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.