8-13 contained insoluble fractions. Lanes 3, 5, 7, 9, 11 and 13 contain IPTG-induced samples, and lanes 2, 4, 6, 8, 10 and 12 were uninduced. Lanes 2, 3, 8 and 9 were at 0 hours post-induction, lanes 4, 5, 10 and 11 were at 8 hours and lanes 6, 7, 12 and 13 were at 16 hours. Lane 14 contained IMAC-purified recombinant malaria TIM.

As the pET15b vector expresses the recombinant protein as a fusion

protein with a N-terminal 6 histidine tag, recombinant *P. falciparum* TIM was purified by His-tag affinity chromatography on a Ni^{2+} chelated column. The flowthrough from the column was constantly monitored at 280nm (Figure 2.7).

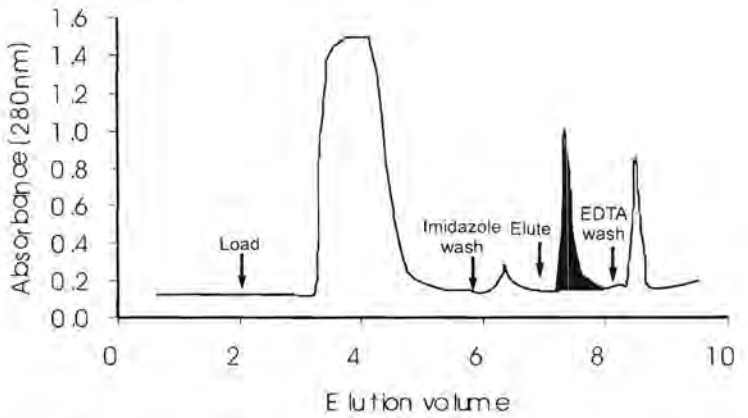


Figure 2.7: His-tag purification of recombinant TIM. The eluted recombinant protein is indicated by black shading.

Following the flowthrough of unbound *E. coli* proteins, a relatively small peak was eluted during the 60mM imidazole wash step which is optimised to elute proteins binding non-specifically to the column, possibly due to some poly-histidine residues. A large, sharp TIM peak was eluted during the 1M imidazole elution step and was collected in a volume of approximately 2ml. The column was then washed with EDTA, stripping the Ni^{2+} and all remaining bound material from the column. Subsequent SDS-PAGE analysis of the eluted TIM showed a sharp band at the expected molecular mass, with two very faint bands (not clearly visible) at a lower molecular mass (Figure 2.6). The two faint bands may either be contaminating proteins or TIM breakdown products. The oligo-histidine motif was not removed, as modeling of the TIM dimer indicated that it would probably not interfere with dimerisation or activity (Figure 2.8).

TIM concentration was determined by amino acid analysis, yielding approximately 3mg of purified TIM per 100ml of culture. The amino acid composition for all amino acids could not be accurately determined due to the background caused by the imidazole elution step. MALDI analysis of TIM was performed to determine the exact mass of the translated product, and showed a major peak of 30,258Da which corresponds to the sum of the denatured m/z TIM monomer peak plus the 1/2m/z

Figure 2.8: A model of malaria TIM with the subunits indicated in red and blue, the active sites in CPK-coloured spheres and the oligo-histidine tag in green.

peak of the remaining native TIM dimer. Some dimer is still visible at m/z=60.606Da (Figure 2.9).

Kinetic analysis was performed in triplicate on the recombinant TIM to determine its K_m and V_{max}, for comparison to that of TIM from other species. These values were used in an inverse reciprocal plot to calculate the kinetic parameters (Figure 2.10).

The K_m was determined as 0.586mM and V_{max} as 0.027μmole/min, which compares well to the values from other species (Table 2.2). From this data, K_{cat} was calculated as $1.05 \times 10^5 min^{-1}$. A specific activity of 3913U/mg was calculated which is similar to those found for other isolates of *P. falciparum* TIM (Gopal *et al.*, 1999). The pH optimum for recombinant malaria TIM (Figure 2.11) was shown to be approximately 8.5 which corresponds to those determined for other species.

The temperature stability was tested by preincubation at temperatures

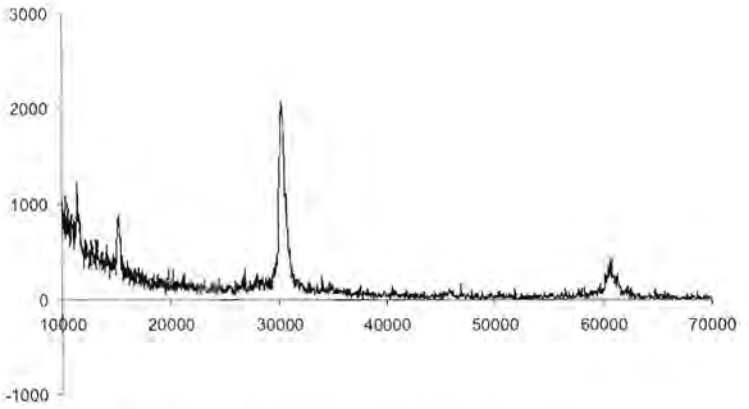


Figure 2.9: MALDI analysis of purified recombinant TIM. The major peak of 30,258Da corresponds to the sum of the m/z monomer peak plus the 1/2m/z peak of the remaining dimer. Some dimer is still visible at m/z=60,606Da.

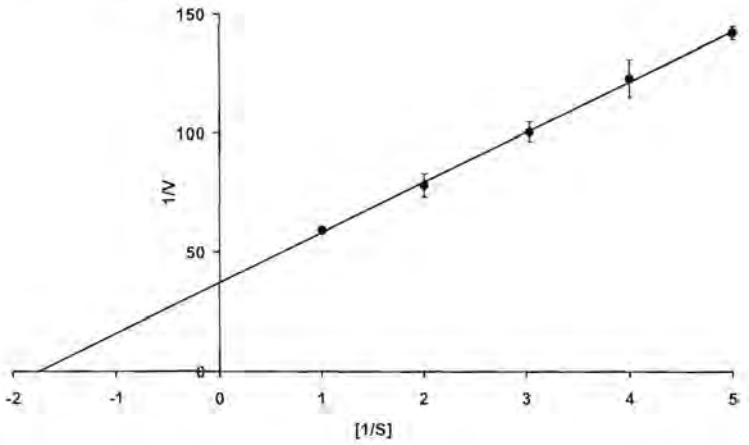


Figure 2.10: Inverse reciprocal plot for recombinant malaria TIM. K_m was determined as 0.586mM and V_{max} as 0.027μmole/min from assays performed in triplicate.

Table 2.2: Properties of TIM from various species.

Source	K_m (mM)	$K_{cat}(min^{-1})$	Reference
P. falciparum	0.586	$1.05x10^5$	
T. brucei	1.2	$6.5x10^4$	Lambeir *et al.* (1987)
Chicken muscle	0.97	$2.59x10^4$	Putman *et al.* (1972)
	1.57	$2.92x10^4$	Plaut and Knowles (1972)
Rabbit Muscle	0.39		Krietsch *et al.* (1970)
	0.32	$5.2x10^4$	Hartman *et al.* (1975)
Yeast	1.22		
	1.45		Hartman and Ratrie (1977)
	1.27	$4.9x10^4$	Krietsch *et al.* (1970)

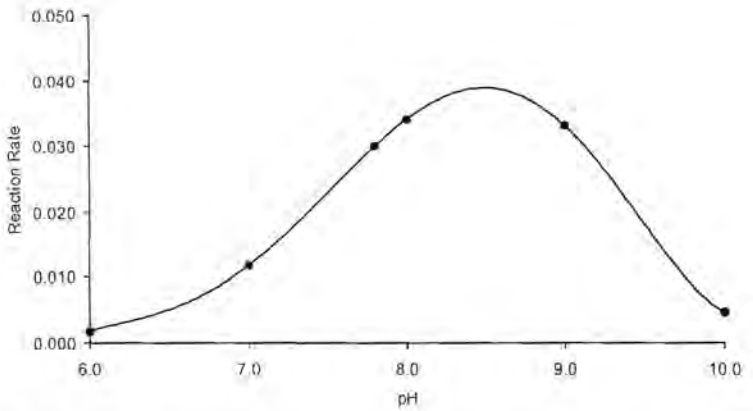


Figure 2.11: A pH optimum plot for recombinant TIM. The optimum was determined as approximately 8.5.

varying from 30 to 70°C. TIM was stable up to 55°C after which a sudden loss of activity occurred (Figure 2.12).

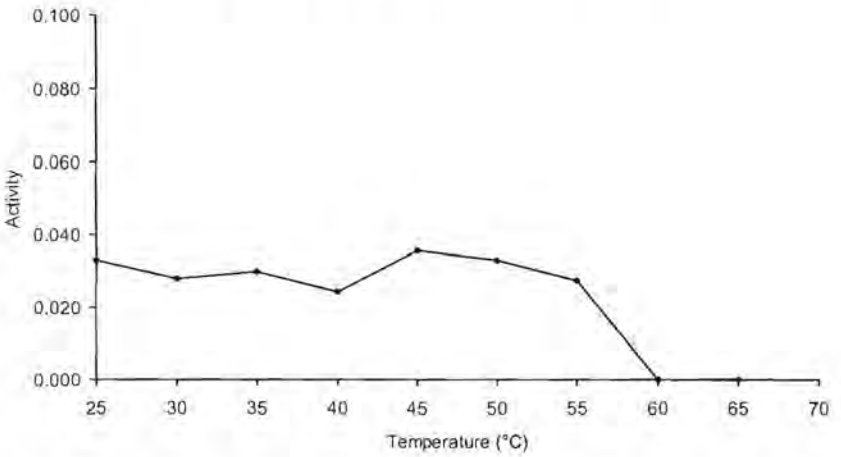


Figure 2.12: Temperature stability plot for recombinant TIM. The enzyme activity was stable to a temperature of 55°C, after which a sudden decrease in activity occurred.

2.4 Discussion

Expression of the synthetic DHFR gene in the pET17 vector showed an accumulation of DHFR in inclusion bodies as described in the literature (Sirawaraporn *et al.*, 1993). Overexpressed proteins often accumulate in the cytoplasm in the form of inclusion bodies that may make up a major fraction of the total cellular protein. These inclusion bodies may be separated from the rest of the cellular components due to their high density, and dissociated by detergents to extract high comcentrations of a recombinant protein. However, the extracted proteins are not always correctly folded and may need denaturation and refolding. In this case, the inclusion bodies could be refolded to give approximately 20% active recombinant DHFR. Unfortunately, these type of refolding procedures are relatively variable and do not lend themselves well to the expression of proteins for complex manipulations. If mutations were to be made, comparison between native and mutated recombinant proteins would be difficult, as measured differences could not be conclusively ascribed to either the mutation or to the refolding procedure. The ideal would be to obtain large quantities of expressed active recombinant protein in soluble form that could be easily purified by employing some sort of a tag.

The pTrxFus system was initially chosen due to its ability to express recombinant proteins as N-terminal fusion proteins with *E. coli* thioredoxin. In this way the solubility of many recombinant proteins have been enhanced to give rise to a correctly folded active soluble protein (LaVallie *et al.*, 1993)(Yasukawa *et al.*, 1995)(Lunn and Pigiet, 1982). The pTrxFus system provides an enterokinase cleavage site between thioredoxin and the recombinant protein, so that the thioredoxin fusion peptide may be removed if necessary (Baratti *et al.*, 1973). pTrxFus employs a tryptophan-inducible bacteriophage λ major leftward promoter. Unfortunately when expression was induced in this system, no recombinant protein was visible at temperatures of 30°C or 37°C.

As expression had taken place under a T7 promoter in the pET17 system, a new expression vector was chosen. This was pET32 which contained a strong IPTG inducible T7lac promoter, and expressed recombinant proteins as N-terminal fusion proteins with thioredoxin and S-tag (a ribonuclease-S binding peptide) (Kim and Raines, 1993). The T7lac promoter carries a lac sequence just downstream from the T7 promoter, as well as the natural promoter and coding sequence for the lac re-

pressor (lacI), oriented such that the T7lac and lacI promoters diverge (Studier *et al.*, 1990). The lac repressor acts both at the lacUV5 promoter in the host chromosome to repress transcription of the T7 RNA polymerase and at the T7lac promoter in the vector to block transcription of the target gene by any synthesized T7 RNA polymerase. Thus strict control may be exercised over the expression of a recombinant protein. Even with the presence of thioredoxin, recombinant DHFR was expressed in the insoluble phase, although at very high quantities. The insolubility may be a combination of incorrect folding and overexpression of the protein.

In this study no expression system was able to express active recombinant malaria DHFR in a soluble form. However, in some recent publications alternative systems have been described. Hekmat-Nejad *et al.* (1997) described the expression of the native malaria gene for DHFR in the pET23d system employing a T7 promoter. Although the protein had to be refolded to regain activity, it was shown that the native malaria gene could be expressed at high quantities in *E. coli* if a T7 promoter was used. Expression of native DHFR-TS was attempted by Sirawaraporn et al. (Sirawaraporn *et al.*, 1990), but low quantities of expression led them to the use of a synthetic DHFR gene. Another publication from the Sirawaraporn group described the chemical synthesis of the complete DHFR-TS gene (Prapunwattana *et al.*, 1996). In this study the complete bifunctional DHFR-TS enzyme was expressed in *E. coli* and a correctly folded fully active enzyme was produced. From these results it thus seems that malaria DHFR is not able to fold correctly in *E. coli* in the absence of the junction and TS chains. The optimal solution seems to be the expression of the complete native DHFR-TS gene in *E. coli* under control of a T7-based promoter. The possibility of expressing proteins such as DHFR in yeast complementation systems was also demonstrated by Wooden *et al.* (1997) where yeast expressing constructs with malaria DHFR alleles were screened in terms of pyrimethamine resistance.

P. falciparum TIM was expressed in pET15b under the control of aT7lac promoter. This system expresses the recombinant protein as a fusion with a N-terminal motif containing 6 histidines followed by a thrombin cleavage site. TIM was expressed in quantities of approximately 5mg per 100ml of culture, with a 1:4 distribution between the soluble and insoluble phase. This may be due to weak accommodation of overexpressed TIM in the *E. coli* cytoplasm. Recombinant TIM was success-

fully purified using a Ni^{2+}chelating column, and yielded approximately 3mg of purified TIM per 100ml culture. SDS-PAGE analysis showed two other barely visible bands in the sample which may be contaminating proteins binding to the column, truncated proteins or TIM breakdown products. The recombinant protein was approximately 95% pure (densitometry) and was deemed suitable for kinetic and inhibition studies. Kinetic measurements were performed with DHAP as substrate in an NADH-linked spectrophotometric assay. The K_m was determined as 0.586mM and V_{max} as 0.027μmole/min, which compares well to those of other species. The increased K_{cat} of $1.05x10^5$ corresponds to the necessary increase in glycolytic activity that occurs upon infection of red blood cells with malaria parasites. A specific activity of 3913U/mg was calculated. TIM had a pH optimum of approximately 8.5 and remained temperature stable up to 55°C. The recombinant TIM was suitable for further inhibitory and mutational studies.

Chapter 3

Homology modelling of putative malaria drug target proteins

3.1 Introduction

For the purposes of inhibitor design, structural models were needed for the DHFR and TIM enzymes. DHFR from *P. falciparum* has as yet not been crystallised, and TIM was crystallised by Velanker *et al.* (1997) only during the course of this study. Thus, other methods for the preparation of the enzyme structural models had to be attempted. Since the first protein structure was solved, people have been speculating about the possibility of predicting protein structures from the amino acid sequences. Homology modelling has become extremely popular together with the significant increase in the amount of available sequences in nucleotide and protein databases. The crystallographic databases have not grown at the same tempo, but many new sequences are similar to existing proteins in the Brookhaven Protein Databank (PDB) and often the structures of new proteins may be predicted in this way. Although there are approximately 500,000 proteins in GenPept, only about 10,000 experimentally determined structures are available in PDB. Of the known protein sequences, 20-50% have segments that are related to one or more known structures. Approximately 900 folds have been defined out of an estimated few thousand. Projects are be-

ing designed to perform homology modelling on genomes from different organisms (Sanchez and Sali, 1998).

Homology modelling comprises of four basic steps (Sali and Blundell, 1993): the first step is the identification of proteins with known 3D structures that are related to the target sequence. The second step is to prepare optimal sequence alignments of the target with the known structure's sequence. The third step is to build a model for the target sequence given its alignment with the template structures. The fourth step is model evaluation. This cycle may then be repeated as necessary. Various systems are currently available and differ mainly in the way the 3D model is calculated. The first group of methods is based on rigid body assembly (Johnson *et al.*, 1994). A model is constructed from core regions, and loops and side-chains are added from related structures. This assembly involves fitting the rigid bodies on the framework. Another family of methods are based on segment matching (Levitt, 1992). Positions of conserved atoms from the template are used to calculate the coordinates of other atoms. This method employs a database of short segments of protein structure, as well as energy or geometric rules. The third group of methods is based on the satisfaction of spatial restraints (Sali and Blundell, 1993). Restraints are calculated from the alignment of the target protein with the template proteins, and satisfaction of these restraints are attempted.

Our study used homology modelling by satisfaction of spatial restraints, specifically the MODELLER 4 package (Sali and Blundell, 1993). This method can use many different types of information about the target sequence making it probably the most effective of the current approaches. All restraint-based methods determine distance and dihedral angle restraints to be applied on the target structure from the alignment with the template structures and stereochemical restraints and then calculate the model by minimising the violations of these restraints. In MODELLER, many distances and dihedral angle restraints are derived, followed by spatial restraints obtained from the statistical analysis of the relationships between various features of protein structure. Databases of family alignment are used to obtain tables quantifying relationships such as $C\alpha$ - $C\alpha$ distances or dihedral angles. The relationships are expressed as conditional probability density distributions and are used directly as spatial restraints. Spatial restraints are thus obtained empirically from a database and not determined theoretically. Next, the homology-derived restraints and energy terms regard-

ing stereochemistry are combined into an objective function. The model is finalised by optimising the objective function in Cartesian space. In the final iterations, molecular dynamics with simulated annealing is used for model refinement. In this way models with a R.M.S lower than 2Å have been obtained. When the target sequence is at least 40% identical to one or more of the templates, a main-chain Root Mean Square (R.M.S.) error as low as 1Å can be achieved for 90% of the residues. When sequence identity is between 30% and 40%, main-chain R.M.S. error rises to approximately 1.5Å for approximately 80% of the residues. Insertions of larger than 8 residues cannot be modelled accurately. Below 40% sequence identity, it can be expected that about 20% of the residues will be misaligned, but it is possible to edit alignments to prevent large errors. The following errors are commonly found in structures prepared by MODELLER: distortion or shifts in correctly aligned regions (loops, helices, strands), side-chain packing errors, distortion or shifts of regions without template equivalents and distortion or shifts in incorrectly aligned regions (Sanchez and Sali, 1997).

Even though comparative modelling is less accurate than high-resolution experimental techniques, it can be helpful in proposing and testing hypotheses. In some cases, high quality homology models have been used for the design of highly efficient inhibitors. Sudbeck *et al.* (1999) used homology models of JAK3 kinase to design specific high affinity inhibitors for possible cancer trials. Homology models of different cytochrome P450 structures have been utilised along with site-directed mutagenesis to elucidate the molecular determinants of substrate specificity (Szklarz and Halpert, 1997). Already in 1991, *Schistosoma mansoni* serine protease homology models were used to elucidate the function of the enzyme as well as design specific peptide inhibitors (Cohen *et al.*, 1991). Homology modelling can also be used to help guide X-ray and NMR refinement processes.

Some malaria protein structures have been prepared by homology modelling. Lactate dehydrogenase (LDH) was modelled by Hewitt *et al.* (1997). The homology structure was tested by protein engineering and found to be of adequate accuracy. They were able to suggest explanations of the unusual properties of malaria LDH compared with all other LDHs. The malaria cysteine proteinase Falcipain was homology modelled by Li *et al.* (1996), followed by the identification of a series of inhibitors for this protein.

In this study either malaria DHFR or TIM was used as target. All

DHFRs or TIMs with available X-ray structures were then used to obtain optimal alignments with the target. Selected templates were then used for the modelling process. The following goals were set:

- Preparation of an optimal alignment for DHFR
- Homology modelling of DHFR
- Preparation of an optimal alignment for TIM
- Homology modelling of TIM

3.2 Methods

3.2.1 Homology modelling of dihydrofolate reductase and triosephosphate isomerase

All entries for dihydrofolate reductases or triosephosphate isomerases were retrieved from the Brookhaven Protein Data Bank (Bernstein *et al.*, 1977). After filtering for duplicates varying only by point mutation or bound ligand, alignments together with the *P. falciparum* DHFR or TIM sequences were prepared with the ClustalW package (Thompson *et al.*, 1994), and optimised by visual inspection. Large insertions were removed as necessary by visual inspection.

Alignments were converted to modified PIR format for input in the MODELLER 4 package (Sali and Blundell, 1993). The crystal structures from *E. coli* (1DRA) and *L. casei* (1AO8) were used as templates for model building of DHFR, while all available species were used in the case of TIM. The MODELLER routine for full model building was used on a Silicon Graphics Power Indigo2 Extreme workstation (Silicon Graphics, Mountainview).

The MODELLER log file was inspected for model quality information as compared to empirical guidelines determined from known proteins. Further quality assessments were done with the Procheck package (Laskowski *et al.*, 1996), yielding data such as total Ramachandran plots, individual Ramachandran plots, summaries of bond lengths, angles, etc. The homology structures for *P. falciparum* DHFR or TIM were superimposed with that of the templates and deviations calculated using the ProFit package (University College London).

3.3 Results

3.3.1 Homology modelling of dihydrofolate reductase

Amino acid sequence alignments were prepared using the ClustalW package. An initial alignment was done with DHFR sequences from all species in the Brookhaven Database (Figure 3.1).

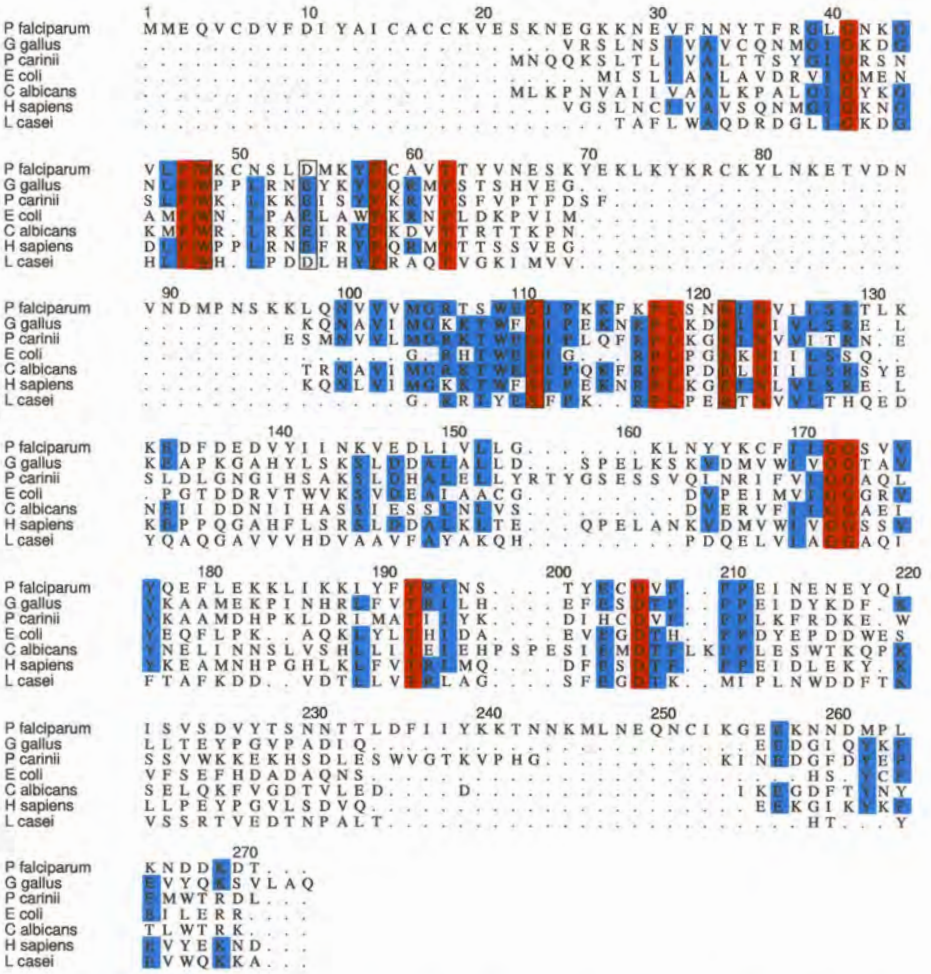


Figure 3.1: Amino acid sequence alignment of malaria DHFR with other species DHFRs from the Brookhaven Protein Data Bank. Homology is indicated in blue and identity in red. Conserved active site residues are boxed.

Malaria DHFR contained a N-terminal segment of 22 amino acids that could not be aligned with any of the other species. Two other large insertions were also present, a 27 residue segment at residue 70-97

and a 38 residue segment at the C-terminal. *P. carinii* DHFR was the only protein also containing an insertion near the C-terminal with another near the middle of the protein. The catalytic residues (*Trp48*, *Glu/Asp54*, *Phe58*, *Ser113*, *Arg122*) were conserved or semi-conserved in all species.

The *E. coli* (1DRA) and *L. casei* (1AO8) DHFR structures were selected as templates being the most homologous to the malaria DHFR. The C-terminal insertion was removed and alignments were reoptimised (Figure 3.2).

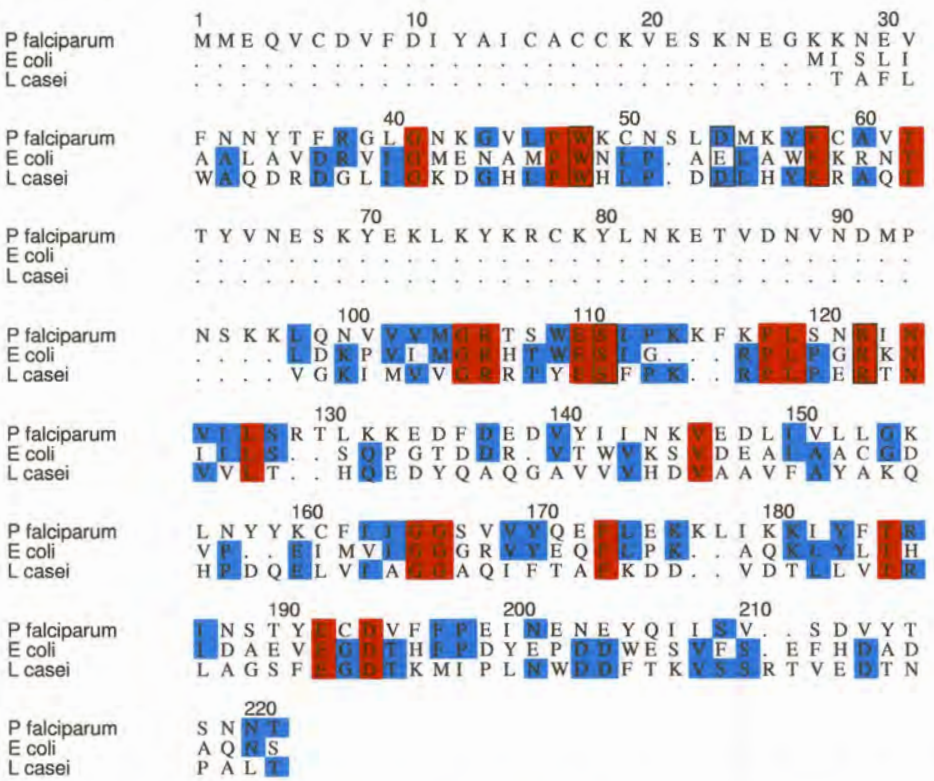


Figure 3.2: Sequence alignment of modified malaria DHFR with *E. coli* and *L. casei* DHFR. Homology is indicated in blue and identity in red. Important catalytic residues are indicated by boxes.

The alignment showed the catalytically important residues to be interrupted by a large insertion of 27 residues. For practical reasons, this insertion together with the N-terminal 22 residues were excised and the alignment re-optimised (Figure 3.3).

The re-optimised alignment was used as input to the MODELLER package. After model building, the output structure was analyzed with the

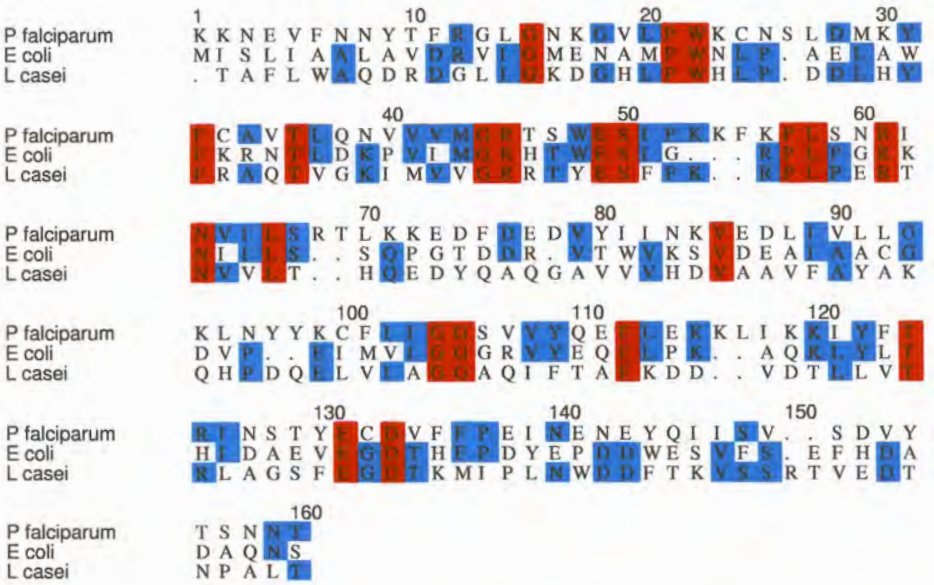


Figure 3.3: Modified sequence alignment of truncated malaria DHFR with *E. coli* and *L. casei* DHFR after removal of large insertions from the malaria sequence. Homology is indicated in blue and identity in red.

Procheck package (not all results shown).

The Ramachandran (Figure 3.4) plot shows 75.2% of residues to be in favoured regions (red), and 16.6% to be in allowed regions (yellow). Of the remaining residues 3.8% were in generously allowed regions (beige) and 4.5% in disallowed regions (white). Based on structures with a resolution of at least 2.0Å, a model of very high quality should show approximately 90% of residues in the favoured regions (Laskowski *et al.*, 1996). None of the catalytic residues were found in problematic regions.

Further analysis of values from the Ramachandran plot showed the main-chain residues to be generally of good quality (Figure 3.5). A bad contact score of 8.8 /100 residues was calculated and an overall G-factor of -0.3. The G-factor is essentially a log-odds score based on the observed distributions of stereochemical parameters, providing a measure of how "normal", or alternatively how "unusual", a given stereochemical property is (Engh and Huber, 1991). When applied to a given residue, a low G-factor indicates that the property corresponds to a low-probability conformation. So, for example, residues falling in the disallowed regions of the Ramachandran plot will have a low (or very negative) G-factor. Similarly for unfavourable chi1-chi2 and chi1

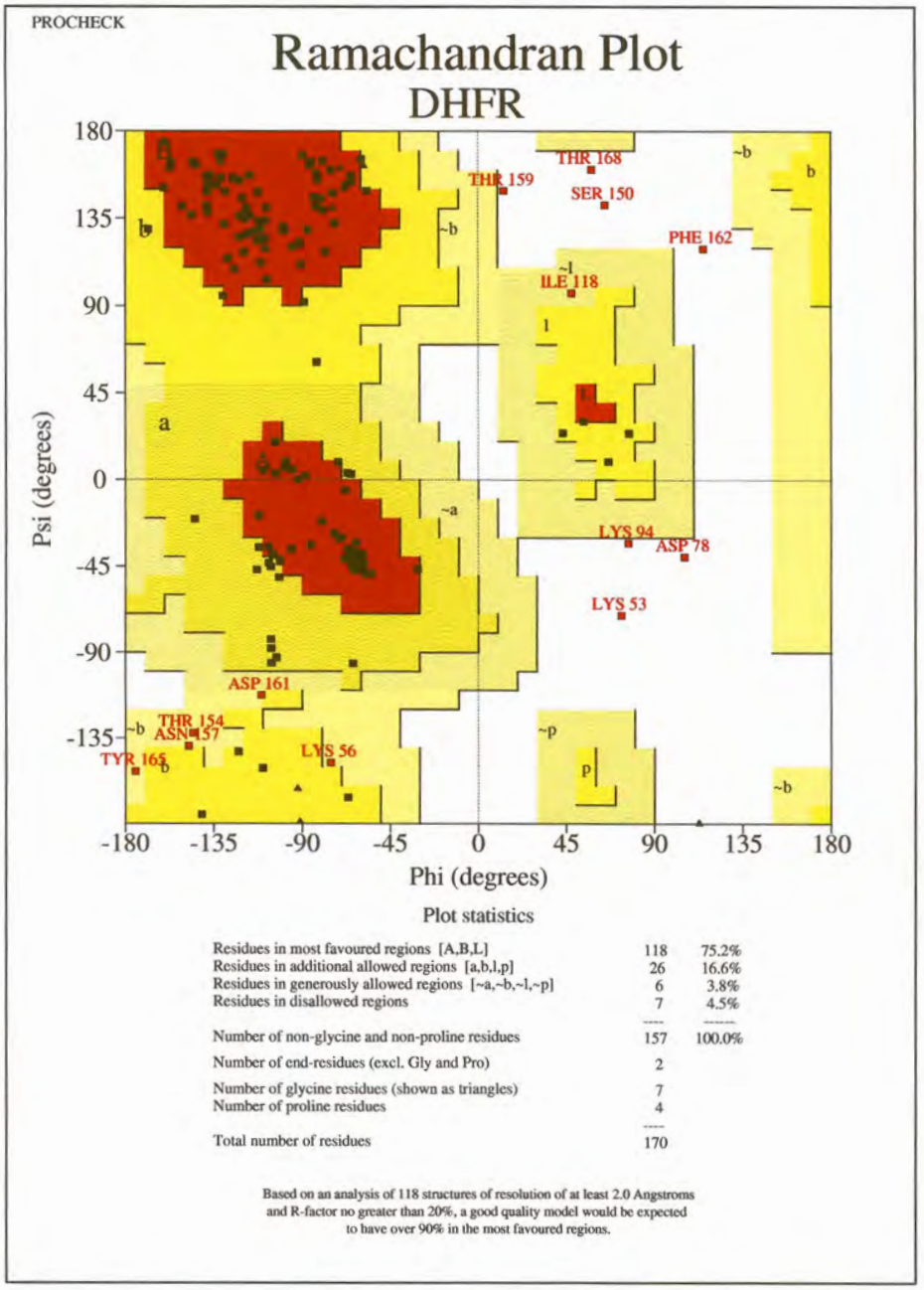
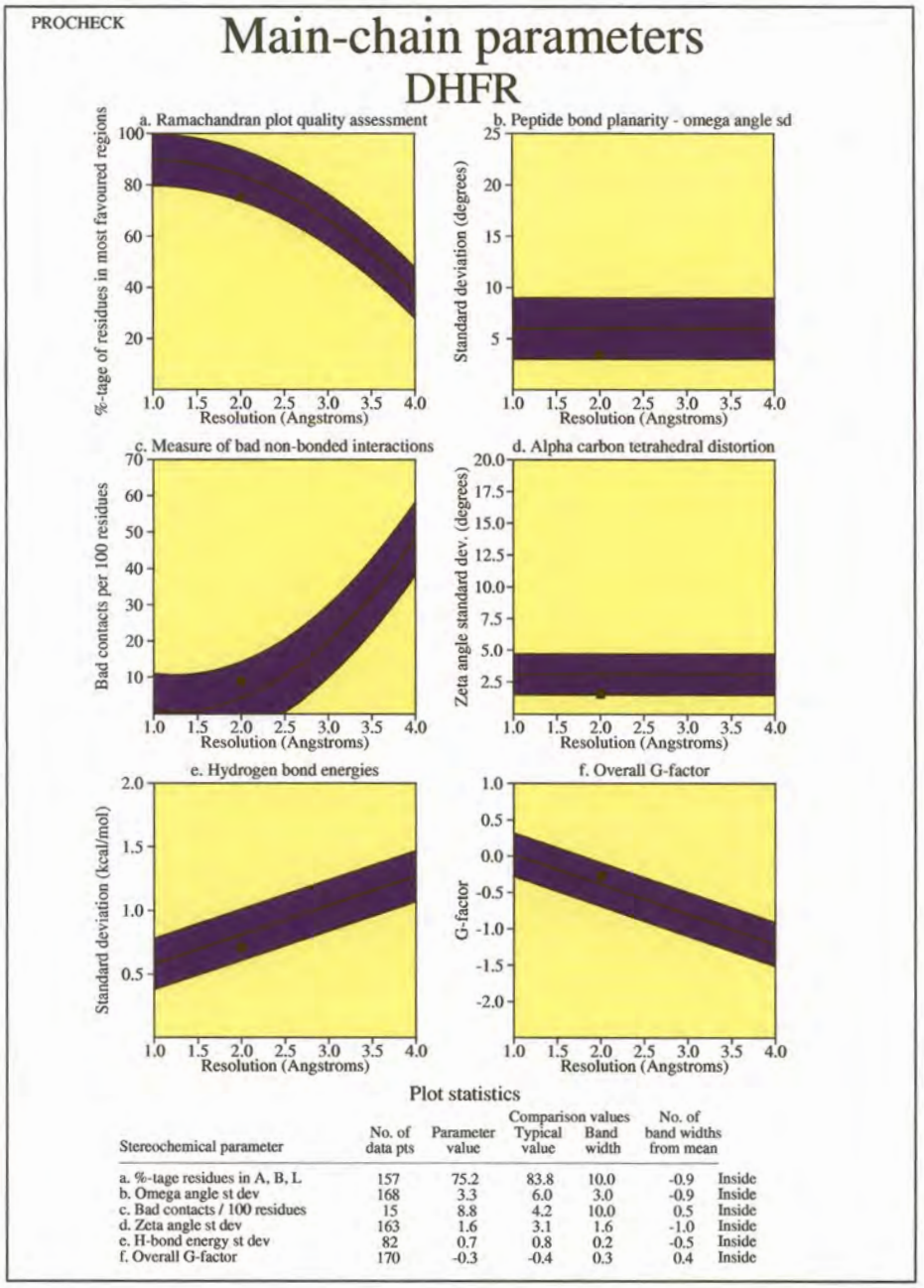


Figure 3.4: Ramachandran plot for the homology model of malaria DHFR.

values.

Side-chain parameters were several deviations away from expected means, and were generally of lesser quality (Figure 3.6). Especially large side-chain chi deviations were calculated.

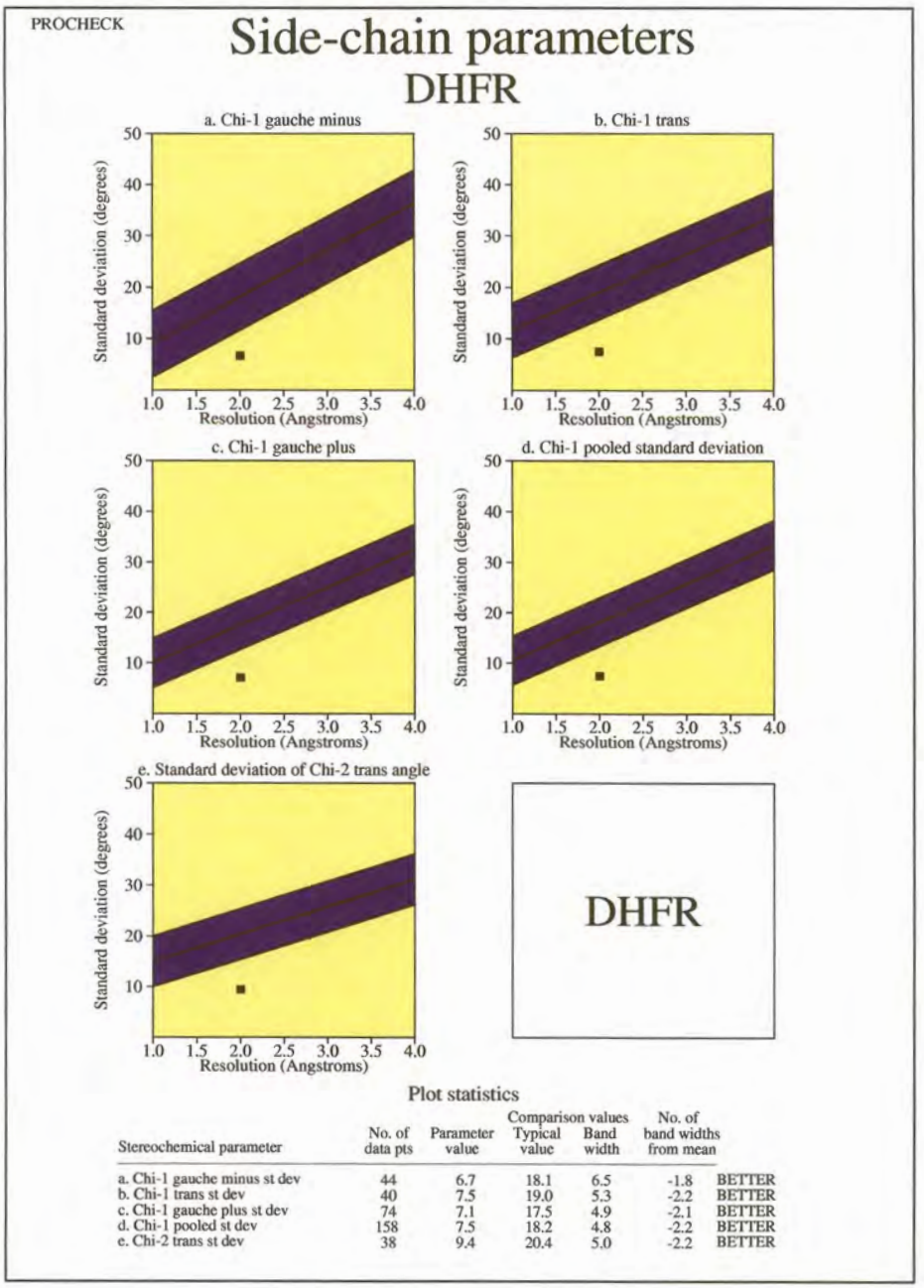


Stereochemical parameter	No. of data pts	Parameter value	Comparison values Typical value	Band width	No. of band widths from mean	
a. %-tage residues in A, B, L	157	75.2	83.8	10.0	-0.9	Inside
b. Omega angle st dev	168	3.3	6.0	3.0	-0.9	Inside
c. Bad contacts / 100 residues	15	8.8	4.2	10.0	0.5	Inside
d. Zeta angle st dev	163	1.6	3.1	1.6	-1.0	Inside
e. H-bond energy st dev	82	0.7	0.8	0.2	-0.5	Inside
f. Overall G-factor	170	-0.3	-0.4	0.3	0.4	Inside

Figure 3.5: Main chain parameter analysis for homology modelled malaria DHFR.

A three-dimensional diagram was prepared indicating the quality values for the model by different colours (Figure 3.7).

A relatively high number of red regions are visible, especially away from the core of the protein and in loop regions. The homology structure was



Stereochemical parameter	No. of data pts	Parameter value	Comparison values Typical value	Band width	No. of band widths from mean	
a. Chi-1 gauche minus st dev	44	6.7	18.1	6.5	-1.8	BETTER
b. Chi-1 trans st dev	40	7.5	19.0	5.3	-2.2	BETTER
c. Chi-1 gauche plus st dev	74	7.1	17.5	4.9	-2.1	BETTER
d. Chi-1 pooled st dev	158	7.5	18.2	4.8	-2.2	BETTER
e. Chi-2 trans st dev	38	9.4	20.4	5.0	-2.2	BETTER

Figure 3.6: Side-chain parameter analysis for homology modelled malaria DHFR.

superimposed with the structure for human DHFR, and a Cα R.M.S. deviation of 3.6Å was calculated (Figure 3.8).

Active site residues were superimposed (Figure 3.9). Catalytic residues are reasonably similar with the major change being *Glu28Asp* and

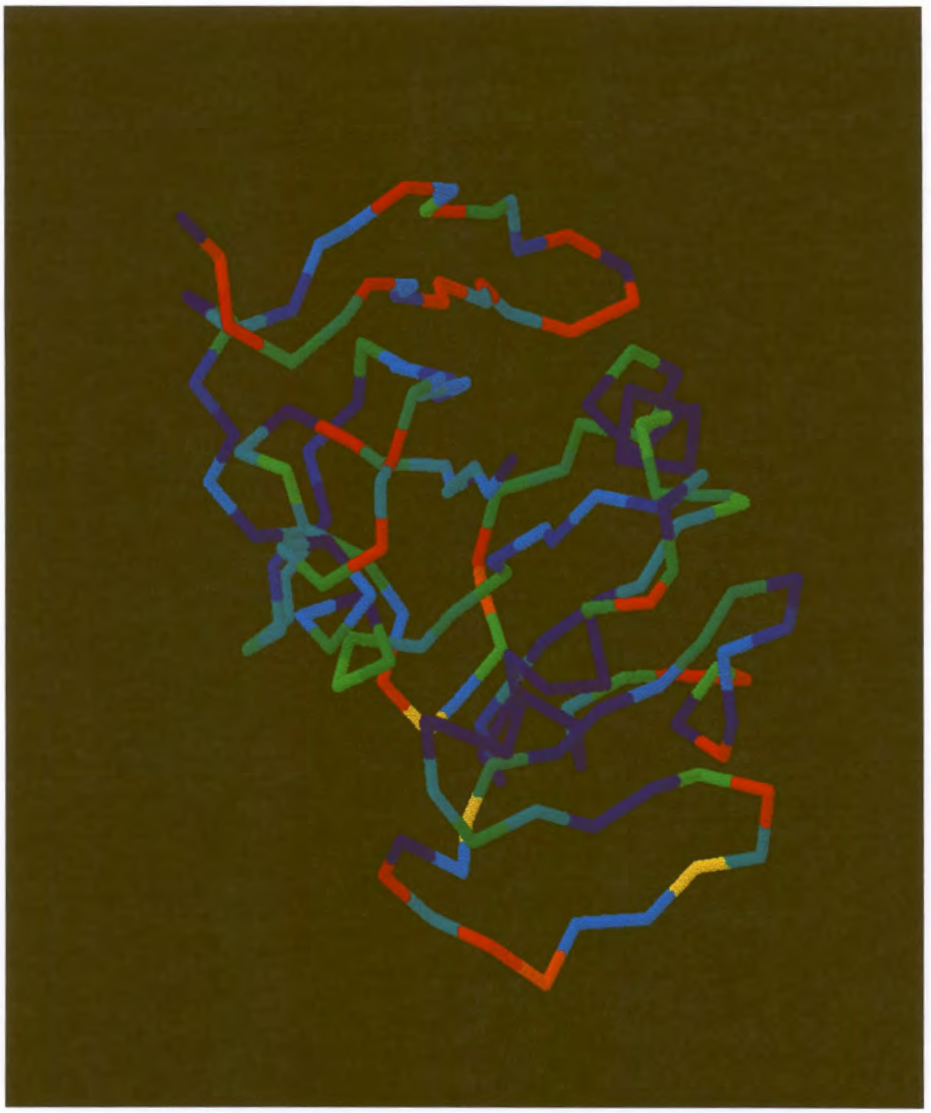


Figure 3.7: Quality score indications of homology-modelled DHFR. Red regions indicate lower quality, green intermediate and blue regions indicate higher quality.

twisted conformations of some side-chains.

The main-chain parameters indicated the correct fold of the malaria DHFR backbone, excepting the putative loop regions excised from the alignments. Side-chain parameters however showed values of inferior quality generating low confidence in the side-chain conformations of many amino acids.

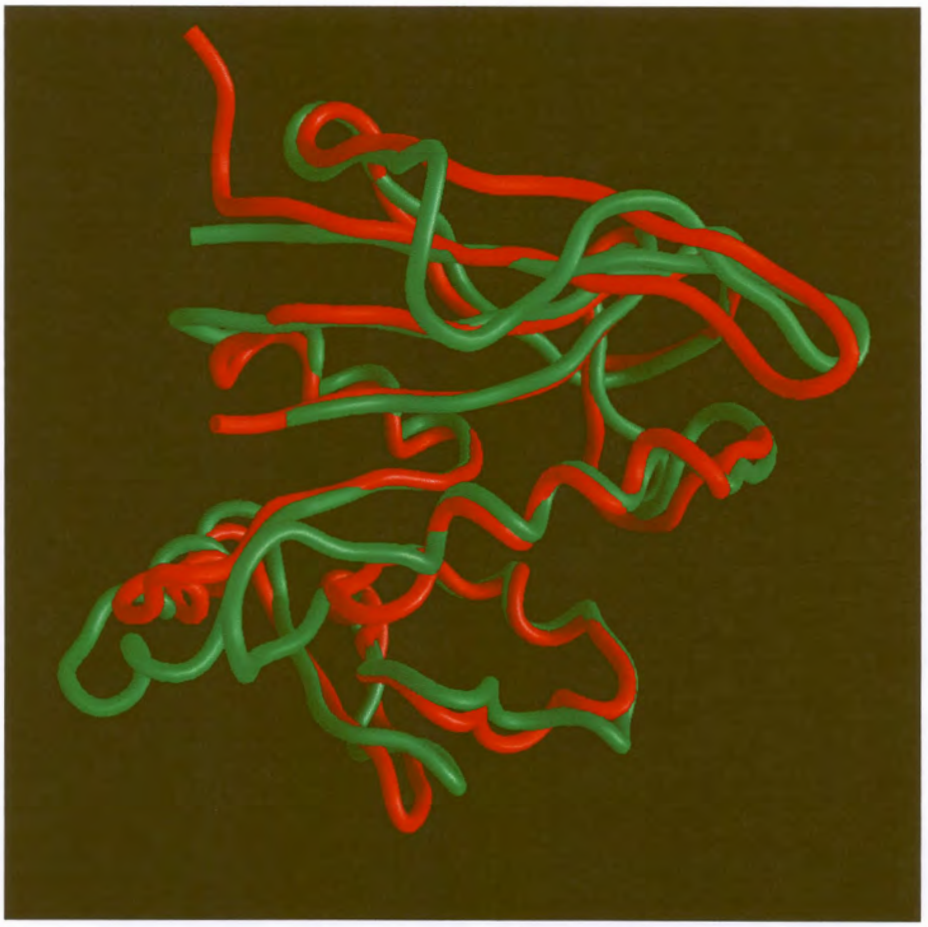


Figure 3.8: Fitted structures of modelled malaria DHFR (red) and human DHFR (green).

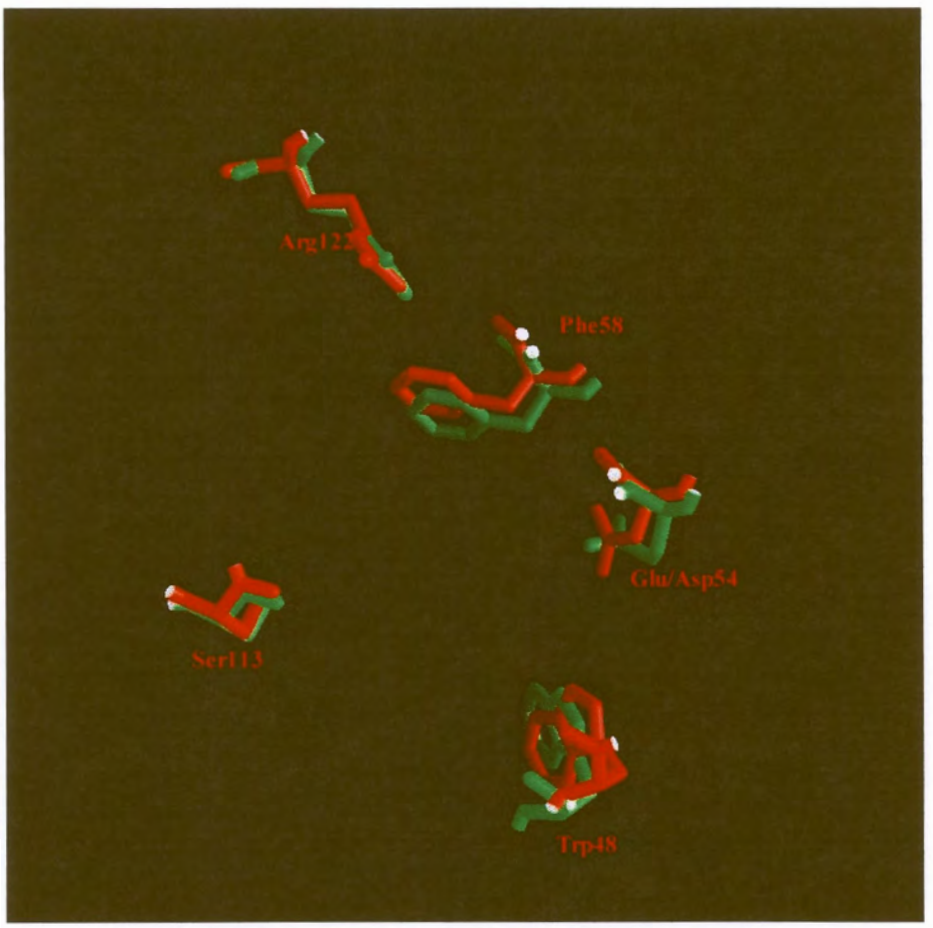


Figure 3.9: Superimposed active site residues of malaria DHFR (red) and human DHFR (green).

3.3.2 Homology modelling of triosephosphate isomerase

The initial amino acid sequence alignment of malaria TIM with those of other species in the Brookhaven Protein Database is shown in Figure 3.10.

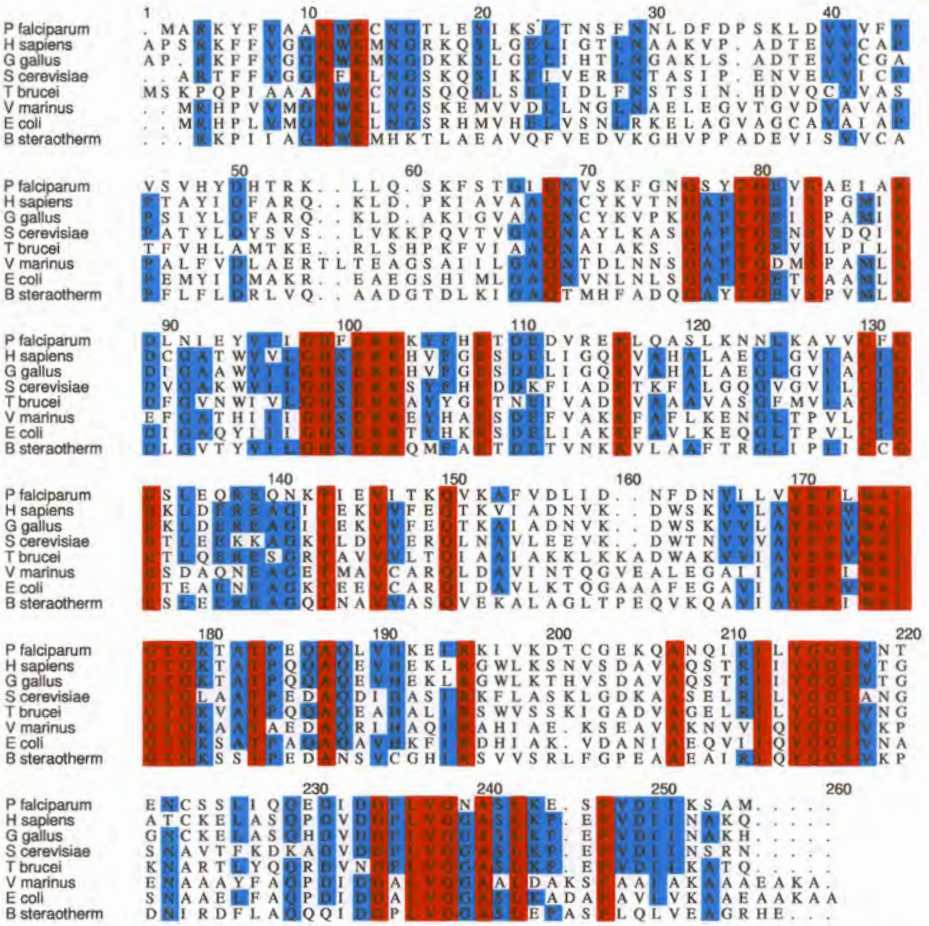


Figure 3.10: Sequence alignments of all TIM structures available in PDB. Homology is indicated in blue and identity in red.

In this alignment, TIM was found to be conserved to a relatively high degree in different species. No large insertions or deletions were observed as was the case with DHFR.

Because of the conserved nature of TIMs, the structure for all TIMs from PDB were used for model building. After model building, the structure was analyzed with the Procheck package (Laskowski *et al.*, 1996). Some results are summarised in Figure 3.11 (results from all

plots not shown).

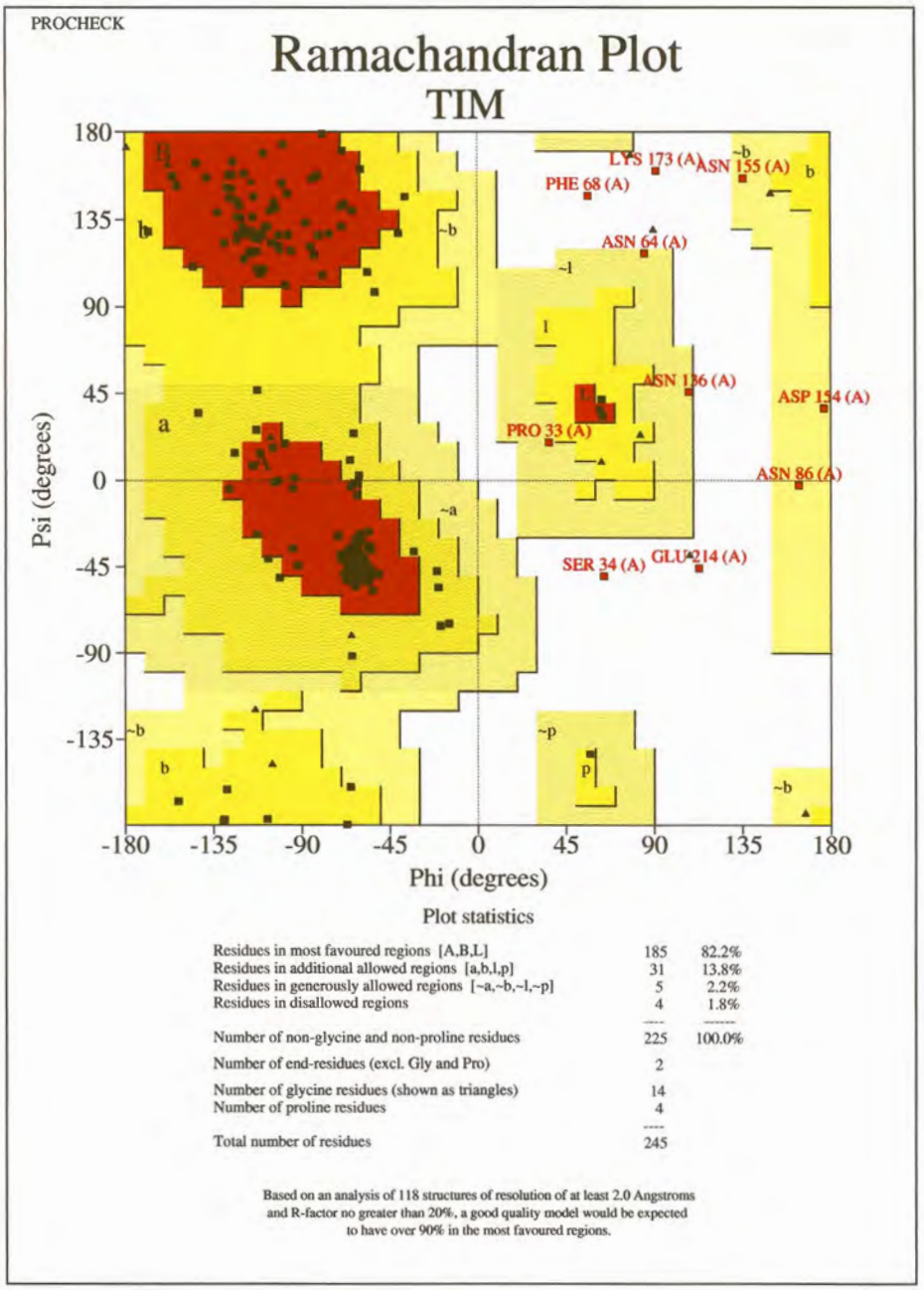
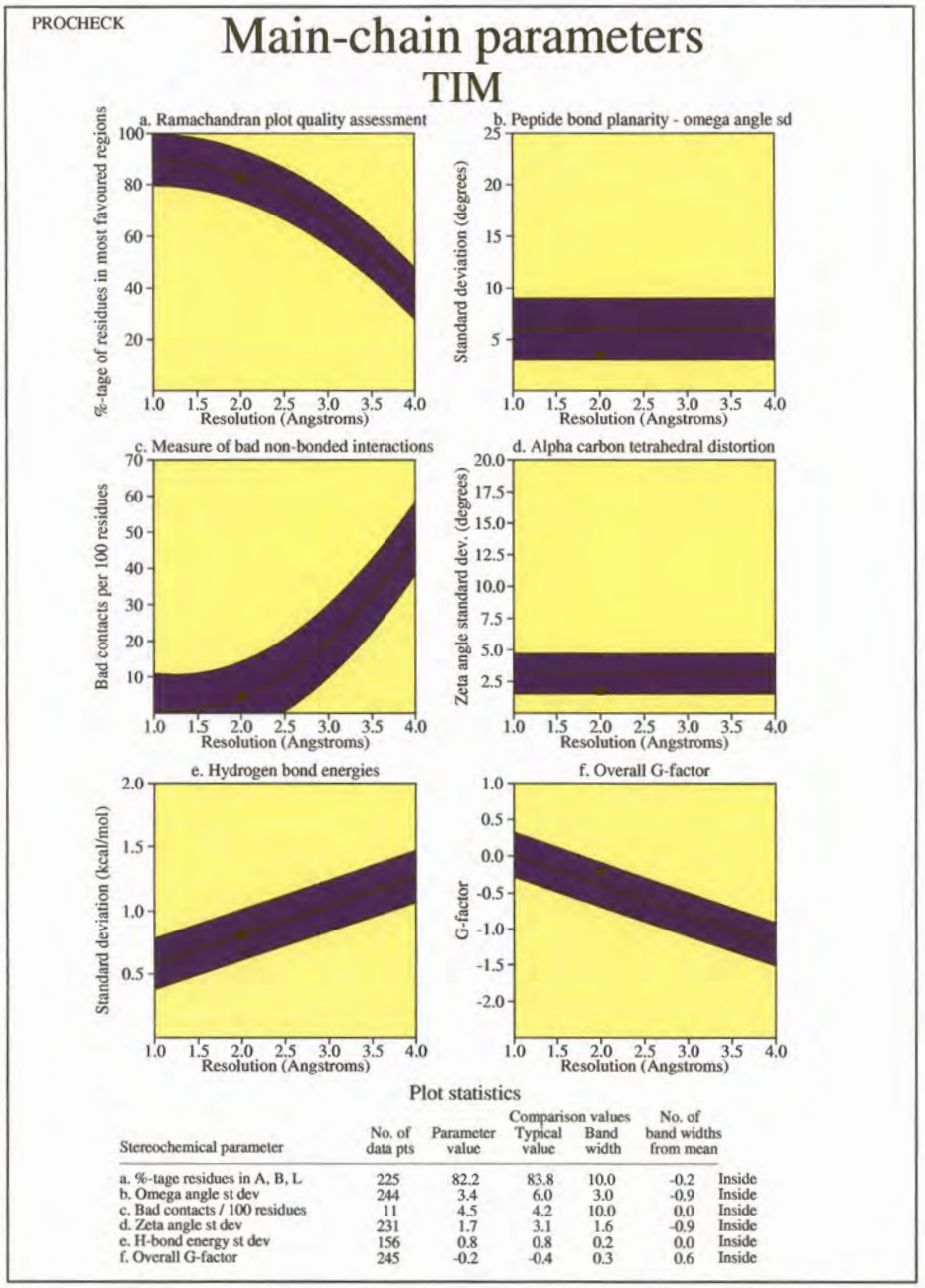


Figure 3.11: Ramachandran plot for the homology model of malaria TIM.

The Ramachandran plot shows 82.2% of residues to be in favoured regions (red), and 13.8% to be in additionally allowed regions (Yellow). Of the remaining residues 2.2% were in generously allowed regions

(beige) and 1.8% in disallowed regions (white). None of the catalytic residues were found in disallowed regions.

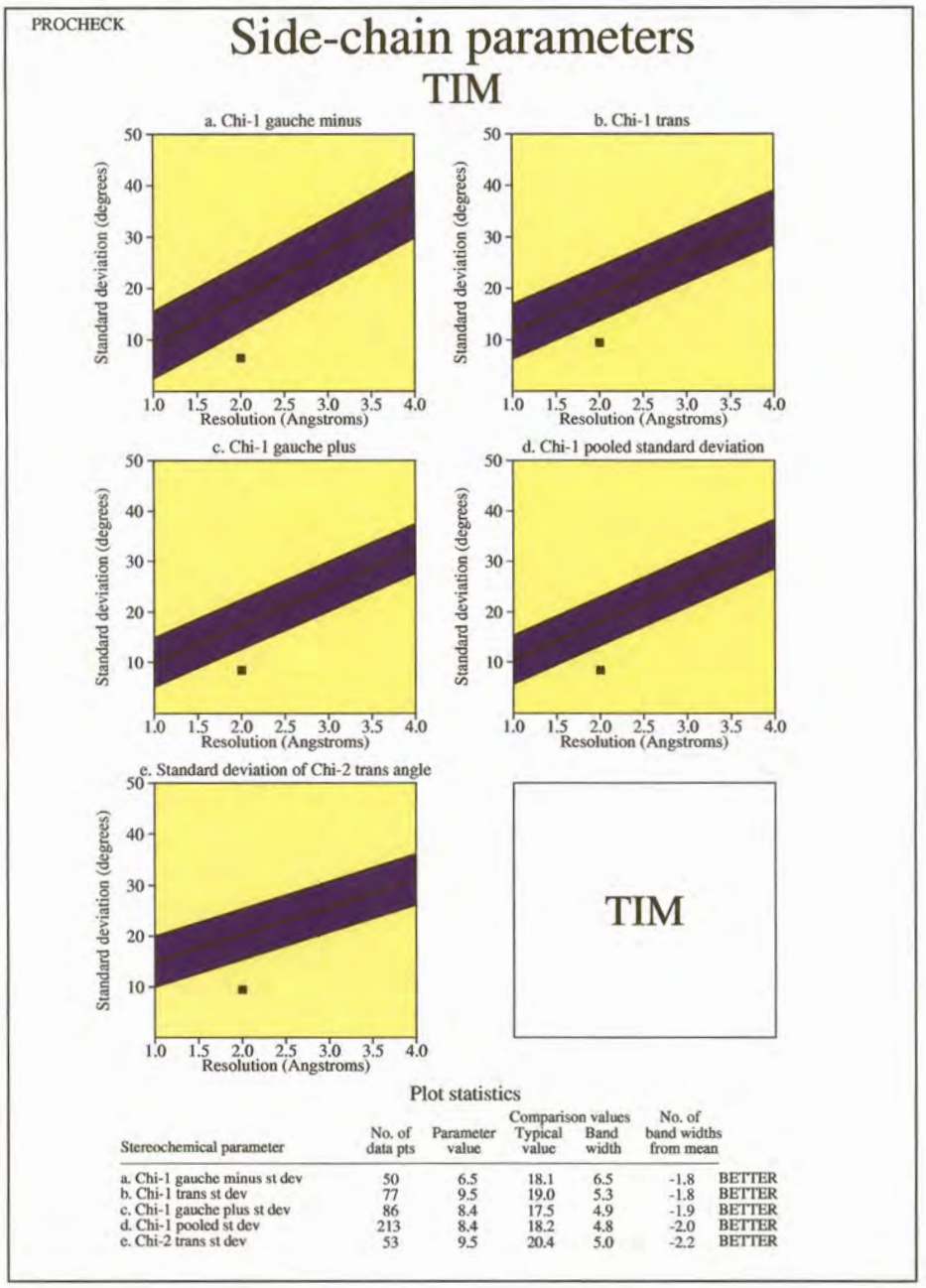


Stereochemical parameter	No. of data pts	Parameter value	Comparison values Typical value	Band width	No. of band widths from mean	
a. %-tage residues in A, B, L	225	82.2	83.8	10.0	-0.2	Inside
b. Omega angle st dev	244	3.4	6.0	3.0	-0.9	Inside
c. Bad contacts / 100 residues	11	4.5	4.2	10.0	0.0	Inside
d. Zeta angle st dev	231	1.7	3.1	1.6	-0.9	Inside
e. H-bond energy st dev	156	0.8	0.8	0.2	0.0	Inside
f. Overall G-factor	245	-0.2	-0.4	0.3	0.6	Inside

Figure 3.12: Main chain parameter analysis for homology modelled malaria TIM.

Further analysis of values from the Ramachandran plot showed the main-chain residues to occur within the allowed deviations from mean

values (Figure 3.12). A bad contact score of 4.5/100 residues was calculated and an overall G-factor of -0.2.



Stereochemical parameter	No. of data pts	Parameter value	Comparison values Typical value	Band width	No. of band widths from mean	
a. Chi-1 gauche minus st dev	50	6.5	18.1	6.5	-1.8	BETTER
b. Chi-1 trans st dev	77	9.5	19.0	5.3	-1.8	BETTER
c. Chi-1 gauche plus st dev	86	8.4	17.5	4.9	-1.9	BETTER
d. Chi-1 pooled st dev	213	8.4	18.2	4.8	-2.0	BETTER
e. Chi-2 trans st dev	53	9.5	20.4	5.0	-2.2	BETTER

Figure 3.13: Side-chain parameter analysis for homology modelled malaria TIM.

Side-chain parameters were not of high quality but were generally closer to the accepted deviation from mean than those of DHFR (Figure 3.13).

A three-dimensional diagram was prepared indicating the quality values for the model by different colours (Figure 3.14).

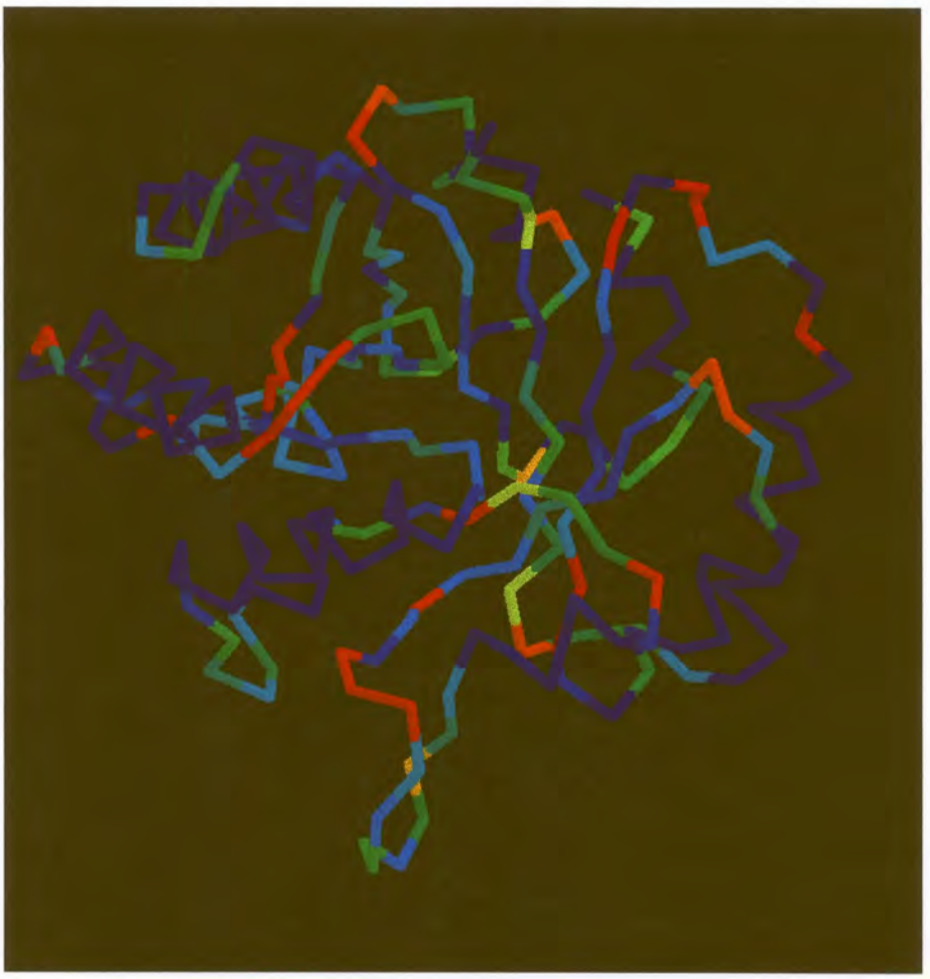


Figure 3.14: Quality score indications of homology-modelled TIM. Red regions indicate lower quality, green intermediate and blue regions indicate higher quality.

Much less red-colored regions are visible than in the case of DHFR. Low quality scores were mostly found in some loops at the surface of TIM. The homology structure was superimposed with the structure for human TIM to investigate the possibility of preparing selective inhibitors, and a R.M.S. deviation of 0.893Å was calculated (Figure 3.15).

Small deviations were visible in some helices and in loops at the surface of the TIM molecule. Active site residues were superimposed (Figure 3.16). Catalytic residues were nearly identical except for the *Ser95Phe* mutation and some slight rotation of the *His94* side-chain.

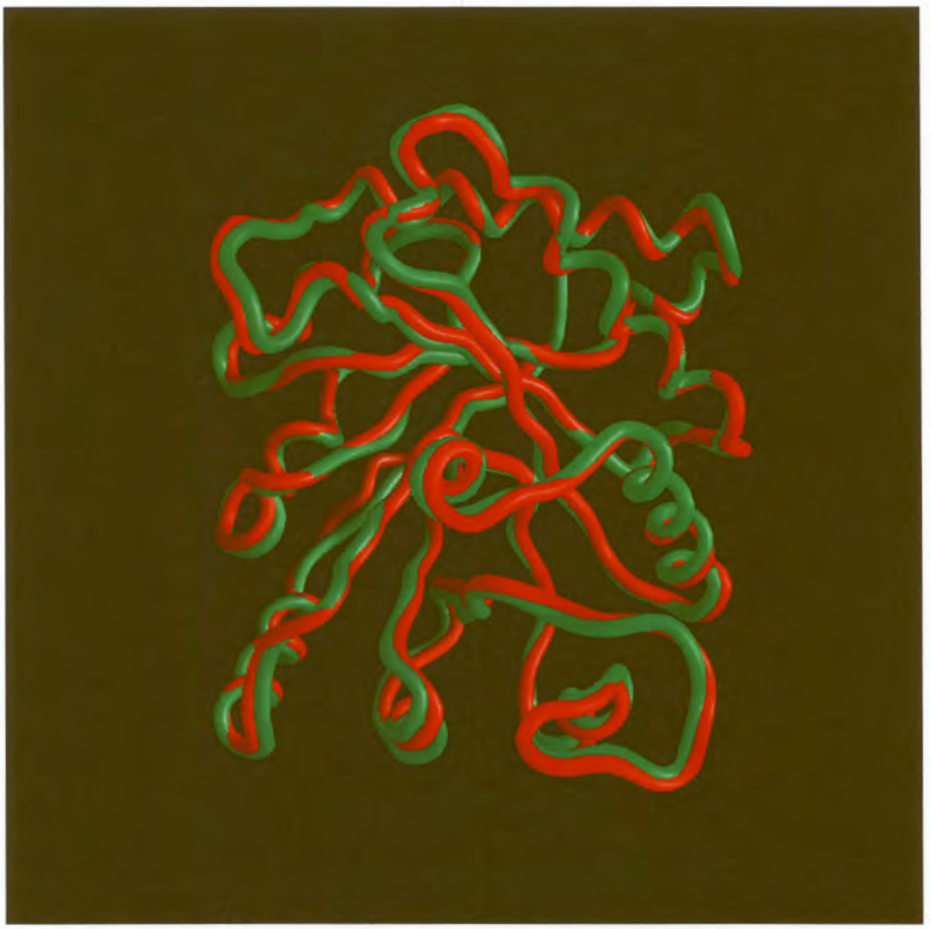


Figure 3.15: Fitted structures of modelled malaria TIM (red) and human TIM (green).

The main-chain parameters are of high quality, possibly due to the high number of templates used for model building and to the very high degree of conservation of the TIM family fold. The side-chain parameters are of lower quality, correlating with the observed lower degree of amino acid identity and high degree of fold homology found for the different TIMs.

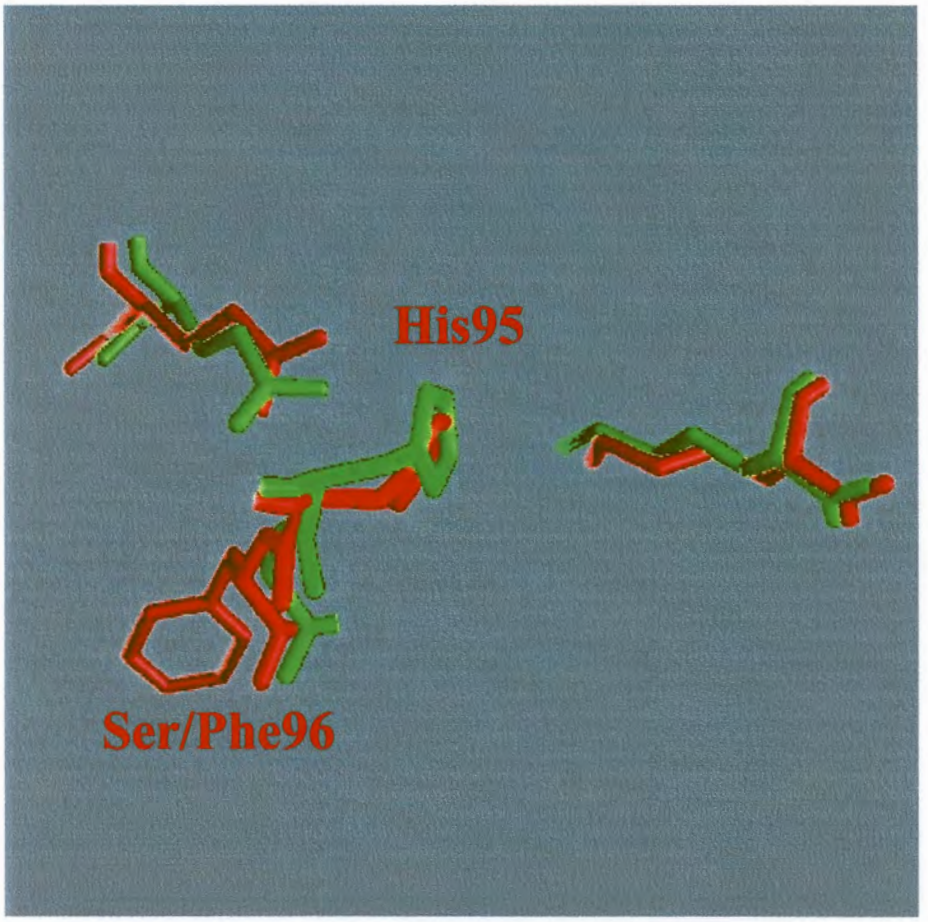


Figure 3.16: Superimposed active sites of malaria TIM (red) and human TIM (green).

3.4 Discussion

The homology modelling of *P. falciparum* DHFR was complicated by the presence of a 27 residue insertion near the N-terminal region of DHFR, as well as a 23 residue insertion near the C-terminal region. Structures from *E. coli* and *L. casei* were chosen as template, and the sequences were re-aligned. The two insertions did not seem to be involved in active site regions, and insertion regions were excised from the malaria sequence in order to facilitate model building for practical reasons, although this is not an ideal situation since the functions of these regions have not been ascertained. A recent publication by Lemcke *et al.* (1999) also addressed the preparation of homology models for malaria DHFR, but no structural models were made available. They suggested the inserted regions occur as loops pointing away from the surface of the proteins, but could not conclusively offer proof regarding the loop conformations, and found relatively low quality values for these regions. These loops are located well away from the active site cleft, and were proposed not to affect the model in terms of active site investigation. They used the conserved regions as template for model building, and added the loops afterwards from libraries of preferred side-chain conformations. The N-terminal extension has been proposed to improve contact between the two domains of DHFR. Lemcke *et al.* (1999) used their model to propose a mechanism for the influence of the Ser108Asn point mutation on pyrimethamine resistance, but could not yet explain the mechanism of the Asn51Ile mutation. Toyoda *et al.* (1997) prepared a homology-based model of DHFR by methods not yet published, and this was used for ligand design purposes (See Chapter 4). They were able to prepare inhibitors with micro- to nanomolar affinities in this way.

An alternative to modifying known drugs would be to find compounds with structures differing from the known inhibitors such as pyrimethamine, cycloguanil and methotrexate, thus decreasing the chance of the rapid development of drug resistance at the same sites. The validity of constructing a homology model of DHFR without the adjoining junction and TS domains remains debatable. *In vitro* studies have shown that DHFR will function normally without these domains (Prapunwattana *et al.*, 1996)(Sano *et al.*, 1994). However, the possibility that excision of insertion loops may affect the overall fold of the protein and also the regions facing the active site, can not be excluded.

The homology modelling of malaria TIM was less difficult than DHFR, as the sequences for TIM are very similar in all species. Quality reports of homology modelled TIM indicated satisfactory main-chain and acceptable side-chain values. Although the sequence identity is not very high, homology is up to 98% (chicken) and no large insertions or deletions are present. Near the end of this study the X-ray structure for *P. falciparum* TIM was published (Velanker *et al.*, 1997). Coordinates were obtained from the authors, and the homology and X-ray structures were compared. Cα backbone comparisons of homology modelled TIM with the X-ray structure showed a Cα R.M.S. deviation of 1.5Å (Figure 3.17). Some differences were visible in coil and loop regions.

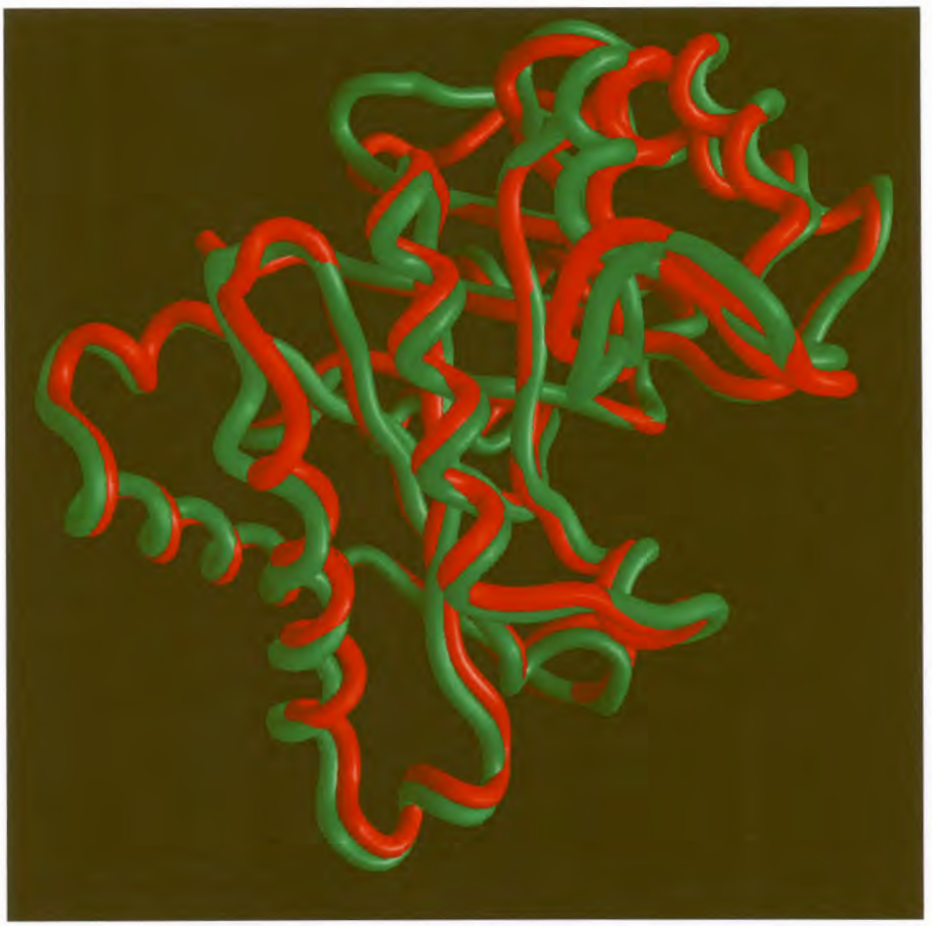


Figure 3.17: Superimposed Cα backbones of homology modelled TIM (red) and the X-ray structure of TIM (green). A R.M.S. deviation for carbon-α of 1.5Å was found.

A detailed analysis of the active sites showed the model to be nearly

identical to the X-ray structure, except for the rotation of Phe96 (Figure 3.18). This residue did however attain a most favoured rating in the model according to the Ramachandran plot (not shown).

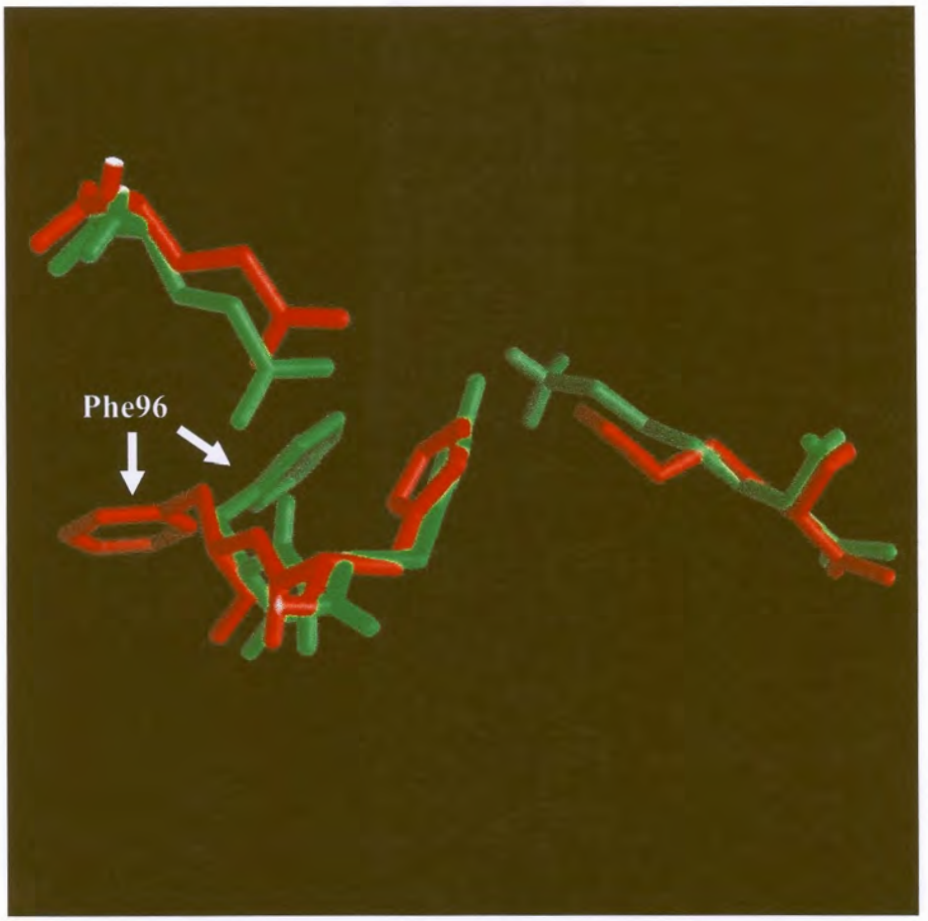


Figure 3.18: Superimposed active site residues of homology modelled TIM (red) and the X-ray structure of TIM (green). The only major difference was the rotation angle of Phe96.

Although the homology model of malaria TIM would probably have been of high enough quality for ligand design, the X-ray structure (1YDV) was used in further docking studies.

Chapter 4

Ligand discovery for malaria triosephosphate isomerase

4.1 Introduction

With current increases in world-wide malaria resistance, investigations into new inhibitors as lead drug compounds are needed. The design of new enzyme inhibitors is a challenging task. With the onset of computerised modelling in drug discovery the development times of successful drugs have decreased dramatically. All proteins bind their ligands in specific conformations. The protein and ligand combining sites determine both the affinity and the specificity of the interaction. In this interaction various forces need to be taken into account. Hydrogen bonds between peptide bonds are perhaps the greatest force in proteins, organising the proteins into secondary structure elements such as α helices and β sheets. Apolar and hydrophobic interactions also contribute to protein stability, together with side-chain to backbone hydrogen bonding.

A series of docking packages have been designed. Some are meant to determine *de novo* structures fitting into active sites, while others are designed for the scoring of existing molecules. Two main types of approaches will now be described.

The AutoDock package is optimised for the automated docking of flexi-

ble small ligands to receptors. AutoDock uses a Monte Carlo simulated annealing approach for the configurational exploration with fast energy calculation based on molecular affinity potentials (Goodsell *et al.*, 1996). Atomic affinity potentials are precalculated as described by Reynolds *et al.* (Reynolds *et al.*, 1989). The protein is embedded in a 3D grid and a probe atom is placed at each grid point. The energy of this single atom with the protein is then assigned to this grid point. This is performed for each type of atom in the ligand molecule, as well as a grid of electrostatic potentials. The docking simulation is carried out by performing a random movement of the ligand through the static protein. At each step a small random displacement is applied to each of the degrees of freedom of the ligand. If the energy is lower than that of the previous configuration, it is accepted. If higher, it is rejected based on a defined limit or a probability expression. Simulated annealing allows exploration of a complex configurational space. A range of molecular interactions may be taken into account such as dielectrics, potential functions, etc. This approach has been used for the docking of ligands to fXa inhibitor (Rao and Olson, 1999), the cannabinoid receptor (Mahmoudian, 1997), antibodies (Gamper *et al.*, 1996) and many other proteins.

The DOCK package (Gschwend *et al.*, 1996) used in this study is designed to find favourable orientations of a ligand in a receptor. The ligand orientation in the binding site is broken down into a series of steps. First a potential site of interest in the protein is identified. Within this site, points are identified where ligand atoms may be located. This is done by generating a set of overlapping spheres to fill the site. This limits the potential number of orientations in the active site. Spheres may overlap each other but not the protein surface. To orient a ligand within the active site some of the sphere centers are matched with ligand atoms, generating many such sets. A set of atom-sphere pairs are used to calculate the orientation of the ligand at the site of interest. Various filters are used to eliminate conformations which would generate poorly scoring orientations, thus only a small subset of possible ligand orientations are generated and scored. The spheres do not take chemical matching into account. These properties may then be introduced by matching characteristics such as hydrogen-bond donors, hydrophobes, electro-positiveness, electro-negativeness, etc. In this way the number of unfavourable operations which are generated and scored may be further reduced. A critical point set may be defined in-

corporating known interactions. This method has been used with great success in cases such as *P. carinii* DHFR (Gschwend *et al.*, 1997), influenza virus hemagglutinin (Hoffman *et al.*, 1997) and double-strand RNA systems (Chen *et al.*, 1997).

Before scoring, a bump filter is applied which eliminates orientations in which ligand atoms occupy space already in use by the protein. The orientation of the ligand is then evaluated with a shape scoring function and/or a function approximating the protein-ligand binding energy. Evaluations are performed on scoring grids to minimise computing time. This approach stores the protein contributions to the score only once according to the grid, and retrieves them as necessary. The shape scoring function is an empirical function similar to the van der Waals attractive energy. The shape score is generated by summing the receptor terms from the grid point nearest to each non-hydrogen ligand, thus the shape score is determined by the position of each ligand atom on the shape scoring grid. The binding energy of the ligand-protein interaction is approximated by the sum of the van der Waal attractive, van der Waal dispersive and Coulombic electrostatic energies. To generate the energy score, ligand atom terms are combined with receptor terms from the nearest or virtual interpolated grid point. Thus the energy score is determined by both the ligand atom types and ligand atom positions on the grid.

The 3D Available Chemicals Database is a good source of ligand structures but is expensively priced. The Cambridge Crystallogaphic Database contains accurate structures but many of them play no role in biological systems. It was decided to screen against the National Cancer Institute (NCI) 3D database. The advantages are that this database contains compounds used in cancer studies, thus the biological side effects of many of these molecules are known. The database is also freely available without any cost. However, the 3D database was theoretically generated with the Corina package (Gasteiger *et al.*, 1990), thus altough the structures are energetically optimal, they are not always possible to synthesise in these conformations. Also, many of the compounds have been specifically synthesized and are not commercially available.

As DHFR was not expressed in an optimised form, and the DHFR homology models were of lower quality, DHFR was not employed for ligand screening. Also, the newly available X-ray structure of TIM increased the chances of finding successful inhibitors.

The following aims were set for this study:

- The screening of a 3D molecular database against the active site region of malaria TIM
- The *in vitro* screening of putatitive inhibitors against recombinant TIM
- The screening of putatitive inhibitors against malaria cultures

4.2 Methods

The X-ray structure (1YDV) for malaria TIM (Velanker *et al.*, 1997) was used for all docking studies. Of the dimer A- and B-chains, the A chain was used, and cleaned for waters and other ions. SYBYL (Tripos Associates, St. Louis) was used for correction of the N- and C-terminal residues and for hydrogen addition and charge assignment. The structure was subsequently converted to MOL2 format. A grid box was manually added to define the region of interest, and the GRID module was used to define the steric and electrostatic properties at each grid point to enhance rapid scoring of ligand orientations during the DOCK run. The molecular surface was created with the QCPE MS package (University of California, Berkeley). The SPHGEN module was used for the calculation of spheres filling the site, and the sphere site points were optimised by visual inspection. The NCI-3D database (Milne *et al.*, 1994) was converted from SDF to MOL2 format, and SYBYL was used to assign atom types and charges. The docking run was performed with the DOCK 4 (Gschwend *et al.*, 1996) package on a Silicon Graphics Challenge S system.

Inhibition kinetics of ligands to recombinant malaria TIM was assayed as described in paragraph 2.3.2. Inhibitors were added in the 10 to 250μM range and assayed for changes in TIM activity. Results were used to prepare inverse reciprocal plots of reaction rate against substrate concentration.

To test the effect of the inhibitors on parasites, malaria cultures were incubated with inhibitors at concentrations of 1000μM, 100μM and 10μM for a period of 24 hours. Cultures were subsequently stained with thiazole orange according to Schulze *et al.* (1997) and parasitemia measured by flow cytometry (Coulter FACSScan, Beckman-Coulter, California).

4.3 Results

After conversion of the database from SDF to MOL2 format and assignment of charges and hydrogens in SYBYL, 107,955 compounds were tested for docking. After removing compounds with high bump scores (thus not being able to fit because of their size), 105029 compounds were suitable for docking in the defined region. The region of interest was optimised by visual inspection to surround the active site region of malaria TIM. The superimposed structure of human TIM bound to PGA was used to determine the center of the TIM active site region (Figure 4.1) and to select a bounding box.

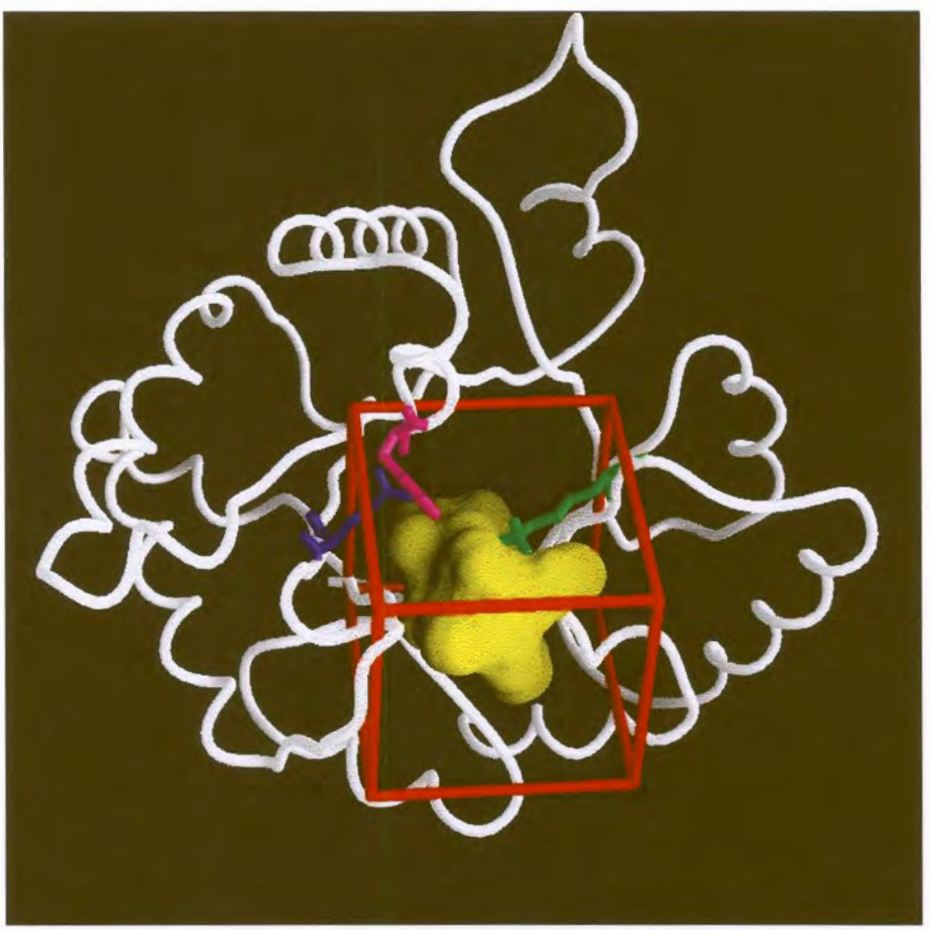


Figure 4.1: Region of malaria TIM surrounding the active site used for ligand docking. The active site was determined by the binding position of the inhibitor PGA. Yellow spheres indicate the active site surface, and the region of interest is bound by a red box. Active site residue *Lys12* is indicated in green, *His96* in magenta and *Glu165* in blue.

The spheres for the docking space were optimised by visual inspection before use (Figure 4.2) and spheres leading away from the desired region were removed.

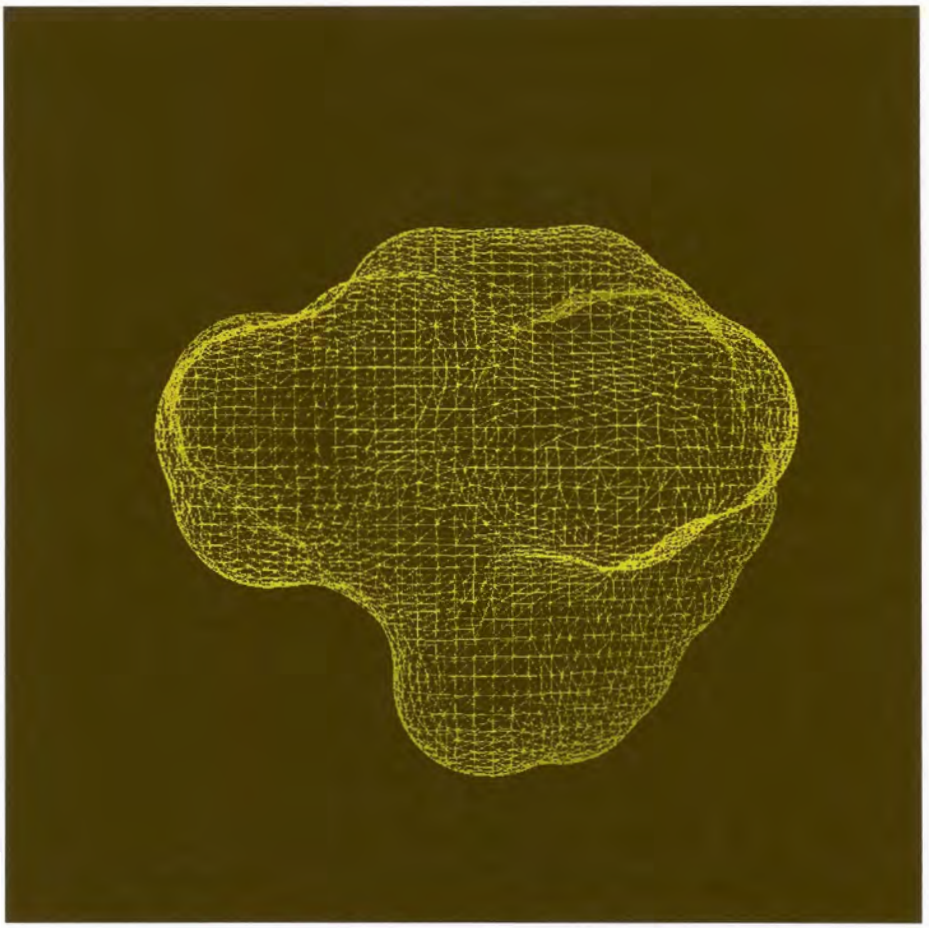


Figure 4.2: Detailed view of spheres defining the inverse of the active site cavity.

Computer screening took approximately 500 hours. The top 100 hits were saved, and manually inspected for suitability in terms of size, physical properties, solubility, etc. Various permutations of the same molecules were discarded, for example compounds differing only in the bound salt ion, etc. Hits were further refined according to commercial availability.

The following compounds were chosen for inhibition studies (Table 4.1 and 4.2):

The commercially unavailable compounds are being synthesized by

Table 4.1: Commercially available compounds tested for TIM inhibition.

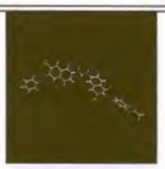
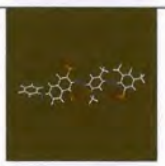
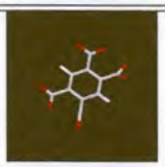
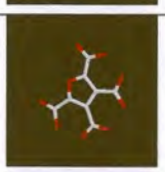
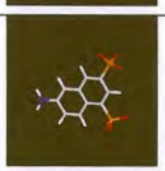
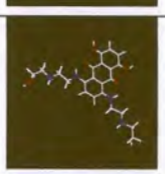
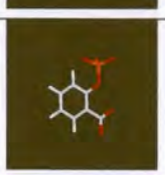
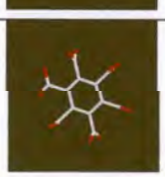
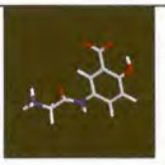
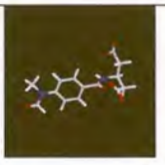
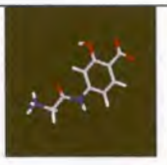
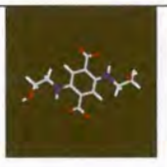
Compound	Structure	Commercial availability	Score
Direct Red 23		Yes	-38.31
Direct Violet 51		Yes	-36.89
Pyromellitic acid		Yes	-36.05
Furanetetracarboxylic acid		Yes	-35.59
Amino I acid		Yes	-35.47
Mitoxanthrone		Yes	-35.11
Tricarballylic acid		Yes	-34.66
Salicyl phosphate		Yes	-34.49
Mellitic acid		Yes	-33.72

Table 4.2: Synthesized compounds tested for TIM inhibition.

Compound	Structure	Commercial availability	Score
Bicyclo[2.2.1]heptane-2,3,5,6-tetracarboxylic acid		No	-41.31
5-(2-aminoacetamido)-salicylic acid		No	-37.37
3-[[4-(formylmethylamino) benzoyl] amino]-pentanedioic acid		No	-36.93
4-(2-aminoacetamido)-salicylic acid		No	-36.66
2,5-bis[(2-hydroxyethyl)amino]-terephthalic acid		No	-36.34

Prof. T van Ree at the University of Venda but were not yet available at the time of writing this thesis.

Inhibition studies on recombinant purified malaria TIM were performed with DHAP as substrate. Of the commercially available inhibitors, only Direct Red 23 and Direct Violet 51 showed any inhibition of TIM activity at concentrations lower than 100μM. Kinetic studies of these two compounds were performed to determine the mode of inhibition (Figure 4.3).

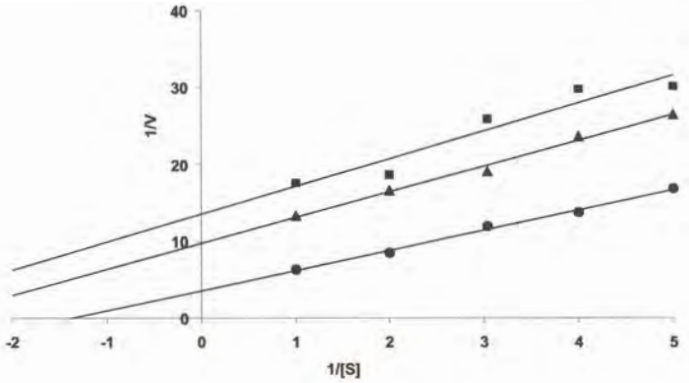


Figure 4.3: Inhibitor kinetic plot for Direct Red and malaria TIM. Squares indicate a concentration of 60μM, triangles 50μM and circles 40μM.

For Direct Red, a mixed type inhibition was found. Whereas competitive inhibition would yield lines converging to the same 1/[V], and uncompetitive inhibition would yield parallel lines, the plots indicated an inhibition mode similar to uncompetitive inhibition, with slightly varying angles to the parallel lines. Direct Violet showed similar values (results not shown).

Inhibition of malaria growth in culture was measured by thiazole orange staining and flow cytometry (Figure 4.4). This method discerns between the malaria parasites and red blood cells based on the fluorescence of genomic DNA when stained with thiazole orange.

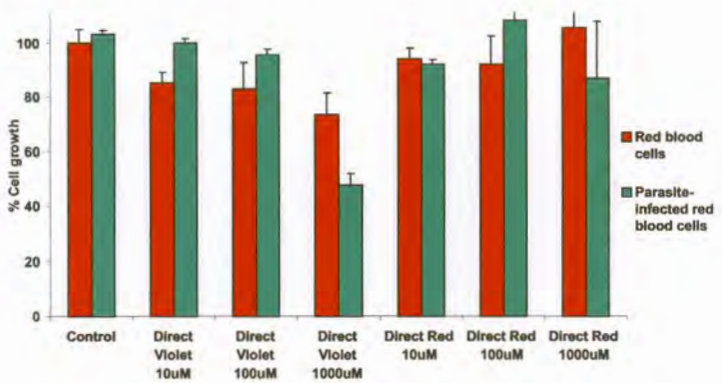


Figure 4.4: Parasite growth in the presence of Direct Red and Direct Violet. Red bars indicate red blood cell growth and cyan bars indicate parasite growth.

Direct Violet showed a reduction in parasitemia at concentrations of 1000μM, but not at lower concentrations. However, a similar decrease was found for red blood cells. No specific effect on malaria parasites could thus be shown. Direct Violet did not yield any significant differences. Other questions remain to be answered, such as the penetrative ability of these compounds through the red blood cell membrane, and subsequently through the parasite membrane.

Investigation into the binding modes of Direct Red and Direct Violet showed the compounds to be bound at the active site of TIM, with one end buried and the other protruding into the solvent (Figure 4.5).

Analysis of intermolecular contacts by LIGPLOT showed common contacts involved in the binding of both compounds (Figure 4.6, 4.7).

Both molecules bound with a diphenyl-phenyl type conformation in the protein, but not at the same depth. A common substructure bound with the exact same conformation is visible when Direct Red and Violet are superimposed (Figure 4.8).

Of interest are the common contacts that both compounds have with actve site residues *Phe96*, *Glu165* and *Val212* . Especially *Phe96* and *Glu165* play a role in substrate contact in TIM from many different species. Lead compounds may possibly be synthesized based on the common structure of Direct Red and Violet showing contact with the above residues.

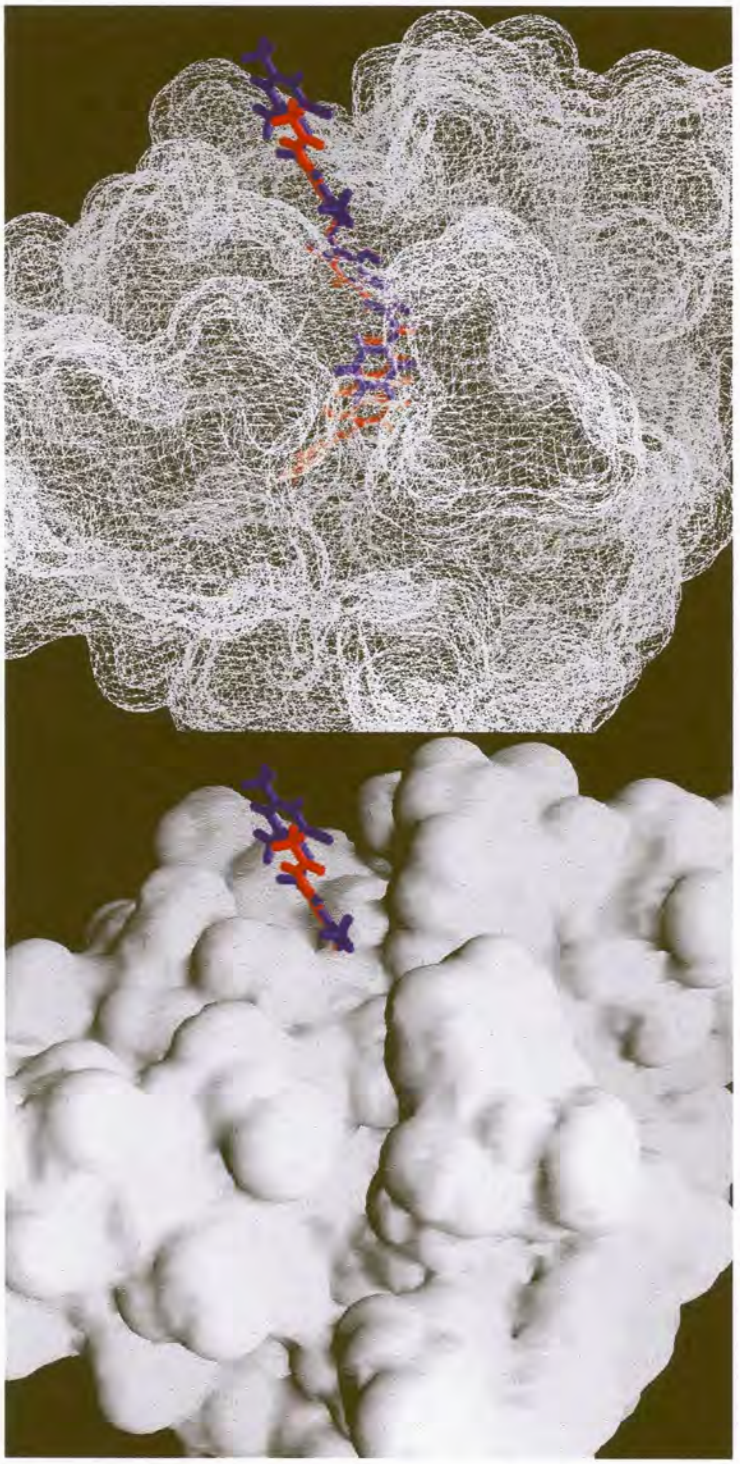
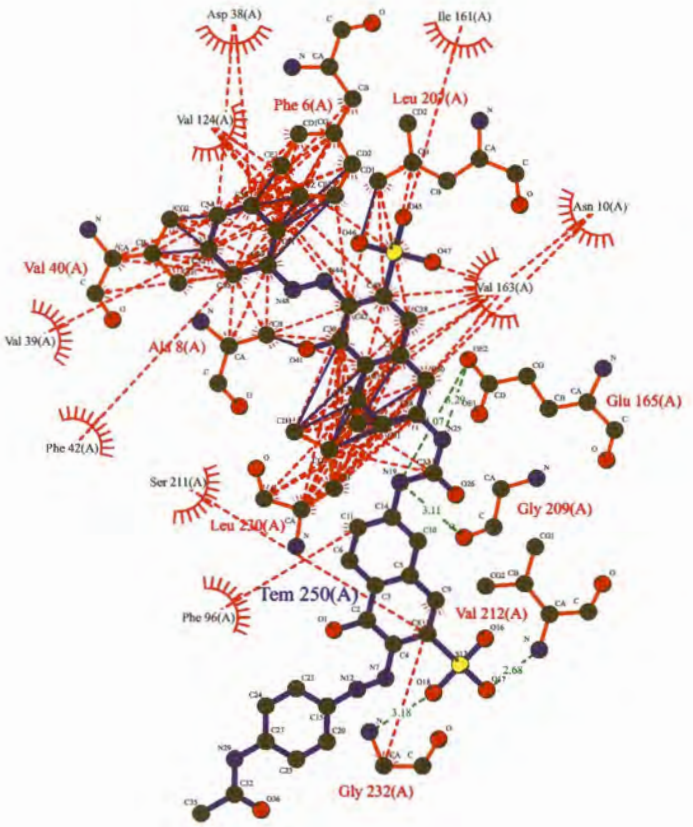


Figure 4.5: Direct Red and Violet bound in the active site of *P. falciparum* TIM. An accessibility surface is shown as mesh (top) or as solid (bottom).



Key

Ligand bond

Non-ligand bond

Hydrogen bond and its length

His 53 Non-ligand residues involved in hydrophobic contact(s)

Corresponding atoms involved in hydrophobic contact(s)

Direct Red

Figure 4.6: Detailed contact map of the malaria TIM-Direct Red complex.

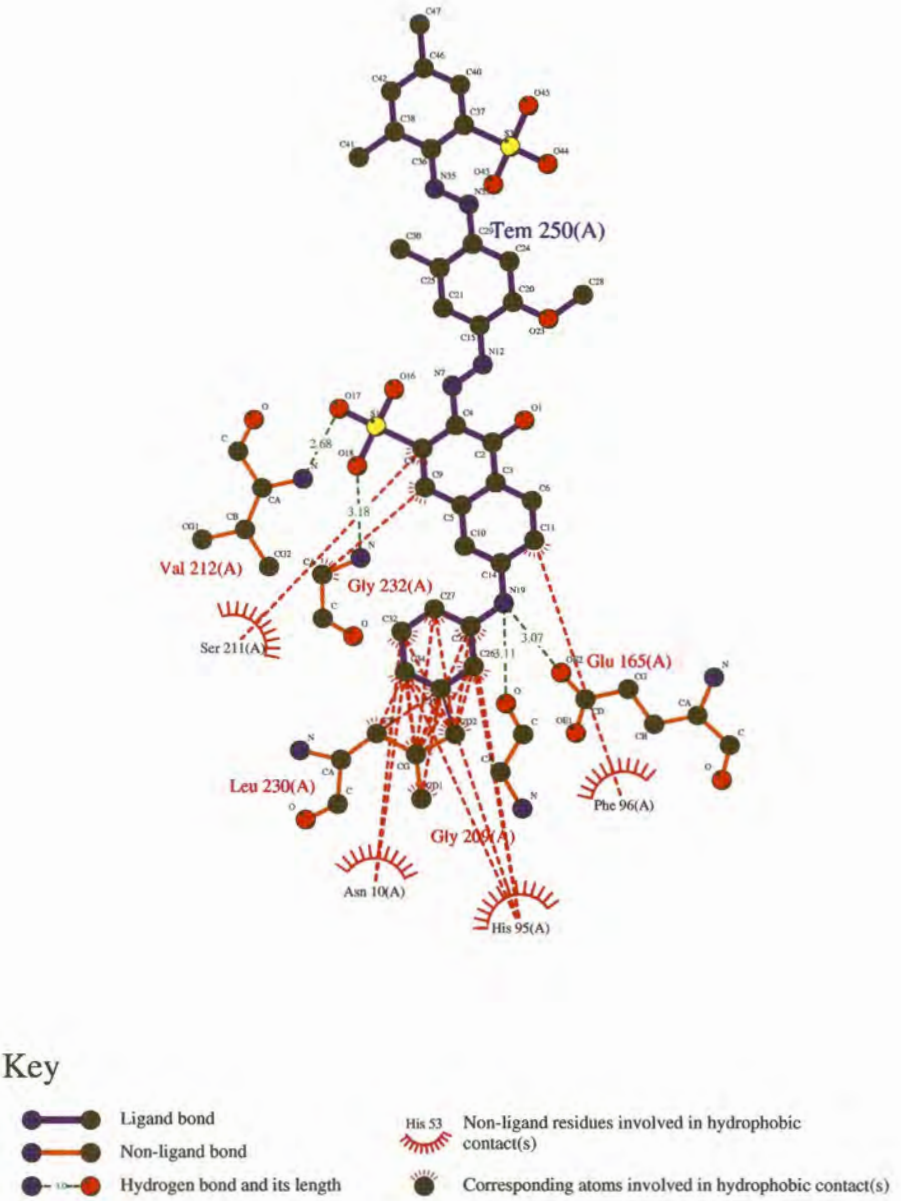


Figure 4.7: Detailed contact map of the malaria TIM-Direct Violet complex.

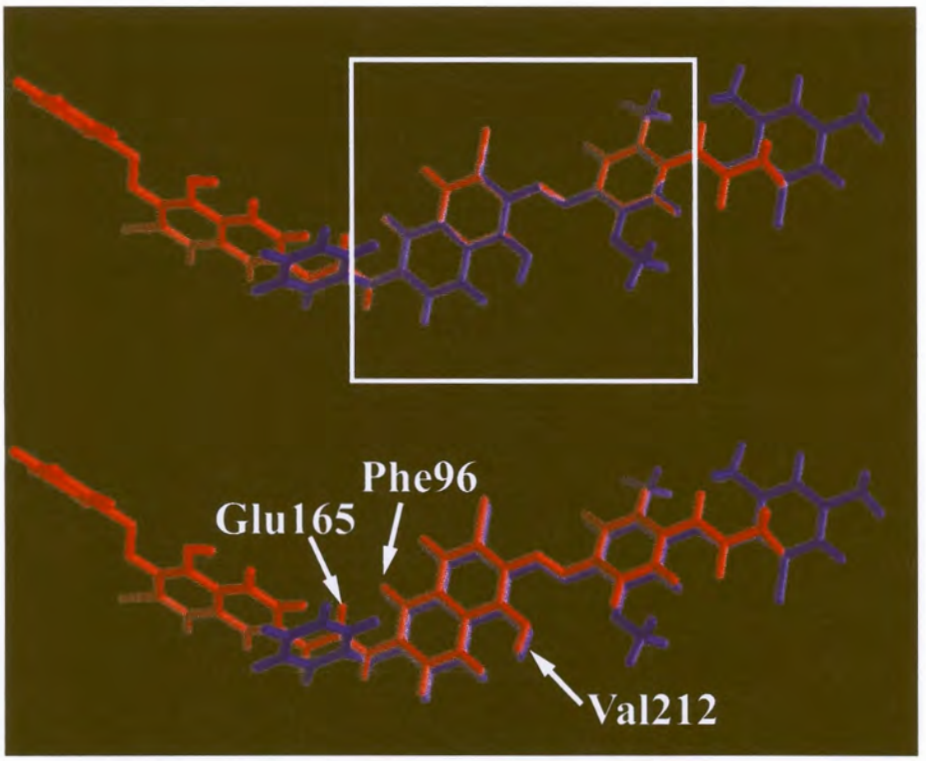


Figure 4.8: Comparison of the spatial positions of bound Direct Red and Violet. The common structural area is indicated by a white box. *Phe96*, *Glu165* and *Val212* show common contact points on both molecules. The top figure shows the exact orientation, while the bottom figure has been shifted slightly to distinguish the two molecules.

4.4 Discussion

Triosephosphate isomerase has long been a target for a series of inhibitors. Crystallisation studies were initially performed with bound PGH which is a substrate analogue (Zhang *et al.*, 1994). Verlinde *et al.* (1992) prepared crystals of *T. brucei brucei* TIM with the competitive inhibitor N-hydroxy-4-phosphono-butanamide, finding that the active site flexible loop was locked in a completely open conformation. TIM has been extensively studied in the parasite *T. brucei brucei* for the design of putative inhibitors. Hydrophobic cyclic hexapeptides were synthesized, and were found to be active and selective inhibitors of TIM (Kuntz *et al.*, 1992). In a study by Schnackerz *et al.* (1991), 3-chloroacetol phosphate (CAP) was shown to be bound to the active site by reaction with *Glu165* to irreversibly fill the binding site of one subunit. The reversibly binding transition state analogue PGA, was used to probe the remaining vacant active site of the heterodimer and showed independent binding at the two active sites. Lolis *et al.* (1990) showed that the inhibitor formed hydrogen bonds to the side-chains of *His95* and *Glu165*. The latter hydrogen bond confirmed that *Glu165* was protonated upon PGA binding. Conformational changes were observed: the side-chain of *Glu165* moved over 2Å and a 10-residue flexible loop moved over 7Å to close over the active site like a lid. The ferrate anion, an analog of orthophosphate anion, was also shown to rapidly inactivate triose phosphate isomerase from chicken muscle. The inactivation can be prevented by the presence of competitive inhibitors (Steczko *et al.*, 1983).

Velanker *et al.* (1997) described the active site of malaria TIM, and found the structures of malaria and human TIM to be very similar. This makes the design of species-specific inhibitors highy challenging. The high specificity of TIM makes this task even more difficult. The only significant differences in the active site are at *Cys13* and *Phe96*. Also, a charged surface at *Glu183* is replaced with the hydrophobic *Leu183*, which has been proposed to play a role in membrane attachment. The other potential site for interference is at the dimerisation interface.

The strategy we employed was to target not only the active site residues of TIM, but also allow ligands having contacts surrounding the active site region. These ligands might possibly have higher selectivity, because TIM is less conserved away from the active site. The ligands would then still have to be capable of providing steric hindrance to

substrate binding.

The two inhibitory compounds Direct Red and Direct Violet are both brilliantly coloured dyes used in the textile industry. They have been used in a recent study (Gao *et al.*, 1998) regarding the effect of Suramin and Suramin-like compounds on *Trypanosoma* TIM. The structures of Suramin and its similar compounds Direct Red and Direct Violet are shown in Figure 4.9, 4.10.

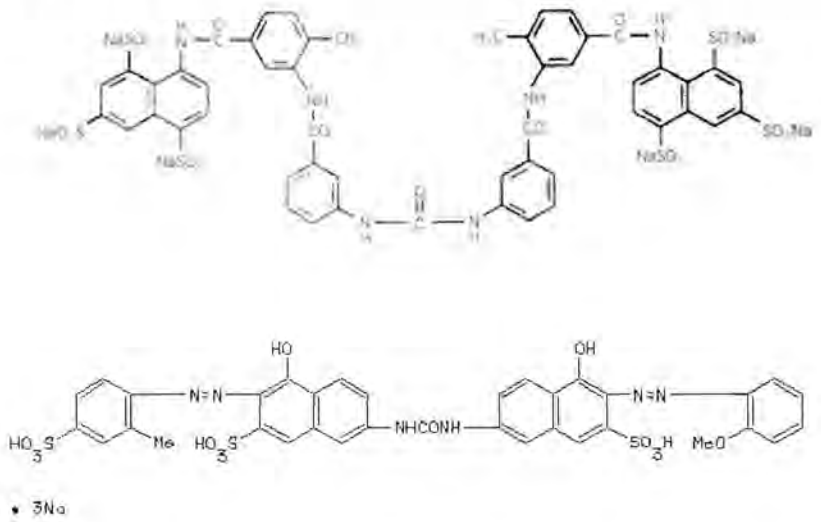


Figure 4.9: Structures of Suramin (top), and Direct Red (bottom).

Suramin is a negatively charged molecule that has been used extensively in the treatment of African sleeping sickness caused by *T. brucei brucei*, but is ineffective in Chagas disease with *T. cruzi* as causative agent. It is known to affect various enzymes including some in the glycolytic pathway of *T. brucei*, and has been found to inhibit the reactivation and dimer formation of *Trypanosoma* and human TIM after unfolding (Gao *et al.*, 1998). TIM is inactive in monomeric form, and the authors proposed that Suramin interfered with TIM dimerisation thus decreasing TIM activity. However, TIM was also shown to be inhibited by Suramin (Misset and Opperdoes, 1984) at concentrations lower than those used for inhibition of refolding. Suramin has been shown to compete with substrates (K_i = 0.1mM), indicating an influence at the active site (Lambeir *et al.*, 1987). This data indicates that it probably binds to TIM in a conformation that affects both the active site and the

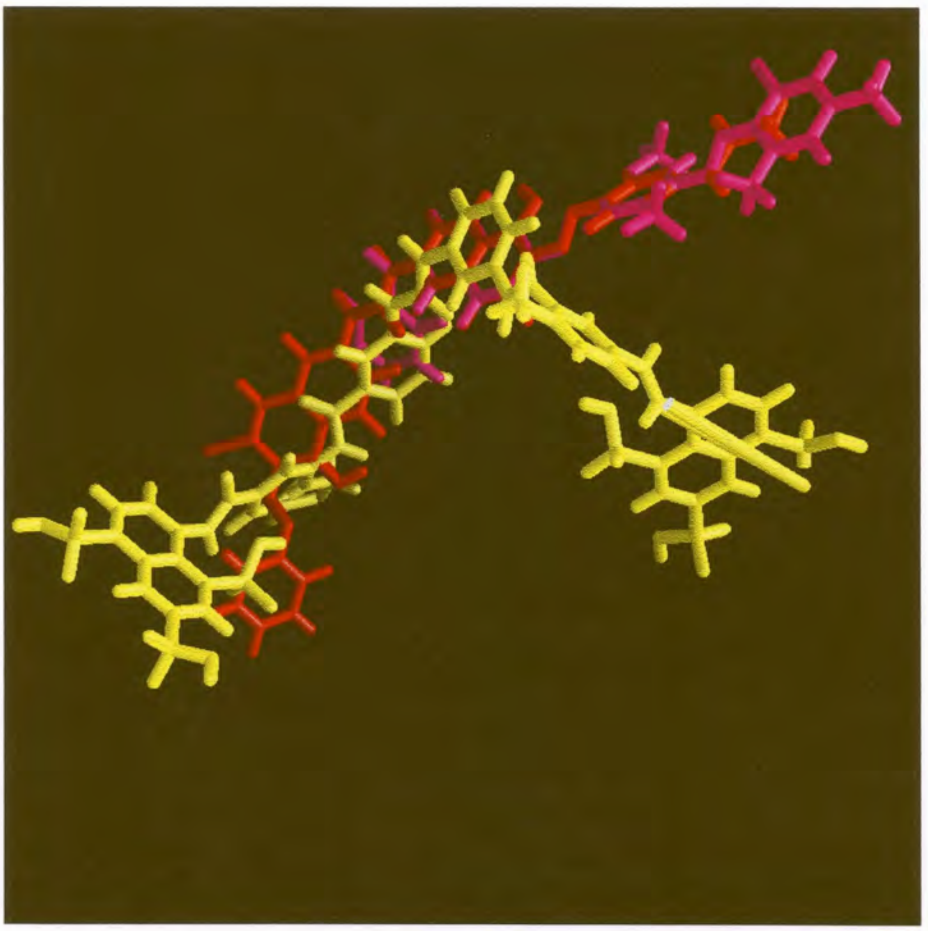
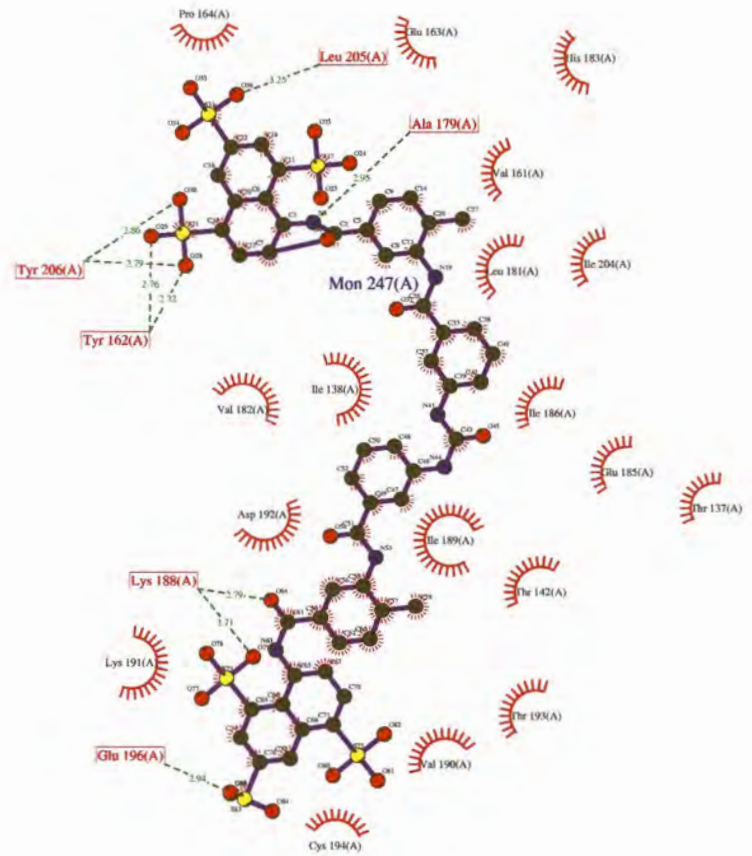


Figure 4.10: A comparison of the three-dimensional structures of Suramin (yellow), Direct Red (red) and Direct Violet (magenta).

dimerisation interface.

The two inhibitors Direct Red and Direct Violet were selected for TIM binding by the DOCK package, totally independent from the article regarding the effect of these compounds on *Trypanosoma* and human TIM (Suramin was not present in the ligand library). The binding scores of these two compounds together with the independent data from Gao *et al.* (1998) suggest that their binding occurs specifically to TIM. This counteracts previous speculations that the promiscuous interactions of Suramin with many different enzymes is simply due to the negative charge of the molecule. Although Suramin can be docked successfully into the TIM active site, no significant contacts of Suramin with the catalytic residues could be indicated (Figure 4.11).



Key

Ligand bond
Non-ligand bond
Hydrogen bond and its length
His 53 Non-ligand residues involved in hydrophobic contact(s)
Corresponding atoms involved in hydrophobic contact(s)

Suramin

Figure 4.11: Detailed contact map for the TIM-Suramin complex.

Mutational analyses of residues in contact with the two dyes may be used to confirm their conformation of binding to TIM. Ideally, the minimal fragment of these dyes capable of inhibiting TIM activity should be determined, and possibly used as a lead compound for anti-malarial

drug design. Lead compounds may be designed based on the common structural fragment as shown in Figure 4.8. Databases may initially be searched for molecules similar to the common fragment of both dyes, and these compounds then tested *in vitro*. Alternatively, packages are available for the *de novo* design of ligands based on a site contact map.

A comparison of the active sites for *Trypanosoma*, malaria and human TIM regarding residues having contact with Direct Red show little differences (Figure 4.12). Except for the mutation of Ser96 to Phe in the case of malaria, few other differences were visible. Both Direct Red and Direct Violet show contact with this residue, and it may be exploited for the design of species specific inhibitors.

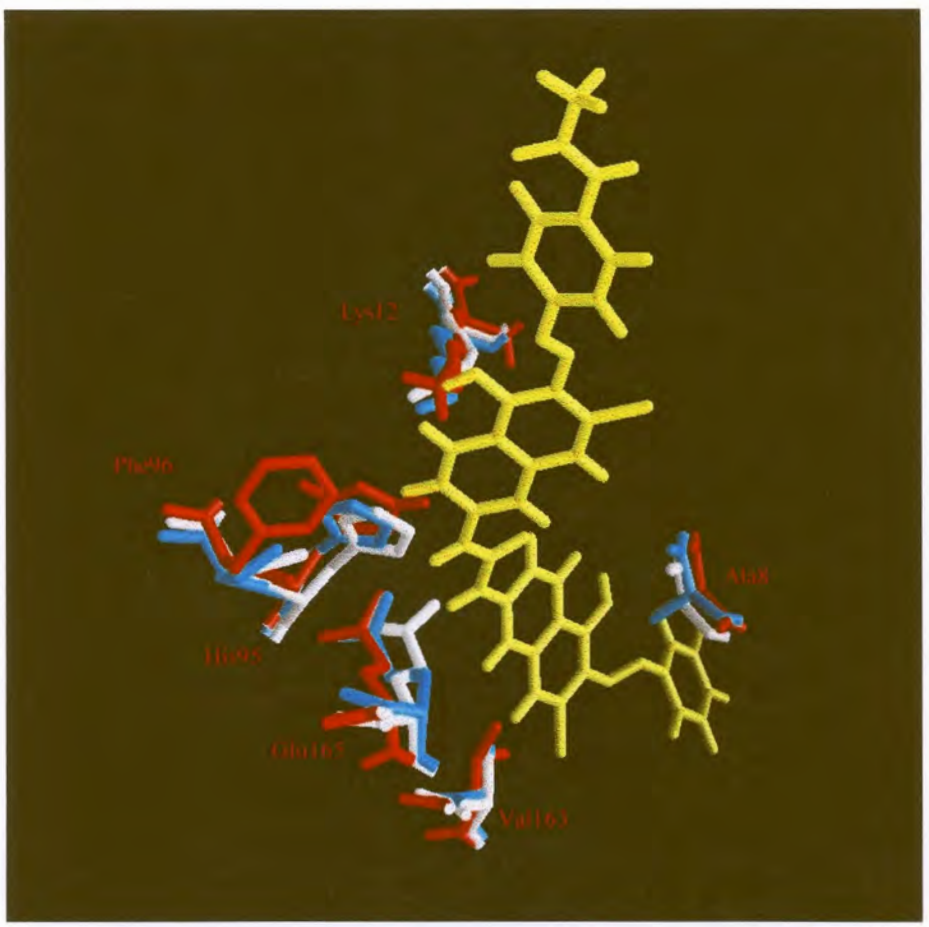


Figure 4.12: Active site comparison for malaria TIM (red), *Trypanosoma* TIM (cyan) and human TIM (white). Direct Red as docked into the malaria TIM active site is shown in yellow.

The mixed type inhibition that was found for Direct Red and Direct

Violet is probably due to a similar binding conformation to that of Suramin affecting both TIM dimerisation and the active site region. Docking studies of Direct Red and Violet do not show direct contact with the dimerisation interface of dimeric TIM, and it is suggested that the inhibitor binding probably causes perturbation of the TIM backbone to affect the dimerisation process.

Chapter 5

Concluding Discussion

The ultimate aim of this study was the generation of novel inhibitors against malaria metabolic enzymes. DHFR was chosen as candidate, since it has been validated as an efficient therapeutic target, and had been characterised in detail in various organisms. TIM has not been used as a drug target enzyme up to now, but plays a key role in the parasite's carbohydrate metabolism and has been studied extensively in many organisms.

To be able to design enzyme inhibitors, detailed knowledge is needed regarding the structure of the protein target. Unfortunately, in the case of malaria, only five enzymes have been crystallised. These are TIM (Velanker *et al.*, 1997), HGPRT (Shi *et al.*, 1999), LDH (Dunn *et al.*, 1996), fructose-1,6-bisphosphate aldolase (Kim *et al.*, 1998), and plasmepsin II (Bernstein *et al.*, 1999). DHFR has been cloned and crystallised from many organisms, and generally shows a high degree of amino acid sequence homology between species. Malaria DHFR was first expressed by Sirawaraporn *et al.* (1990). They were able to express low amounts of the DHFR-TS complex in *E. coli*. Speculating that the AT-rich malaria codons were inhibiting overexpression, a synthetic gene incorporating *E. coli* codon preferences was constructed. When overexpression of this DHFR domain was attempted, the protein was found only in inclusion bodies (Sirawaraporn *et al.*, 1993). A synthetic gene of DHFR-TS was recently prepared, which overexpressed in *E. coli* as an active correctly folded enzyme (Prapunwattana *et al.*, 1996). Up to now, this enzyme has not been crystallised. Initially, we set out to obtain DHFR in an active, correctly folded form. The malaria DHFR gene could be expressed using a T7 promoter, but no soluble expression of the DHFR domain

could be achieved despite the presence of solubility-enhancing fusion peptides. This study showed that a synthetic gene was not necessary for the overexpression of AT-rich genes, and that extremely high levels could be expressed under a T7-lac promoter. It was concluded that the presence of the TS-domain was crucial for the correct folding of malaria DHFR.

Homology modelling of malaria DHFR was complicated by the large insertions present in the DHFR gene compared to that of other species. These insertions could not be modelled from segments of other proteins and seem to be unique to malaria DHFR. Similar insertions have been found in various other malaria proteins, but the function for the inserted regions is unknown (LeBlanc and Wilson, 1993). Homology modelling of the active site was performed, but low quality scores did not yield sufficient confidence in the model to use it in ligand docking studies. Two recent articles discussed the use of homology modelled DHFR active sites in drug discovery and mechanisms of resistance. Lemcke *et al.* (1999) suggested the inserted regions occur as loops pointing away from the surface of the proteins. However, they could not conclusively offer proof regarding the loop conformations, and found relatively low quality values for these regions. They used their model to propose a mechanism for the influence of the *Ser108Asn* point mutation on pyrimethamine resistance, but could not yet explain the mechanism of the *Asn51Ile* mutation. Toyoda *et al.* (1997) prepared a homology-based model of DHFR by methods not yet published, and this was used for ligand design purposes. They were able to prepare inhibitors with micro- to nanomolar affinities in this way.

TIM has been extensively studied due to its role in diseases, as well as its structural properties. Although sequence identity is not high, the TIM fold is conserved in the enzyme from all species, and catalytic residues are conserved. Malaria TIM was first cloned by Ranie *et al.* (1993), and expressed in a TIM deficient *E. coli* strain. TIM activity was shown, but no isolations were performed. We obtained the TIM cDNA from local parasite isolates, and expressed TIM in the pET15b system, which supplies a N-terminal His-Tag for purification purposes. TIM was expressed in active form, and could be purified to homogeneity by immobilised metal affinity chromatography.

Due to the highly conserved nature of the TIM backbone, homology modelling was not as problematic as in the case of DHFR. The modelled backbone of malaria TIM showed very little deviation from that of

other species. The model was subsequently compared to the recently published X-ray structure of malaria TIM (Velanker *et al.*, 1997). In the active site, the only difference was the rotation angle of *Phe96*. This indicates the value of homology models in the drug design process when numerous templates for a protein with a conserved fold are available. Due to the discrepancy in the *Phe96* side-chain angle, the X-ray structure 1YDV was used in the rest of the study for ligand docking studies.

A ligand search was performed only for TIM, as DHFR could not be expressed in a soluble form and the DHFR homology models were not of sufficient quality. TIM is active only in dimeric form, thus lead compounds may be targeted againt the dimerisation interface as well as the active site. Our approach was to prepare a grid of the active site, whereupon a series of positions of each ligand could be scored. This approach minimises computational time by evaluating the protein chacteristics at each grid position only once. A set of spheres was then constructed to limit the ligand binding to the active site area. The NCI-3D database was screened against the binding site with bump filtering to eliminate molecules which would not fit, before scoring took place. Ideally, a chemical library such as the ACD should be used during screening, as all ligands are commercially available. Many of the NCI-3D compounds have been custom synthesized, and are extremely difficult to obtain. It should also be taken into account that the 3D-version of the NCI database has been constructed by means of the Corina packages, and the theoretically optimal conformation is not always the major isomere found during synthesis. For example, the highest ranking ligand, bicyclo[2.2.1]heptane-2,3,5,6-tetracarboxylic acid, is found in an endo-endo conformation in the NCI library, but is very difficult to synthesise in that conformation.

The 100 top scoring compounds were inspected, and duplicates eliminated which differed only in bound salts, and some similar substituents. Water insoluble compounds were also eliminated together with highly toxic or carcinogenic compounds. Of the remaining compounds, 7 were purchased commercially and 5 were chosen for synthesis. When tested for *in vitro* inhibition of TIM activity, the compounds Direct Red 23 and Direct Violet 51 (hereafter referred to as Direct Red and Direct Violet) which are commercial dyes, were the only compounds showing a decrease in activity in the $<100\mu$M range. Upon literature investigation of these and related compounds, it was found that a recent study had been performed by Gao *et al.* (1998), investigating the effect of these

compounds on reactivation of *Trypanosoma* and human TIM. The primary compound used in this study was Suramin, a drug used for sleeping sickness which was originally based on the trypanocidal effect of Trypan Red. Suramin is known to also affect a series of other enymes in the glycolytic pathway (Brunner *et al.*, 1996)(Gonzalez and Cazzulo, 1989)(Misset *et al.*, 1986) as well as tumor necrosis factor α (Alzani *et al.*, 1993). It was speculated that Suramin interfered either with TIM monomer folding or with monomer association. Although Suramin has been shown to inhibit TIM (Misset and Opperdoes, 1984), refolding was affected at concentrations $<1\mu$M, a concentration shown not to affect TIM activity. Preincubation with Suramin followed by dilution did not affect TIM activity. It was also found that the effect on TIM renaturation was protein concentration dependent, at concentrations above 1μg/ml the effect vanished. It was suggested that Suramin acted by the formation of Suramin-mTIM complexes which failed to dimerise, and occured only when the monomers had existed for a substantial time. When testing the related compound Direct Red, a 10x more potent inhibitory effect was found. They could show no common structural feature between all the inhibitors.

Our results indicate that Direct Red and Direct Violet do in fact bind in the active site of TIM, leading to uncompetitive inhibition of TIM activity. This suggests that these compounds may bind to the active site in such a way as to also disrupt the overall TIM conformation, thus interfering with successful dimerisation. A common TIM-bound three-dimensional structural motif could be shown for Direct Red and Direct Violet (Figure 4.8), and a similar region occurs on the other compounds shown to affect TIM reactivation, with Suramin and Direct Yellow showing the least similarity as well as the lowest activity (Figure 5.1). It is hoped that this may be the basis for the design of species-specific inhibitors.

Once the synthetic compounds have been screened, those showing inhibition of TIM activity will also be inspected for a structural motif similar to the one indicated above. This motif may then be consolidated into the minimal fragment capable of inhibiting TM activity. Chemical searches may be performed to find the existing compounds with structures closest to the desired motif. These may then be chemically modified to optimise a TIM inhibitor showing the highest inhibition of malaria TIM and the smallest effect on human TIM.

There are two basic metabolic methods for killing an organism, either

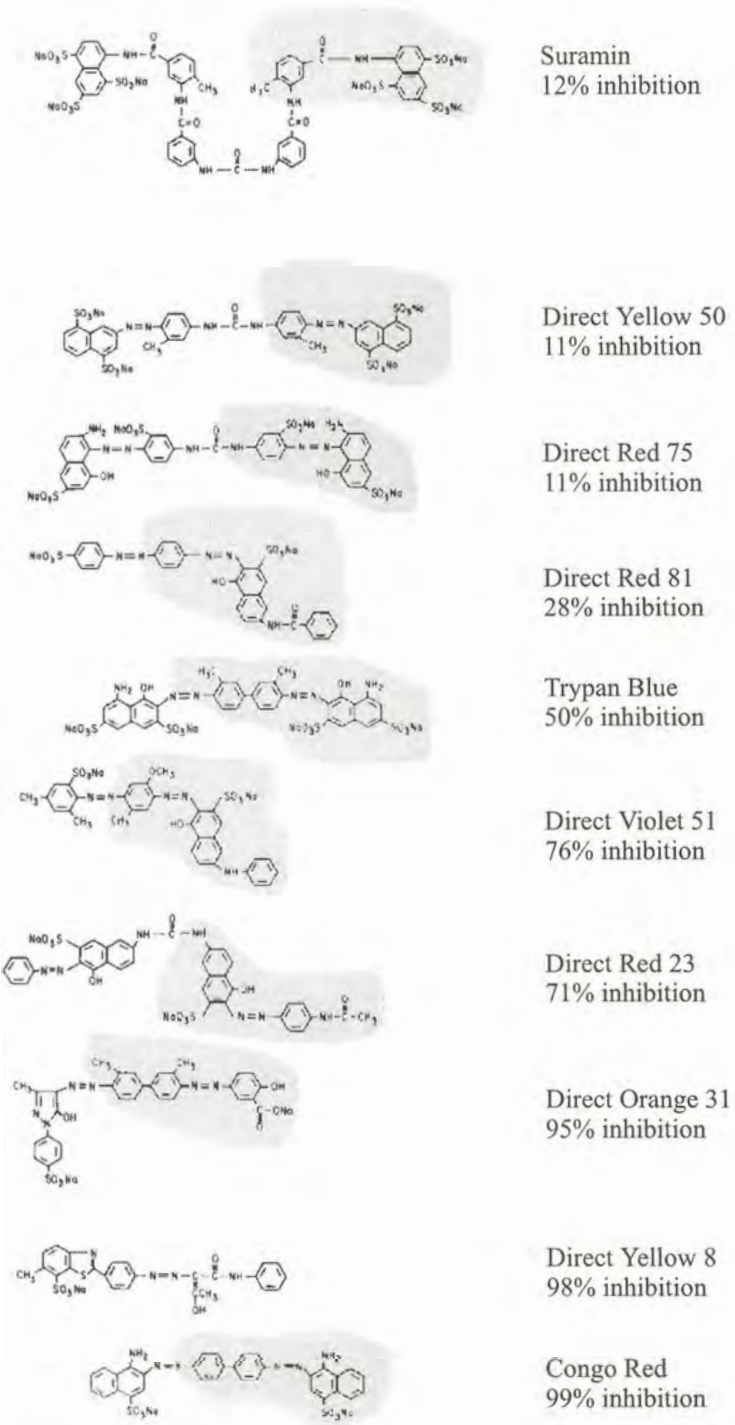


Figure 5.1: Structures of Suramin and related compounds. Percentage inhibition of *Trypanosoma* TIM reactivation at 10μM is indicated. Similar structural motifs are shaded in grey (Gao *et al.*, 1998).

the flux through an essential metabolic pathway may be decreased, or a metabolite concentration can be increased to toxic levels. Decrease of flux will normally require a tight-binding inhibitor with a significant flux control coefficient, while an uncompetitive inhibitor with a small flux control coefficient will lead to the increase of metabolites to toxic levels (Eisenthal and Cornish-Bowden, 1974). Thus, uncompetitive-type inhibitors are also of value as lead drug compounds. Of the glycolytic enzymes in *Trypanosoma*, Bakker *et al.* (1999) identified the glucose transporter to be the most important, followed by fructose-1.6-bisphosphate aldolase (ALD), glycerol-3-phosphate dehydrogenase (GDH), glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and phosphoglycerate kinase (PGK). In red blood cells, deficiencies of ALD, GAPDH and PGK did not cause any clinical symptoms. It has been shown that no organism can survive with a serious TIM deficiency, making it a good therapeutical target, but also indicating the dangers of TIM inhibitors for the human patient.

With the need for novel malaria drugs increasing on a daily basis, all studies showing promise of lead drugs should be followed up to as great an extent as possible. As drug companies are generally not interested in potential therapeutical targets until the *in vivo* value has been proven, compounds with a history of clinical applications such as the dyes described in this thesis should be regarded as serious contenders for the design of clinically useful agents.

Bibliography

Ahrweiler, P. M. and Frieden, C. (1991) Effects of point mutations in a hinge region on the stability, folding, and enzymatic activity of *Escherichia coli* dihydrofolate reductase. *Biochemistry* **30**(31), 7801–9.

Alzani, R., Corti, A., Grazioli, L., Cozzi, E., Ghezzi, P. and Marcucci, F. (1993) Suramin induces deoligomerization of human tumor necrosis factor alpha. *J Biol Chem* **268**(17), 12526–9.

Aqvist, J. and Fothergill, M. (1996) Computer simulation of the triosephosphate isomerase catalyzed reaction. *J Biol Chem* **271**(17), 10010–6.

Bakker, B. M., Michels, P. A., Opperdoes, F. R. and Westerhoff, H. V. (1999) What controls glycolysis in bloodstream form *Trypanosoma brucei*? *J Biol Chem* **274**(21), 14551–9.

Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., Wilson, I. A., Corran, P. H., Furth, A. J., Milman, J. D., Offord, R. E., Priddle, J. D. and Waley, S. G. (1975) Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 angstrom resolution using amino acid sequence data. *Nature* **255**(5510), 609–14.

Baratti, J., Maroux, S. and Louvard, D. (1973) Effect of ionic strength and calcium ions on the activation of trypsinogen by enterokinase. A modified test for the quantitative evaluation of this enzyme. *Biochim Biophys Acta* **321**(2), 632–8.

Bardosi, A., Eber, S. W., Hendrys, M. and Pekrun, A. (1990) Myopathy with altered mitochondria due to a triosephosphate isomerase (TPI) deficiency. *Acta Neuropathol* **79**(4), 387–94.

Basco, L. K., Eldin de Pecoulas, P., Wilson, C. M., Le Bras, J. and Mazabraud, A. (1995) Point mutations in the dihydrofolate reductase-

thymidylate synthase gene and pyrimethamine and cycloguanil resistance in *Plasmodium falciparum*. *Mol Biochem Parasitol* **69**(1), 135–8.

Bathurst, I. C. (1994) Protein expression in yeast as an approach to production of recombinant malaria antigens. *Am J Trop Med Hyg* **50**(4 Suppl), 20–6.

Beard, W. A., Appleman, J. R., Huang, S. M., Delcamp, T. J., Freisheim, J. H. and Blakley, R. L. (1991) Role of the conserved active site residue tryptophan-24 of human dihydrofolate reductase as revealed by mutagenesis. *Biochemistry* **30**(5), 1432–40.

Berman, P. A., Human, L. and Freese, J. A. (1991) Xanthine oxidase inhibits growth of *Plasmodium falciparum* in human erythrocytes *in vitro*. *J Clin Invest* **88**(6), 1848–55.

Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. E., J., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. and Tasumi, M. (1977) The Protein Data Bank: a computer-based archival file for macromolecular structures. *J Mol Biol* **112**(3), 535–42.

Bernstein, N. K., Cherney, M. M., Loetscher, H., Ridley, R. G. and James, M. N. (1999) Crystal structure of the novel aspartic proteinase zymogen proplasmepsin II from *Plasmodium falciparum*. *Nat Struct Biol* **6**(1), 32–7.

Bhasin, V. K. and Nair, L. (1996) *In vitro* selection of *Plasmodium falciparum* lines resistant to dihydrofolate-reductase inhibitors and cross resistance studies. *Jpn J Med Sci Biol* **49**(1), 1–14.

Birkholtz, L., Visser, L., Louw, A. I., van der Linde, R. and Brink, A. J. (1998) The prevalence of mixed-species and antifolate-resistant malaria infections in Mpumalanga. *S Afr Med J* **88**(1), 58–60.

Borchert, T. V., Abagyan, R., Jaenicke, R. and Wierenga, R. K. (1994) Design, creation, and characterization of a stable, monomeric triosephosphate isomerase. *Proc Natl Acad Sci U S A* **91**(4), 1515–8.

Brown, J. R., Daar, I. O., Krug, J. R. and Maquat, L. E. (1985) Characterization of the functional gene and several processed pseudogenes in the human triosephosphate isomerase gene family. *Mol Cell Biol* **5**(7), 1694–706.

Brueckner, R. P., Coster, T., Wesche, D. L., Shmuklarsky, M. and Schuster, B. G. (1998) Prophylaxis of *Plasmodium falciparum* infection in a human challenge model with WR 238605, a new 8-aminoquinoline antimalarial. *Antimicrob Agents Chemother* **42**(5), 1293–4.

Brunner, G., Zalkow, L., Burgess, E., Rifkin, D. B., Wilson, E. L., Gruszecka-Kowalik, E. and Powis, G. (1996) Inhibition of glycosylphosphatidylinositol (GPI) phospholipase D by suramin-like compounds. *Anticancer Res* **16**(5A), 2513–6.

Bzik, D. J., Fox, B. A. and Gonyer, K. (1993) Expression of *Plasmodium falciparum* lactate dehydrogenase in *Escherichia coli*. *Mol Biochem Parasitol* **59**(1), 155–66.

Bzik, D. J., Li, W. B., Horii, T. and Inselburg, J. (1987) Molecular cloning and sequence analysis of the *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase gene. *Proc Natl Acad Sci U S A* **84**(23), 8360–4.

Certa, U., Ghersa, P., Dobeli, H., Matile, H., Kocher, H. P., Shrivastava, I. K., Shaw, A. R. and Perrin, L. H. (1988) Aldolase activity of a *Plasmodium falciparum* protein with protective properties. *Science* **240**(4855), 1036–8.

Chen, Q., Shafer, R. H. and Kuntz, I. D. (1997) Structure-based discovery of ligands targeted to the RNA double helix. *Biochemistry* **36**(38), 11402–7.

Chunduru, S. K., Cody, V., Luft, J. R., Pangborn, W., Appleman, J. R. and Blakley, R. L. (1994) Methotrexate-resistant variants of human dihydrofolate reductase. Effects of Phe31 substitutions. *J Biol Chem* **269**(13), 9547–55.

Cody, V. and Ciszak, E. (1991) Computer graphic modeling in drug design–conformational analysis of antifolate binding to avian dihydrofolate reductase: crystal and molecular structures of 2,4-diamino-5-cyclohexyl-6-methylpyrimidine and 5-cyclohexyl-6-methyluracil. *Anticancer Drug Des* **6**(2), 83–93.

Cohen, F. E., Gregoret, L. M., Amiri, P., Aldape, K., Railey, J. and McKerrow, J. H. (1991) Arresting tissue invasion of a parasite by protease inhibitors chosen with the aid of computer modeling. *Biochemistry* **30**(47), 11221–9.

Craig, L. C., Pirtle, I. L., Gracy, R. W. and Pirtle, R. M. (1991) Characterization of the transcription unit and two processed pseudogenes of chimpanzee triosephosphate isomerase (TPI). *Gene* **99**(2), 217–27.

Creedon, K. A., Rathod, P. K. and Wellems, T. E. (1994) *Plasmodium falciparum* S-adenosylhomocysteine hydrolase. cDNA identification, predicted protein sequence, and expression in *Escherichia coli*. *J Biol Chem* **269**(23), 16364–70.

Davenport, R. C., Bash, P. A., Seaton, B. A., Karplus, M., Petsko, G. A. and Ringe, D. (1991) Structure of the triosephosphate isomerase-phosphoglycolohydroxamate complex: an analogue of the intermediate on the reaction pathway. *Biochemistry* **30**(24), 5821–6.

Deacon, H. E., Freese, J. A. and Sharp, B. L. (1994) Drug-resistant *Plasmodium falciparum* malaria in the eastern Transvaal. *S Afr Med J* **84**(7), 394–5.

Dobeli, H., Itin, C., Meier, B. and Certa, U. (1991) Is *Plasmodium falciparum* aldolase useful for rational drug design? *Acta Leiden* **60**(1), 135–40.

Dobeli, H., Trzeciak, A., Gillessen, D., Matile, H., Srivastava, I. K., Perrin, L. H., Jakob, P. E. and Certa, U. (1990) Expression, purification, biochemical characterization and inhibition of recombinant *Plasmodium falciparum* aldolase. *Mol Biochem Parasitol* **41**(2), 259–68.

Dunn, C. R., Banfield, M. J., Barker, J. J., Higham, C. W., Moreton, K. M., Turgut-Balik, D., Brady, R. L. and Holbrook, J. J. (1996) The structure of lactate dehydrogenase from *Plasmodium falciparum* reveals a new target for anti-malarial design. *Nat Struct Biol* **3**(11), 912–5.

Eisenthal, R. and Cornish-Bowden, A. (1974) The direct linear plot. A new graphical procedure for estimating enzyme kinetic parameters. *Biochem J* **139**(3), 715–20.

Elliott, J. F., Albrecht, G. R., Gilladoga, A., Handunnetti, S. M., Neequaye, J., Lallinger, G., Minjas, J. N. and Howard, R. J. (1990) Genes for *Plasmodium falciparum* surface antigens cloned by expression in COS cells. *Proc Natl Acad Sci U S A* **87**(16), 6363–7.

Engh, R. and Huber, R. (1991) Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Cryst* **A47**, 392–400.

Enomoto, S., Chen, G. and Berman, J. (1998) Vectors for expressing T7 epitope- and His6 affinity-tagged fusion proteins in S. cerevisiae. *Biotechniques* **24**(5), 782–6, 788.

Fasel, N., Begdadi-Rais, C., Bernard, M., Bron, C., Corradin, G. and Reymond, C. D. (1992) *Dictyostelium discoideum* as an expression host for the circumsporozoite protein of *Plasmodium falciparum*. *Gene* **111**(2), 157–63.

Foley, M., Corcoran, L., Tilley, L. and Anders, R. (1994) *Plasmodium falciparum*: mapping the membrane-binding domain in the ring- infected erythrocyte surface antigen. *Exp Parasitol* **79**(3), 340–50.

Gamper, A. M., Winger, R. H., Liedl, K. R., Sotriffer, C. A., Varga, J. M., Kroemer, R. T. and Rode, B. M. (1996) Comparative molecular field analysis of haptens docked to the multispecific antibody IgE(Lb4) [published erratum appears in J Med Chem 1997 Mar 14;40(6):1047-8]. *J Med Chem* **39**(20), 3882–8.

Gao, X. G., Garza-Ramos, G., Saavedra-Lira, E., Cabrera, N., De, G. m.-P. M. T., Perez-Montfort, R. and A, G. m.-P. (1998) Reactivation of triosephosphate isomerase from three trypanosomatids and human: effect of Suramin. *Biochem J* **332**(Pt 1), 91–6.

Garza-Ramos, G., Perez-Montfort, R., Rojo-Dominguez, A., de Gomez-Puyou, M. T. and Gomez-Puyou, A. (1996) Species-specific inhibition of homologous enzymes by modification of nonconserved amino acids residues. The cysteine residues of triosephosphate isomerase. *Eur J Biochem* **241**(1), 114–20.

Gasteiger, J., Rudolph, C. and Sadowski, J. (1990) Automatic generation of 3-D atomic coordinates for organic molecules. *Tetrahedron Comp. Method.* **3**, 537–547.

Gonzalez, N. S. and Cazzulo, J. J. (1989) Effects of trypanocidal drugs on protein biosynthesis *in vitro* and *in vivo* by *Trypanosoma cruzi*. *Biochem Pharmacol* **38**(17), 2873–7.

Goodsell, D. S., Morris, G. M. and Olson, A. J. (1996) Automated docking of flexible ligands: applications of AutoDock. *J Mol Recognit* **9**(1), 1–5.

Gopal, B., Ray, S. S., Gokhale, R. S., Balaram, H., Murthy, M. R. and Balaram, P. (1999) Cavity-creating mutation at the dimer interface

of *Plasmodium falciparum* triosephosphate isomerase: restoration of stability by disulfide cross-linking of subunits. *Biochemistry* **38**(1), 478–86.

Grall, M., Srivastava, I. K., Schmidt, M., Garcia, A. M., Mauel, J. and Perrin, L. H. (1992) *Plasmodium falciparum*: identification and purification of the phosphoglycerate kinase of the malaria parasite. *Exp Parasitol* **75**(1), 10–8.

Gschwend, D. A., Good, A. C. and Kuntz, I. D. (1996) Molecular docking towards drug discovery. *J Mol Recognit* **9**(2), 175–86.

Gschwend, D. A., Sirawaraporn, W., Santi, D. V. and Kuntz, I. D. (1997) Specificity in structure-based drug design: identification of a novel, selective inhibitor of *Pneumocystis carinii* dihydrofolate reductase. *Proteins* **29**(1), 59–67.

Hartman, F. C., LaMuraglia, G. M., Tomozawa, Y. and Wolfenden, R. (1975) The influence of pH on the interaction of inhibitors with triosephosphate isomerase and determination of the pKa of the active- site carboxyl group. *Biochemistry* **14**(24), 5274–9.

Hartman, F. C. and Ratrie, H. d. (1977) Apparent equivalence of the active-site glutamyl residue and the essential group with pKalpha 6.0 in triosephosphate isomerase. *Biochem Biophys Res Commun* **77**(2), 746–52.

Hassoun, H., Hanada, T., Lutchman, M., Sahr, K. E., Palek, J., Hanspal, M. and Chishti, A. H. (1998) Complete deficiency of glycophorin A in red blood cells from mice with targeted inactivation of the band 3 (AE1) gene. *Blood* **91**(6), 2146–51.

Hekmat-Nejad, M., Lee, P. C. and Rathod, P. K. (1997) *Plasmodium falciparum*: direct cloning and expression of pyrimethamine- sensitive and pyrimethamine-resistant dihydrofolate reductase domains. *Exp Parasitol* **85**(3), 303–5.

Hewitt, C. O., Sessions, R. B., Dafforn, T. R. and Holbrook, J. J. (1997) Protein engineering tests of a homology model of *Plasmodium falciparum* lactate dehydrogenase. *Protein Eng* **10**(1), 39–44.

Hoffman, L. R., Kuntz, I. D. and White, J. M. (1997) Structure-based identification of an inducer of the low-pH conformational change in the influenza virus hemagglutinin: irreversible inhibition of infectivity. *J Virol* **71**(11), 8808–20.

Holder, A. A. (1999) Malaria vaccines. *Proc Natl Acad Sci U S A* **96**(4), 1167–9.

Johnson, M. S., Srinivasan, N., Sowdhamini, R. and Blundell, T. L. (1994) Knowledge-based protein modeling. *Crit Rev Biochem Mol Biol* **29**(1), 1–68.

Joseph, D., Petsko, G. A. and Karplus, M. (1990) Anatomy of a conformational change: hinged "lid" motion of the triosephosphate isomerase loop. *Science* **249**(4975), 1425–8.

Joseph-McCarthy, D., Rost, L. E., Komives, E. A. and Petsko, G. A. (1994) Crystal structure of the mutant yeast triosephosphate isomerase in which the catalytic base glutamic acid 165 is changed to aspartic acid. *Biochemistry* **33**(10), 2824–9.

Juckett, G. (1999) Malaria prevention in travelers. *Am Fam Physician* **59**(9), 2523–30, 2535–6.

Kaslow, D. C. and Hill, S. (1990) Cloning metabolic pathway genes by complementation in *Escherichia coli*. Isolation and expression of *Plasmodium falciparum* glucose phosphate isomerase. *J Biol Chem* **265**(21), 12337–41.

Keough, D. T., Ng, A. L., Winzor, D. J., Emmerson, B. T. and de Jersey, J. (1999) Purification and characterization of *Plasmodium falciparum* hypoxanthine-guanine-xanthine phosphoribosyltransferase and comparison with the human enzyme. *Mol Biochem Parasitol* **98**(1), 29–41.

Kim, H., Certa, U., Dobeli, H., Jakob, P. and Hol, W. G. (1998) Crystal structure of fructose-1,6-bisphosphate aldolase from the human malaria parasite *Plasmodium falciparum*. *Biochemistry* **37**(13), 4388–96.

Kim, J. S. and Raines, R. T. (1993) Ribonuclease S-peptide as a carrier in fusion proteins. *Protein Sci* **2**(3), 348–56.

Krietsch, W. K., Pentchev, P. G., Klingenburg, H., Hofstatter, T. and Bucher, T. (1970) The isolation and crystallization of yeast and rabbit liver triose phosphate isomerase and a comparative characterization with the rabbit muscle enzyme. *Eur J Biochem* **14**(2), 289–300.

Krungkrai, J., Krungkrai, S. R. and Phakanont, K. (1992) Antimalarial activity of orotate analogs that inhibit dihydroorotase and dihydroorotate dehydrogenase. *Biochem Pharmacol* **43**(6), 1295–301.

Kumar, S. (1999) WHO gives southeast Asia a health warning. *Lancet* **354**(9183), 1010.

Kuntz, D. A., Osowski, R., Schudok, M., Wierenga, R. K., Muller, K., Kessler, H. and Opperdoes, F. R. (1992) Inhibition of triosephosphate isomerase from *Trypanosoma brucei* with cyclic hexapeptides. *Eur J Biochem* **207**(2), 441–7.

Lambeir, A. M., Opperdoes, F. R. and Wierenga, R. K. (1987) Kinetic properties of triose-phosphate isomerase from *Trypanosoma brucei brucei*. A comparison with the rabbit muscle and yeast enzymes. *Eur J Biochem* **168**(1), 69–74.

Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R. and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**(4), 477–86.

LaVallie, E. R., DiBlasio, E. A., Kovacic, S., Grant, K. L., Schendel, P. F. and McCoy, J. M. (1993) A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. *Biotechnology (N Y)* **11**(2), 187–93.

LeBlanc, S. B. and Wilson, C. M. (1993) The dihydroorotate dehydrogenase gene homologue of *Plasmodium falciparum*. *Mol Biochem Parasitol* **60**(2), 349–51.

Lemcke, T., Christensen, I. T. and Jorgensen, F. S. (1999) Towards an understanding of drug resistance in malaria: three- dimensional structure of *Plasmodium falciparum* dihydrofolate reductase by homology building. *Bioorg Med Chem* **7**(6), 1003–11.

Levitt, M. (1992) Accurate modeling of protein conformation by automatic segment matching. *J Mol Biol* **226**(2), 507–33.

Li, C. M., Tyler, P. C., Furneaux, R. H., Kicska, G., Xu, Y., Grubmeyer, C., Girvin, M. E. and Schramm, V. L. (1999) Transition-state analogs as inhibitors of human and malarial hypoxanthine-guanine phosphoribosyltransferases. *Nat Struct Biol* **6**(6), 582–7.

Li, L., Falzone, C. J., Wright, P. E. and Benkovic, S. J. (1992) Functional role of a mobile loop of *Escherichia coli* dihydrofolate reductase in transition-state stabilization. *Biochemistry* **31**(34), 7826–33.

Li, R., Chen, X., Gong, B., Selzer, P. M., Li, Z., Davidson, E., Kurzban, G., Miller, R. E., Nuzum, E. O., McKerrow, J. H., Fletterick, R. J., Gillmor, S. A., Craik, C. S., Kuntz, I. D., Cohen, F. E. and Kenyon, G. L. (1996) Structure-based design of parasitic protease inhibitors. *Bioorg Med Chem* **4**(9), 1421–7.

Lin, K. H. and Cheng, S. Y. (1991) An efficient method to purify active eukaryotic proteins from the inclusion bodies in *Escherichia coli*. *Biotechniques* **11**(6), 748, 750, 752–3.

Lodi, P. J., Chang, L. C., Knowles, J. R. and Komives, E. A. (1994) Triosephosphate isomerase requires a positively charged active site: the role of lysine-12. *Biochemistry* **33**(10), 2809–14.

Lodi, P. J. and Knowles, J. R. (1993) Direct evidence for the exploitation of an alpha-helix in the catalytic mechanism of triosephosphate isomerase. *Biochemistry* **32**(16), 4338–43.

Lolis, E., Alber, T., Davenport, R. C., Rose, D., Hartman, F. C. and Petsko, G. A. (1990) Structure of yeast triosephosphate isomerase at 1.9-A resolution. *Biochemistry* **29**(28), 6609–18.

Lolis, E. and Petsko, G. A. (1990) Crystallographic analysis of the complex between triosephosphate isomerase and 2-phosphoglycolate at 2.5-A resolution: implications for catalysis. *Biochemistry* **29**(28), 6619–25.

Lunn, C. A. and Pigiet, V. P. (1982) Localization of thioredoxin from *Escherichia coli* in an osmotically sensitive compartment. *J Biol Chem* **257**(19), 11424–30.

Mahmoudian, M. (1997) The cannabinoid receptor: computer-aided molecular modeling and docking of ligand. *J Mol Graph Model* **15**(3), 149–53, 179.

Marchionni, M. and Gilbert, W. (1986) The triosephosphate isomerase gene from maize: introns antedate the plant-animal divergence. *Cell* **46**(1), 133–41.

Mathews, C. and Van Holde, K. (1990) *Biochemistry*. Benjamin Cummings Publishing Company, New York, 1st edition.

Matsuoka, H., Kobayashi, J., Barker, G. C., Miura, K., Chinzei, Y., Miyajima, S., Ishii, A. and Sinden, R. E. (1996) Induction of anti-malarial transmission blocking immunity with a recombinant

ookinete surface antigen of *Plasmodium berghei* produced in silkworm larvae using the baculovirus expression vector system. *Vaccine* **14**(2), 120–6.

Milne, G. W., Nicklaus, M. C., Driscoll, J. S., Wang, S. and Zaharevitz, D. (1994) National Cancer Institute Drug Information System 3D database. *J Chem Inf Comput Sci* **34**(5), 1219–24.

Misset, O., Bos, O. J. and Opperdoes, F. R. (1986) Glycolytic enzymes of *Trypanosoma brucei*. Simultaneous purification, intraglycosomal concentrations and physical properties. *Eur J Biochem* **157**(2), 441–53.

Misset, O. and Opperdoes, F. R. (1984) Simultaneous purification of hexokinase, class-I fructose-bisphosphate aldolase, triosephosphate isomerase and phosphoglycerate kinase from *Trypanosoma brucei*. *Eur J Biochem* **144**(3), 475–83.

Moon, R. P., Tyas, L., Certa, U., Rupp, K., Bur, D., Jacquet, C., Matile, H., Loetscher, H., Grueninger-Leitch, F., Kay, J., Dunn, B. M., Berry, C. and Ridley, R. G. (1997) Expression and characterisation of plasmepsin I from *Plasmodium falciparum*. *Eur J Biochem* **244**(2), 552–60.

Mukhopadhyay, A. (1997) Inclusion bodies and purification of proteins in biologically active forms. *Adv Biochem Eng Biotechnol* **56**, 61–109.

Nakano, T., Spencer, H. T., Appleman, J. R. and Blakley, R. L. (1994) Critical role of phenylalanine 34 of human dihydrofolate reductase in substrate and inhibitor binding and in catalysis. *Biochemistry* **33**(33), 9945–52.

Olafsson, P. and Certa, U. (1994) Expression and cellular localisation of hexokinase during the bloodstage development of *Plasmodium falciparum*. *Mol Biochem Parasitol* **63**(1), 171–4.

Olafsson, P., Matile, H. and Certa, U. (1992) Molecular analysis of *Plasmodium falciparum* hexokinase. *Mol Biochem Parasitol* **56**(1), 89–101.

Old, S. E. and Mohrenweiser, H. W. (1988) Nucleotide sequence of the triosephosphate isomerase gene from *Macaca mulatta*. *Nucleic Acids Res* **16**(18), 9055.

Pearson, R. D. and Hewlett, E. L. (1987) Use of pyrimethamine-sulfadoxine (Fansidar) in prophylaxis against chloroquine-resistant

Plasmodium falciparum and *Pneumocystis carinii*. *Ann Intern Med* **106**(5), 714–8.

Pennisi, E. (1999) Malarial genome comes into view. *Science* **286**(5443), 1263; 1265.

Penny, J. I., Hall, S. T., Woodrow, C. J., Cowan, G. M., Gero, A. M. and Krishna, S. (1998) Expression of substrate-specific transporters encoded by *Plasmodium falciparum* in *Xenopus laevis* oocytes. *Mol Biochem Parasitol* **93**(1), 81–9.

Peterson, D. S., Milhous, W. K. and Wellems, T. E. (1990) Molecular basis of differential resistance to cycloguanil and pyrimethamine in *Plasmodium falciparum* malaria. *Proc Natl Acad Sci U S A* **87**(8), 3018–22.

Peterson, D. S., Walliker, D. and Wellems, T. E. (1988) Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase confers resistance to pyrimethamine in *falciparum* malaria. *Proc Natl Acad Sci U S A* **85**(23), 9114–8.

Plaut, B. and Knowles, J. R. (1972) pH-dependence of the triose phosphate isomerase reaction. *Biochem J* **129**(2), 311–20.

Pollack, Y., Shemer, R., Metzger, S., Spira, D. T. and Golenser, J. (1985) *Plasmodium falciparum*: expression of the adenine phosphoribosyltransferase gene in mouse L cells. *Exp Parasitol* **60**(3), 270–5.

Prapunwattana, P., Sirawaraporn, W., Yuthavong, Y. and Santi, D. V. (1996) Chemical synthesis of the *Plasmodium falciparum* dihydrofolate reductase- thymidylate synthase gene. *Mol Biochem Parasitol* **83**(1), 93–106.

Putman, S. J., Coulson, A. F., Farley, I. R., Riddleston, B. and Knowles, J. R. (1972) Specificity and kinetics of triose phosphate isomerase from chicken muscle. *Biochem J* **129**(2), 301–10.

Queen, S. A., Vander Jagt, D. and Reyes, P. (1988) Properties and substrate specificity of a purine phosphoribosyltransferase from the human malaria parasite, *Plasmodium falciparum*. *Mol Biochem Parasitol* **30**(2), 123–33.

Ranie, J., Kumar, V. P. and Balaram, H. (1993) Cloning of the triosephosphate isomerase gene of *Plasmodium falciparum* and expression in *Escherichia coli*. *Mol Biochem Parasitol* **61**(2), 159–69.

Rao, M. S. and Olson, A. J. (1999) Modelling of factor Xa-inhibitor complexes: a computational flexible docking approach. *Proteins* **34**(2), 173–83.

Raynes, K. (1999) Bisquinoline antimalarials: their role in malaria chemotherapy. *Int J Parasitol* **29**(3), 367–79.

Read, M., Hicks, K. E., Sims, P. F. and Hyde, J. E. (1994) Molecular characterisation of the enolase gene from the human malaria parasite *Plasmodium falciparum*. Evidence for ancestry within a photosynthetic lineage. *Eur J Biochem* **220**(2), 513–20.

Reynolds, C. A., Wade, R. C. and Goodford, P. J. (1989) Identifying targets for bioreductive agents: using GRID to predict selective binding regions of proteins. *J Mol Graph* **7**(2), 103–8, 100.

Richards, F. M., Wyckoff, H. W., Carlson, W. D., Allewell, N. M., Lee, B. and Mitsui, Y. (1972) Protein structure, ribonuclease-S and nucleotide interactions. *Cold Spring Harb Symp Quant Biol* **36**, 35–43.

Ritter, K., Kuhlencord, A., Thomssen, R. and Bommer, W. (1993) Prolonged haemolytic anaemia in malaria and autoantibodies against triosephosphate isomerase. *Lancet* **342**(8883), 1333–4.

Robson, K. J. and Jennings, M. W. (1991) The structure of the calmodulin gene of *Plasmodium falciparum*. *Mol Biochem Parasitol* **46**(1), 19–34.

Roth, E., J., Joulin, V., Miwa, S., Yoshida, A., Akatsuka, J., Cohen-Solal, M. and Rosa, R. (1988a) The use of enzymopathic human red cells in the study of malarial parasite glucose metabolism. *Blood* **71**(5), 1408–13.

Roth, E. F., J., Calvin, M. C., Max-Audit, I., Rosa, J. and Rosa, R. (1988b) The enzymes of the glycolytic pathway in erythrocytes infected with *Plasmodium falciparum* malaria parasites. *Blood* **72**(6), 1922–5.

Salas, F., Fichmann, J., Lee, G. K., Scott, M. D. and Rosenthal, P. J. (1995) Functional expression of falcipain, a *Plasmodium falciparum* cysteine proteinase, supports its role as a malarial hemoglobinase. *Infect Immun* **63**(6), 2120–5.

Sali, A. and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* **234**(3), 779–815.

Sanchez, R. and Sali, A. (1997) Evaluation of comparative protein structure modeling by MODELLER-3. *Proteins* **Suppl 1**, 50–8.

Sanchez, R. and Sali, A. (1998) Large-scale protein structure modeling of the *Saccharomyces cerevisiae* genome. *Proc Natl Acad Sci U S A* **95**(23), 13597–602.

Sano, G., Morimatsu, K. and Horii, T. (1994) Purification and characterization of dihydrofolate reductase of *Plasmodium falciparum* expressed by a synthetic gene in *Escherichia coli*. *Mol Biochem Parasitol* **63**(2), 265–73.

Sarma, P. S., Mandal, A. K. and Khamis, H. J. (1998) Allopurinol as an additive to quinine in the treatment of acute complicated *falciparum* malaria. *Am J Trop Med Hyg* **58**(4), 454–7.

Schapira, A., Solomon, T., Julien, M., Macome, A., Parmar, N., Ruas, I., Simao, F., Streat, E. and Betschart, B. (1993) Comparison of intramuscular and intravenous quinine for the treatment of severe and complicated malaria in children. *Trans R Soc Trop Med Hyg* **87**(3), 299–302.

Schlagenhauf, P. (1999) Mefloquine for malaria chemoprophylaxis 1992-1998: a review. *J Travel Med* **6**(2), 122–33.

Schliebs, W., Thanki, N., Eritja, R. and Wierenga, R. (1996) Active site properties of monomeric triosephosphate isomerase (monoTIM) as deduced from mutational and structural studies. *Protein Sci* **5**(2), 229–39.

Schnackerz, K. D. and Gracy, R. W. (1991) Probing the catalytic sites of triosephosphate isomerase by 31P-NMR with reversibly and irreversibly binding substrate analogues. *Eur J Biochem* **199**(1), 231–8.

Scott, H. V., Rieckmann, K. H. and O'Sullivan, W. J. (1987) Synergistic antimalarial activity of dapsone/dihydrofolate reductase inhibitors and the interaction of antifol, antipyrimidine and antipurine combinations against *Plasmodium falciparum in vitro*. *Trans R Soc Trop Med Hyg* **81**(5), 715–21.

Sherman, I. W. (1979) Biochemistry of *Plasmodium* (malarial parasites). *Microbiol Rev* **43**(4), 453–95.

Shi, W., Li, C. M., Tyler, P. C., Furneaux, R. H., Cahill, S. M., Girvin, M. E., Grubmeyer, C., Schramm, V. L. and Almo, S. C. (1999) The

2.0 A structure of malarial purine phosphoribosyltransferase in complex with a transition-state analogue inhibitor. *Biochemistry* **38**(31), 9872–80.

Sirawaraporn, W., Prapunwattana, P., Sirawaraporn, R., Yuthavong, Y. and Santi, D. V. (1993) The dihydrofolate reductase domain of *Plasmodium falciparum* thymidylate synthase-dihydrofolate reductase. Gene synthesis, expression, and anti- folate-resistant mutants. *J Biol Chem* **268**(29), 21637–44.

Sirawaraporn, W., Sirawaraporn, R., Cowman, A. F., Yuthavong, Y. and Santi, D. V. (1990) Heterologous expression of active thymidylate synthase-dihydrofolate reductase from *Plasmodium falciparum*. *Biochemistry* **29**(48), 10779–85.

Sirawaraporn, W. and Yuthavong, Y. (1986) Potentiating effect of pyrimethamine and sulfadoxine against dihydrofolate reductase from pyrimethamine-sensitive and pyrimethamine- resistant *Plasmodium chabaudi*. *Antimicrob Agents Chemother* **29**(5), 899–905.

Smith, L. S., Lewis, T. L. and Matsui, S. M. (1995) Increased yield of small DNA fragments purified by silica binding. *Biotechniques* **18**(6), 970–2, 974–5.

Song, J. K., Kim, M. K. and Rhee, J. S. (1999) Cloning and expression of the gene encoding phospholipase A1 from *Serratia* sp. MK1 in Escherichia coli. *J Biotechnol* **72**(1-2), 103–14.

Srivastava, I. K., Schmidt, M., Grall, M., Certa, U., Garcia, A. M. and Perrin, L. H. (1992) Identification and purification of glucose phosphate isomerase of *Plasmodium falciparum*. *Mol Biochem Parasitol* **54**(2), 153–64.

Steczko, J., Hermodson, M., Axelrod, B. and Dziembor-Kentzer, E. (1983) Identification of the target amino acids in the site-specific inactivation of triose phosphate isomerase by ferrate anion. *J Biol Chem* **258**(21), 13148–54.

Steketee, R. W., Wirima, J. J., Slutsker, W. L., Khoromana, C. O., Breman, J. G. and Heymann, D. L. (1996) Objectives and methodology in a study of malaria treatment and prevention in pregnancy in rural Malawi: The Mangochi Malaria Research Project. *Am J Trop Med Hyg* **55**(1), 8–16.

Straus, D. and Gilbert, W. (1985) Genetic engineering in the Precambrian: structure of the chicken triosephosphate isomerase gene. *Mol Cell Biol* **5**(12), 3497–506.

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 Enzymol* **185**, 60–89.

Subbayya, I. N., Ray, S. S., Balaram, P. and Balaram, H. (1997) Metabolic enzymes as potential drug targets in *Plasmodium falciparum*. *Indian J Med Res* **106**, 79–94.

Sudbeck, E. A., Liu, X. P., Narla, R. K., Mahajan, S., Ghosh, S., Mao, C. and Uckun, F. M. (1999) Structure-based design of specific inhibitors of Janus kinase 3 as apoptosis-inducing antileukemic agents. *Clin Cancer Res* **5**(6), 1569–82.

Szklarz, G. D. and Halpert, J. R. (1997) Use of homology modeling in conjunction with site-directed mutagenesis for analysis of structure-function relationships of mammalian cytochromes P450. *Life Sci* **61**(26), 2507–20.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**(22), 4673–80.

Thompson, P. D. and Freisheim, J. H. (1991) Conversion of arginine to lysine at position 70 of human dihydrofolate reductase: generation of a methotrexate-insensitive mutant enzyme. *Biochemistry* **30**(33), 8124–30.

Touze, J. E., Fourcade, L., Peyron, F., Heno, P. and Deharo, J. C. (1997) Is halofantrine still advisable in malaria attacks? *Ann Trop Med Parasitol* **91**(7), 867–73.

Toyoda, T., Brobey, R. K., Sano, G., Horii, T., Tomioka, N. and Itai, A. (1997) Lead discovery of inhibitors of the dihydrofolate reductase domain of *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase. *Biochem Biophys Res Commun* **235**(3), 515–9.

Vasanthakumar, G., Davis, R. L., J., Sullivan, M. A. and Donahue, J. P. (1990) Cloning and expression in *Escherichia coli* of a hypoxanthine-guanine phosphoribosyltransferase-encoding cDNA from *Plasmodium falciparum*. *Gene* **91**(1), 63–9.

Velanker, S. S., Ray, S. S., Gokhale, R. S., S, S., Balaram, H., Balaram, P. and Murthy, M. R. (1997) Triosephosphate isomerase from *Plasmodium falciparum*: the crystal structure provides insights into antimalarial drug design. *Structure* **5**(6), 751–61.

Verlinde, C. L., Witmans, C. J., Pijning, T., Kalk, K. H., Hol, W. G., Callens, M. and Opperdoes, F. R. (1992) Structure of the complex between trypanosomal triosephosphate isomerase and N-hydroxy-4-phosphono-butanamide: binding at the active site despite an "open" flexible loop conformation. *Protein Sci* **1**(12), 1578–84.

Volz, K. W., Matthews, D. A., Alden, R. A., Freer, S. T., Hansch, C., Kaufman, B. T. and Kraut, J. (1982) Crystal structure of avian dihydrofolate reductase containing phenyltriazine and NADPH. *J Biol Chem* **257**(5), 2528–36.

Waley, S. G. (1973) Refolding of triose phosphate isomerase. *Biochem J* **135**(1), 165–72.

Warren, M. S., Brown, K. A., Farnum, M. F., Howell, E. E. and Kraut, J. (1991) Investigation of the functional role of tryptophan-22 in *Escherichia coli* dihydrofolate reductase by site-directed mutagenesis. *Biochemistry* **30**(46), 11092–103.

Weber, J. L. (1987) Analysis of sequences from the extremely A + T-rich genome of *Plasmodium falciparum*. *Gene* **52**(1), 103–9.

Webster, H. K., Whaun, J. M., Walker, M. D. and Bean, T. L. (1984) Synthesis of adenosine nucleotides from hypoxanthine by human malaria parasites (*Plasmodium falciparum*) in continuous erythrocyte culture: inhibition by hadacidin but not alanosine. *Biochem Pharmacol* **33**(9), 1555–7.

White, N. J. and Olliaro, P. (1998) Artemisinin and derivatives in the treatment of uncomplicated malaria. *Med Trop (Mars)* **58**(3 Suppl), 54–6.

Wierenga, R. K., Hol, W. G., Misset, O. and Opperdoes, F. R. (1984) Preliminary crystallographic studies of triosephosphate isomerase from the blood parasite *Trypanosoma brucei brucei*. *J Mol Biol* **178**(2), 487–90.

Wooden, J. M., Hartwell, L. H., Vasquez, B. and Sibley, C. H. (1997) Analysis in yeast of antimalaria drugs that target the dihydrofolate

reductase of *Plasmodium falciparum*. *Mol Biochem Parasitol* **85**(1), 25–40.

Yasukawa, T., Kanei-Ishii, C., Maekawa, T., Fujimoto, J., Yamamoto, T. and Ishii, S. (1995) Increase of solubility of foreign proteins in *Escherichia coli* by coproduction of the bacterial thioredoxin. *J Biol Chem* **270**(43), 25328–31.

You, J., Cohen, R. E. and Pickart, C. M. (1999) Construct for high-level expression and low misincorporation of lysine for arginine during expression of pET-encoded eukaryotic proteins in *Escherichia coli*. *Biotechniques* **27**(5), 950–4.

Zhang, Z., Sugio, S., Komives, E. A., Liu, K. D., Knowles, J. R., Petsko, G. A. and Ringe, D. (1994) Crystal structure of recombinant chicken triosephosphate isomerase- phosphoglycolohydroxamate complex at 1.8-A resolution. *Biochemistry* **33**(10), 2830–7.

Zolg, J. W., Plitt, J. R., Chen, G. X. and Palmer, S. (1989) Point mutations in the dihydrofolate reductase-thymidylate synthase gene as the molecular basis for pyrimethamine resistance in *Plasmodium falciparum*. *Mol Biochem Parasitol* **36**(3), 253–62.