PURIFICATION AND CHARACTERIZATION OF *Plasmodium falciparum* HYPOXANTHINE PHOSPHORIBOSYLTRANSFERASE

Edwin Kimathi Murungi

A minithesis submitted in partial fulfilment of the requirements for the degree of Magister Scientiae (Structural Biology) at the University of the Western Cape.

Supervisors:

Associate Professor David B. McIntosh
Associate Professor Trevor B. Sewell

April 2007

KEYWORDS

Plasmodium falciparum

Hypoxanthine Phosphoribosyltransferase

Bioinformatics

Photoaffinity Labeling

Protein

Oligomeric Structure

Crystallization

Homology Modeling

ABSTRACT

PURIFICATION AND CHARACTERIZATION OF *Plasmodium falciparum* HYPOXANTHINE PHOSPHORIBOSYLTRANSFERASE

Edwin Kimathi Murungi

MSc minithesis, Faculty of Natural Sciences, University of the Western Cape and Faculty of Health Sciences, University of Cape Town.

Malaria remains the most important parasitic disease worldwide. It is estimated that over 500 million infections and more than 2.7 million deaths arising from malaria occur each year. Most (90%) of the infections occur in Africa with the most affected groups being children of less than five years of age and women. This dire situation is exacerbated by the emergence of drug resistant strains of *Plasmodium falciparum*. As a result, novel antimalarials are urgently needed to combat the disease.

The purine salvage enzyme, hypoxanthine phosphoribosyltransferase of *Plasmodium falciparum* (PfHPRT) catalyses the conversion of phosphoribosyl pyrophosphate and hypoxanthine (or guanine) into inosine (or guanosine) monophosphate and pyrophosphate. The nucleotides are needed for DNA and RNA synthesis. As such, the enzyme is essential for parasite survival and is a recognized drug target. Previous work in this laboratory developed methodology for expressing the protein in *Escherichia coli* and partial purification. Also, the laboratory has identified a group of chalcones that accelerate catalytic turnover, activate an inactive to active transition of the enzyme, and potently inhibit enzyme turnover under select conditions. One objective has been to crystallize the chalcone/enzyme complex.

The work reported in this thesis focuses on improving the purification of PfHPRT by investigating the characteristics of anion exchange DE-52 chromatography (the first stage of the purification), developing an HPLC gel filtration method for examining the quaternary structure of the protein and possible end stage purification, and initial

crystallization trials. A homology model of the open, unliganded PfHPRT is constructed using the atomic structures of human, *T. cruzi* and *S. typhimurium* HPRT as templates. A method for the synthesis of [γ-^{32}P]TNP-8N_3-ITP is developed for use as a photoaffinity probe of the active site.

The original purification protocol involved a relatively short DE-52 column placed in tandem with a Reactive Red 120 column. PfHPRT does not bind to the first column, but binds to the second and can be eluted with pyrophosphate. This study found that about half of the PfHPRT was eluted during the wash stage and that the pH rose to about 11. A well-equilibrated column maintained the pH at 8.9, but had little effect on the protein elution characteristics. Increasing the length of the column and thoroughly equilibrating the resin resulted in the elution of a rather pure PfHPRT component without retardation on the column. The subsequent Reactive Red 120 chromatography was found to be effective in eliminating key contaminants and undesirable forms of PfHPRT.

HPLC gel filtration was found to be effective in distinguishing between tetrameric and dimeric forms of PfHPRT, the equilibrium of these two forms dependent on salt concentration, with the tetramer stable at low salt and the dimer predominating at high salt. Preparations of PfHPRT were shown to contain significant amounts of undesirable dimer even at low salt concentrations. HPLC gel filtration was shown to be an effective method of purifying tetrameric PfHPRT on a preparative scale.

Attempts were made to crystallize PfHPRT alone, in complex with a chalcone, GMP or IMP. Despite varying the concentrations of ligands, precipitant, pH, divalent cations, and protein, and additionally employing various crystallization methods such as vapour diffusion (both the hanging drop and the sitting drop) and microbatch, no protein crystals were obtained. Instead, numerous salt/ligand crystals were obtained as judged by methylene blue dye test, prick test and diffraction patterns.

Only one crystal structure of PfHPRT complexed with a transition state analog inhibitor, ImmucillinHP, PPi, and two Mg^{2+} has been solved. The substrates and Mg^{2+} are tightly

enclosed by several protein loops. A model of an open, unliganded form of PfHPRT was constructed based on sequence homology to the human, *T. cruzi* and *S. typhimurium* HPRT using the program MODELLER. The active site of the open structure is much larger and will be useful for docking of inhibitors, such as the chalcone.

Synthesis of 2',3',4'-*O*-(2,4,6-trinitrophenyl)-8-azidoinosine triphosphate (TNP-8N_3-ITP) was achieved by nitrous acid-induced deamination of 2',3',4'-*O*-(2,4,6-trinitrophenyl)-8-azidoadenosine triphosphate (TNP-8N_3-ATP). The yield was good enough to effect the synthesis of highly radioactive [γ-^{32}P]TNP-8N_3-ITP with minimal purification. Photoaffinity labeling experiments showed that both the adenine and inosine probes bound with high affinity ($K_{0.5}$ = 2-5 μM), but derivatized the protein with rather poor efficiency. Nevertheless, the probes could be useful for determining the K_d values for inhibitor binding to the active site through competitive inhibition of probe binding.

Overall, the impurity of the protein may explain why no crystals were obtained.

April 2007

DECLARATION

I declare that ***Purification and Characterization of Plasmodium falciparum Hypoxanthine Phosphoribosyltransferase*** is my own work, that it has not been submitted before for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged by complete references.

Edwin Kimathi Murungi 16 April 2007

Signed:

ACKNOWLEDGEMENTS

I wish to sincerely thank my supervisor, Assoc. Prof. David B. McIntosh, for his teaching and guidance throughout the course of this work. Thank you too for financial assistance.

My sincerest thanks also to my co-supervisor, Assoc. Prof. Trevor B. Sewell, for introducing me to structural biology.

I wish to acknowledge my indebtedness to the following:
Mr. David Woolley for help with the photoaffinity experiments and excellent technical support.
My undergraduate mentor Faith Okalebo for inspiring me to pursue postgraduate studies.
My parents Mr. and Mrs. Murungi for your love and the many sacrifices you've made to get me here.
My siblings Wambui, Mbae, Gaiti and Mercy for urging me on when things were tough.
Moni, for all things wonderful.
My cousin Foi for being the best *cuzo* ever!
Dee, for your prayers, understanding and presence.
Lenganji, for your unwaving friendship and prayers.
Godfrey Wanjohi, Mohammed Wanyoike and Dan Mwaura for endless support and encouragement.
Kariuki, Davis, Timo, Ng'ang'a, Jane, Carol, Owino, Babra, Evelyn, Ndoriah, Muchai and Gitari for being the best buddies I would have wish for in South Africa.
My support crew back in Kenya, Torosh, Ongarora, Jennifer, and Steff. You rock!
My fellow students in the Division of Chemical Pathology; Naadia, Jackson, Vuyo and Gabz. Thank you for your encouragement.
Am so thankful to my Lord and Saviour Jesus Christ for many blessings.
Financial support from Carnegie Corporation of New York is gratefully acknowledged.

ABBREVIATIONS

Ala	Alanine
AMP	adenosine 5′-monophosphate
Arg	Arginine
Asn	Asparagine
Asp	Aspartate
ATP	adenosine 5′-triphosphate
cDNA	complementary DNA
CQ	Chloroquine
Cys	Cysteine
DDT	Dichloro-Diphenyl-Trichloroethane
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease 1
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EPPS	(N-[2-Hydroxyethyl]piperazine-N′-[3-propane-sulfonic acid]
Gln	Glutamine
Glu	Glutamate
Gly	Glycine
GMP	guanosine 5′-monophosphate
His	Histidine
HPLC	high performance liquid chromatography
HPRT	hypoxanthine (-guanine) phosphoribosyltranferase
PfHPRT	*Plasmodium falciparum* hypoxanthine (-guanine) phosphoribosyltransferase
Ile	Isoleucine
IPTG	β-D-isopropyl-thiogalactopyranoside
KPi	KH_2PO_4/K_2HPO_4
LB	Luria-Bertani medium (Luria broth)
Leu	Leucine

Lys	Lysine
Met	Methionine
MgPPi	magnesium pyrophosphate
MOPS	3-[N-morpholino]propanesulfonic acid
Mw	molecular weight
N_3	azide
PDB	Protein Data Bank
Phe	Phenylalanine
PMSF	phenylmethylsulfonyl fluoride
PPi	inorganic pyrophosphate
Pro	Proline
PRPP	5-phosphorylribose 1-pyrophosphate
PSIPRED	Protein Structure Prediction
RMSD	root mean square deviation
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
Ser	Serine
Thr	Threonine
TMAH	Tetramethylammonium hydroxide
TNP	2′,3′,-O-(2,4,6-trinitrophenyl)
TNP-ATP	TNP-adenosine 5′-triphosphate
TNP-GMP	TNP-guanosine 5′-monophosphate
TNP-ITP	TNP-inosine 5′- triphosphate
Trp	Tryptophan
Tyr	Tyrosine
UV	Ultraviolet
Val	Valine

TABLE OF CONTENTS

CHAPTER ONE: LITERATURE REVIEW

1.1 Introduction

Malaria is the most important of the parasitic diseases of humans. More than three billion people live in malarious areas and the disease causes between one and three million deaths each year (Breman *et al.,* 2004). Snow and co-workers (2005) further estimate the global falciparum malaria morbidity burden to be about 515 million cases. Sub-Saharan Africa suffers the vast majority (90%) of this toll with children below 5 years of age and pregnant women being the most affected groups (Breman, 2001). Pregnant women are highly susceptible largely because the transient depression of their cell-mediated immunity that allows retention of the fetal allograft also interferes with resistance to infections (Ismaili *et al.,* 2003). Resurgence of the disease in many parts of the tropics, mainly due to the parasites' resistance to drugs and the vectors' resistance to insecticides, together with the ever looming danger of importation to nonmalarious countries aggravates the situation. Other factors implicated in the increase in malaria burden, especially in Africa, include war and civil disturbance, environmental changes, climatic changes and population increase (Greenwood and Mutabingwa, 2002).

The disease remains a growing threat to the economic development of countries in the tropical and subtropical regions of the world. Majority of countries with endemic malaria are poor with low annual economic growth rates. Miller *et al.* (2002) argue that malaria is a disease of poverty and underdeveloped countries. This view is echoed by Sachs and Malaney (2002) who show a correlation between malaria and poverty. Results obtained

earlier by Gallup and Sachs (2001) show that between 1965 and 1990, countries with a high intensity of malaria experienced an average growth in per-capita GDP of 0.4% per year compared with 2.3% for non-malarious countries.

Human malaria is caused by infection with intracellular parasites of the genus *Plasmodium*, namely *Plasmodium falciparum, Plasmodium vivax, Plasmodium ovale and Plasmodium malariae*, which are transmitted by female *Anopheles gambiae* mosquitoes (Clark *et al.,* 2004). The *Anopheles gambiae* complex of species has a remarkable tendency towards human biting (anthropophily). *Plasmodium falciparum* and, to a much lesser extent, *Plasmodium vivax* accounts for the majority of instances of morbidity and mortality (Miller *et al.,* 2002). Exposure to infected blood products (transfusion malaria) and congenital transmission (Trampuz *et al.,* 2003) are alternative routes of infection.

1.2 Life cycle of Plasmodium

Although the different species of *Plasmodium* exhibit genetic and phenotypic differences, their life cycles are similar (Khan and Waters, 2004). The life cycle of *P. falciparum*, shown in Figure 1.0, involves two hosts. During a blood meal, mosquitoes inject sporozoites into the human subcutaneous tissue, and less-frequently directly into the blood stream from where they travel to the liver (Miller *et al.,* 2002). In the liver, sporozoites move through several hepatocytes before infecting the ultimate one with the formation of a vacuole around the sporozoites (Mota *et al.,* 2001).

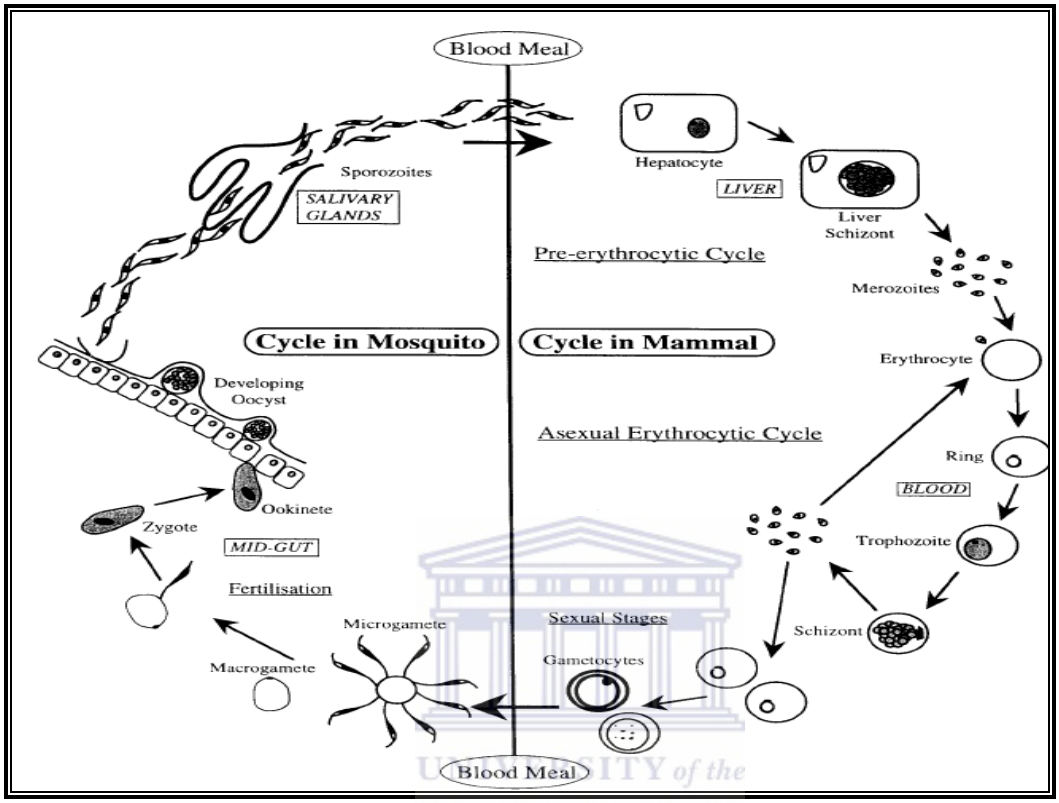


Figure 1.0: The life cycle of the malaria parasite (adapted from Phillips, 2001)

The migration through host cells induces the Ca^{2+} dependent exocytosis of sporozoite's apical organelles (rhoptries, micronemes and dense granules), a prerequisite for infection (Mota *et al.,* 2002). Host cells traversed by sporozoites would in turn be recognized by the host's immune system as being infected, thus creating false targets. This mechanism may aid the parasite to escape the host's immune response (Mota and Rodriguez, 2001). Inside the hepatocytes, sporozoites develop into mature schizonts (tissue schizogony) which rupture and release merozoites that infect erythrocytes. In *P.vivax* and *P.ovale* a

dormant hepatocyte stage (hypnozoites) can persist and cause relapses weeks or years later. Binding of parasite ligands to erythrocytic receptors is a prerequisite to invasion of the erythrocytes by the parasite. Numerous molecules, (including glycophorin A, B and C, EBA-175), implicated in the invasion process have been identified on both the merozoite and erythrocyte surfaces (Gaur *et al.,* 2004). The invasion process proceeds as follows; first, merozoites reversibly attach to the erythrocyte surface and reorient their apices. This is followed by the formation of an irreversible junction, a parasitophorous vacuole, entry into the vacuole by movement of the junction and finally resealing of the vacuolar and erythrocytic membranes (Pasvol, 2003; Gaur *et al.,* 2004). Within the erythrocytes, merozoites undergo asexual multiplication (erythrocytic schizogony) whereby they develop through the ring, trophozoite and schizont stages. Upon rupture, schizonts release merozoites that infect fresh erythrocytes. Approximately 15 hours into this asexual reproductive phase, there is a gradual increase in the permeability of the erythrocyte plasma membrane to a range of small molecules and ions. (Kirk, 2004; Staines *et al.,* 2005). This occurs due to the induction of new permeation pathways (NPP) by the internal parasite. These pathways aid in the uptake of nutrients and removal of both lactate and excess amino acids from the erythrocyte (Staines *et al.,* 2005).

Some merozoites differentiate into gametocytes (Trampuz *et al.,* 2003). Male (microgametocytes) and female (macrogametocytes) are ingested by the insect vector during a blood meal. Microgametes differentiate into several individual gametes

(exflagellating of microgametocyte). Gamete formation is followed by fertilization, resulting in the diploid zygote. After one round of meiotic division, the zygote develops into a motile ookinete that penetrates the cells of the midgut and traverses the midgut wall to form an oocyst on the basolateral lamina. Sporozoites develop within the oocyst and, on release, migrate via the haemolymph to invade the salivary glands. When the slender, haploid sporozoites are inoculated into a new human host, they invade the hepatocytes within minutes and develop into exoerythrocytic forms (liver schizonts) which rupture and release merozoites that infect erythrocytes and perpetuate the cycle.

1.3 The pathogenesis of malaria

Malaria presents as a complex syndrome affecting many organs. Central to the pathology of the disease are the repeated cycles of parasite invasion and destruction of human erythrocytes. In response, host mononuclear cells secrete cytokines, such as tissue necrotic factor (TNF), pro-inflammatory interleukins, oxygen free radicals and nitric oxide (Maitland and Marsh, 2004) which result in damage to host endothelium and tissues.

Microvascular occlusion in falciparum malaria is related to two properties of the parasite: adhesion of parasitized erythrocytes to endothelial cells-cytoadherence, and agglutination of nonparasitized erythrocytes around erythrocytes containing trophozoites and schizonts of *P. falciparum*- rosetting (Udomsangpetch *et al.,* 1992). Several host ligands have been identified as potential mediators of cytoadhesion and include thrombospondin (TSP)

(Roberts *et al.,* 1985) endothelial leukocyte adhesion molecule 1 (ELAM-1), vascular cell adhesion molecule (VCAM-1) (Ockenhouse *et al.,* 1992), platelet GP1V (CD36), platelet endothelial cell adhesion molecule 1 (CD31), chrondroitin sulfate A and hyaluronic acid and intercellular adhesion molecule 1 (ICAM-1) (Berendt *et al.,* 1989). Infected erythrocytes may rosette uninfected ones using complement receptor 1 (CD35). Cytoadherence and rosetting result in the formation of erythrocyte aggregates and intra vascular sequestration of erythrocytes in vital organs such as the brain, heart and placenta. This interferes with microcirculation and metabolism and allows parasite development away from the host defense system, a phenomenon referred to as splenic processing and filtration (Reeder *et al.,* 1999).

Infection results in a wide spectrum of clinical manifestations which are primarily due to schizont rupture and destruction of erythrocytes. Symptoms include fever, chills, headaches, diaphoresis, dizziness, malaise, myalgia, abdominal pain, nausea, vomiting, mild diarrhea and dry cough (Trampuz *et al.,* 2003). Additional signs include tachycardia, jaundice, pallor, orthostatic hypotension, hepatomegally and splenomegally. The major complications of severe malaria include cerebral malaria, pulmonary edema, acute renal failure, severe anemia, metabolic acidosis, hypoglycemia (Garg, 2000), circulatory collapse, repeated generalized convulsions, hyperbilirubinemia and hypovolaemia (WHO, 2000).

1.4 Drugs used in malaria treatment

Preventive and/or curative drug therapies play an important role in the control malaria (Gregson and Plowe, 2005). Although drugs in use target different stages of the malaria life cycle, majority of them act on the asexual, intra-erythrocytic stage: the phase responsible for the clinical symptoms of disease (Maitland *et al.,* 2004). The choice of an anti-malarial agent largely depends on the patient's level of immunity, the drug's side effect profile, cost and the site where the infection was acquired - an indicator of a particular drug's resistance probability (White, 2004). Quinolines, folate antagonists, antibiotics, artemisinin derivatives and hydroxynapthoquinones are the main classes of antimalarial drugs. Each class of drugs exhibit unique properties as explained below.

1.4.1 Quinolines

This class is comprised of: (i) the 4-aminoquinolines (chloroquine and amodiaquine), (ii) the quinoline methanols (quinine, quinidine and mefloquine), (iii) the phenathrene methanols (halofantrine and lumefantrine). Drugs in this class exert their action on the parasite's food vacuole (Maitland *et al.,* 2004). Predominantly in the trophozoite and early schizont stages, hemoglobin is ingested with the cytoplasm of the host erythrocyte by a phagocytosis-like mechanism into the food vacuole where it is degraded to generate free heme (Ginsburg *et al.,* 1999). Free heme is potentially toxic through an oxidative mechanism and is therefore disposed of by conversion into long insoluble polymers of hemozoin (heme polymerization). An alternative heme degradation pathway is through it's reaction with hydrogen peroxide generated by the spontaneous oxidation of the

released heme from a ferrous to a ferric state (Loria *et al.,* 1999). Famin *et al.* (1999) have also found out that some of the heme diffuses into the cytoplasm of the parasite where it is likely destroyed by reduced glutathione. Heme polymerization and the oxidative and glutathione-dependent heme degradation pathways are inhibited by the quinolines. They bind to heme through a π-π stacking of their planar aromatic structures (Ridley *et al.,* 1997). Moreover, mefloquine and quinine have been shown to block the uptake of hemoglobin from the host cell (Famin and Ginsburg, 2002). The activity of quinolines seems to depend on the weak base effect whereby presence of a basic amino function enables concentration of the drug in the acidic food vacuole in its membrane-impermeable protonated form (Ginsburg *et al.,* 1989).

Members of this class are widely used in malaria treatment as exemplified by the use of chloroquine as a first line agent in many African countries for a long time before the emergence of chloroquine resistance (Winstanley, 2000). Although generally safe, several adverse effects which include pruritis, retinopathy, shock, cardiac arrhythmias, psoriasis, leukopenia and aplastic anaemia have been reported for chloroquine (Khoo *et al.*, 2005). Amodiaquine through its active metabolites has been shown to cause hepatitis and agranulocytosis (Winstanley *et al*., 1990).

Parenteral quinine is the drug of choice for severe malaria as compliance with oral quinine is poor due its bitter taste (Newton and White, 1999). The therapeutic index for quinine is narrow and severe side effects such as hypoglycemia, coma, hemolytic-

uraemic syndrome and cardiovascular disorders may arise. Cinchonism, (characterized by tinnitus, deafness, dizziness, nausea and visual problems) is, however, the commonest side effect.

Mefloquine, mostly used in prophylaxis due to its long terminal half-life (14-21 days) has been linked with neuropsychotic side effects (Mai *et al*., 1996). Halofantrine by prolonging the QT interval causes ventricular arrhythmias (Monlun *et al*., 1993). By targeting the dormant liver forms (hypnozoites) of *P. vivax* and *P. ovale* infections, the 8-amino quinoline, primaquine, prevents relapses of these infections. Primaquine is, however, contraindicated in glucose 6-phosphate dehydrogenase deficient patients in whom it causes hemolysis.

1.4.2 Folate antagonists

Inhibition of enzymes of the folate pathway results in decreased pyrimidine synthesis and consequently reduced DNA, serine and methionine formation (Olliaro, 2001). Available antifolates include synergistic mixtures such as pyrimethamine-sulfadoxine (Fansidar®), chlorproguanil-dapsone (LapDap®), sulfalene-pyrimethamine (metakelphin®) and proguanil-atovaquone (Malarone®). These combinations are carefully selected for matching pharmacokinetics to maximize the synergy.

Classification of the antifolates is based on the enzyme inhibited in the folate synthesis pathway. Type 1 antifolates (sulfonamides and sulfones) by mimicking *p*-aminobenzoic

acid (PABA), compete for the active site of the bifunctional enzyme dihydropteroate synthase (DHPS) and block the formation of dihydropteroate from hydroxymethyldihydropterin, a reaction catalyzed by DHPS. This enzyme is encoded by the same gene as dihydro-6-hydroxymethylpterin pyrophosphokinase. Type 2 antifolates (pyrimethamine, biguanides and quinazolines) on the other hand inhibit dihydrofolate reductase (DHFR), also a bifunctional enzyme in plasmodia coupled with thymidylate synthase, and inhibit the NADPH-dependent reduction of dihydrofolate to tetrahydrofolate (Olliaro, 2001). Tetrahydrofolate serves as an essential methyl donor in the conversion of 2′-deoxyribosyluracil monophosphate to 2′-deoxyribosylthymine monophosphate. Malaria parasites are particularly susceptible to inhibition of DHFR because unlike mammalian cells, transcriptional inhibition (mediated by the protein binding to its own message) is not relieved by the accumulation of substrate that occurs in the presence of inhibitor (Zhang and Rathod, 2002).

Apart from the severe allergic reactions caused by the sulfonamides, few side effects have been reported for the antifolates. The classical condition of antifolates toxicity is the Stevens - Johnson syndrome, also referred to as toxic epidermal necrolysis.

1.4.3 Artemisinins

This class consists of a unique family of sesquiterpene lactone endoperoxides. The parent compound, artemisinin, was first extracted from the Chinese herb *Artemisia annua* (qinghao). Several semisynthetic derivatives which include artemether, arteether,

artelinate and artesunate are in use. They are metabolized to dihydroartemisinin which is the main active agent in the body (Ridley, 2002). Members of this group act on all phases of the asexual intra-erythrocytic schizogonic cycle and also possess gametocytocidal activity (Ridley, 2002; Newton and White, 1999). Although the mechanism of action is not known with certainty, the prevailing hypothesis is that the essential pharmacophore, an endoperoxide bridge, undergoes reductive cleavage by ferroheme ferrous-protoporphyrin IX (Fe(II)PPIX) to generate carbon-centered free radicals that alkylate protein and damage parasites' microorganelles and membranes (Meshnick et al, 1996). Recent studies show that artemisinin may also inhibit the SERCA orthologue (PfATP6) of *Plasmodium falciparum* as evidenced by thapsigargin (a sesquiterpene lactone and highly specific SERCA inhibitor) antagonism of parasiticidal activity of artemisinin (Eckstein-Ludwig *et al.,* 2003). Chelation of iron by desferrioxamine abrogates the antimalarial activity of artemisinins and correspondingly attenuates inhibition of PfATP6 (Wei and Sadrzadeh, 1994; Eckstein-Ludwig *et al.,* 2003).

As a class, artemisinins are fast acting and potent but due to their short half-lives, they are currently used in combination with longer half-life drugs. This strategy forms the basis of artemisinin combined therapies (ACTs) and is hoped to improve individual compliance and prevent or retard the development of drug resistance. Several combinations are already in the market and include artemether-lumefantrine (Coartem®), dihydroartemisinin-piperaquine (Artekin®) and the triple combination of chlorproguanil-dapsone-artesunate undergoing phase II trials in Malawi (Maitland *et al.,* 2004).

Although relatively lacking in adverse effects, embryonic toxicity has been reported for this class of antimalarials in China although a later study allayed fears of artemisinin related birth defects (McGready *et al*., 2001).

1.4.4 Antibiotics

Several antibacterial agents also exhibit antimalarial activity. This activity stems from the fact that malarial parasites, like other *Apicomplexian* parasites, possess a plastid-like organelle, the apicoplast (Kohler *et al.,* 1997) which fulfills some metabolic functions such as the synthesis of isoprenoids, fatty acids and probably heme (Wilson, 2002). The apicoplast contains a residual genome that encodes tRNA's, rRNA's, RNA polymerases and ribosomal proteins all of which ensure self-replication of this organelle (Wiesner *et al.,* 2003). Antibiotics act through inhibition of the prokaryote-like RNA and protein synthesis in the apicoplast (Ralph *et al.,* 2001). The tetracyclines may also block mitochondrial protein synthesis in the parasite (Ralph *et al.,* 2001). Rifampicin, tetracyclines, lincosamides and macrolides are being increasingly used in combination with other antimalarials to augment their activity. However, due to their slow effect, antibiotics are mostly used for prophylaxis in which case doxycycline is the most popular agent.

1.4.5 Hydroxynapthoquinones

The prototype of this class is atovaquone which blocks mitochondrial electron transport through inhibition of the cytochrome bc_1 complex (Fry and Pudney, 1992). When used alone though, resistance emerges rapidly due to a point mutation of cytochrome *b* gene localized in the mitochondrial genome (Srivastava *et al.,* 1999). As a result, atovaquone is used as a fixed-dose combination with proguanil which is thought to enhance the atovaquone-induced collapse of the mitochondrial membrane potential (Srivastava *et al.,* 1999). In prophylaxis, atovaquone-proguanil activity against liver stages is advantageous as the parasites are killed before an infection of erythrocytes can be established (Berman *et al.,* 2001).

1.4.6 Novel antimalarials

Numerous new antimalarials, some targeting novel molecular entities are under development (Rosenthal, 2003). Pyronaridine, a 9-anilinoacridine is generally active against chloroquine-resistant parasites while tafenoquine, a derivative of 8-amino quinolines primaquine and pamaquine, is active against liver and erythrocytic stages of *P. vivax* and has a better toxicological profile (Wiesner *et al.,* 2003). Fosmidomycin is an inhibitor of 1-deoxy-D-xylulose 5-phosphate (DOXP) reductoisomerase, a key enzyme in the mevalonate-independent pathway of isoprenoid biosynthesis (Wiesner *et al.,* 2000). Kinetic studies with *E.coli* DOXP reductoisomerase revealed that fosmidomycin alters the enzyme conformation after binding thus enhancing affinity (Koppisch *et al.,* 2002).

The indoloquinoline alkaloid cryptolepine activity probably depends on the inhibition of heme polymerization (Wright *et al.,* 2001). Aspartic proteases (plasmepins), cysteine proteases (falcipain-2), metalloproteases (falcilysin) and histoaspartic protease (HAP) involved in hemoglobin digestion in the food vacuole are potential antimalarial targets and several specific protease inhibitors have been developed (Haque *et al.,* 1999; Moon *et al.,* 1997).

Of note, with regards to our work, is that attempts to develop nonpeptide falcipain-2 inhibitors lead to the identification of chalcones (Li *et al.,* 1996). Type II fatty acid synthesis pathway in the parasite's apicoplast is inhibited by triclosan (Surolia and Surolia, 2001) whereas the antifungal imidazoles form complexes with heme (Huy *et al.,* 2002) and inhibit *P. falciparum* in vitro (Chong and Sullivan, 2003).

1.5 Antimalarial drug resistance

Lack of an effective malaria vaccine has resulted in the reliance upon vector control measures and chemotherapy as the two main ways of combating the disease. The former involves removal of mosquito breeding sites, use of residual insecticides such as DDT to kill adult mosquitoes and prevention of human-mosquito contact by use of repellant screens and insecticide-treated bed nets while the latter entails administering drugs for prophylaxis or cure of malaria (Olliaro, 2001). The effectiveness of drugs in malaria control has depreciated considerably over the last few years due the development of drug resistant strains of the parasite. This has been attributed to widespread misuse of common

antimalarials together with their extensive deployment (particularly chloroquine (CQ)) to the tropical regions of the world. This has provided a tremendous selection pressure for the parasites to develop resistance mechanisms (White, 2004).

Globally therefore, malaria control programmes such as the WHO funded Roll Back Malaria campaign, which aims to halve malaria mortality cases by the year 2010, are under the threat of failing unless measures are put in place to curtail resistance progression (Yeung *et al*., 2004). Resistance to antimalarials is mainly due to point mutations and changes in steady-state transcript levels (Arav-Boger and Shapiro, 2005). The mutations in, or changes in, the copy number of genes encoding the drug's parasite target e.g. membrane transporters or enzymes affect the intraparasitic drug accumulation (White, 2004).

CQ resistance has mainly been attributed to the accumulation of multiple point mutations in the gene coding for *Plasmodium falciparum* chloroquine resistant transporter *(PfCRT)*, a putative transporter protein located in the parasite's digestive vacuole membrane. This 424 amino acid protein with ten predicted transmembrane domains (Fidock *et al*., 2000) is thought to be involved in drug fluxes and/or pH regulation in the digestive vacuole (Wellems and Plowe, 2001). Two mutations, K76T and A220S, have been shown to be critical to the evolution of resistance (Fidock *et al*., 2000; Wellems and Plowe, 2001), a fact that lead Hastings *et al* (2002) to propose that either these two are the pharmacologically relevant mutations or that the other mutations serve to compensate for

impaired protein function after acquisition of the critical mutations. Sidhu and co-workers (2002), by swapping the (*PfCRT)* gene in chloroquine (CQ) sensitive strain with the one from CQ-resistant strain, also demonstrated the role of K76T mutation in CQ resistance and further proved that up-regulation of *PfCRT* is not obligatory for resistance to occur. Although the actual resistance mechanism remains unclear, several hypotheses have been posited and include drug expulsion through an acquired efflux system (Sanchez *et al*., 2003), leaking of the drug out of the digestive vacuole (Bray *et al*., 2005) and variation of vacuolar pH due to altered chloride conductance across the vacuolar membrane (Zhang *et al.,* 2002). These fit in with the presumed role of *PfCRT*. Reversal of chloroquine resistance has been observed with several agents among them verapamil, a calcium channel blocker and chlopheniramine, a histamine H_1 receptor antagonist. These agents do not however restore full chloroquine sensitivity (Bray *et al*., 1996). Verapamil has been proposed to act via its hydrophobic binding to the mutated PfCRT protein. By so doing, it replaces the lost positive charge and, therefore, repels the access of 4-aminoquinoline cations to PfCRT (Henry *et al*., 2006; Warhurst, 2003).

Apart from *PfCRT*, the multidrug resistance gene *Pfmdr1* has also been implicated in chloroquine resistance. This gene encodes the ATP-dependent P-glycoprotein homolog 1 (Pgh-1), a transmembrane protein also found in the parasite digestive vacuole (Cowman *et al*., 1991; Wilson *et al.,* 1989). A widely reported defect in Pgh-1 associated with chloroquine resistance is the N86Y mutation (Dorsey *et al*., 2001). A further four Pgh-1 mutations (Y184F, S1034C, N1042D and D1246Y) have been shown to play a role in

other quinoline antimalarials resistance in addition to modulating the parasite's sensitivity to chloroquine and artemesinin (Reed *et al*., 2000).

Resistance to the antifolate antimalarials is conferred by single point mutations of the gene encoding for the respective enzyme, resulting in substitutions in the amino acid side chains. High-level pyrimethamine resistance results from the accumulation of mutations in DHFR, principally at codons 108, 59, 51 and 164, where allelic variation gives rise to the mutations S108N, C59R, N51I and I164L. Occasionally, mutations also occur at codons 50, 140 and the 'Bolivian repeat' of the DHFR gene. Mutations occurring in DHPS at codons 436, 437, 540, 581 and 613 correlate with estimated levels of sulphadoxine resistance (Ridley, 2002). Highest level of clinical resistance result from parasites with four mutations in DHFR and two in DHPS, which may suggest the maximum number of mutations that can be tolerated in competition with less affected strains (Arav and Shapiro, 2005). Dihydrofolate reductase –thymidylate synthase crystal structure provides evidence that the critical mutations mediating clinical drug resistance map to the dihydrofolate reductase active site (Yuvaniyama *et al.,* 2003).

Resistance to atovaquone is due to mutations affecting five amino acids clustered in a highly conserved fifteen amino acid sequence of cytochrome *b* from *P. yoelii* (Arav-Boger and Shapiro, 2005). Analysis of *P. falciparum* isolated from patients who failed atovaquone monotherapy confirmed predilection for mutations at residue Y268 (Korsinczky *et al.,* 2000).

.

1.6 Purine metabolism

Purine nucleotides are essential for the synthesis of nucleic acids and proteins as well as for energy-requiring reactions in living organisms (el Kouni, 2003). The metabolic pathway for the production of purine nucleotides is conspicuously different between protozoan parasites and their mammalian hosts (Allen and Ullman, 1993). Whereas mammalian cells obtain purine nucleotides primarily by *de novo* synthesis and/or the purine salvage pathways, parasitic protozoa are auxotrophic for purines and rely on the host to provide them with purines or purine ribosides (Shi *et al.,* 1999). The *de novo* synthesis of purine nucleotides requires a minimum of ten enzymes for the production of inosine monophosphate (IMP), the major intermediate of purine metabolism and thus the salvage pathway provides an energy-saving alternative (You *et al.,* 2003).

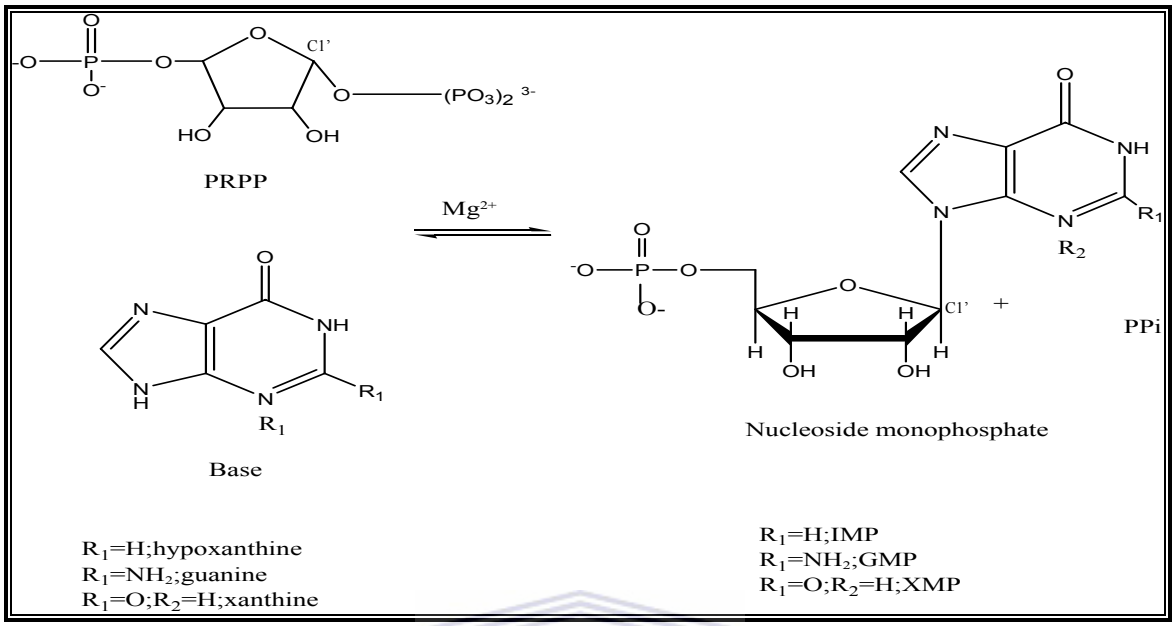
P. falciparum proliferation in human erythrocytes requires purine salvage by hypoxanthine (-guanine) phosphoribosyltransferase (HPRT) (Shi *et al.,* 1999). It has been shown that depletion of hypoxanthine from the parasite cultures by xanthine oxidase leads to the death of the organism indicating that IMP formation from this purine, catalyzed by HPRT, is not by-passed by an alternate reaction (Berman *et al.,* 1991). Furthermore, the relative abundance of adenosine deaminase and purine nucleoside phosphorylase, two enzymes that catalyze the conversion of adenine to hypoxanthine, in *P. falciparum* suggests that the salvage of purine bases in this parasite proceeds mainly through hypoxanthine (Craig and Eakin, 2000).

Because of the central role played by HPRT in providing the parasite with ribonucleotides, it is a promising target in the design of novel anti-malarial chemotherapeutics. Such drugs could potentially act by two mechanisms: (i) as inhibitors of *P. falciparum* HPRT and (ii) as substrates which after conversion to their respective monophosphate nucleotides by HPRT would inhibit downstream enzymes and/or be incorporated into parasite DNA or RNA resulting in eventual cell death (Keough *et al.*, 1999).

Selective inhibition of the *P. falciparum* HPRT is however critical since its complete deficiency in humans is associated with a debilitating neurological disease known as the Lesch-Nyhan syndrome which is clinically characterized by irregular, spastic and athetoid movements, mental retardation, hyperuricemia, uric acid nephrolithiasis and a unique compulsion for self-mutilation (Musick, 1981). Partial deficiency leads to gouty arthritis (Sculley *et al*., 1992). Inhibition of the human enzyme for the usual duration of drug therapy is unlikely to have major deleterious effects.

1.7 The HPRT reaction mechanism

As shown in scheme 1.0, HPRT catalyzes the Mg^{2+}-dependent reversible transfer of the 5-phosphoribosyl group from α-D-5-phosphoribosyl-1-pyrophosphate (PRPP) to a nitrogen atom of the imidazole ring of the purines hypoxanthine, guanine or xanthine to form the corresponding nucleotides and inorganic pyrophosphate (PPi).

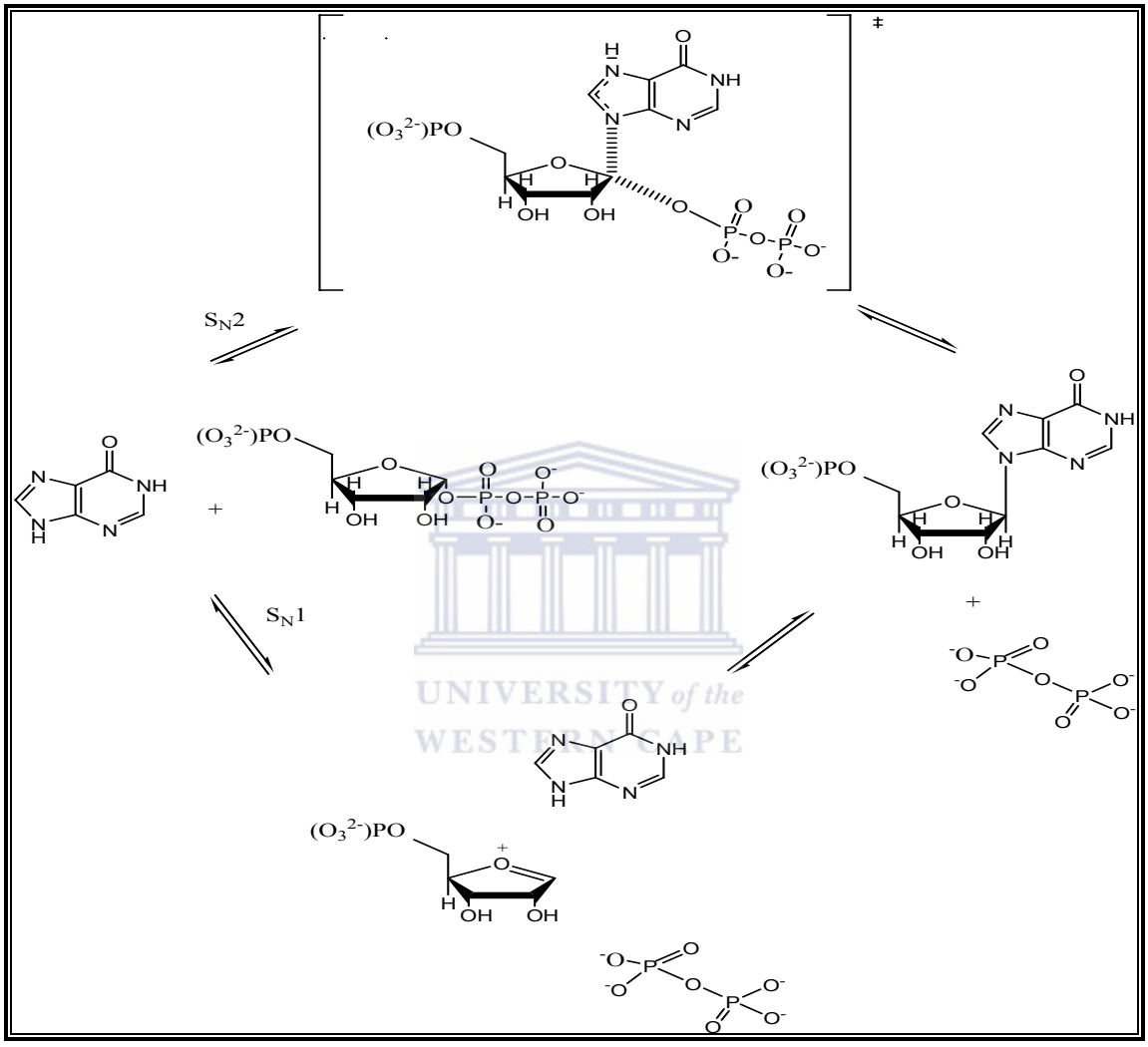


Scheme 1.0: Schematic diagram of the reaction catalyzed by HPRT (Thomas and Field, 2002)

The kinetic mechanisms for the HPRT's from several organisms have been investigated. The forward reaction is ordered and sequential, with PRPP binding first followed by the purine base. The proposed nucleophilic attack by the purine base at the PRPP ribose C1′, which is associated with the inversion of configuration at this anomeric position (Tao *et al.,* 1996), is followed first by the release of pyrophosphate and then that of the ribonucleotide (IMP, GMP or XMP). Kinetic studies with the human HPRT by Xu *et al.* (1997) have further demonstrated that the binding of substrates is effectively ordered in both the forward and the reverse reactions and that product release is the rate-limiting step in catalysis for both ribonucleotide formation and pyrophosphorolysis. The same

probably applies to *P. falciparum* HPRT which shares 44% sequence identity and 76% similarity with human HPRT (Raman *et al.,* 2005).

The reaction mechanism involves electrostatic re-arrangement during the transition state as the PRPP C1´-O1 bond breaks and the C1´-purine N9 bond forms (Focia *et al.,* 1998). Two possible mechanisms have been proposed for the reaction. There is a dissociative, S_N1-like mechanism that first produces PPi and a charged ribooxacarbenium ion and only afterward the nucleotide, and there is an associative, S_N2-like mechanism in which the attack of the purine N9 atom and the departure of the pyrophosphate from ribose C1´ occur simultaneously (Scheme 1.1) (Thomas and Field, 2002). Kinetic isotope effect analysis of orotate phosphoribosyltransferase (OPRT) demonstrated a ribooxacarbenium ion transition state (Tao *et al.,* 1996) similar to that found for purine nucleoside phosphorylase (PNP) (Kline and Schramm, 1995). In the human HGPRT with bound immucillinG-5′-PO_4 and MgPPi, the purine is thought to be anchored throughout the reaction, allowing the ribooxacarbenium to migrate between the purine base and the pyrophosphate nucleophile (Schramm and Shi, 2001).



Scheme 1.1: Two possible mechanisms for the HGPRT reaction (modified from Thomas and Field, 2002).

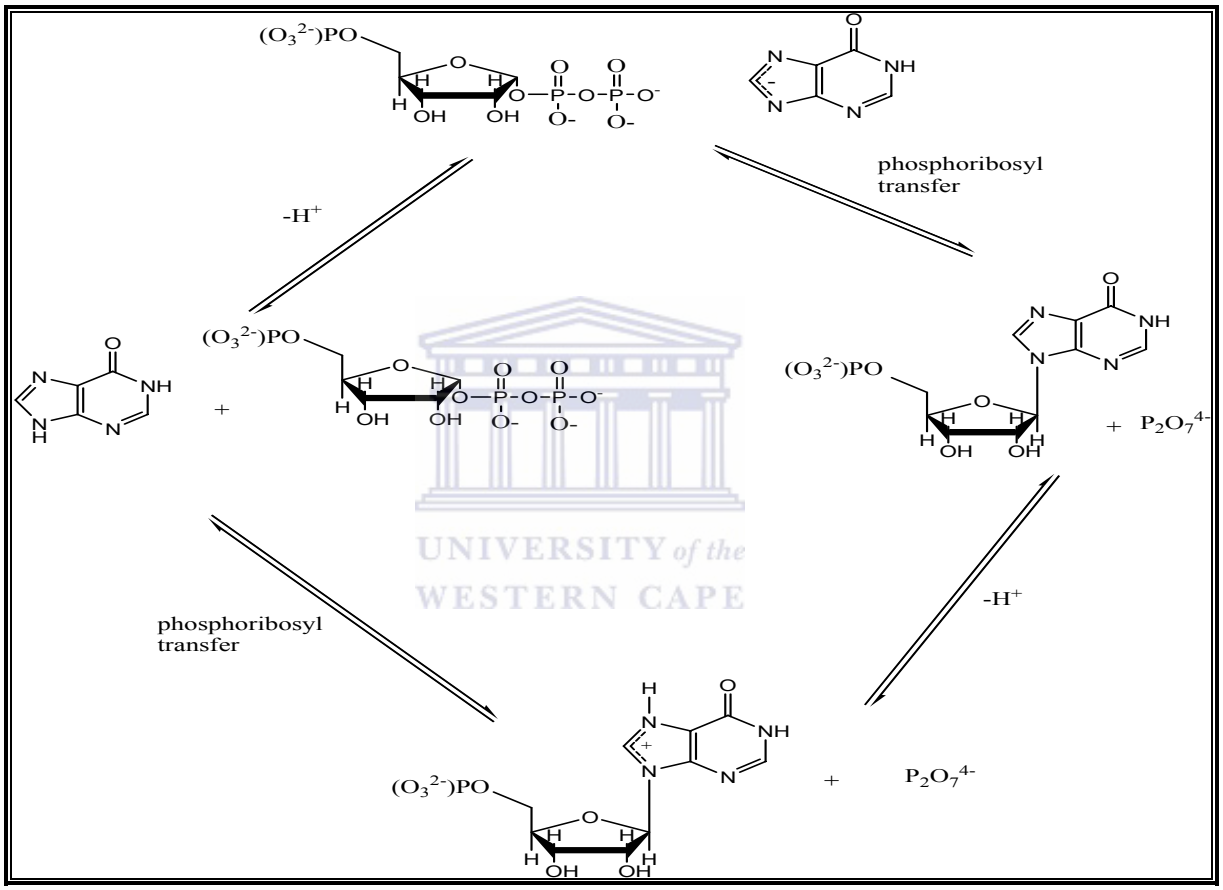
In the case of an S_N1-type chemical reaction, the unstable ribooxocarbenium intermediate would need to be protected from the bulk solvent. It has been proposed that a large,

flexible loop (loop II), which contains an absolutely conserved Ser-Tyr dipeptide, closes over the active site to shield the highly reactive ribooxocarbenium from nonproductive attack by the bulk solvent water molecules. The deletion of seven residues from active

site loop II of trypanosomal HGPRT does not however prevent either the forward or the reverse reaction catalyzed by the enzyme although it lowers the catalytic efficiency (Craig and Eakin, 2000). Furthermore, apart from the possible stability provided to the positively charged intermediate by the 5′-O of the ribose moiety, X-ray crystal structures of the closed active site of the human, malarial and trypanosomal HPRT's reveal that the only residue near enough to provide further stabilization is an invariant tyrosine (human Tyr104) (Shi *et al.,* 1999a; Shi *et al.,* 1999b; Focia *et al.,* 1998). These enzymes therefore provide only minimal electrostatic stabilization for a ribooxocarbenium ion formed during an S_N1-type reaction (Craig and Eakin, 2000). The above mentioned observations suppose the reaction as being associative (S_N2) rather than dissociative (S_N1). This is in agreement with Fedorov and co-workers (2001) who classify the reaction mechanism of phosphoribosyltransferases as nucleophilic displacements by electronic migration, in which the D_NA_N transition state has a partial bond between the purine and sugar and van der Waals contact between the pyrophosphate and sugar.

Recent studies (Crehuet *et al*., 2005) have shown that the preferred reaction chemistry in HPRT proceeds via an associative, S_N2-type nucleophilic displacement in which rapid phosphoribosyl transfer occurs via a ternary complex. Two stepwise mechanisms are possible. In one, there is transfer of the phosphoribosyl group followed by a proton

transfer from the hypoxanthine to the side chain of Asp148 residue. In the other mechanism, proton transfer occurs before the phosphoribosyl transfer. This is illustrated in Scheme 1.2.



Scheme 1.2: Possible steps for S_N2 mechanism in HGPRT (Adapted from Crehuet *et al*., 2005).

1.8 Three dimensional structure of HPRT

The HPRT's are members of a larger family of phosphoribosyltransferases (PRT's) which are involved in the biosynthesis of purine, pyrimidine and pyridine nucleotides, as well as the aromatic amino acids histidine and tryptophan (Musick, 1981).

Type 1 PRT's, of which HGPRT belongs, share a common core tertiary structure resembling a typical Rossman dinucleotide-binding fold (Figure 1.1). This common core

region consists of four or five parallel β-strands flanked by three or four α-helices. A short sequence of 13 amino acids present within this fold, referred to as the PRPP-binding motif, is located near the C-terminal end of the third β-strand in the core β-sheet and is the conserved signature sequence of type 1 PRT's (Lee *et al.,* 1998). This motif is made up of four hydrophobic amino acids, two acidic amino acids and seven variable amino acids (Sinha and Smith, 2001). The C-terminal ends of the β-sheets of the core domain form the floor of the active site of these enzymes while a poorly conserved hood domain contributes residues that complete the active site and participate in binding purine substrates. The active site is partly at the interface of the core and hood domains.

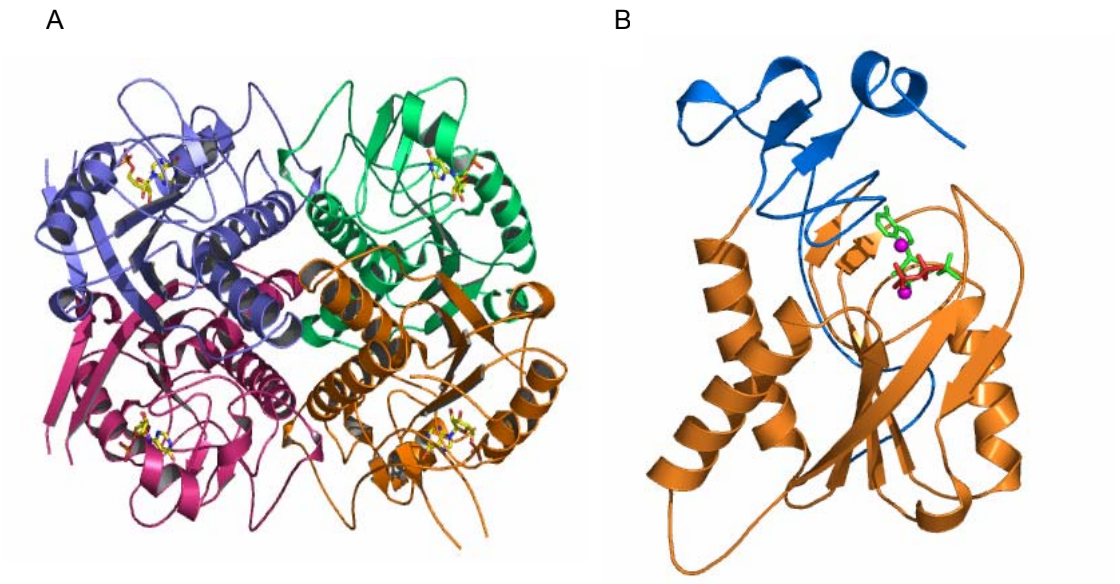


Figure 1.1: The tetramer of *Toxoplasma gondii* HGPRT with bound IMP demonstrates the interfaces of the tetramer (A). View of the monomer of malarial HGPRT with transition state analog imucillinHP bound (B) shows the flexible hood domain (blue) and the core domain (orange). The catalytic site is at the interface of the two domains. Images coordinates taken from Heroux *et al*., 1999 and Shi *et al*., 1999 and rendered with PyMol (DeLano, 2002).

1.8.1 The active site

HPRTs of distantly related organisms share moderate primary sequence homology. The aligned sequences of several parasite HPRTs show between 28-49% identities with their human counterpart more so in the active site loops (Focia *et al*., 1998). Analysis of a number of high-resolution HPRT crystal structures from humans and parasites reveal that several amino acid residues are invariant and play an important role in enzyme function. The invariant residues immediately flank or are near the enzyme's active site which is similar in human and parasites' enzymes (Heroux *et al*., 1999). This is illustrated in Figure 1.2.

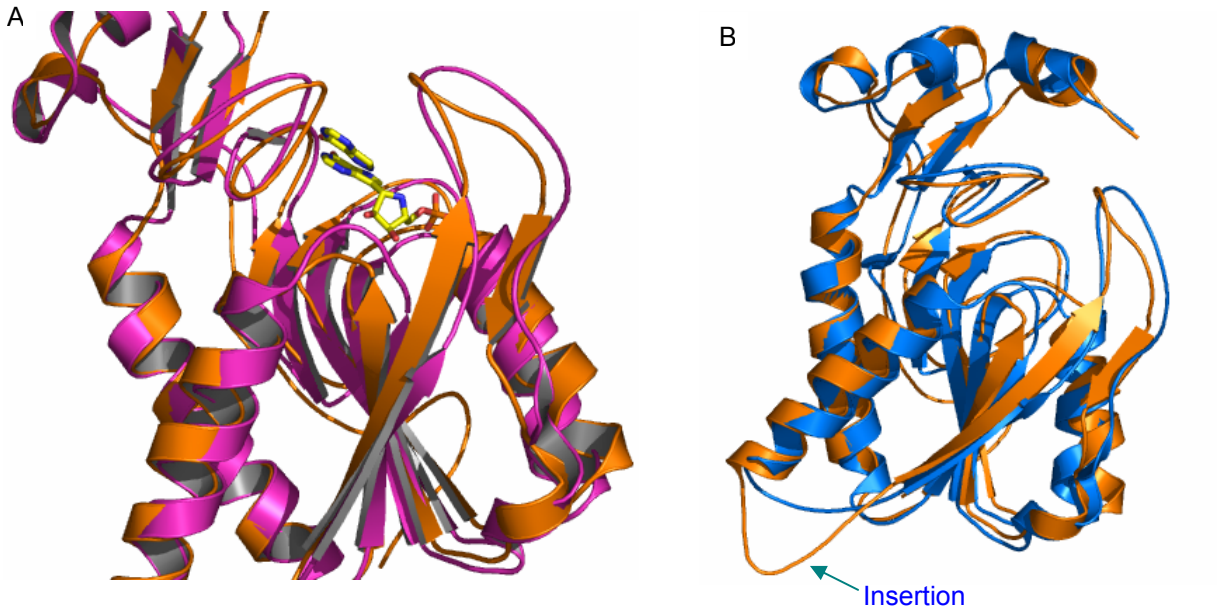


Figure 1.2: A structural alignment of malarial (1CJB) and human (1HMP) HGPRT with bound substrates (immucillinHP and GMP, respectively). The conformation of the base in the active site is similar in both the malarial (orange) and human (purple) enzymes (A). The extended loop (arrow) in (B) reflects the difference in chain length between the two enzymes. The alignment was performed with ALIGN (Cohen *et al*., with image coordinates from Shi *et al*., 1999 and Eads *et al*., 1994.

Four segments of the protein polypeptide backbone (loops I, II, III, IV) define a 3D space unique to the active site of the enzyme (Freymann *et al*., 2000). The smallest of these is the active-site loop I, which is composed of residues Leu67, Lys68, and Gly69. Gly69 is invariant and a G69E mutation inactivates the human HGPRT resulting in Lesch-Nyhan syndrome (Sculley *et al*., 1992). Lys68 participates in an unusual non-proline *cis*-peptide linkage with the preceding leucine residue (Focia *et al*., 1998; Heroux *et al*., 1999; Heroux *et al*., 2000). Non-proline *cis* peptides are infrequently seen in published structures (<0.05%) (Stewart *et al*., 1990). Where they have been observed in other enzymes, structural roles assigned to them have included metal-binding sites (Hohenester

et al., 1996; Dominguez *et al*., 1996). Heroux and co-workers (1999) have hypothesized that *cis-trans* peptide isomerization may assist the catalytic mechanism. The invariant glycine may thus be essential for the formation of a tight turn and the unusual non-proline *cis*-peptide (Vos *et al*., 1997; Shi *et al*., 1999a; Shi *et al*., 1999b). The *cis*-peptide linkage geometry allows the main-chain nitrogen atom of Lys68 to form a hydrogen bond with the β-phosphate group of bound Mg-PRPP or Mg-PPi while the main-chain carbonyl oxygen of Leu67 interacts with Mg-coordinated water molecule (Canyuk *et al*., 2004). This is illustrated in Figure 1.3. The side chain of Lys68 also forms multiple hydrogen bonds with residues in the opposing subunit in the dimeric structure of *Trypanosoma cruzi* HPRT (Focia *et al*., 1998).

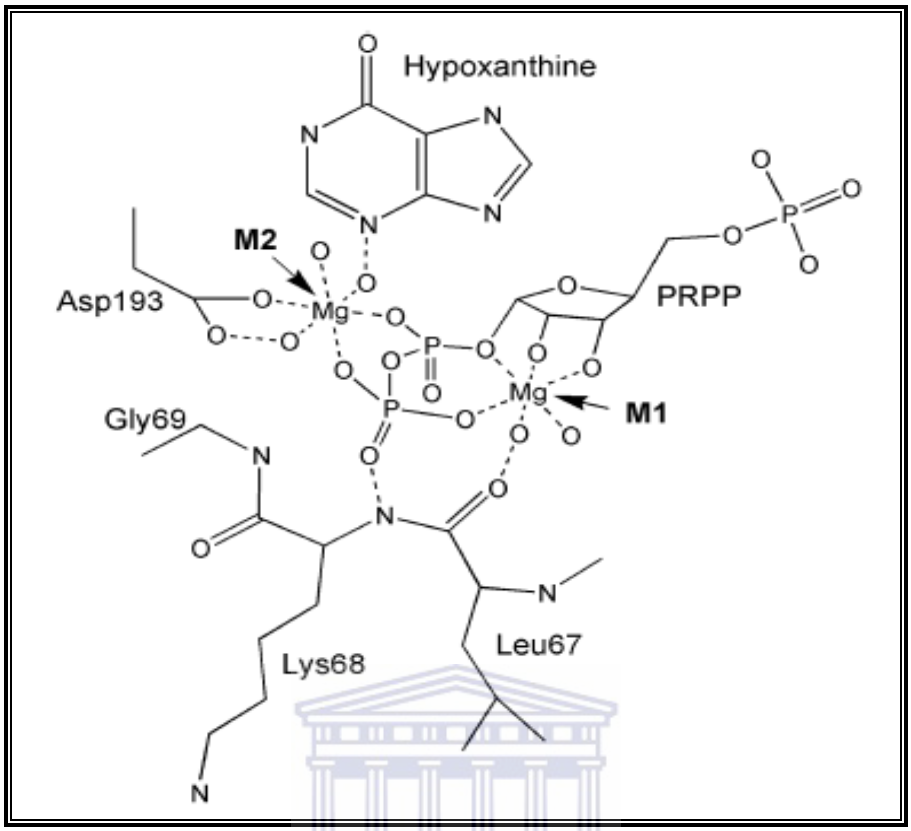


Figure 1.3: Depiction of the interactions involving the cis-peptide bond observed in loop I of the X-ray crystal structure of *T. cruzi* in ternary complex with substrates. One magnesium ion (M1) is associated with two water molecules while the seond (M2) is associated with three. Residue numbering is according to structurally analogous residues in the human HGPRT. Adapted with permission from Canyuk *et al. (*2004).

The primary PRT substrate PRPP is cradled between the active site loops I and III. Residues in these loops form extensive networks of hydrogen bonds with the phosphate and pyrophosphate groups mostly via main chain atoms. One metal ion (usually Mg^{2+}) binding site is located on either side of the PRPP pyrophosphate group (M1 and M2). Each Mg^{2+} ion has an octahedral coordination involving two terminal oxygen atoms of the pyrophosphate which adopt a cis orientation around the Mg^{2+}. M1 is bonded to four

water molecules while M2 is bonded to three water molecules and is also linked to the side chain of the conserved Asp193. This is the only direct contact between HGPRT and the Mg^{2+} ions. The eclipsed conformation imposed upon the oxygen atoms of pyrophosphate by the Mg^{2+} ions is a common configuration in structures where pyrophosphate is coordinated to either one or two divalent metal ions (Heroux *et al*., 1999a). Additionally, Mg^{2+} ions have a preference for interacting with oxygen over nitrogen.

Interaction of Arg199 with Asp193 may help in positioning the substrates for in-line nucleophilic attack at the C1 carbon of PRPP or a nucleotide. A D193N mutation which drastically alters the kinetic parameters of HGPRT results in Lesch-Nyhan syndrome in humans (Sculley *et al*., 1992). A structural explanation for this is may be revealed by the presence of a hydrogen bond linking the side chain carboxyl group of Asp193 to the main chain nitrogen of Asp196.

Loop III forms the typical region of conserved sequence in type 1 PRT's, the PRPP motif (^{129}VLIVEDIIDTGKT^{141}). The invariant Glu133 and Asp134 in this loop form hydrogen bonds with the 2′- and 3′-hydroxyls of PRPP (Vos *et al*., 1998). Moreover, Glu133 forms a hydrogen bond with a metal-associated (M1) water molecule. The two acidic residues are also predicted to participate in stabilizing the transition state of the reaction (Scapin *et al*., 1994). Additionally, they may be involved in networking other water molecules bound in the active site to aid further hydration of the pyrophosphate leaving group

during or following the reaction (Focia *et al*., 1998b). Mutagenesis studies have shown that Asp137, which is also conserved in the PRPP fingerprint, acts as a catalytic base which deprotonates the purine base thus activating it as a nucleophile (Xu and Grubmeyer, 1998). It may also contribute to tight binding of nucleotide analog inhibitors as evidenced by its interaction with the N^7 of the iminoribitols (Li *et al*., 1999; Shi *et al*., 1999). The concluding residues (^{138}TGKT^{141}) of the PRPP motif surround and bind the 5′-monophosphate group of the ribose ring. Minor differences in sequence are found for this part of loop III across type 1 PRT families but usually two threonine residues and a glycine are present (Focia *et al*., 1998b; Sinha and Smith, 2001). This loop participates in up to seven hydrogen bonds with the 5′-phosphate group of bound substrate or product through main chain nitrogens and threonine side-chain oxygen (Somoza *et al*., 1996).

Purine bases in the active site are recognized by a convergent array of hydrogen bonds and by several van der Waals interactions. N^7 of the purine is in van der Waals contact with the side chain carboxyl group of Asp137 while O^6 is hydrogen bonded to the side chain amino group of Lys165 and also main chain nitrogen of Val187. Val187 peptide oxygen interacts with N^1 atom of both guanine and hypoxanthine and also N^2 of guanine which is hydrogen bonded to the main chain oxygen atom of Asp193 in loop IV (Eads *et al*., 1994; Shi *et al*., 1999). Crystal structure of *T. gondii* HGPRT (Heroux *et al*., 1999) shows that in order to bind xanthine favorably, loop IV shifts such that O2 hydrogen bonds to the main chain nitrogen of Asp193. To further stabilize its orientation, the purine ring is clamped in place by a π-π stacking interaction with Phe186 above it. Other

HGPRT's contain a similarly located aromatic residue. On the bottom and top, the purine ring is held by hydrophobic interactions, sandwiched between the side chain of Val187 on one side of the planar aromatic ring and the π-π stacking interaction with Phe186 on the other side.

With both substrates bound, the active site architecture is completed by the closure of active site loop II (catalytic loop) which appears to be highly disordered when only a single substrate is bound (Eads *et al*., 1994; Somoza *et al*., 1996). This loop, contributed by the core domain, generally consists of 11-14 amino acids with only a Ser-Tyr dipeptide conserved among them (Munagala *et al*., 2001). The main chain nitrogen of Ser103 forms a hydrogen bond with O1A atom of PRPP whereas the side chain hydroxyl interacts with a water molecule which is hydrogen bonded to both Ser109 and the 5′-phosphate group either PRPP or a nucleotide analog (Focia *et al*., 1998). The aromatic ring of Tyr104 stacks over the PRPP molecule with its main chain nitrogen forming a hydrogen bond with PRPP O1A atom while its hydroxyl group is bonded to the 5′-phosphate O3P. Tyr104 interactions span the PRPP molecule positioning its aromatic ring almost perpendicularly to the purine ring.

The extensive interactions of the invariant Ser103 and Tyr104 with PRPP secure the base of the closed catalytic loop over the active site and also offer electrostatic stabilization during the transition state. Moreover, this loop may indirectly aid in binding of purine

substrates through its van der Waals interactions with conserved aromatic residue (Phe186) that stack above the purine ring in the active site.

Changes in the interactions involving loops I and II with the active site ligands before and after closure of the active site provide a possible mechanism for the liberation of the pyrophosphate leaving group after catalysis.

1.9 *Plasmodium falciparum* HPRT.

This section aims to analyze the structural features of *P. falciparum* hypoxanthine phosphoribosyltransferase (PfHPRT) mainly, unless otherwise stated, based on the crystal structure of PfHPRT in complex with a transition state analog (Shi *et al*., 1999) which is the only PfHPRT atomic structure available.

PfHPRT is a tetrameric enzyme with 231 residues per subunit and a native molecular mass of 26 232 Daltons per monomer (Keough *et al*., 1999). The subunit incorporates a core domain consisting of five parallel β-strands surrounded by four α-helices (residues 46-194) and a hood domain composed of a four stranded anti-parallel β-sheet and two α-helices (residues 1-45 and 195-231). As seen in the human HGPRT (Eads *et al*., 1994), a mobile loop dissects the core β-sheet into two halves. Three β-strands in the C-terminal portion of the core β-sheet extend to form a secondary β-sheet which is located above and oriented perpendicularly to the central β-sheet. The active form of the enzyme is a tetramer.

When compared, the four independent subunits are essentially identical with a root mean square (RMS) deviation between all C_α atoms, excluding several loop residues, of less than 0.15 Å in pair wise comparison. Superimposing the subunits reveal that the active sites are nearly identical with an RMS deviation of less than 0.1 Å between C_α atoms of the active site residues. The active site loops of PfHGPRT close more tightly over the active site with a distance of 2.8 Å between N7 of the purine base and the side chain carboxyl of Asp148 (human Aps137) compared with 3.3 Å in the crystal structure of *T. cruzi* (Focia *et al*., 1998).

The properties of malarial HGPRT differ significantly from those of other HGPRT's. One difference is in the substrate specificity. While malarial HGPRT can utilize xanthine, human HGPRT lacks that ability (Queen *et al*., 1988). Interestingly however, a chimeric HGPRT in which 49 N-terminal residues of the human HGPRT were replaced with the corresponding residues from PfHGPRT exhibited specificity for xanthine even though none of the switched residues forms part of the purine or PRPP binding region in the available crystal structures of HGPRT's (Subbayya *et al*., 2000). Another unique feature of PfHGPRT lies in its behaviour upon purification. Recombinant PfHGPRT unlike the human and *T. cruzi* enzymes (Xu and Grubmeyer, 1998; Eakin *et al*., 1997) usually has negligible activity on purification and requires incubation with the substrates hypoxanthine and PRPP or Mg^{2+} and PRPP to activate it (Keough *et al*., 1999; Phehane, 2002; Mbewe *et al*., 2006). Although oligomerization into the tetrameric form is a requisite for activity, the active form is less stable (Raman *et al*., 2005). The presence of

substrates does not lead to activation when the enzyme is a dimer (Keough *et al*., 1999). Raman *et al*. (2005) further showed that PfHGPRT activation was actually effected by the product, inosine monophosphate.

In spite of its unique properties, binding of substrates in PfHGPRT is similar to that in other HGPRT's and is analyzed subsequently. The purine ring analog binding exhibits a low antiglycosidic torsion angle while the imminoribitol ring pucker is 3′-exo 5′O cis to N4′. C2′ and C3′ hydroxyl groups of the imminoribitol ring interact with Glu144 and Asp145 and also M1 magnesium ion which slots below the plane of the imminoribitol ring. The magnesium ions associated with pyrophosphate is also located below the imminoribitol ring and its geometry is such that its O2A oxygen is optimally positioned for a nucleophilic attack on the carbenium carbon in case of a ribooxocarbenium ion intermediate formation.

Main chain nitrogen atoms of Asp148, Thr149, Gly150 and Thr152 together with the hydroxyl groups of Thr149, Thr162 and Tyr116 fix the position of the 5-phosphate group of imminoribitol moiety. The aromatic ring of Tyr116 through π-cation, together with the 5-phosphate group interaction with N4′ of the iminoribitol enhances the stabilization of the oxacarbenium ion during catalysis. It should be noted that the backbone nitrogen of Tyr116 together with that of Ser115 form favorable hydrogen bonds to O3A of pyrophosphate. Lys176 interacts with the exocyclic oxygen of the deazahypoxanthine ring while Asp148 acts as the catalytic base by interacting with N7. The peptide carbonyl

group of Val198 hydrogen bonds to N1 of deazahypoxanthine whereas Asp204 directly interacts with M2 magnesium ion. In addition to orienting the peptide amides of Leu76 and Lys77 towards the pyrophosphate, the *cis* bond between these two residues enable the side chain amino nitrogen of Lys77 to interact with the main chain carbonyl oxygen of Glu108 in the neighboring subunit.

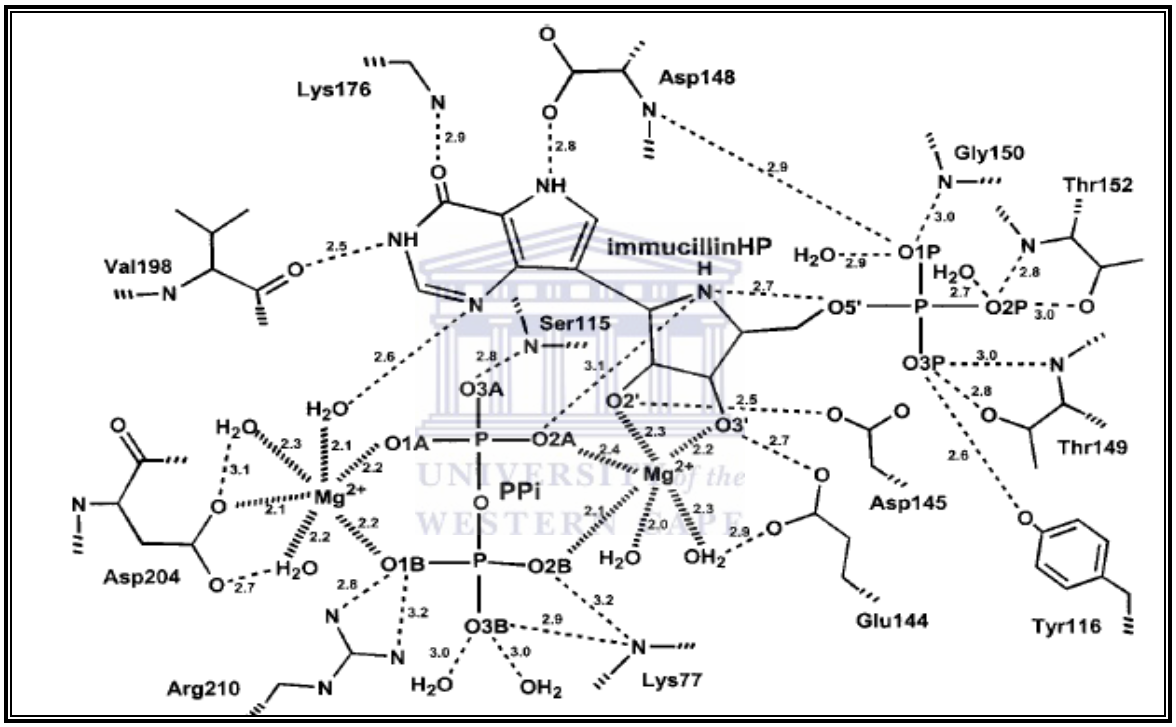


Figure 1.4: Schematic PfHPRT active site residues involved in non covalent interactions with the bound immucillinHP, Mg^{2+} and pyrophosphate (adapted with permission from Shi *et al*., 1999).

2.0 Macromolecular x-ray crystallography

Three-dimensional structures provide important mechanistic insights into the various molecular and cellular processes mediated by proteins. These insights are critical for practical applications such as structure based drug design and the engineering of proteins with improved properties for industrial processes (Saridakis and Chayen, 2003). Over the years, X-ray crystallography has become the preferred method for protein structure solution as reflected by the number of structures solved by this method in the Protein Data Bank (PDB) - the international repository for protein structures (Berman *et al*., 2001).

The complete solution of a three dimensional crystal structure consists of several distinct but interlinked (Dauter, 2006) stages which can be summarized as crystallization, data collection and processing, model building, refinement and structure validation. The availability of high-quality crystals is crucial to the structure determination of proteins by X-ray diffraction (Chayen, 2005).

2.1 Structure based drug design

Innovations in molecular and structural biology techniques have radically transformed the opportunities to use protein three-dimensional structures to accelerate drug discovery (Blundell *et al*., 2006; Congreve *et al*., 2005). By defining the topographies of the complementary surfaces of ligands and their protein targets, protein crystal structures utility ranges from target validation to lead discovery and optimization (Jhoti, 2001). Indeed, several drugs on the market have been developed based on the atomic structures

of their targets. Examples include Agenerase® and Viracept® developed using the crystal structure of HIV-1 protease (Hardy and Malikayil, 2003) and the flu drug Relenza® based on the crystal structure of neuraminidase (Varghese, 1999). Since the majority of successful drugs exert their effects by modifying the activity of proteins (Bourne and Weissig, 2003), attempts being made to understand the structure and function of entire proteomes are expected to increase validated therapeutic targets by up to 10-fold thus paving the way for new opportunities in structure-based drug design (Kuhn *et al*., 2002). A major advantage of X-ray crystallography which is making it increasingly attractive to the pharmaceutical industry is that it defines ligand-binding sites with more certainty. It is also more sensitive than the bio-assays used in high throughput screening thus it can detect lower affinity and therefore lower molecular weight bound ligands (Tickle *et al.*, 2004). The role of X-ray crystallography in the drug design industry is actually set to appreciate with the advent of high throughput crystallography (reviewed in Blundell and Patel, 2004). An important caveat the use of crystal structures in drug design, however, is the static nature of these structures. Receptor–ligand structures solved by X-ray crystallography do not reflect the flexible nature of ligand-binding pockets and as such errors may arise in the design of the chemical libraries.

Structure-based drug design (rational drug design) proceeds via several phases which can be summarized as follows: target identification, lead identification, structural assessment and optimization of the lead, and finally drug candidate validation (Figure 1.6). Structural information plays a vital role in target identification and selection, screening of small

molecules and identification of hits, and more importantly in engineering improved affinity and selectivity during lead optimization. In the absence of experimental structural data of putative targets, homology modeling of the 3D structures also provides insights into the ligand binding sites of these targets.

In target identification, macromolecules unique to the organism and possessing altered properties from those of the host are good candidates (Subbaya *et al.,* 1997). It is however, important to validate the target by proving that modulating its activity affects the disease process. This is done in a number of ways which include gene knockouts and RNA antisense technology to inactivate the gene (Kurreck, 2003).

Due to its vital role in plasmodium, HGPRT inhibition is detrimental to the survival of the parasite and offers an avenue for the design of chemotherapeutic molecules. This is the premise for our work.

After target identification, lead compounds are identified. Knowledge of natural ligands, high throughput *in silico* screening and combinatorial chemistry are ways in which leads are generated and ultimately a chemical library built. The role of protein structure lies in focusing the design and synthesis of leads toward a set of compounds more likely to bind to the target's active site (Salemme *et al.,* 1997; Bourne and Weissig, 2003). Hits-to-leads chemistry is required to assess and prioritize initial results and to synthesize close analogs

of active structures while establishing outline structure activity relationship (SAR). All this is aimed at improving potency by lowering the K_i to the nanomolar range.

Lead optimization is the longest pre-clinical phase and involves improving the biopharmaceutical properties of the chemical lead to enhance its suitability as a drug. In this, Lipinski's rule of five (Lipinski *et al.*, 1997) examines the drug-like properties of a small molecule. Once a drug candidate is selected, pre-clinical animal safety studies are then carried out after which clinical trials are undertaken.

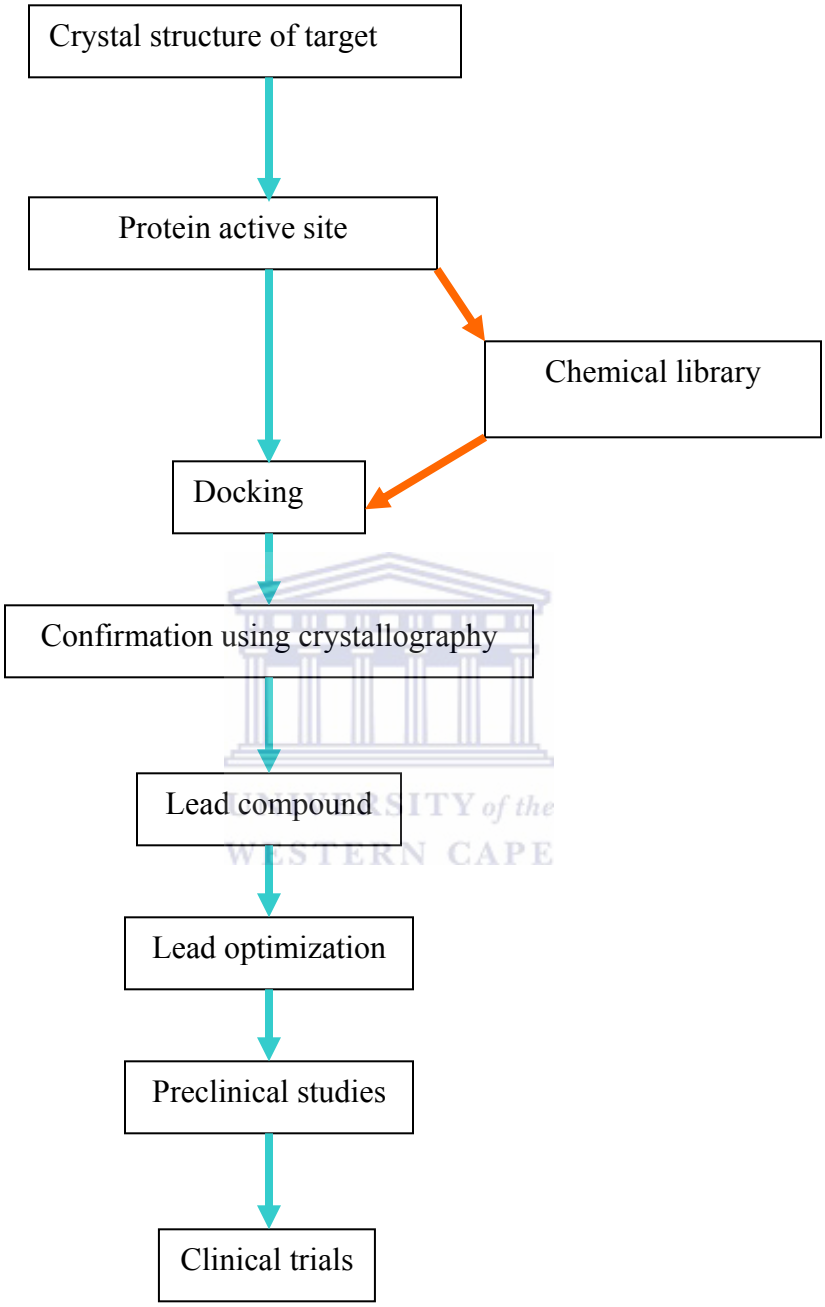


Figure 1.6: Schematic of the steps involved in structure-based drug design (Modified from Jhoti, 2001).

2.2 Photoaffinity labeling

Photoaffinity labeling (PAL) is a biochemical method which utilizes photoactivable ligand analogs to identify and characterize the binding sites of target proteins. Discovery of this technique is credited to Westheimer, Singh and Thornton (1962) who used diazoesters and carbene intermediates to label the active site amino acids of α-chymotrypsin.

Several functional groups, mostly azido and diazo groups, diazonium ions and diazirines, are incorporated into the photoprobes to serve as precursors for the highly reactive nitrenes, carbenes and carbocations intermediates upon photolysis (Brunner, 1993). Formation of these intermediates which in turn form covalent linkages with the target protein is accompanied by chemical changes, for example, extrusion of molecular nitrogen in the case of azido compounds. The photolabeling is, however, preceded by the formation of a reversible protein-photoprobe complex as indicated by the rate saturation effect and competitive inhibition by the natural substrate. Detailed mechanisms are elaborated below.

The success of a photolabeling experiment largely depends on the efficiency and specificity of the photoprobe employed. This imposes the condition that the photoprobe used should give rise to intermediates with a very short half life and no re-arrangement into more stable, but nevertheless, reactive compounds as that will lead to indiscriminate labeling. Paradoxically, highly reactive reagents tend to self-stabilize through

intramolecular rearrangements to form stable reactive products (Brunner, 1993) and as such, design of photoprobes involves a compromise between efficiency and specificity. Additional requirements for the photoprobe are: stability in the dark to allow targeting to the intended site, small size to avoid steric obstruction in the binding site, photolysis at wavelengths that do not damage the protein and formation of easily isolable reaction products. The photoprobes must also carry an easily detectable tag bearing radioactivity, fluorescent or immunogenic marker (Dorman and Prestwich, 2000).

Although a number of photoprobes have been developed over the years, the aryl azides introduced by Bayley and Knowles (1977) are the most popular. This is possibly because of their relatively easier synthesis and stability in the dark (Rizk *et al*., 2006). Substituents on the aromatic ring influence the characteristics and reactivity of aryl azides as exemplified by the perfluorophenylazides (Keana and Cai, 1989). Young and Platz (1991) have shown that fluorinated aryl azides covalently bind to the target sites via the singlet fluorinated arylnitrene which, compared to the unsubstituted aryl azides, rearranges more slowly to the corresponding didehydroazepine. The slow rearrangement allows the intermediate to react with many chemical bonds including C-H bonds by insertion. In 3-nitrophenyl azide, shifting of the absorption wavelength beyond the protein absorption region by the nitro group permits the selective excitation of this azide in protein mixtures (Rizk *et al*., 2005). Other photoprobes include azidonucleotides, aryldiazirines, benzophenones and aryldiazonium salts.

2.2.1 General photochemistry of photoprobes

Arylazides, upon irradiation, are promoted to the singlet excited state which readily loses molecular nitrogen to form singlet nitrene. The singlet species generated are powerful electrophiles that react with carbon-carbon double bonds, heteroatom-hydrogen bonds and can even insert into carbon-hydrogen bonds as is the case in perfluorophenylazides. The most common phenomena for the singlet species, however, is the rapid ring expansion into cyclic ketenimines (didehydroazepines) via azirine intermediates or relaxation to lower energy triplet nitrenes through intersystem crossing (ISC). Both didehydroazepines and triplet nitrenes are known to react with the target site amino acid residues. The fate of singlet nitrene is schematically illustrated in Figure 1.7 below.

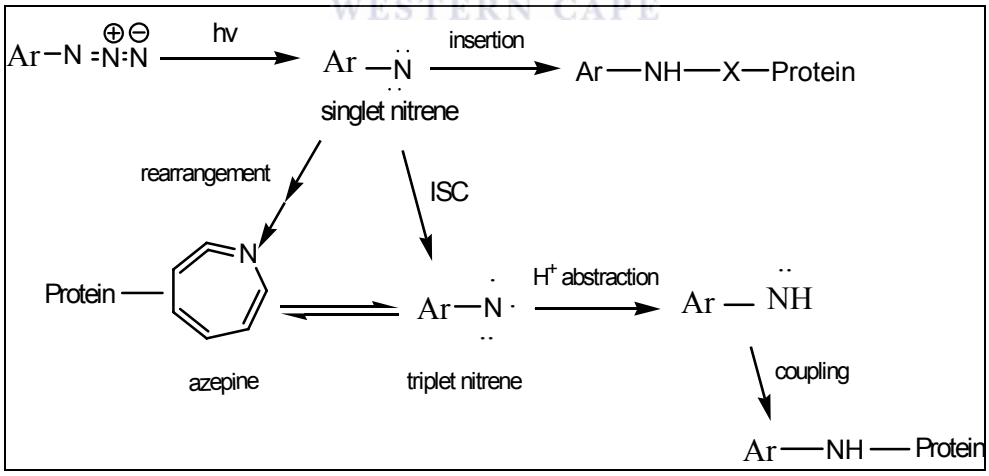


Figure 1.7: Possible mechanisms for aryl azides binding to proteins

Although the notion that the short-lived singlet nitrenes are the general active intermediates for the aryl azides has been challenged before (Bayley and Staros, 1984), a recent study by Rizk *et al*. (2006) further proves that the didehydroazepines and azepinone intermediates produced are actually responsible for photolabeling.

Radiolabeled azidonucleotides, normally containing an azido group on the C_8 position of the heterocyclic ring, have been extensively used as photoprobes. The introduction of a substituent at the C_8 position may, however, cause the compound to adopt a syn conformation as opposed to the anti conformation present in the parent nucleotide (Scheit, 1980) and this may affect binding onto the target site. The presence of a nitrogen atom ortho to the azido group, as in 2-azido nucleotides may also be problematic due to isomerization of the azide into tetrazole which is less sensitive to photolysis.

Like in the case of aryl azides, it has been assumed that the photochemistry of this group proceeds via formation of short lived nitrenes (Hanna *et al*., 1989; Hanna *et al*., 1999). This view has, however, been challenged by Polshakov *et al*. (2005) who, employing laser flash photolysis, show that diazaquinodimethane derived from the singlet nitrene form covalent linkages with nucleophiles in photolabeling experiments involving azidonucleotides.

In this study, we explore the possible use of radiolabeled 2′, 3′-*O*-(2, 4, 6-trinitrophenyl)-8-azido-adenosine triphosphate (TNP-8N_3-ATP) originally developed by Seebregts and

McIntosh (1988) to probe the active site of sarcoplasmic reticulum Ca^{2+} -ATPase, in photolabeling the active site of PfHPRT. Previous studies in our laboratory have indicated that TNP-GMP is a potent inhibitor of PfHPRT (Phehane, 2002) where it was

hypothesized that the TNP moiety occupied the position of Mg^{2+}PPi. In the case of the 8-azido nucleotides the tendency to adopt the syn orientation of the purine about the glycosidic bond would alter the purine binding mode and possibly allow the adenine base to bind in spite the fact that adenine is not normally a substrate. Atomic structures of HGPRTs with GMP, IMP indicate a rather open active site where loop II has not closed down and parts of the phosphate is exposed to the medium suggesting that a triphosphate moiety could be accommodated. This compound is an attractive photoprobe because in addition to the radioactive tag, it carries the fluorescent trinitrophenyl group making its visualization easy. We rationalize that it will bind to the nucleotide binding region in PfHPRT. In a further development we have deaminated TNP-8N_3-ATP to TNP-8N_3-ITP and explored its potential as a photoprobe of PfHPRT.

2.2.2 Application of photoaffinity labeling in drug discovery

PAL provides valuable insights regarding drug-target interactions in structure-based drug design (reviewed in Dorman and Prestwich, 2000). At the primary level, photoprobes are

used to identify target proteins of interest. At the next level, information pertaining to the ligand binding mode, affinity and specificity can be obtained by competitive photolabeling. This is extremely useful when other binding methods are problematic e.g. lack of a radiolabeled ligand. The site derivatization may be obtained by sequencing the short labeled peptide fragments. Figure 1.8 summarizes these steps schematically.

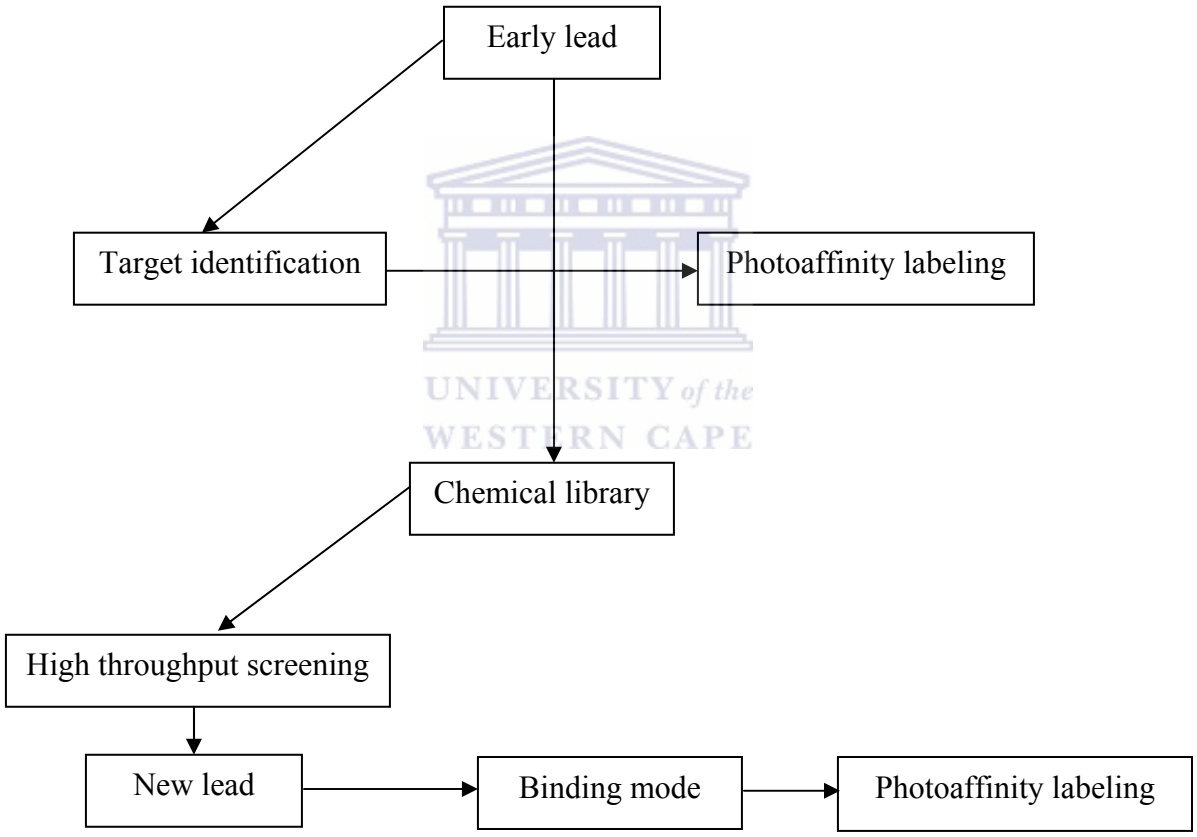


Fig 1.8: The various stages that PAL can be utilized in a drug discovery process (modified from Dorman and Prestwich, 2000).

2.3 Structural bioinformatics

Full understanding of the biological role of proteins requires knowledge of their structure which in turn helps to understand their function. Despite the fact that the number of experimentally solved structures has increased exponentially over the last few years due to advances in X-ray crystallography and NMR methodologies, not many structures are known when compared with the number of elucidated sequences. This sequence-structure gap is largely attributed to the success of the various genome sequencing projects. One focus now is on decoding structural information from sequences through experimental and computational methods. The latter, though prone to inaccuracies (Bonneau *et al*., 2002), provides an avenue for narrowing the sequence-structure gap as it saves time and human effort.

Computational structure prediction methods can be grouped into two classes: (i) homology modeling and (ii) *ab initio* methods (Baker and Sali, 2001). Homology modeling is based on the observation that evolutionarily related sequences tend to have similar three-dimensional structures (Chothia and Lesk, 1986). Thus, given a sequence (target), a model of its three-dimensional structure can be predicted based on a known structure (template) with which it shares considerable sequence identity. Some exceptions to this rule of sequence-structure similarity have, however, been reported (Panchenko *et al.,* 2005) and in such cases, threading comes into play. *Ab initio* structure prediction, on the other hand, solely relies on the sequence information and is usually employed in cases

where no known structure for the sequence in question exists. Two hypotheses, local and global, have been projected to explain how the amino acid sequence of a protein determines its fold (Lattman and Rose, 1993). In the local hypothesis it is argued that only a few critical residues in a protein sequence determine its fold while the global hypothesis proposes that each amino acid position ultimately contributes to protein folding. Studies support both hypotheses (Thomas *et al*., 1995; Bowie *et al*., 1990) but the finding by Russell and co-workers (1997) that very low sequence identities result in striking structural variation seem to give weight to the local hypothesis in that critical residues are likely to be altered at low sequence identity (Wood and Pearson, 1999). This also ties in with the observation that useful protein models can only be built by comparative modeling if the sequence identity between the target and template sequence is over 40% (Chakrabarti *et al*., 2003).

2.3.1 Homology modeling

This knowledge-based strategy can be split into two approaches: comparative modeling and threading. In the first approach, three-dimensional model building heavily relies on sequence similarity between the target and the template. The basic assumption here is that with increasing sequence identity, two proteins are likely to have a similar fold. This is nevertheless not an exhaustive approach since most similar structure pairs seem to have less than 12% sequence identity (Rost, 1997).

Threading (Jones *et al*., 1994; Jones, 1999) on the other hand unambiguously includes template structural information in the modeling procedure (Fischer and Eisenberg, 2003) and as a result can identify distant homologues. Fold assignment, in this case, is by mapping the unknown sequence through each structure in a library of experimentally solved structures. The ability to recognize distant homologues arises from the fact that structure is more conserved than sequence during evolution (Lesk and Chothia, 1980). Additionally, threading can identify unrelated proteins which share a common fold. Considering the estimation that 75% of homologous enzymes share below 30% sequence identity (Todd *et al*., 2001), threading is commonly employed nowadays in homology modeling. The applicability of homology modeling thus depends on finding a suitable template i.e. one with at least some detectable similarity with the target sequence.

2.3.2 Homology modeling steps

2.3.2.1 Searching for homologs

Comparative modeling is a multi-step process which commences with the identification of known structures (templates) related to the target sequence. Computer programs which utilise the target sequence as input are used to query sequence and/or structural databases such as the National Centre for Biotechnology Information (NCBI), TrEMBL/SWISS-PROT, the Protein Data Bank (PDB, Berman et al., 2000), DALI (Holm and Sander, 1998), CATH (Orengo *et al*., 1997) and the Structural Classification of Proteins (SCOP, Murzin *et al*., 1995). These programs can either be exhaustive, heuristic, threading or the Hidden Markov Model (HMM) algorithms. Exhaustive programs compare the unknown

sequence with each entry in a database as opposed to the heuristic programs which scan the database for regions of strong homology. While this latter option compromises sensitivity i.e. detection of distantly related sequences, it saves on the time taken to search a database and also minimizes computational hardware requirements. The obvious pitfall to this method, thus, is that by not comparing the target with all the database entries, weak similarities are possibly missed. Also, since the final statistics are based on a subset of and not the whole database entries, their significance could be doubtful.

The original dynamic programming algorithm (Needleman and Wunsch, 1970) and its modification (Smith and Waterman, 1981) are examples of exhaustive algorithms while the Basic Local Alignment Search Tool ((BLAST) Altschul *et al*., 1990; Altschul *et al*., 1997), Position-Specific Iterated BLAST (PSI-BLAST) and FAST-All ((FASTA) Pearson and Lipman, 1988) represent heuristic methods. The Needleman and Wunsch algorithm produces an optimal global alignment of sequences and even allows the introduction of gaps. BLAST and FASTA both perform pairwise sequence alignments while searching the data base. BLAST has been developed further into specialized forms such as BLAST2 and BLASTp which both query a protein database using a given protein sequence, PHI-BLAST (pattern hit initiated blast), RPS-BLAST used for searching for protein domains in a conserved Domains Database and the Position-Specific Iterated BLAST (PSI-BLAST). A PSI-BLAST search progresses via profile-sequence alignment in which the weighted multiple sequence alignment generated after the initial search is used to construct a position specific matrix that further searches the database for

additional homologs. This iterative procedure makes PSI-BLAST more sensitive. Newer programs include SAM-T02 which employs the Hidden Markov Models method, SALIGN which employs a profile-profile alignment procedure and several threading programs among them GenThreader (Jones, 1999a; McGuffin and Jones, 2003), 3D-PSSM (now phyre; Kelley *et al*., 2000) and FUGUE (Shi *et al*., 2001). A thorough template search involves the use of an assortment of these programs.

The output of a given program depends on the scoring matrix used by the program. The matrix defines the similarity score, penalties for opening and extending gaps and the mismatch penalty. Point Accepted Mutation (PAM; Dayhoff, 1978) and BLOSUM (Henikoff and Henikoff, 1992) matrices are employed by most search programs.

2.3.2.2 Choosing the template

After homologs to the unknown sequence are identified, the next step is choosing the most suitable template for alignment. Although high sequence similarity between the target and template results in superior models, several other factors are also considered. When considering templates solved by X-ray crystallography, the resolution and the reliability factor (R-factor) both of which determine the quality of the structure are important. In the case of NMR structures, the number of restraints per residue is vital. Since protein families can further be grouped into subfamilies, multiple sequence alignments together with phylogeny trees aid in selecting template(s) from the same subfamily as the target sequence (Retief, 2000). The target-template environment (Marti-

Renom *et al*., 2000; Fiser, 2006) should also be factored. For example, when modeling a protein bound to a ligand, a template complexed with a similar ligand or its analog is more appropriate.

The use of multiple templates generally improves the model accuracy (Sanchez and Sali 1997). In modeling using multiple templates, one can choose to align the templates with different domains of the target with little overlap between in which case the entire target sequence is modeled, or the templates can be aligned to the same region of the target and the model built based on the best local template (Fiser, 2006).

Ultimately, the best template picked depends on the quality of the model arising from the target-template alignment. Model quality, determined by energy or scoring functions will be elaborated further below.

2.3.2.3 Aligning the target to the template

The accuracy of the target-template(s) alignment forms the core of homology modeling since it describes the correspondence between the target and template residues. Considering a misalignment by a single residue results in an estimated error of 4 Å in the model (Fiser, 2004), the accuracy of the alignment dictates the quality of the resulting model structure. The importance of correctly aligning a given sequence to the chosen template has been particularly overemphasized in a number of reviews (Dunbrack, 2006; Ginalski, 2006; Sali, 1995).

With high sequence identity (> 40%) between the target and the template, alignment is uncomplicated and can be confidently performed using any of the available automatic programs (see below) with minimal manual intervention. This range of sequence similarity has few gaps and also, the structural differences that exist are confined to the loop regions and the sidechains (Sali, 1995). Below 30% sequence identity however, alignment errors increase making this a major cause of incorrect models in comparative modeling (Fiser *et al*., 2002). In such cases, threading is employed. It is noteworthy to point out though that any given alignment should be manually inspected to ensure that the conserved residues, insertions or deletions are within the appropriate secondary structure elements.

Several sequence-sequence and sequence-structure (threading) alignment programs have been developed. Two most popular ones are ClustalW (Thompson *et al*., 1994) and T-Coffee (Notredame *et al*., 2000). These programs are both based on the progressive alignment strategy (Feng and Doolittle, 1987; Taylor 1988) in contrast to earlier rigorous programs such as MSA (Lipman et al., 1989) and DCA (Stoye et al., 1997). While ClustalW and T-Coffee perform multiple sequence alignments, they differ in the fact that ClustalW performs a global optimal alignment of the sequences while T-Coffee does a local-global alignment. Since global alignments with ClustalW deteriorate at low sequence identity (Sauder *et al*., 2000), T-Coffee is becoming even more favoured.

Once the sequence alignments have been optimized, the JOY program (Mizuguchi *et al*., 1998) can be used to correlate the alignment information with the three-dimensional structural features therefore extracting useful information in a given sequence alignment which aids in understanding the conservation of amino acids in their specific local environments.

Sequence-structure alignment programs include FUGUE (Shi *et al*., 2001), GenThreader (Jones, 1999b) and 3D-PSSM (Kelley *et al*., 2000). Generally these programs match a query sequence (target) with a known structure (template) and assess the compatibility using mean force potentials (Sippl, 1995). The way in which each program executes its function, however, varies. The flexible hybrid of profile and hidden Markov model (HMM) sequence-structural alignment strategy implemented in FUGUE selects the most optimal alignment based on the length difference between the sequence and the structure to be aligned. The algorithm takes into account the residue environment in the template when comparing it with the target sequence equivalent. Thus, the substitution tables are environment specific.

GenThreader, on the other hand, approaches the alignment problem in two ways: (i) using the target sequence, a sequence profile alignment is undertaken using the global-local dynamic programming algorithm and scanned through a structural database or (ii) a profile of the template sequences is built and then matched against the target sequence.

The program uses both approaches concurrently but takes the highest scoring alignment as the preferred one. The alignment is then evaluated using a statistical pairwise potential of mean force [(Hendlich *et al*., 1990; Jones *et al*., 1992) and a solvation potential (Jones *et al*., 1992).

Lastly, in 3D-PSSM, the multiple sequence profile generated by aligning the target sequence with its homologs using the PSI-Blast routine is combined with a novel structural alignment of remote template homologues which includes their secondary structure information and solvation potentials. The resulting alignment is therefore coded for by multiple sequences and structures.

The best possible alignment from either the sequence based method or the fold recognition method is then used for modeling the probable structure of the given sequence.

2.3.2.4 Model building

There are three ways in which the actual modeling proceeds. In the first approach, superposition of the conserved regions of the templates forms the three dimensional outline of the model (Browne *et al*., 1969; Blundell *et al*., 1987). Once this assemblage of template core regions is in place, loops and side chains are then modeled (Greer, 1990). In this case, the use of multiple templates with considerable sequence identity increases

the model accuracy (Srinivasan and Blundell, 1993). This method is implemented in an early model building program COMPOSER (Sutcliffe *et al*., 1987).

The second approach, termed segment matching or co-ordinate reconstruction (Fiser, 2006), uses a subset of the template(s) atomic positions, usually the C_α atoms of residues in the conserved fragments of the target-template alignment, to calculate the co-ordinates of the target sequence atoms (Jones and Thirup, 1986; Levitt, 1992). This method is attractive in that it can model main chain and side chain atoms, gaps and loops.

The last and most widely used method is modeling by satisfaction of spatial restraints (Sali and Blundell, 1993). Here, the target-template alignment is used to derive restraints to be imposed on the model structure by assuming that the distances between the aligned residues in the target and the template are equal. In addition to these homology based restraints, stereochemical restraints from the molecular mechanics force field (Brooks *et al.,* 1983) are also applied to the model. Model restraints violation is then minimized by real space optimization or distance geometry.

Modeling by satisfaction of spatial restraints is implemented in the popular program MODELLER (Sali *et al*., 1990; Sali and Blundell, 1993; Fiser *et al*, 2000). As a first step, MODELLER derives distance and dihedral angle restraints from the target-template alignment and expresses them as probability density functions. These are then converted into an objective function which is then optimized in Cartesian space using the variable

target function which uses conjugate gradient and molecular dynamics with simulated annealing.

2.3.3 Loop modeling

Loops connect the regular secondary structure elements (i.e. alpha helices and beta sheets) of proteins and frequently constitute the variable regions among similar structures (Chothia and Lesk, 1986). Although mostly found on the hydrophilic surface of proteins, they form an important part of the protein structure by contributing to the active site and thus may determine functional specificity (Todd *et al*., 2001). Even though not a trivial task, the accurate modeling of loop conformation is a critical step in homology modeling since this determines the model's usefulness in deciphering protein-ligand interactions and examining active and binding sites (Fernandez-Fuentes *et al*., 2006). Therefore, like in the case of alignment, the modeler's input is essential at this stage and serious thought on which of the seemingly similar choices to pick determines whether one ends up with the correct model.

For the modeling of loops, two main methods have been developed. One involves searching the databases of solved structures for known loop fragments (canonical loops) that fit into the two stem regions cradling the loop (Deanne and Blundell, 2000; Shepherd *et al*., 1999; Krieger *et al*., 2003; Fiser *et al*., 2000). Once obtained, the fragments are evaluated and then chosen based on geometrical properties and sequence matching to both the target and template(s) loop sequences. The chosen fragments are then

superposed and annealed onto the stems and subsequently refined by energy function optimization (Fiser *et al*., 2000). Here, the secondary structure elements bracing the loop to be modeled and the spatially close regions influence the fragment choice (Chothia and Lesk, 1987; Fernandez-Fuentes *et al*., 2006).

Although limited by the exponential increase in the number of possible geometrical conformations as the loop lengthens, and the considerably small number of solved structures, this popular approach has been incorporated in a number of modeling programs and servers such as ArchPRED (Fuentes *et al*., 2006), MODELLER (Sali and Blundell, 1993), WHAT IF (Vriend, 1990) and Swiss-Model (Peitsch *et al*., 2000).

The exponential increase in the number of possible geometrical conformations seen with the direct database search method led to the development of a slightly different technique which uses the CHARMM non-bonded energy function (Brooks *et al*., 1983) for optimization and ranking of probable loop fragments (van Vlijmen and Karplus, 1997). This offers the advantage of slowly increasing the number of fragments to be minimized with the CHARMM function as the loop length increases as opposed to the exponential increment seen with method one.

The second approach originated with Go and Scheraga (1970) who first proposed a method for predicting the conformation of a segment joining two core regions in a solved structure. Since then, numerous studies on *ab initio* loop modeling have been conducted

(Pedersen and Moult, 1995; Fiser *et al*., 2000; Deanne and Blundell, 2000). This method involves a conformational search for the global energy minima. Different energy function provisions and optimization algorithms have been employed in the search and include genetic algorithms (Ring and Cohen, 1994), Monte Carlo and molecular dynamics (Rapp and Friesner, 1999) and self-consistent field optimization (Koehl and Delarue, 1995). An attractive feature of the method is its ability to carry out modeling of several loops simultaneously (Fiser *et al*., 2000).

Apart from the two main approaches, some programs implement a combined algorithm. An example is CODA (Deanne and Blundell, 2001) which integrates a database search program, FREAD, with an *ab initio* method, PETRA (Deanne and Blundell, 2000).

2.3.4 Side chain modeling

The importance of correctly modeling side chains in a protein structure arises from the central role that side chains play in molecular recognition. For example, they are involved in the packing of protein hydrophobic cores, binding of ligands and regulation of protein-protein interactions (Vasquez, 1996).

Side chain modeling is, however, difficult due to the large number of possible combinations of conformations encountered and as a result, methods commonly utilized perform limited sampling of side-chain torsion angle space. Generally, two procedures to

tackle this problem have been developed. The popular prediction strategy is to choose a set of torsion angles which represent the most frequent orientations i.e. rotamers observed in solved structures (Swain and Kemp, 2001; Petrella and Karplus, 2002) and model the unknown side chains based on this. The rotamer search may consider main-chain-side-chain interactions or may just focus on side-chain-side-chain interactions (Tanimura *et al*., 1994). The suitability of the rotamers is then measured using some scoring function (Peterson *et al.,* 2004). Several algorithms such as the genetic algorithm (Jones, 1994), simulated annealing (Koehl and Levitt, 1999) and the dead-end elimination algorithm (Voigt et al., 2000) have been developed for the rotamer search and placement.

Despite being relatively inexpensive in terms of computational speed, this cannot always map all model side chains accurately onto a given rotamer library. Such was the case in a study conducted by Bower and co-workers (1997) in which approximately 6 % of the more than 40,000 side-chains studied exhibited a variance of ~40 ° from any rotamer in their library. Furthermore, an analysis of high resolution crystal structures conducted by Lovell and co-workers (2000) found out that about 5 % of residues in these structures did not match any of the 153 rotamers in their library. These 'off-rotamer' (Petrella and Karplus, 2002) side chain conformations can be modeled using an alternative strategy referred to as *ab initio* modeling.

In *ab initio* side chain modeling, a search for the global minimum energy, is performed empirically by optimization of an energy function. This approach, however, is computationally intensive and thus not widely applied.

Ultimately the two methods yield comparable results since they both involve energy minimization to find the most suitable side chain conformation. They differ from one another primarily in two aspects; the selection of an initial conformation and the survival criteria used to produce the next generation in an iteration procedure.

The complexity of side chain modeling is exemplified by the fact that only a limited number of computer programs are publicly available (Canutescu *et al.,* 2003). The program side-chains with a rotamer library ((SCWRL) Bower *et al.,* 1997) is one such program that is widely used because of its speed, accuracy, and simplicity. The algorithm building of an initial residue placement is strongly backbone dependent. After the initial placement, methodical searches are carried out to minimize steric clashes. Overall, the program assumes that the local main-chain conformation of each residue largely determines the side-chain conformation and that a search strategy to mitigate steric clashing is also essential for precise placements.

2.3.5 Model quality assessment and usefulness

Once a three-dimensional model has been built, the final step of evaluating its accuracy is then carried out. During this stage, errors that might have occurred during earlier stages are highlighted. Target-template(s) mis-alignments, mistakes in loop modeling and error in side chain conformation assignment are frequently encountered causes of improper models. Usually, when assessing a model, sequence identity between the target and the template(s) is a good indicator of the expected accuracy. With more than 50% identity, models with up to 1 Å random mean standard deviation (rmsd) of the backbone carbon atoms compared to solved structures can be obtained (Ginalski, 2006).

Evaluating the model stereochemistry is an obvious starting point in quality assessment. In doing this, bond lengths, bond angles, torsion angles and hydrogen bonding energies are some of the parameters checked for comparability with those normally encountered in solved structures. The Ramachandran plot (Ramachandran *et al.,* 1963), a plot of phi versus psi main-chain angles, is routinely used to check the stereochemical quality of a model. If the quality of the model is good, most of the residues will fall in the favoured regions of the plot. Glycines and prolines are, however, exempted from the clustering rule due to their unique side chain properties. Software programs based on the Ramachandran plot have been written and include the popular PROCHECK (Laskowski *et al.,* 1993), WHAT-CHECK (Hooft *et al*., 1996) in the WHAT-IF suite (Vriend, 1990).

An alternative model evaluation protocol involves the use of energy functions such as the statistical potential functions implemented in PROSA (Sippl, 1995). This approach is, however, not very common and is mostly undertaken when authenticating whether the correct template was used for the modeling procedure.

2.4 Research aims

Several aims were set for this project. One was to purify the malarial enzyme and attempt co-crystallizing it with a chalcone synthesized in our laboratory. This chalcone has been shown to accelerate the enzyme turnover, activate an inactive to active transition and also inhibit the enzyme in the presence of Ca^{2+} (Mbewe, 2005). We wanted to validate whether this chalcone binds to the active site or an effector site on the enzyme. In case of the former, successful structure solution of PfHPRT with the chalcone would then facilitate development of better binding chalcone inhibitors.

We also set out to model the structure of the malarial enzyme in the open conformation based on the structures of other phosphoribosyltransferases in the same state. To our knowledge, neither the crystal structure of the unliganded PfHPRT nor its homology model has been elucidated. Availabilty of a free PfHPRT model would then facilitate docking of ligands onto the enzyme's active site.

We also aimed to photalabel the active site of PfHPRT using [^{32}P]TNP-8N_3-ATP. In addition, we set out to synthesize a new azidonucleotide derivative, [^{32}P]TNP-8N_3-ITP,

from [^{32}P]TNP-8N_3-ATP and evaluate whether it was useful as a photoprobe of the malarial enzyme.

CHAPTER TWO: MATERIALS AND METHODS

2.0 Reagents and Chemicals

Reagent or Chemical	Supplier
Ammonium persulfate	Bio-Rad, USA
Disodium Hydrogen Phosphate ($Na_2HPO_4. 2H_2O$)	Merck, Germany
Sodium dihydrogen phosphate ($NaH_2PO_4. H_2O$)	Merck, Germany
Hepes	Sigma, USA
3-Morpholinopropanesulfonic acid (MOPS)	Sigma, USA
Sodium pyrophosphate	Sigma, USA
Polyethylene glycol 4000	Merck, Germany
Tris(hydroxymethyl)aminomethane ($H_2NC(CH_2OH)_3$)	Merck, Germany
Sodium chloride (NaCl)	Merck, Germany
Sodium dodecyl sulfate (SDS)	Bio-Rad, USA
Acrylamide/Bis solution, 19:1 (5% C)	Bio-Rad, USA
Dnase1	Roche, Switzerland
Dithiothreitol (DTT)	Sigma, USA
Diethylaminoethyl cellulose (DE-52)	Whatman, USA
Reactive Red 120 agarose	Sigma, USA
Tetramethylammonium hydroxide (TMAH)	Fluka, USA

Coomassie Brilliant Blue R-250	Sigma, USA
Ethylenediaminetetraacetic acid (EDTA)	BDH AnalR, England
Magnesium chloride hexahydrate ($MgCl_2$. $6H_20$)	Merck, Germany
Yeast Extract	Merck, Germany
Tryptone	Merck, Germany
Chloramphenicol	Roche, Switzerland
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	Sigma, USA
Ampicillin	Roche, Switzerland
Crystal screen 1 & 2	Hampton Research, USA
Structural screen 1 & 2	MDL, USA
Silicon oil	Hampton Research, USA
Paraffin oil	Hampton Research, USA
Ethanol (EtOH) [99.7-100% v/v)	BDH AnalR, England
Methanol	BDH AnalR, England
Acetic acid, Glacial	BDH AnalR, England
Adenosine monophosphate (AMP)	Sigma, USA
Guanosine monophosphate (GMP)	Sigma, USA
Dimethyl sulfoxide (DMSO)	Merck, Germany
Sodium hydroxide (NaOH)	Merck, Germany

Molecular weight marker kit for gel filtration chromatography (MW-GF-200 Kit) was from Sigma-Aldrich, VDX plates, 72 well microbatch plate, seeding tool, cryo-loops, glass sitting drop rods, 22 mm thick siliconized circle cover slides and Dow Corning vacuum grease were all from Hampton Research.

2.1 Culture media

2.1.1 Luria Bertani (LB) Broth

Luria Bertani broth, supplemented with 10 mM sodium phosphate, was prepared as follows:

Constituents, per litre	
Tryptone	10 g
Yeast extract	5 g
NaCl	5 g
Na_2HPO_4. $2H_2O$	1.78 g
NaH_2PO_4. H_2O	1.38 g

pH adjusted to 7.3 ± 0.2 with 5 N NaOH at 25 °C before autoclaving.

2.2 Protein expression

The glycerol stocks of BL21(DE3)pLysS *Escherichia coli* cells transformed with pET-17b expression plasmid harboring the malarial HPRT cDNA were prepared by Dr. Boniface Mbewe in our laboratory (2005) and stored in the -80 ºC freezer.

1 L of Luria-Bertani medium supplemented with 10 mM sodium phosphate, 100 µg/mL ampicillin and 34 µg/mL chloramphenicol was inoculated with ~20 µL of BL21(DE3)pLysS *E. coli* cells. The cells were grown overnight at 37 ºC without shaking until an OD_{600} of 0.4-0.6 was attained. PfHPRT expression was induced by the addition of IPTG to a final concentration of 0.4 mM. Expression was allowed to continue for three hours at 37 ºC with vigorous shaking at 200 rpm. Cells were then harvested in an Avanti J-30I high-speed centrifuge (Beckman Coulter) using a JA-14 rotor rotating at 8 000 rpm at 4 ºC for 10 minutes. The supernatant was decanted, pellet rinsed gently with de-ionised water and stored -80 ºC.

The bacterial pellets were then re-suspended in an equal volume (~25 mL) of ice-cold lysis buffer (50 mM Tris-HCl, pH 8.0, 25 mM NaCl, 2 mM EDTA containing 1 mM DTT, 1 mM PMSF) then disrupted gently mechanically using a Potter-Elfeljen glass homogeniser. To the pooled homogenate (100 mL), 5 mM $MgCI_2$ was added followed by 0.05 mg/mL DNAse1 (Sigma) to break down the DNA strands released upon cell lysis. The cell homogenate was then incubated for 30 min at 25 ºC while shaking at 160 rpm in a C24KC refrigerated incubator shaker (New Brunswick Scientific, NJ, USA). After, it

was divided into four 25 mL fractions and clarified by spinning at 21 000 x *g* for 30 minutes at 4 ºC in an Avanti® J-30I centrifuge with a JA-30.50 rotor (Beckman Coulter). The supernatant containing the soluble PfHPRT was then subjected to purification procedures.

2.3 Protein purification

Chromatographic steps were routinely performed at 4 ºC unless otherwise stated. Prior to use, buffers for HPLC experiments were filtered through 0.22-micron filters (Millipore). Protein samples were concentrated by ultrafiltration to a final volume of 1 mL using an Amicon 8010 stirred cell fitted with a 25 mm diameter Ultracel YM-10, NMWL, 10 kDa membrane (Millipore). Further concentration to microliter volumes was routinely carried out using 0.5 mL Amicon centrifugal filter devices fitted with Ultracel YM-10, NMWL, 10 kDa membrane (Millipore). Buffer exchange between chromatographic steps was achieved by dialysis using a 6-8 kDa NMWL dialysis tubing (Merck).

2.3.1 Tandem anionic and dye affinity chromatography

2.3.1.1 DEAE-cellulose column preparation

Approximately 10 g of DEAE cellulose resin (DE-52, Whatman) was suspended in ~ 250 mL of de-ionised water. The fines were decanted and the resin washed with more water, each time removing the fines. The slurry was loaded onto a 2.0 x 6.0 cm column which was then washed with 10 mL of cell lysis buffer.

2.3.1.2 Reactive Red 120 column preparation

Reactive Red 120-agarose suspension (type 3000-CL) in 0.5 M NaCl, 0.02% thimerosal (Sigma) was packed onto a 2 x 5 cm column which was then thoroughly washed with de-ionised water.

2.3.1.3 Chromatography on DE-52 and Reactive Red 120 columns.

The two columns were then placed in tandem with the anionic DE-52 column on top of the Reactive Red column. The supernatant was applied and allowed to flow through by gravity at a flow rate of 30 mL/h. The columns were then washed with 50 mL of 50 mM Tris-HCl, pH 8.0, 1 mM DTT, and 1 mM PMSF and detached. The Reactive Red 120 column was developed with 50 mL of 10 mM sodium phosphate, pH 7.0/ 200 mM NaCl and subsequently eluted with 50 mM sodium pyrophosphate in 10 mM sodium phosphate, pH 7.0 (PPi), collected in 10 mL fractions. The column was finally washed with 50 mL of 1 M NaCl. The purity of PfHPRT obtained from this and subsequent preparations were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE; Coomassie Blue stain, see below). Fractions containing PfHPRT were pooled, dialysed overnight in 10 mM sodium phosphate, pH 7.0, 1 mM DTT, concentrated to 1 mL using an Amicon 8010 stirred cell (Millipore). Further concentration for crystallization was carried out using 0.5 mL Amicon centrifugal devices.

2.3.1.4 Modifications to the chromatographic procedure

To further understand the separation behaviour on the two columns, several modifications were tested. Firstly, the two columns were run separately. The DE-52 flow through followed by the Tris-HCl wash was collected in 10 mL fractions, analysed on SDS-PAGE and fractions containing PfHPRT loaded onto the Reactive Red 120 column which was developed as described in section 2.1.1.3. Fractions containing PfHPRT, determined by SDS-PAGE gel, were treated as outlined earlier. In the next preparation, the pH of the DE-52 flow through and the wash fractions was monitored at room temperature to ascertain the highest pH to which the target enzyme was exposed to. Fractions containing PfHPRT in this preparation were again purified further on a Reactive Red 120 column. In line with the method of Mbewe *et al*. (2006) for the purifications described so far, the DE-52 column was washed with 10 mL of the lysis buffer before loading of the supernatant. For the subsequent experiment, the DE-52 column was equilibrated as follows: approximately 6 g of the resin was suspended in 60 mL of 10X 75 mM Tris-HCl, pH 8.9 and the fines washed off. This was followed by further suspension in 120 mL 1X 75 mM Tris-HCl, pH 8.9. The slurry was then packed onto the column and the pH of the effluent measured. The supernatant pH was adjusted to 8.9 before loading onto the column. The flow-through was collected in 10 mL fractions. The column was then washed with 50 mL of 50 mM Tris-HCl, pH 8.0, 1 mM DTT, and 1 mM PMSF also collected in 10 mL. Fractions containing PfHPRT were purified further on a Reactive Red 120 column as described earlier.

Finally, two purifications using a longer DE-52 column (2 x 17.5 cm) were carried out. In one purification, the column was only washed with 30 mL lysis buffer prior to applying the supernatant. In the other purification, the column was equilibrated with 10 column volumes of the lysis buffer. In the two cases, the column was washed with 80 mL of 50 mM Tris-HCl, pH 8.0, 1 mM PMSF, and 1 mM DTT after collecting the supernatant flow-through. The eluate was collected in 10 mL fractions. Other contaminating proteins were washed off the column with 100 ml of 1 M NaCl. Fractions containing PfHPRT as judged by SDS-PAGE were pooled, purified further on a Reactive Red 120 column as described previously.

2.4 Alternative purification strategies

2.4.1 Ammonium sulphate fractionation and dye affinity chromatography

Solid ammonium sulphate was added to 100 mL of the supernatant with gentle stirring to achieve 40-80% saturation. After the salt was completely dissolved, the solution was stirred for an additional 1 h before centrifugation at 10 000 x *g* for 30 min. The resulting pellet was re-suspended in a minimal volume of 50 mM Tris-HCl, pH 8.0, 1 mM DTT, and 1 mM PMSF, dialysed overnight against 10 mM MOPS, pH 7.0, 1 mM DTT and applied to a Reactive Red column 120 (2 x 5 cm) equilibrated with 10 mM MOPS, pH 7.0, 1 mM DTT. The column was washed with 20 mL of 10 mM MOPS, pH 7.0 and protein eluted with 50 mM sodium pyrophosphate in 10 mM sodium phosphate, pH 7.0. Fractions containing PfHPRT were pooled, dialysed against 10 mM sodium phosphate,

pH 7.0, 1 mM DTT, concentrated in an Amicon stirred cell and assayed for activity as described below.

2.4.2 Anion and cation exchange chromatography

The supernatant was loaded onto a DE-52 column (2 x 15 cm) equilibrated with 50 mM Tris-HCl, pH 8.0, 25 mM NaCl, 2 mM EDTA, 1 mM DTT, and 1 mM PMSF. Following collection of the flow through, the column was washed with 80 mL of 50 mM Tris-HCl, pH 8.0, 1 mM DTT, and 1 mM PMSF collected in 10 ml fractions. A final wash of the column was with 100 mL of 1M NaCl. SDS-PAGE gel was then run to determine the purity of the fractions containing PfHPRT. Fractions containing PfHPRT were pooled and passed through two cationic Sep-Pak cartridges held in tandem. The cartridges were then detached and each eluted with 5 mL of 10 mM MOPS, pH 7.0, 400 mM NaCl, and 0.05% NaN_3 collected in 1 mL fractions. SDS-PAGE (15%) gel was then run to determine purity attained. Pure fractions were pooled, dialysed against 10 mM sodium phosphate, pH 7.0, 1 mM DTT, concentrated by ultrafiltration in an Amicon stirred cell and assayed.

2.5 HPLC gel filtration chromatography

Size exclusion chromatography was used to examine the oligomeric status of PfHPRT under various conditions and also for further purification of PfHPRT. Experiments were performed at ambient temperature using a Tosohaas TSKgel G 3000 SW column (7.5 mm x 30 cm, 10 μm) attached to a Beckman System Gold 126 pump mechanism/168 diode array detector HPLC system. Elution was monitored at 280 nm. Molecular mass was

determined from a calibration curve using Alcohol dehydrogenase (Mw 150.0 kDa), albumin (Mw 67.0 kDa), ovalbumin (Mw 43.0 kDa), carbonic anhydrase (Mw 29.0 kDa) and cytochrome C (Mw 12.4 kDa) (MW-GF-200 Kit, Sigma-Aldrich) as protein markers. Blue dextran (2 000 kDa) and tyrosine (185 Da) were used to determine the excluded (V_o) and included void (V_i) respectively. Partition coefficient (K_m) values for the calibration curve were calculated as per Equation 1 and plotted against the logarithm of the molecular weight for the markers.

$$V_r = V_o + K_m V_i \qquad \text{(Equation 1)}$$

Where V_r = retention volume of protein

V_o = excluded volume

K_m = partition coefficient

V_i = included volume

2.5.1 Standards preparation and running buffer evaluation

Standard stocks (1 mg/mL) were prepared by dissolving the appropriate amount in de-ionized water. 20 µL of each standard was then taken and mixed to give the final standard cocktail from which 20 µL was routinely injected into the column. Three running buffer systems, 10 mM MOPS/TMAH, pH 7.0, 0.05% NaN_3, 10 mM MOPS/TMAH, pH 7.0, 100 mM NaCl, 0.05% NaN_3, and 10 mM MOPS/TMAH, pH 7.0, 400 mM NaCl, 0.05%

NaN_3 were tested at a flow rate of 1 mL/min. Duplicate runs were performed for each buffer system.

2.6 Analytical techniques

2.6.1 SDS-PAGE gel electrophoresis

All SDS-PAGE analyses were carried out according to the discontinuous method of Laemmli (1970). Resolving gels consisted of a 15% (w/v) polyacrylamide matrix (arylamide: bisacrylamide, 29:1) in 375 mM Tris-HCl and 0.1% (w/v) SDS, and were overlaid with a 3% (w/v) polyacrylamide stacking gel (acrylamide: bisacrylamide, 29:1) in 75 mM Tris-CI and 0.1% (w/v) SDS. After stacking, gels were allowed to polymerize for 2 h at room temperature. In some cases gels were prepared a day before electrophoresis, tightly wrapped in a clear plastic wrap and stored at 4 ºC. Samples were diluted 1:1 with SDS-PAGE sample application buffer (0.5 M Tris-CI, pH 6.8, 10% (w/v) SDS, 0.1% (w/v) bromophenol blue, 10% (v/v) glycerol and 2% (v/v) β-mercaptoethanol) and vortexed for ~10 sec before loading 10 µL into the wells. Gels were electrophoresed at 20 mA per gel for 1 h and 30 min in a Hoefer dual gel caster (Armesharm) electrophoresis system. Following electrophoresis, gels were stained with 0.25% (w/v) Coomassie Brilliant Blue R-250 (Sigma) stain solution overnight and destained for 1 h in 10% acetic acid, 30% ethanol followed by 24 h destaining in 7% acetic acid.

2.7 Determination of protein concentration

2.7.1 Bradford Assay

Protein concentration was determined by the method of Bradford (1976) using bovine serum albumin (BSA) as a standard. The assay was carried out as follows: 50 µL of sample followed by 750 µL of dH_20 were pipetted into 1.5 ml Eppendorf tubes into which 200 µl of Bradford reagent (Bio-Rad) was added. The mixture was incubated for 5 min at ambient temperature before measuring absorbance at 595 nm. A mixture containing 800 µl dH_2O and 200 µl Bradford reagent was used as a blank. The standard calibration curve obtained is shown in the appendix section.

2.7.2 Spectroscopic assay

During purification HPRT estimation was performed spectrophotometrically on an Agilent 8453 UV-visible spectrophotometer (Agilent technologies) by measuring the absorbance at 280 nm. The program PROTPARAM on the ExPASy web server (http://au.expasy.org/tools/protparam.html) which computes protein molar extinction coefficients (E_M) based on their amino acid sequence as per the method of Gill and von Hippel (1989) was used in calculating HPRT E_M. The value returned was 23, 755 M^{-1} cm^{-1}.

2.8 PfHPRT activation

Routinely, prior to setting up activation experiments, the protein was dialysed against 10 mM sodium phosphate buffer pH 7.0, 1 mM DTT, and 1 mM $MgCI_2$. Typical protein concentration for activation was 1 mg/mL. To aid solubility, base stock solutions were prepared in 0.4 M NaOH. PRPP stock solution was prepared in either de-ionised water or in 10 mM sodium phosphate, pH 7.0. 50 µL of purified protein was routinely incubated overnight at 4 ºC with PRPP and hypoxanthine at concentrations indicated in the table below.

PRPP (mM)	Hypoxanthine (mM)
1	0.06
0.2	0.06

2.8.1 PfHPRT activity assay

Enzyme activity for the purine bases guanine and hypoxanthine was determined in a continuous spectrophotometric assay on an Agilent Chemstation spectrophotometer at room temperature according to the methods of Keough *et al.* (1999) and Subbayya and Balaram (2000). The change in absorbance at 257.5 nm and 245 nm for guanine and hypoxanthine respectively was monitored. The 1 mL total assay mixture consisted of 100 mM Tris-HCI, pH 8.5, 110 mM $MgCI_2$ or 100 mM Tris-CI, pH 7.5, 12 mM $MgCI_2$. To the mixture, 1 mM PRPP and 0.06 mM hypoxanthine/guanine were added immediately followed by the addition of an appropriate amount protein (10-1000 µg) to initiate the

reaction. The values of the extinction coefficients were measured under assay conditions for the guanine and hypoxanthine reactions at pH 8.5 were 5817 and 2283 M^{-1} cm^{-1}, respectively.

2.9 Crystallization experiments

A number of strategies were explored in an attempt to crystallize free PfHPRT or in complex with IMP, GMP or chalcone.

2.9.1 Vapour diffusion screening

Crystal Screen kits I and II (Hampton Research) based on the sparse matrix sampling method (Jancarick and Kim, 1991) were used to test for probable crystallization conditions and also to develop cogent strategies for subsequent screens. The hanging drop vapour diffusion method was employed. To 200 µL of 10-15 mg/mL PfHPRT in 10 mM sodium phosphate, pH 7.0, 1 mM DTT, IMP or GMP and $MgCl_2$ were added to a final concentration of 1 mM for each. The trials for PfHPRT with the chalcone were set up in the presence of $CaCl_2$ also at 1 mM final concentration for both. The mixtures were then incubated on ice for 30 min to allow the substrate to bind. For each screen condition, 2 µL of protein plus substrate was dispensed onto a 22 mm diameter circular siliconized cover slide and mixed with an equal amount of precipitant. The cover slide was then gently but rapidly inverted and sealed (using Dow Corning® vacuum grease) over 1 mL of the precipitant held in the wells of a VDX plate (Hampton Research). The hanging drops were observed under the microscope immediately after set up, every day for a week

and once weekly and thereafter until equilibration was assumed complete. For each substrate, parallel screens for the 100 individual conditions were set up at 4 ºC and 18 ºC.

2.9.2 pH and precipitant concentration grid search.

Using the hanging drop vapour diffusion method, a grid search of pH versus precipitant concentration was performed for the promising conditions obtained from screening. An example of the grid screens set up is shown in Figure 2.0. The range for both pH and precipitant concentration was chosen to be on either side of the promising screen conditions. Protein and substrate concentrations were not adjusted. The incubation temperature was 18 °C

PEG 4000 / pH	20%	22%	24%	26%	28%	30%
7.0						
7.5						
8.0						
8.5						

Figure 2.0: An example of pH versus precipitant concentration grid search.

2.9.3 Vapour diffusion rate control

This method aids in slowing down the rate of nucleation and subsequent crystal growth. Hanging drop vapour diffusion method using VDX plates (Hampton Research) was employed. Paraffin oil, Silicon oil and Al's oil (a 1:1 mixture of paraffin and silicon oils) were separately tested. Trials were set up as follows: 1 mL of precipitant at the desired concentration and pH was dispensed into the wells. To a 1 µL drop of 15-20 mg/mL PfHPRT pipetted onto square siliconized cover slides, 0.5 µL of IMP, GMP or chalcone (10 mM stock for each) was added followed by 0.5 µL of $CaCI_2$. The drop was gently mixed with 2 µl of precipitant from an appropriate well. To the reservoir, 400 µL of oil was then gently layered on top. The cover slide was sealed over the well and the crystallization plate incubated at 18 ºC. Subsequently, trials were performed with 200-700 µL of oil layered onto the reservoir.

2.9.4 Modified microbatch screening

In this modified microbatch method, the drop containing the protein of interest combined with the precipitant is dispensed under a layer of oil which acts as sealant by preventing evaporation of water from the drop. Protein and precipitant concentrations thus remain almost constant. This method was used in an attempt to crystallize PfHPRT with GMP or IMP and was performed as follows:

Paraffin oil (6 mL) was pipetted into a 72 well microbatch plate (Hampton Research) so that the cone shaped wells on the plate were submerged. 2 µL of PfHPRT (incubated

beforehand with 1 mM of either IMP or GMP plus 1 mM $MgCI_2$ on ice for 30 min) were carefully pipetted into the appropriate well followed by 2 µL of the precipitant (Crystal Screen I and II conditions; Hampton Research) and incubated at 18 ºC.

3.0 X-ray data collection

Diffraction data for the crystals tested were collected on the in-house X-ray diffraction machine at the University of Western Cape, South Africa. The diffractometer consists of a Rigaku RUH3R copper rotating-anode X-ray source operated at 40 kV, 22 mA; a Rigaku R-axis IV+ image plate camera (Rigaku MSC, Tokyo, Japan); an X-stream 2000 low-temperature system (Rigaku MSC, Houston TX, USA); and an AXCO PX50 glass capillary optic with a 0.1 mm focus (Australian X-Ray Capillary Optics, Parkville VIC, Australia).

Two initial images with an oscillation 0.5° per frame were collected to ascertain whether the crystals were protein or salt crystals. The detector-to-crystal distance was set at 120 mm.

3.1 Comparative modeling

The first step of the modeling procedure was to align PfHPRT sequence with the structures of human HPRT (PDB code 1Z7G (Keough *et al*., 2005)), *Tritichomonas foetus* HPRT (PDB code 1HGX (Somoza *et al.,* 1996)) and *Salmonella typhimurium* HPRT (PDB code 1J7J (Lee *et al.,* 2001)) using 3D-Coffee (Poirot *et al*., 2004). The

template structures selected were of the free form of the enzymes. All the templates share moderate sequence similarity with PfHPRT. The sequence-structure alignment was manually edited to ensure no gaps in the core regions, guided by PSIPRED (McGuffin *et al*., 2000), and that the aligned template sequences matched the corresponding PDB entries. Missing residues in the PDB templates were omitted from the alignment. Modeling was performed using the external program MODELLER v0.99 (Sali and Blundell, 1993) which implements comparative modeling by satisfaction of spatial restraints. These template-derived restraints are then combined with CHARMM energy terms (Brooks *et al.,* 1983) to obtain a full objective function which is optimized by conjugate gradients and molecular dynamics with simulated annealing.

An iterative approach comprising sequence-structure alignment and model building was followed. 50 models were generated and one with the fewest restraints' violation and lowest energy was chosen for further evaluation. Side chains for this model were optimized using SCWRL (Canutescu *et al.,* 2003). Model quality was evaluated using PROCHECK (Laskowski *et al.,* 1993) and MOLPROBITY (Davis *et al.,* 2004). Three-dimensional images of PfHPRT model structure were generated using PyMOL v0.99 (DeLano, 2002).

3.2 Synthesis of TNP-8N_3-ITP and [γ^{32}P]TNP-8N_3-ITP

A pilot synthesis of TNP-8N_3-ITP from TNP-8N_3-ATP was first carried out to ascertain the time required for the reaction to proceed to completion. To form TNP-8N_3-ITP, the 6-amino group of TNP-8N_3-ATP is converted into an oxo-group using sodium nitrite ($NaNO_2$) and acetic acid. The reaction was carried out as follows: To 140 µL of ice-cold glacial acetic acid (99.8%), in a 2 mL Eppendorf, 263 µL of ice-cold 2 M $NaNO_2$ (≡0.526 millimoles) was added. To the mixture, 200 µL of 150 µM TNP-8N_3-ATP (≡30 nanomoles) was added followed by 97 µL of ice-cold distilled water. The reaction mix was incubated on ice in the dark. Each aliquot (25 µL, taken after 3 min, 1 h, 2 h, 3 h and 4 h) was mixed with 50 µL of 10 mM potassium phosphate, pH 6.0, 60% (v/v) acetonitrile and analyzed on a Vydac C_{18} column attached to a Beckman System Gold 126 pump mechanism/168 diode array detector HPLC system at room temperature. Absorption was monitored at 210 nm and 408 nm.

After 4 h the reaction mix was purified on a C_{18} Sep-Pak cartridge (Waters) previously conditioned with 6 mL methanol followed by 10 mL of distilled water. After sample application, the cartridge was washed with 10 mL of 10 mM potassium phosphate, pH 6-7 to displace the unattached TNP-8N_3-ITP 10 mL. The cartridge was then washed with 0.8 mL of distilled water and sample eluted with 10 mM potassium phosphate, pH 6-7, 60% (v/v) acetonitrile in two fractions of 0.5 mL and 1.5 mL. The first fraction was discarded and the acetonitrile content of the latter fraction reduced by blowing N_2 over the surface.

Synthesis and purification of [γ^{32}P]TNP-8N_3-ITP from [γ^{32}P]TNP-8N_3-ATP was carried out in a similar manner.

3.3 [γ^{32}P]TNP-8N_3-ATP and [γ^{32}P]TNP-8N_3-ITP photoaffinity labeling of PfHPRT

PfHPRT photolabeling conditions included 25 mM EPPS/TMAH, pH 8.5, 2 mM EDTA (or 1 mM $MgCl_2$, or 1 mM $CaCl_2$) , and 20% glycerol. The experiments were performed as follows: To 0.01 mg/mL of PfHPRT in 1.5 mL Eppendorf tubes placed on ice, 18.8 µL of 4X buffer (100 mM EPPS/TMAH, pH 8.5) and an appropriate volume of distilled water were added. This was followed by the addition of 15 µL of glycerol and an appropriate volume of [γ^{32}P]TNP-8N_3-ATP or [γ^{32}P]TNP-8N_3-ITP to give a concentration range of 0.01 to 30 µM. The final reaction mix (75 µL) was transferred to a 1 mL quartz cuvette and placed between two quartz cuvettes with fresh toluene to protect the protein from ultraviolet radiation damage (the only transmissible light was above 290 nm). Samples were then irradiated using a 150 W Xenon lamp (Hi-Tech). Following the labeling reaction, samples were transferred into 1.5 mL Eppendorf tubes and placed on ice. To the mixtures, 8 µL of solubilization buffer (1.2%, 3.6% β-mercaptoethanol and bromophenol blue) was added. After vortexing, 50 µL aliquots were taken and analyzed on 12% SDS-PAGE gels, run at 40 mA for 3-4 h, according the method Laemmli (1970). Once electrophoresis was complete, the bottom section of the gels (containing the majority of the radioactivity) was excised and the remainder dried by heating under vacuum. To the dried gels, reference spots, 1 µL of 10 µM [γ^{32}P]TNP-8N_3-ATP) were

positioned alongside the gel bands for stoichiometry measurement. Gels were then placed on a MS phosphor screen in film cassette (in a lead-lined box which protects the gels against cosmic radiation) for ~24 h after which they were scanned for radioactivity on a Cyclone Storage Phosphor Screen (Packard Bioscience). Quantitative analysis of the visible bands was performed using OptiQuant software which allows calculation of an average background value. By quantitating the counts per band in digital light units (dlu), moles of [γ^{32}P]TNP-8N_3-ATP (or [γ^{32}P]TNP-8N_3-ITP) bound per mole of PfHPRT could be calculated using the formula described below.

3.3.1 Amount of PfHPRT in gel band

Mw of PfHPRT = 26230 g/mol >> 0.0262 mg/nmol

Protein on gel = 0.01 mg/mL x 0.05 mL = 0.0005 mg

Nmol protein on gel = 0.0005/0.0262 = 0.0191 nmol

3.3.2 Amount of [γ^{32}P]TNP-8N_3-ATP in gel band

a = dlu in gel band **b** = dlu in mean background

c = dlu in 1 µL spot of [γ^{32}P]TNP-8N_3-ATP (1 µL 1 µM [γ^{32}P]TNP-8N_3-ATP ≡ 0.001 nmol)

c dlu / 0.001 nmol TNP-8N_3-ATP

therefore **c** * 1000 dlu / nmol TNP-8N_3-ATP

nmol probe in gel band = (**a- b**) / **c** * 1000 nmol

mol probe bound / mol protein = (**a-b**) / (**c** * 1000 * 0.0191) mol / mol

3.3.3 Curve fitting

Graphs were drawn using SigmaPlot 3.0 (Jandel Scientific Software). TNP-8N_3-ATP and TNP-8N_3-ITP concentration-dependent photolabeling of PfHPRT data were fitted to the binding equation

$RC = RC_{max}[S]/ (K_{0.5} + [S])$

Where RC = relative counts,

RC_{max} = relative maximum counts,

[S] = photolabel concentration,

$K_{0.5}$ = photolabel concentration at which half-maximal labeling is obtained.

CHAPTER THREE: RESULTS

3.0 Protein purification: Ionic exchange and dye-affinity chromatography

In order to obtain PfHPRT of the highest possible purity for crystallization experiments, considerable effort was devoted to improving and optimizing the purification process.

Our starting point was the purification procedure developed by B. Mbewe during his PhD programme in this laboratory (Mbewe, 2005; Mbewe *et al.*, 2006). The method involves starting with 1 L culture, to produce 100 mL of supernatant in 50 mM Tris-HCl, 25 mM NaCl, 2 mM EDTA, 1 mM PMSF and 1 mM DTT (lysis buffer), which was passed by gravity through a DE-52 cellulose anionic exchanger column (2 x 6 cm, treated with 10 mL lysis buffer) placed in tandem with a Reactive Red 120-agarose dye affinity column (2 x 5 cm). A further 50 mL of lysis buffer was passed through, the columns separated, the second column washed with 50 mM sodium phosphate, pH 7.0, 200 mM NaCl followed by PfHPRT elution with 10 mM sodium phosphate, 50 mM sodium pyrophosphate, pH 7.0 (RR elution buffer or PPi buffer). The first 8 mL were discarded and the next 15 mL collected, pooled, dialysed against 10 mM sodium phosphate, 1 mM DTT, lyophilized, and re-suspended to approximately 3 mL at around 3 mg of protein /mL. Under the conditions used, PfHPRT came through the anion exchange column, but bound to the Reactive Red 120 column. The preparation was inactive as isolated, but could be activated by incubation overnight at 0 ^{o}C with 0.06 mM hypoxanthine plus 1 mM PRPP. It is well-known that isolated recombinant PfHPRT is inactive and requires

activation with partial substrates or products to produce activity (Keough *et al*., 1999; Phehane, 2002; Mbewe, 2005; Mbewe *et al*., 2006; Phehane, 2002; Raman *et al.*, 2005).

An attempt to reproduce this procedure is shown in Figures 3.0 and 3.1.

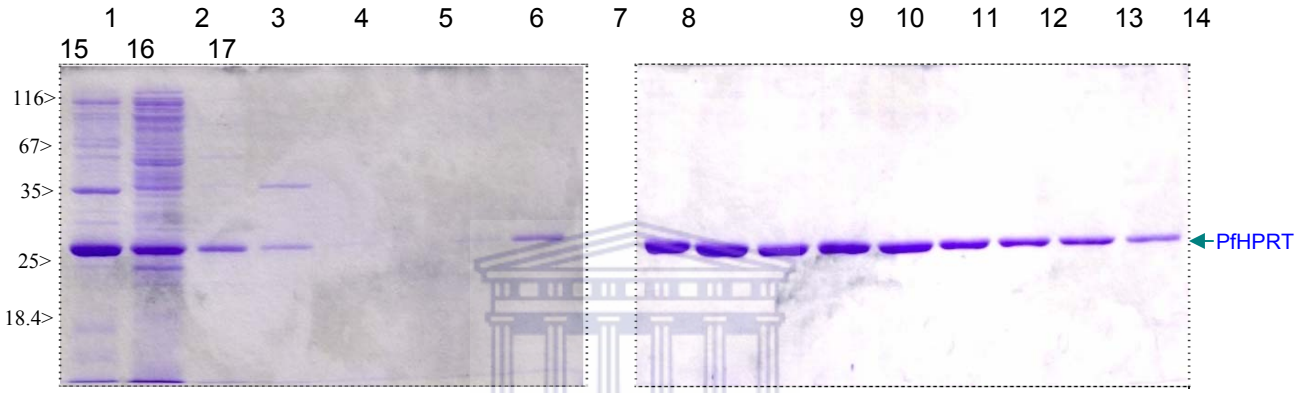


Figure 3.0: SDS–PAGE (15%) analysis of fractions eluting from a short DE-52 column (2 x 6 cm) placed in tandem with a Reactive Red 120 column (2 x 5 cm). Lane 1, pellet fraction (re-suspended to 100 mL); lane 2, supernatant fraction (100 mL); lane 3, Reactive Red flow through fraction (150 mL); lane 4, 10 mM sodium phosphate, pH 7.0, 200 mM NaCl wash (40 mL); lanes 5-16, RR elution buffer (10 mM sodium phosphate, 50 mM sodium pyrophosphate, pH 7.0) (2 mL each); lane 17, Reactive Red 1 M NaCl wash (50 mL). Equal volume aliquots (10 μL) were applied to each well, and the gel was stained with Coomassie Brilliant Blue. Approximate positions of molecular weight standards (in kDa) are shown on the left hand side.

In our case, after elution, the RR elution buffer (PPi) fractions were pooled and dialysed overnight against two changes of 10 mM sodium phosphate, pH 7.0, 1 mM DTT and afterwards concentrated to about 2 mL in an Amicon stirred cell at 4 °C fitted with a YM-10 membrane (Millipore). The concentrate was afterward re-dialyzed against 10 mM sodium phosphate, pH 7.0, 1 mM DTT. For activation, aliquots were incubated with substrates (1 mM PRPP + 0.06 mM hypoxanthine) for 16 h, and subsequently assayed for

activity. PfHPRT eluted from the Reactive Red column with 1 M NaCl was treated in the same manner. The protein content of both concentrates is shown in Figure 3.1. It can be seen from a comparison of lanes 1 and 2 in Figure 3.0 that approximately 40% of the recombinant PfHPRT was in the supernatant fraction, which is consistent with what Mbewe obtained (Mbewe *et al*., 2006). Presumably the majority of the expressed enzyme is in an insoluble form in inclusion bodies. Lane 3 shows that approximately half of the protein in the supernatant came through both columns (a comparison of the band in lanes 2 and 3 needs to take into account the larger volume of the eluate). The 200 mM NaCl wash removed a bit of the target protein but effectively removed a contaminant of ~ 35 kDa (lane 4). PfHPRT started eluting in the 4th fraction with PPi (lane 8), and continued through to the 12th fraction (lane 16). The column was finally washed with 50 mL 1 M NaCl, and the contents are shown in lane 17. Approximately 40% of the protein came off with the high salt, and probably consists of PfHPRT that could have eluted with PPi if the elution had been continued and some tightly bound PfHPRT, which may be a different form (see later). Because there was a significant amount of PfHPRT in this 1 M NaCl fraction, it was processed further along with the PPi fraction

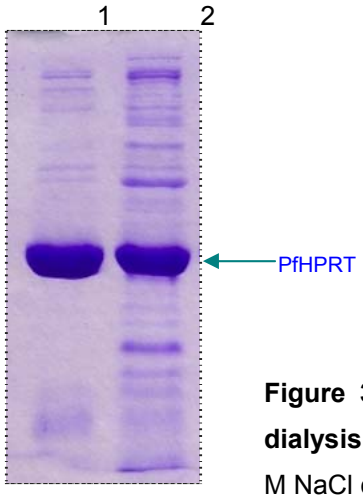


Figure 3.1: SDS-PAGE (15%) analysis of PfHPRT fractions after dialysis and concentration. Lane 1, PPi concentrate (2 mL); lane 2, 1 M NaCl concentrate (2 mL). Gel stained with Coomassie Blue.

Analysis of the protein content of the concentrates in Figure 3.1 reveals that the RR elution buffer concentrate was ~80% pure and the 1 M NaCl concentrate ~40%, the first figure is consistent with what Mbewe *et al*., 2006 obtained.

Both the PPi and NaCl concentrates were capable of being activated with substrates (specific activity of 1.7 and 0.6 µmols/min/mg, respectively, activated by overnight incubation at 0 ^{o}C with 1 mM PRPP + 0.06 mM hypoxanthine, and measured with guanine as substrate at pH 7.5 with 10 mM $MgCl_2$ at room temperature; the figure of 1.7 approximately equates to the maximum activity of 5 µmoles/min/mg at pH 8.5 with 100 mM $MgCl_2$ at room temperature described by Mbewe *et al*. (2006) and Keough *et al*. (1999). The lower activity of the salt fraction can be attributed to the protein contamination. The fact that PfHPRT contained in the fraction eluted with 1 M NaCl had high activity was surprising considering that it is an unstable enzyme (Mbewe, 2005; Raman *et al.*, 2005) and is inactive in high salt due to a conversion from the active tetramer to an inactive dimer (Keough *et al*., 1999). This rather impure salt fraction was

used to investigate whether HPLC gel filtration might be used for further purification, and the experiments and results will be described in a later section (Section 3.2). The purer PPi fraction was used for crystallization trials, which will also be described later (Section 3.3).

In conclusion, although significant purification was achieved by the tandem anion exchange and affinity chromatography with the procedure described, it could be better both in terms of purification and minimization of losses, and improvements were explored. Also, the retention of significant amounts of PfHPRT on the Reactive Red 120 column after the PPi elution was not what Mbewe reported (about 90% eluted with PPi), and needed verification.

We first checked whether the entire flow-through and eluate from the DE-52 column needed to be passed through the next Reactive Red 120 column (i.e., if both beginning and end contained PfHPRT). Fractions from the DE-52 column were collected individually (10 mL each), and analysed for their protein content prior to placing on the Reactive Red 120 column, and the results are shown in Figure 3.2.

It is evident that PfHPRT begins to elute in fraction 2 (lane 2), continues through the entire flow-through period, and into the wash period, but gradually diminishing in amount with more volume. The majority elutes during the flow-through, but a significant amount is in the first 40 mL of wash.

This substantiates what was done in the previous experiment; the Mbewe method (100 mL supernatant plus 50 mL wash) captured most of the target protein for the Reactive Red 120 column. However, it is also evident that the protein eluting in the flow-through was significantly purer than that coming off with the wash. In the flow-through fractions there is virtually no high molecular weight species, only smaller molecular weight proteins, and these latter we now know to be mostly proteolytic digestion products of PfHPRT (see later). In contrast, the fractions eluting in the wash contained significant amounts of high molecular weight species.

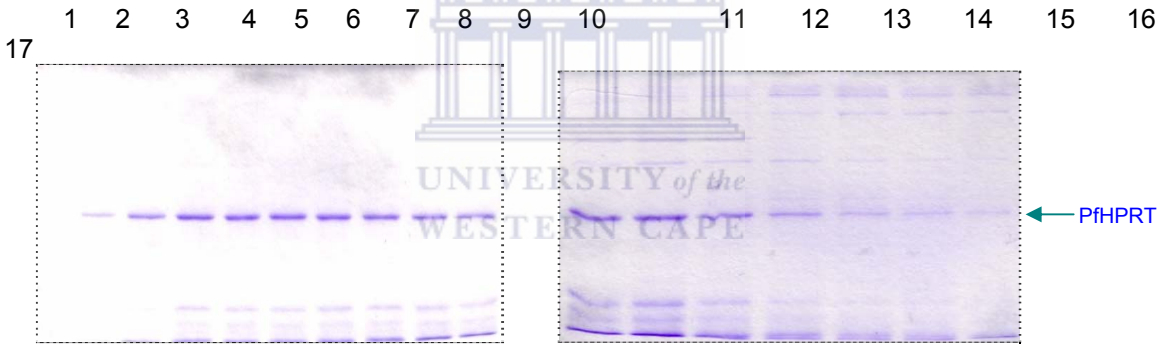


Figure 3.2: Elution profile from a poorly equilibrated, short DE-52 column. Lanes 1-10 represent the 10 fractions (10 mL each) of the 100 mL flow-through while lanes 11-17 are the wash of 70 mL of 50 mM Tris-HCl, pH 8.0, 1 mM PMSF, and 1 mM DTT (10 mL each). SDS-PAGE analysis with 15% acrylamide and staining with Coomassie Blue.

The flow-through fractions 2-10 and wash fractions 1-4 were pooled and loaded onto a Reactive Red 120 column (2 x 5 cm). The column was developed in the usual way with 40 mL 10 mM sodium phosphate, pH 7.0/ 200 mM NaCl and then eluted with 50 mL 10 mM sodium phosphate, 50 mM sodium pyrophosphate, pH 7.0 (collected in 10 mL fractions), and finally washed with 50 mL 1 M NaCl (Figure 3.3).

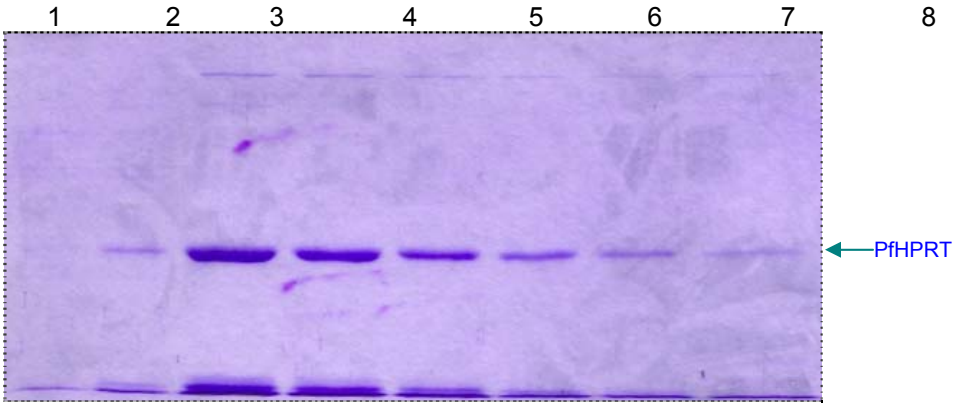


Figure 3.3: Elution profile from a Reactive Red 120 column. Lane 1, flow through fraction (140 mL); lane 2, sodium phosphate/NaCl wash (40 mL); lanes 3-7, pyrophosphate eluate (10 mL each); lane 8, 1 M NaCl wash (50 mL).

It is evident that all the PfHPRT was retained by the column (the flow-through, lane 1, contained no PfHPRT), and the overwhelming majority of the PfHPRT was eluted with sodium pyrophosphate (lanes 3-7), and the high salt wash produced rather little protein (lane 8). It is also evident in this gel that the amount of small molecular weight species (running with the front in this poorly resolved electrophoresis run) correlates with the proportion of PfHPRT, suggesting that the former are proteolytic digestion products of the latter. This was confirmed later by HPLC gel filtration in which the primary protein and digestion products co-eluted (evidently the nicked protein still retained nearly native secondary, tertiary and quaternary structure). The digestion seen here was peculiar to this preparation, and the eluate was not processed further, but it did prompt us to explore the use of additional protease inhibitors.

The other aspect of the purification that concerned us was the significant elution of contaminant proteins during the DE-52 wash period (Figure 3.2), which led us to investigate the pH of the individual fractions. It seemed that the presence of the extra proteins could possibly be caused by the pH drifting from 8 down to acid pH values.

To examine this, a new preparation was made, incorporating a protease inhibitor cocktail of PMSF, AEBSF, aprotinin, bestatin hydrochloride, E-64, EDTA and leupeptin and the expression and purification are shown in Fig 3.4. Again, the supernatant contained less PfHPRT than the pellet fraction (not shown). Lane 1 shows the pooled flow-through fractions and lanes 2-6 the wash fractions, which corresponds very nicely to the results obtained for the previous preparation. Again it is clear that many contaminating proteins emerge with the extra 50 mL wash. A 1 M NaCl wash of the DE-52 column is shown for the first time, and it is evident that none of the PfHPRT remains bound (lane 7).

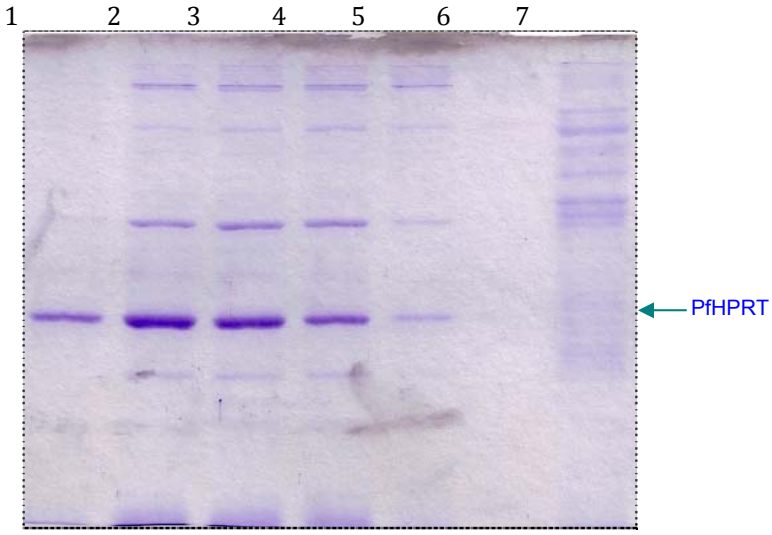


Figure 3.4: Elution profile from a poorly equilibrated DE-52 column. Lane 1, DE-52 flow-through fraction (100 mL); lanes 2-6, wash fractions (10 mL each); lane 7, 1 M NaCl wash (100 mL).

The pH measurements, shown in Figure 3.5, of each fraction coming through the DE-52 column were very revealing. The first fractions were approximately pH 11, until about 80 mL of supernatant had passed through the column, at which stage the pH began to drop, reaching approximately 6.5 at the time the wash was applied, and slowly returning to 8 after about 50 mL. In the original procedure of Mbewe, and in all the above purifications, the DE-52 column was pre-treated with 10 mL of lysis buffer (50 mM Tris-HCl, pH 8.0, 2 mM EDTA, 25 mM NaCl, 1 mM DTT, and 1 mM PMSF) which obviously failed to equilibrate the column. The other interesting thing is that the pH went down below that of the Tris-HCl buffer.

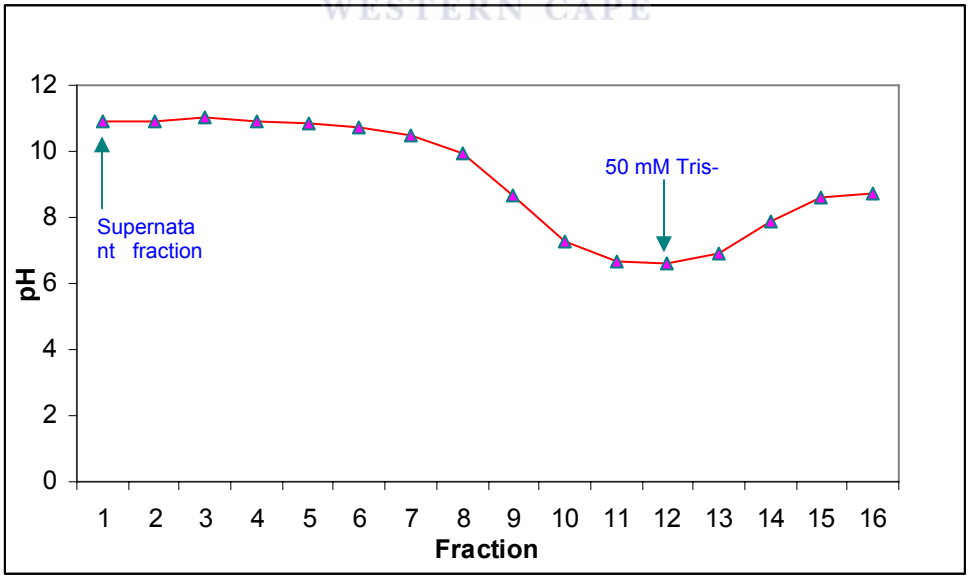


Figure 3.5: pH profile of fractions eluting from the poorly equilibrated, short DE-52 column. Supernatant (100 mL) was applied to the column pretreated with 10 mL lysis buffer (50 mM Tris-HCL, pH 8.0, 25 mM NaCl, 2 mM EDTA, 1 mM DTT, and 1 mM PMSF) followed by 60 mL of

wash (50 mM Tris-HCl, pH 8.0, 1 mM DTT, and 1 mM PMSF). Fractions (10 mL each) were collected and the pH of each determined. The protein content is shown in Figure 3.4.

The DE-52 flow-through and the first 3 wash fractions were combined and loaded onto a Reactive Red 120 column. As usual, the column was then washed with 10 mM sodium phosphate, pH 7.0, 200 mM NaCl, and then developed with the PPi elution buffer, and the results are shown in Fig 3.6. The 200 mM salt wash again removed high molecular weight contaminants, and PPi eluted a relatively pure form of PfHPRT. Unfortunately, some proteolysis of PfHPRT is evident, perhaps to the extent of more than 50% after concentrating (Figure 3.6, gel B), and the enzyme failed to gain activity after activation with substrates.

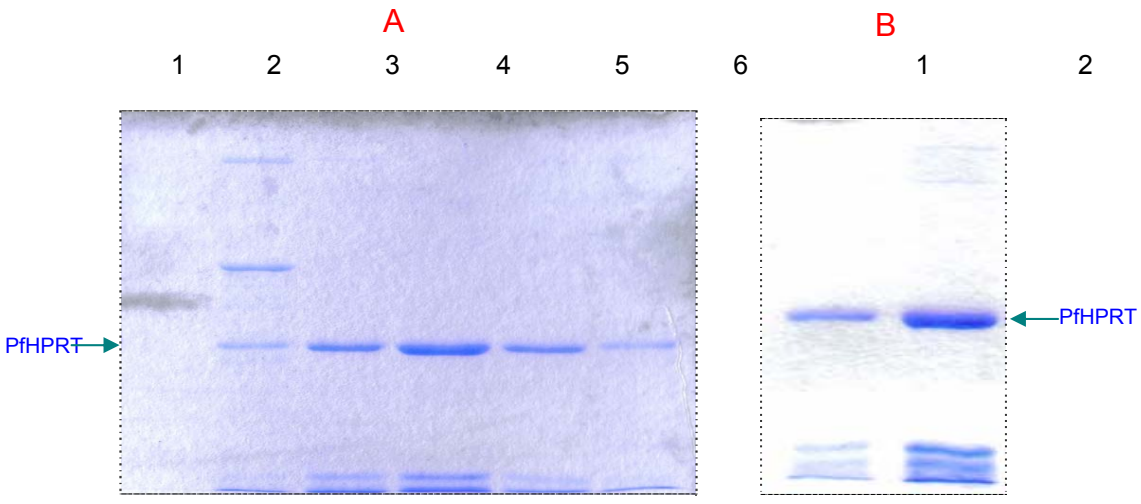


Figure 3.6: Elution profile from a Reactive Red 120 column. In the experiment shown in Gel A, the flow-through and 3 wash elutions from DE-52 chromatography (Figure 3.4) were loaded onto a Reactive Red 120 column. The column flow-through is shown in lane 1; the 10 mM sodium

phosphate/200 mM NaCl wash (30 mL) is in lane 2; 10 mM sodium phosphate/50 mM sodium pyrophosphate, pH 7.0 eluate (10 mL each) is in lanes 3-5 and 1 M NaCl wash (30 mL) is in lane 6. Gel B shows the 3 pooled PPi fractions (30 mL) after dialysis (lane 1) and concentration (8 mL) (lane 2).

In the next preparation, we used a properly equilibrated DE-52 column and a stronger buffer in an attempt to achieve better separation and also to ascertain whether the pH variation seen before could be controlled and subsequently the effect on PfHPRT activity. The column was well equilibrated with 10x strength 75 mM Tris-HCl, pH 8.9 (i.e. 750 mM), and finally with 1x buffer, as directed by the manufacturer. Prior to loading the supernatant onto the column, its pH was also adjusted to 8.9. The flow-through followed by 50 mL of 50 mM Tris-HCl, pH 8.0, 1 mM DTT, and 1 mM PMSF, were all collected in 10 mL fractions and the pH measured. After that, the column was washed with 80 mL 1 M NaCl to bring off all other bound contaminants. Figure 3.7 shows the results obtained.

Protein expression levels for this preparation were relatively low compared with previous preparations. Like in the preceding preparations, most of the target protein was in the insoluble inclusion bodies (lane 2). The first two fractions of the flow-through (lanes 4 and 5) contained no PfHPRT consistent with earlier results. PfHPRT starts eluting in the third fraction (lane 6) and continues till all the supernatant has passed through although the last two fractions contain very little protein (results not shown). At the start of Tris-

HCl wash (lane 12), a significant amount of the high molecular weight contaminants start coming off with more PfHPRT. Again this is consistent with earlier results.

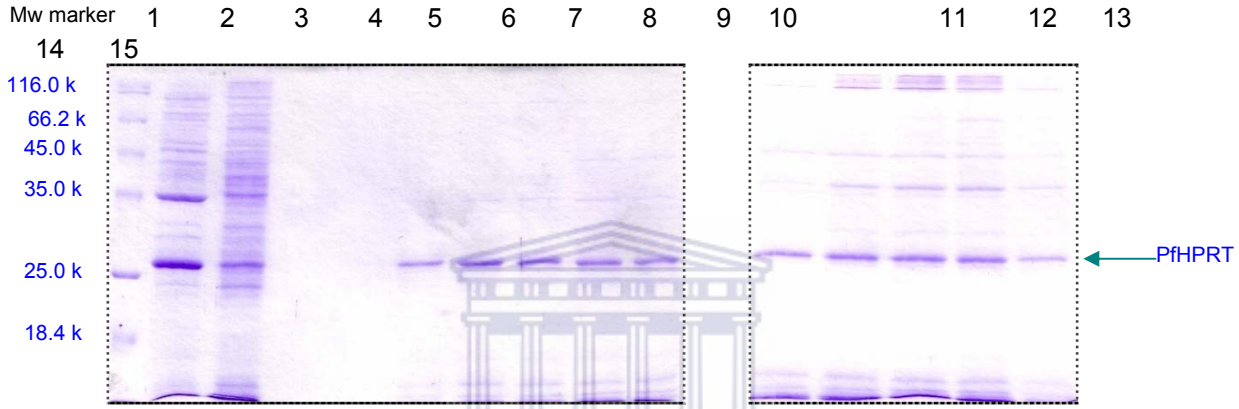


Figure 3.7: Elution profile from a short DE-52 column well equilibrated with 75 mM Tris-HCl, pH 8.9. Lane 1, molecular weight marker; lane 2, pellet fraction (re-suspended to 100 mL); lane 3, supernatant fraction (100 mL); lanes 4-11, DE-52 flow-through fractions (10 mL each); lanes 12-15, wash fractions (10 mL each of 50 mM Tris-HCl, pH 8.0, 1 mM PMSF, 1 mM DTT).

The 1 M NaCl wash brought off other bound contaminants without much PfHPRT as is shown in Figure 3.8 (lane 1). Despite doubling the concentration of the protease inhibitor cocktail, PfHPRT proteolysis was evident as some low molecular weight species in this preparation.

The flow-through fractions containing PfHPRT were pooled, an aliquot taken and concentrated and the rest loaded onto a Reactive Red 120 column. The column was then washed in the usual manner with 10 mM sodium phosphate, pH 7.0, 200 mM NaCl and

developed with the PPi elution buffer. The results are shown in Fig 3.8 A. The concentrate of the pooled DE-52 flow-through fractions (lane 2) shows that a lot of impurities come off in these fractions and thus the need for further purification. Most high molecular weight contaminants flow straight through the Reactive Red 120 column (lane 3) and in the 10 mM sodium phosphate, pH 7.0/200 mM NaCl wash (lane 4). The

PPi eluate (lanes 5-9) is significantly purer with the major impurity apparently being low molecular weight proteolytic products. Satisfyingly, very little target protein was brought off by the 1 M NaCl wash, probably due to well-controlled pH (see below). The PPi fractions were then pooled, dialysed overnight in 10 mM sodium phosphate, pH 7.0, 1 mM DTT, concentrated (shown in Figure 3.8 C) and activated with PRPP and hypoxanthine as previously described. Unfortunately, on assay, the specific activity was very low (0.2 µmole/min/mg). This is most likely due to PfHPRT proteolysis.

The Tris-HCl eluted fractions were also pooled, taken through a Reactive Red 120 column and eluted in the usual manner. The results, though less visible due to the low protein content, are shown in Figure 3.8B and are similar to the flow-through results shown on the left hand side.

A B C

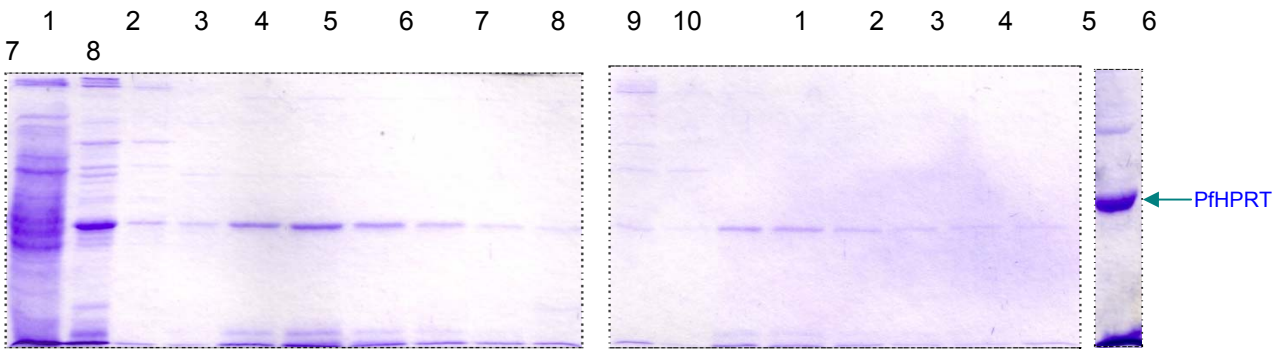


Figure 3.8: Elution profile from a Reactive Red 120 column. In Gel A, lane 1 shows the bound contaminants coming off the DE-52 column with the 1 M NaCl wash (continued from Fig 3.7); lane 2 shows the concentrated DE-52 fraction (0.2 mL); lane 3 shows Reactive Red 120 column flow-through (70 mL); The 10 mM sodium phosphate, 200 mM NaCl wash is shown in lane 4 (40 mL); 10 mL aliquots of 10 mM sodium phosphate, 50 mM sodium pyrophosphate, pH 7.0 are in lanes 5-9 while the final 1 M NaCl wash in shown in lane 10 (40 mL). Gel B shows the elution profile of the pooled DE-52 Tris-HCl wash fractions from a Reactive Red 120 column. The flow-through fraction (50 mL) in lane 1 is followed in lane 2 by the 10 mM sodium phosphate, 200 mM NaCl wash (40 mL). Lanes 3-7 show the PPi eluate (10 mL each) while lane 8 shows the 1 M NaCl wash (40 mL). Gel C shows a concentrate of pooled PPi fractions (continued from gel A).

The pH measurements for the eluted DE-52 fractions, shown in Figure 3.9, indicate that the pH was reasonably well controlled, and the pH of the pooled flow-through fractions was 8.9.

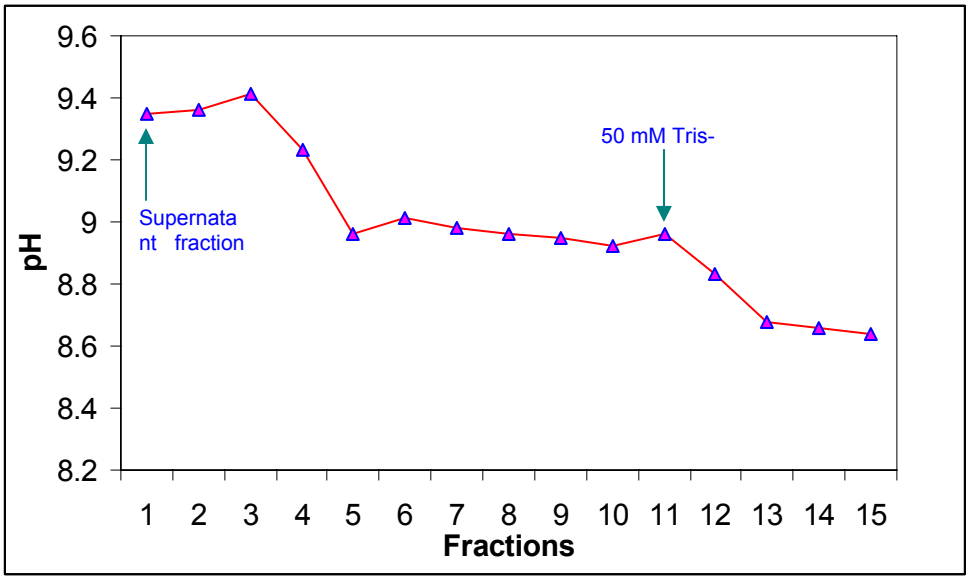


Figure 3.9: pH of fractions eluting from a short DE-52 column, well equilibrated at pH 8.9.

In conclusion, the results in terms of elution characteristics and purification were very similar for both the poorly and well-equilibrated short DE-52 columns, even the elution profiles of the Reactive Red 120 column were the same, except that the well-equilibrated column failed to yield active enzyme, which may be due to proteolysis. It, therefore, seems that exposure to pH 11 is not harmful to PfHPRT.

We then decided to employ a longer DE-52 cellulose column (2.0 x 17.5 cm) in an attempt to effect better separation by increasing the number of theoretical plates.

In the first attempt the column was equilibrated with 30 mL of the lysis buffer before loading the supernatant, knowing that this would not control the pH, but influenced by the fact that the purity of the DE-52 eluate of the poorly equilibrated column (Figure 3.2) was somehow better than that of the well-equilibrated one (Figure 3.7), and the

knowledge that the former produced active PfHPRT. The results are shown in Figure 3.10.

Passage of the supernatant through the DE-52 column resulted in all the proteins being retained (lane 1). The column was developed with an extra 100 mL of 50 mM Tris-HCI, pH 8.0, 1 mM DTT, and 1 mM PMSF collected in 10 mL fractions. The first 4 fractions did not contain any protein (not shown), but the next 3 did (lanes 5-7) and with reduced amounts in the following fractions (not shown), and the 1 M salt wash showed that rather little was retained (lane 8). In terms of purification and yield the results were quite remarkable – in lanes 5 and 6 there is a very strong band associated with PfHPRT and little else. Pooling the fractions from which the lanes 5, 6, and 7 were derived produced 30 mL of 0.2 mg protein/mL, i.e. 6 mg total.

The sample was dialysed overnight against two changes of 10 mM MOPS, 1 mM DTT, pH 7.0, and tested for activity after overnight activation with substrates. Unfortunately, the enzyme was inactive. This was perhaps not surprising as the pH of the sample prior to dialysis was approximately 13! It still, however, remains a puzzle that PfHPRT behaved so differently from all the other proteins.

Out of interest, the protein was further processed through a Reactive Red 120 column, and the results are shown in Figure 3.11. Before loading the protein, the column was

equilibrated with 10 mM MOPS, pH 7.0, 1 mM DTT, whereas before it was simply washed with water.

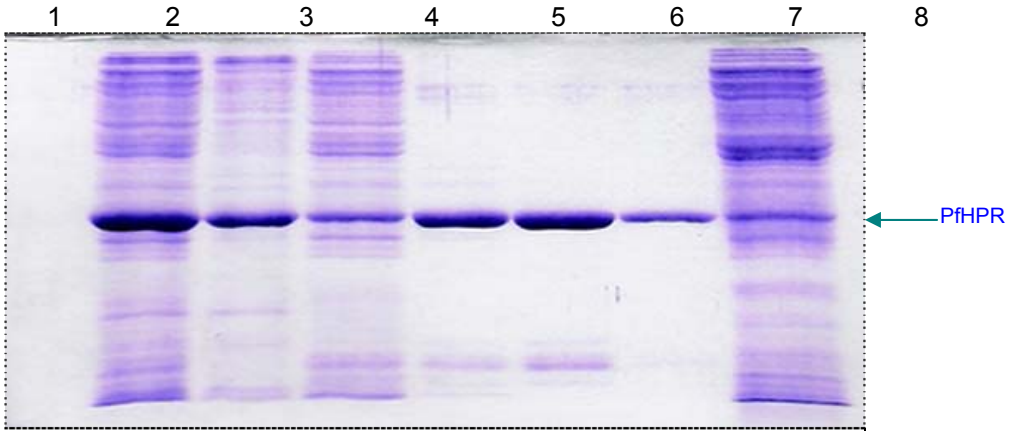


Figure 3.10: Elution profile from a poorly equilibrated, long DE-52 column (2.0 x 17.5 cm). Lane 1, flow-through fraction (100 mL); lane 2, crude lysate (100 mL); lane 3, pellet fraction (re-suspended to 100 mL); lane 4, supernatant fraction (100 mL); lanes 5-7, 50 mM Tris-HCl, pH 8.0 fractions (10 mL each); lane 8, 1 M NaCl wash (100 mL).

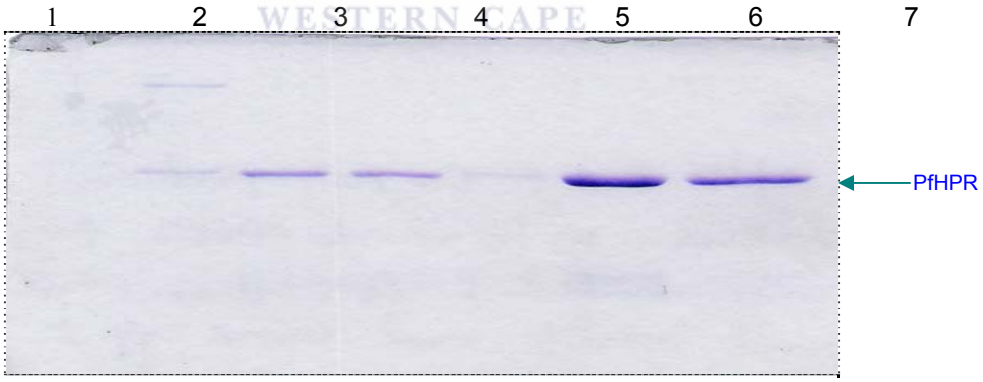


Figure 3.11: Elution profile from a Reactive Red 120 column. Lane 1, flow- through fraction (30 mL); lane 2, 10 mM MOPS, pH 7.0, 1 mM DTT wash (10 mL); lanes 3-5, PPi eluate (10 mL each); lanes 6-7, 1 M NaCl wash (10 mL each).

All protein bound to the affinity column as is evident from the lack of protein bands in lane 1. The prominent high molecular weight impurity and also some PfHPRT came off with 10 mM MOPS, pH 7.0, 1 mM DTT (lane 2). Very little protein eluted with the PPi eluate, (lanes 3-5), and the majority appeared in the 1 M NaCl wash (lanes 6-7). It is clear that most of the protein is behaving very differently from before, binding extremely tightly to Reactive Red 120, which is consistent with it being in a different conformation or state and incapable of being activated with substrates.

Finally, a long DE-52 column (2 x 17.5 cm) well-equilibrated with the lysis buffer was run and the results analysed (Figure 3.12). Although the pellet fraction still had a considerable amount (lane 1) of the target protein, this preparation contained more soluble PfHPRT (lane 2) in comparison with previous ones. Starting in the third fraction, PfHPRT came off the column almost entirely in the flow-through (lanes 5-12) and only required a 30 mL rinse with 50 mM Tris-HCl, pH 8.0, 1 mM DTT, and 1 mM PMSF to bring off most of the remainder (lanes 13-15). A further 20 mL wash (lanes 16-17) brought off a negligible amount of PfHPRT.

If the small molecular weight species are ignored (probably proteolytic digestion product of PfHPRT), the purification achieved with this long DE-52 column seems rather similar to that of the short column (compare with Figure 3.7), although the wash fractions seem worse with the shorter column and the latter contain more PfHPRT.

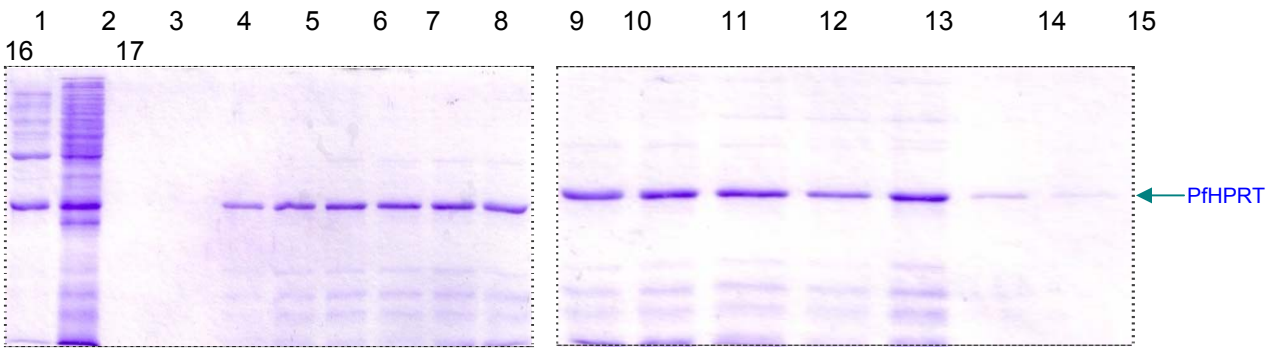


Figure 3.12: Elution profile from a well equilibrated, long DE-52 column (2 x 17.5 cm). Lane 1, pellet fraction (re-suspended to 100 mL); lane 2, supernatant fraction (100 mL); lanes 3-12, flow-through fractions (10 mL each); lanes 13-17, wash fractions (10 mL each). Equal volume aliquots (10 μL) were applied to each well. Approximate position of PfHPRT is shown.

The flow-through fractions together with the first three wash fractions were pooled and further purified on a Reactive Red 120 column as described earlier. Although this step eliminated some contaminating proteins, two high molecular weight bands plus the proteolysis band were visible on SDS-PAGE after concentrating the PPi fractions (lane 10). This preparation exhibited a very low specific activity (0.7 μmole/min/mg protein).

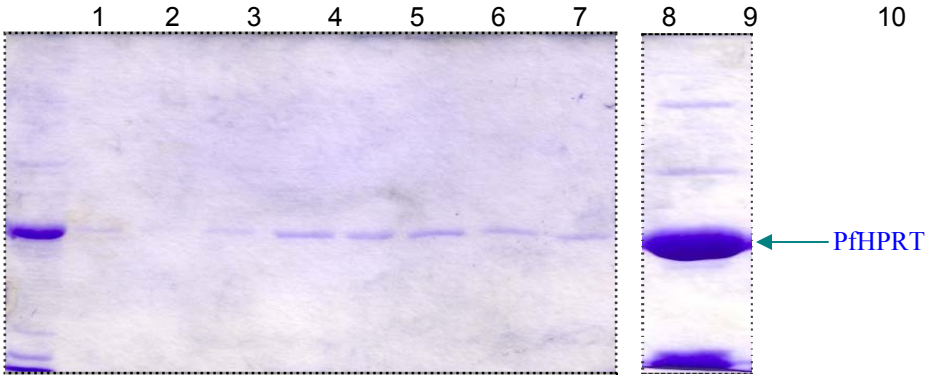


Figure 3.13: Elution profile from a Reactive Red 120 column followed by concentration. Lane 1, pooled fractions from DE-52 (110 mL); lane 2, flow-through fraction (110 mL); lane 3, 10 mM sodium phosphate, 200 mM NaCl wash (40 mL); lanes 4-8, 10 mM sodium phosphate, 50 mM sodium pyrophosphate, pH 7.0 eluate (10 mL each); lane 9, 1 M NaCl (40 mL); lane 10, concentrated PPi fractions (2 mL).

3.1 HPLC gel filtration: PfHPRT oligomeric state analysis and further purification

The oligomeric state of PfHPRT during and after purification, and also on activation with substrates was analysed by analytical gel filtration on a Tosohaas TSKgel G 3000 SW column (7.5 mm x 30 cm, 10 μm). A major motivation for performing these experiments was to determine whether the "pure" PfHPRT was a single species or not. For crystallization, it is desirable that the protein exists as a single oligomeric species in solution.

The column characteristics were determined by running standards at different salt concentrations as shown in Figure 3.14. The exclusion and inclusion marker compounds were Dextran Blue (2 000 000 Daltons) and tyrosine (185 Daltons), respectively. The resolution of the column was better with increasing concentration of salt. At 0 mM NaCl, three proteins co-eluted and cytochrome c remained bound to the column.

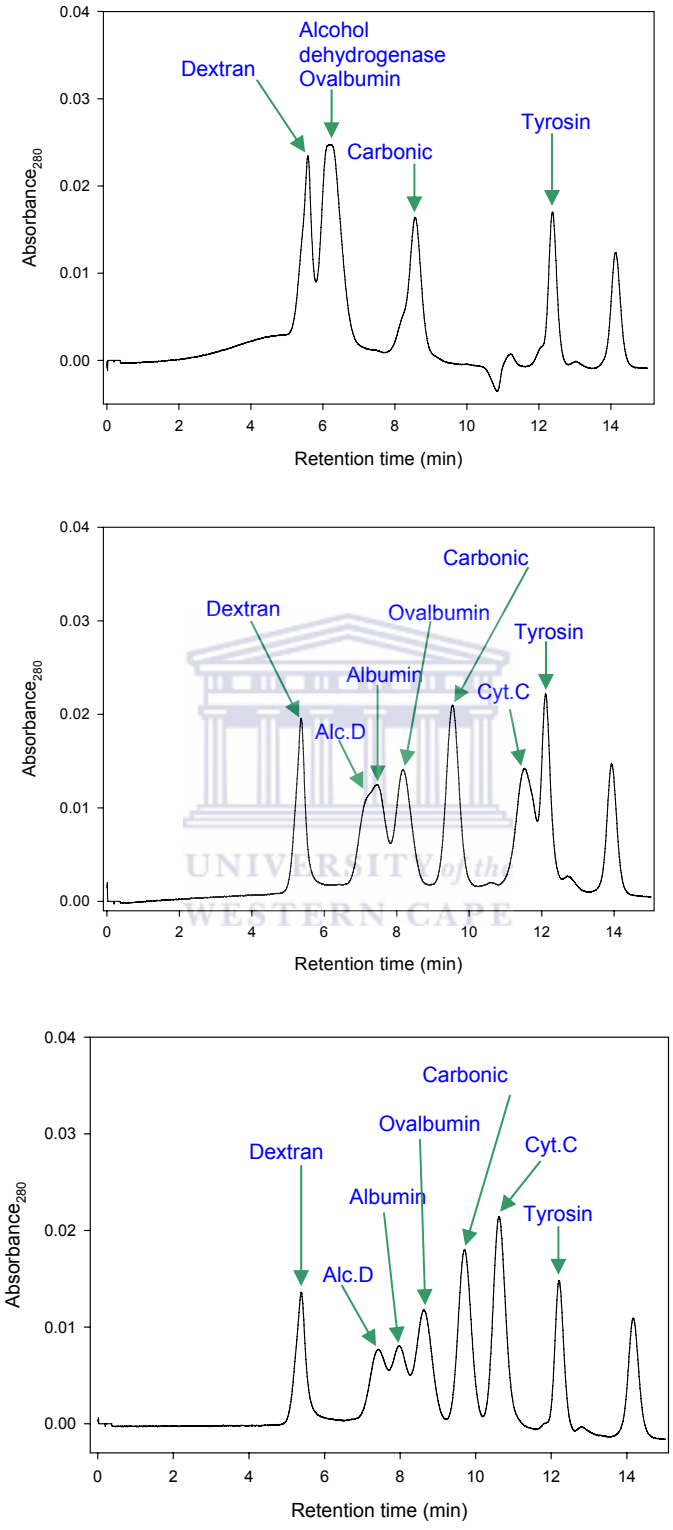


Figure 3.14: Calibration of the Tosohaas TSKgel G SW 3000 HPLC column (7.5 mm x 30 cm, 10 μm). A mixture of standards containing Dextran Blue (MW 2 000 000 Daltons), Alcohol dehydrogenase (150 000 Daltons), Albumin (67 000 Daltons), ovalbumin (44 000 Daltons), carbonic anhydrase (29 000 Daltons), cytochrome c (11 700 Daltons) and tyrosine (185 Daltons) was injected (20 μL) at time zero, and the mobile buffer was 10 mM MOPS/TMAH, pH 7.0, 0.05% NaN_3 with 0 (top panel), 100 (middle panel), or 400 mM NaCl (bottom panel). The flow rate was 1 mL/min.

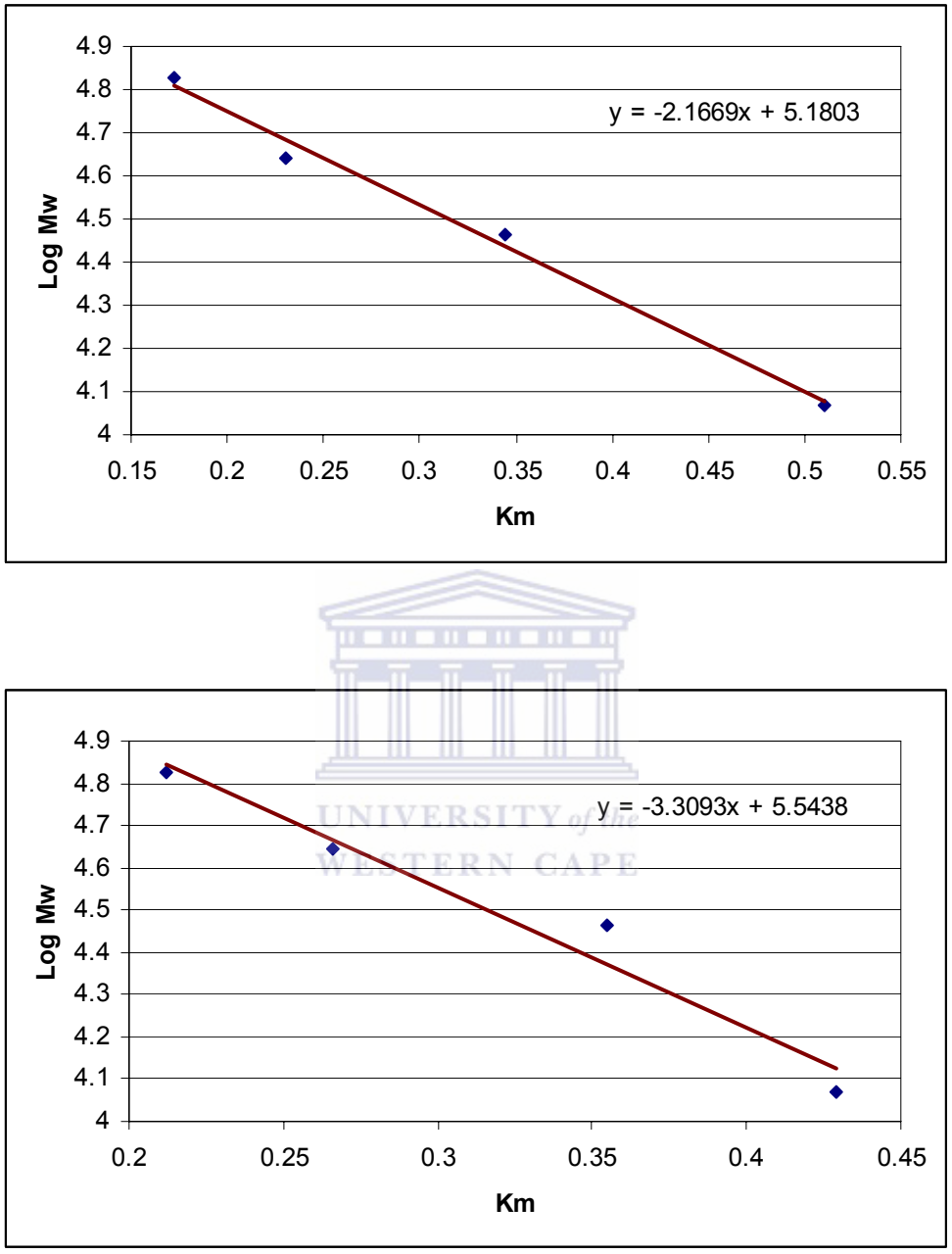


Figure 3.15: Calibration curves for protein standards at 100 mM NaCl (top) and 400 mM NaCl (bottom). The chromatography of the standards is shown in Figure 3.14. The equation used in the plot is provided in Chapter Two: Materials and Methods. Mw refers to the molecular weight and Km the Partition Coefficient.

The quaternary structure of PfHPRT is known to be dependent on salt concentration, and has been identified as tetrameric at 10 mM phosphate, pH 6.9, in the presence of 60 μM hypoxanthine, 0.2 mM PRPP, and 5 mM DTT (active preparation), and dimeric if 1.2 M KCl is included (inactive) by sedimentation equilibrium (Keough *et al*., 1999). We tested the quaternary structure of PfHPRT by gel filtration HPLC using the first preparation described above following elution from Reactive Red 120 column with 1 M NaCl, dialysis, and concentration, to have a final preparation of Figure 3.1. The preparation was inactive, but capable of being activated with substrates to approximately 3 μmole/min/mg of protein. The results at several concentrations of NaCl are shown in Figure 3.16. It is simplest to start with the profile obtained at 100 mM NaCl. Here major peaks are found at approximately 8 (labeled A), 9 (labeled B), 11 (labeled X), and 13 min. Peaks A and B were identified as PfHPRT by collecting the peaks and analyzing by SDS-PAGE, and those at 11 and 13 min were not identified, but did not appear to be protein. Peaks A and B correspond to molecular weights of 63 000 and 40 000 Daltons, respectively, according to the calibration curve at 100 mM NaCl (Figure 3.15), and 57 000 and 36 000 Daltons at 400 mM NaCl, and have tentatively been assigned as tetramer (should be 104 000 Daltons) and dimer (52 000 Daltons), respectively. As can be seen in the other frames, the proportion of tetramer increases at low salt and decreases at high salt. Quantification of the PfHPRT peaks is shown in Figure 3.17. It is evident that the proportion of tetramer at low salt is approximately 70%, and this could have profound implications for obtaining crystals.

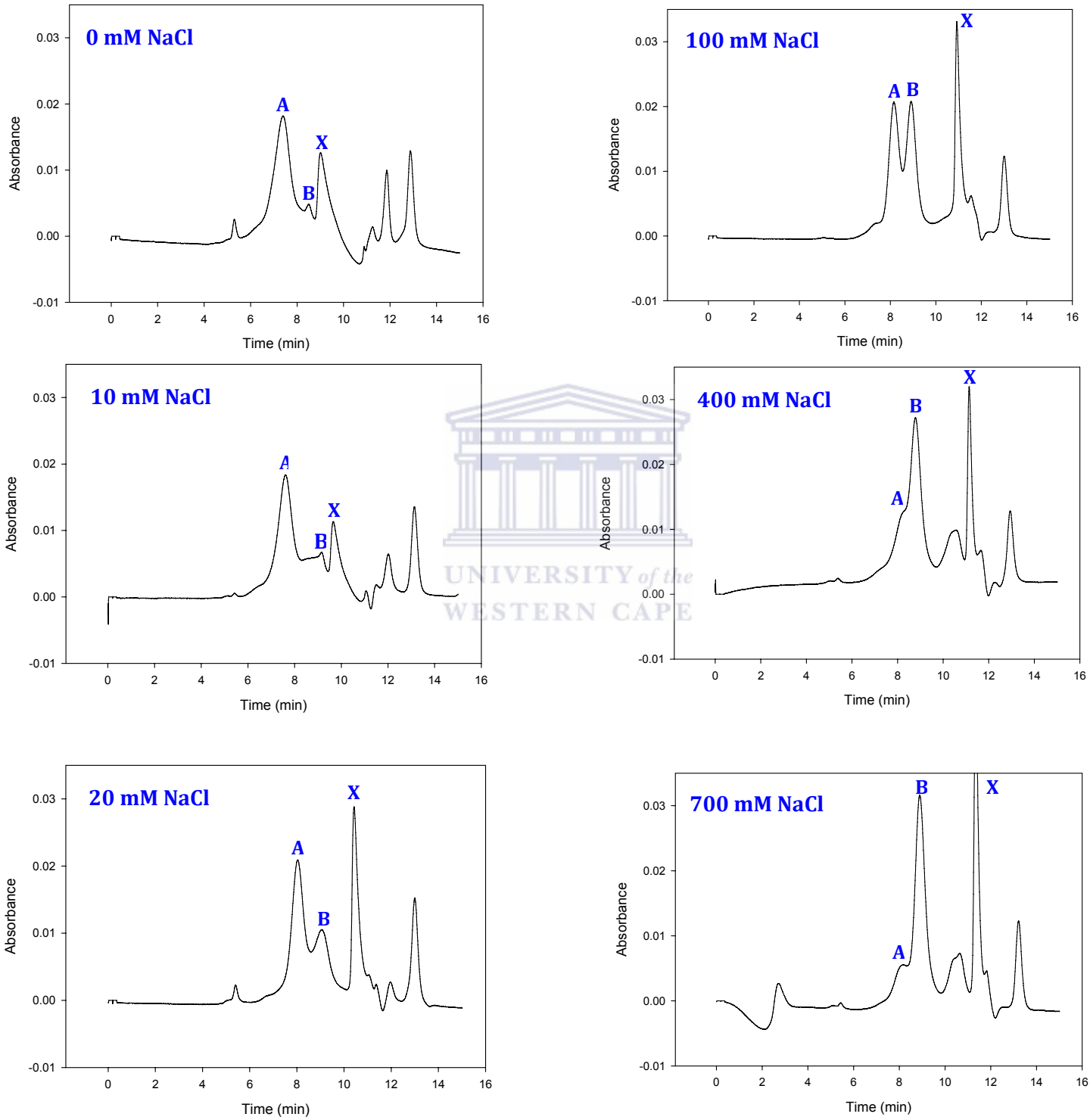


Figure 3.16: Effect of NaCl concentration on PfHPRT quaternary structure. Aliquots of protein (50 µL; the preparation had been dialysed against 10 mM sodium phosphate for ~24 h, and concentrated to ~1 mg/mL; it could be activated by substrates to yield an activity of ~3 µmol/min/mg of protein) were injected onto HPLC Tosohaas TSKgel G 3000 SW column (7.5 mm x 30 cm, 10 µm) at time zero with running buffer of 10 mM MOPS/TMAH, pH 7.0, 0.05% NaN_3, and the concentrations of NaCl shown. A and B identify two putative quaternary forms of PfHPRT, and X denotes an unidentified peak that shifts position with salt concentration.

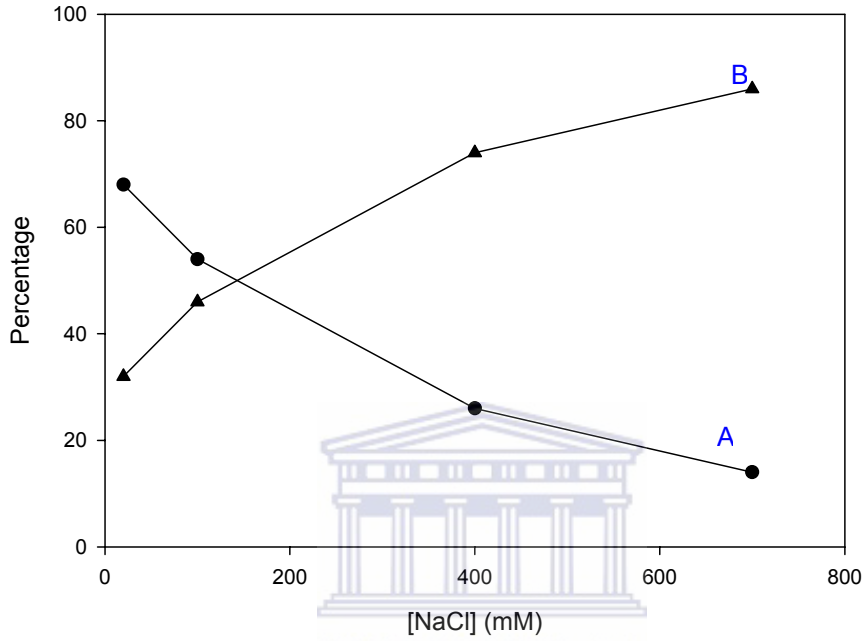


Figure 3.17: Effect of NaCl concentration on the proportion of A and B quaternary forms of PfHPRT. Details of conditions and HPLC elution profiles are shown in Figure 3.16. Quantification of peaks A and B was by gravimetric analysis of peak cut-outs. Profiles at 0 and 10 mM NaCl were not

We investigated whether HPLC gel filtration could be used to further purify PfHPRT. For this purpose the preparation that was capable of being activated and obtained from the 1 M NaCl elution from Reactive Red 120 column detailed in Figures 3.0 and 3.1 was used. The gel filtration was carried out at 10 mM MOPS, pH 7.0, 20 mM NaCl without NaN_3 in order to maximize the proportion of tetramer and a representative elution profile is shown in Figure 3.18. Peaks 1, 2, and 3 were collected from 4 injections of 200 µL

each, concentrated, and re-analyzed by HPLC gel filtration (Figure 3.19) and by SDS-PAGE (Figure 3.20).

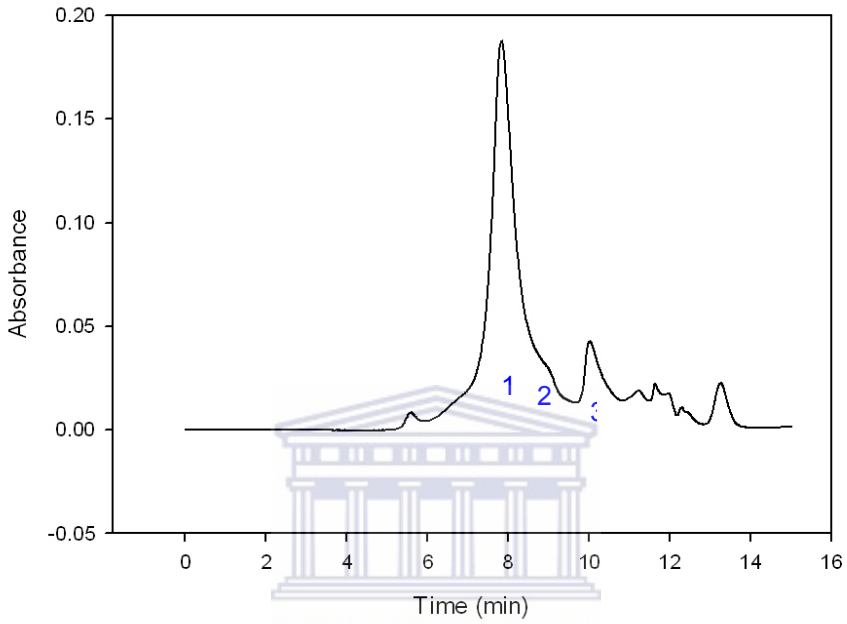


Figure 3.18: Preparative isolation of peaks 1, 2, and 3 by HPLC gel filtration. The HPLC trace shows an example of preparative scale purification by HPLC gel filtration on a Tosohaas TSKgel G3000 SW column (7.5 mm x 30cm, 10 μm) with 10 mM MOPS/TMAH, pH 7.0, and 20 mM NaCl (flow rate 1 mL/min), where 200 μl of PfHPRT, purified by ion exchange and Reactive Red 120 chromatography, dialyzed and concentrated to ~0.3 mg/ml, was injected and peaks 1, 2 and 3 collected. Similar HPLC chromatography was performed 4 times and

The gel filtration in 20 mM NaCl revealed that peak 1 was almost entirely tetrameric PfHPRT with one or two very minor peaks of smaller molecular weight species (the peak at ~13 min is not protein), and SDS-PAGE confirmed that the major species was PfHPRT plus a protein of approximately 30 000 Daltons (lanes 3 and 4). The contaminant species is likely a trimer or tetramer, to account for its elution with tetrameric PfHPRT during gel

filtration. Peak 2 contained tetrameric and dimeric PfHPRT in roughly equal amounts by gel filtration as expected, and the presence of PfHPRT was verified in the gel (lanes 5

and 6; there are also faint higher molecular weight bands because of the high concentrating effect). Peak 3 showed the presence of the species eluting at 10 min as expected, but what was difficult to understand was that the gel (lanes 7 and 8) showed the presence of a small amount of PfHPRT and little else. We tentatively explain this by the peak eluting at 10 min not to be protein, and a small amount of contamination by PfHPRT, although where it came from is not clear. Another possibility is that it is monomeric PfHPRT.

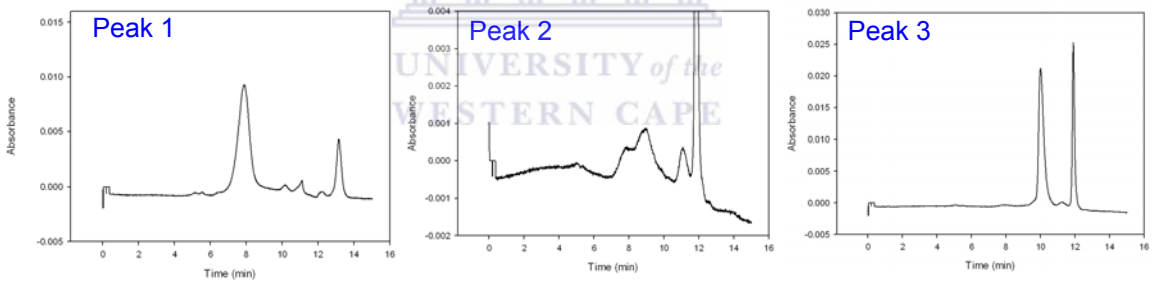


Figure 3.19: Rechromatography of Peaks 1, 2, 3 by HPLC gel filtration. Aliquots (50 μl, peak 1 with volume of 1.4 ml from approximately 10 mL, ~7x concentration; 30 μl for peaks 2 and 3 with volumes of 0.2 ml from 6 mL each, ~30x concentration) from the above experiment were analysed as in Figure 5.0 after concentrating. The peak eluting at 10-12 min (included volume) is a small molecular species that was not

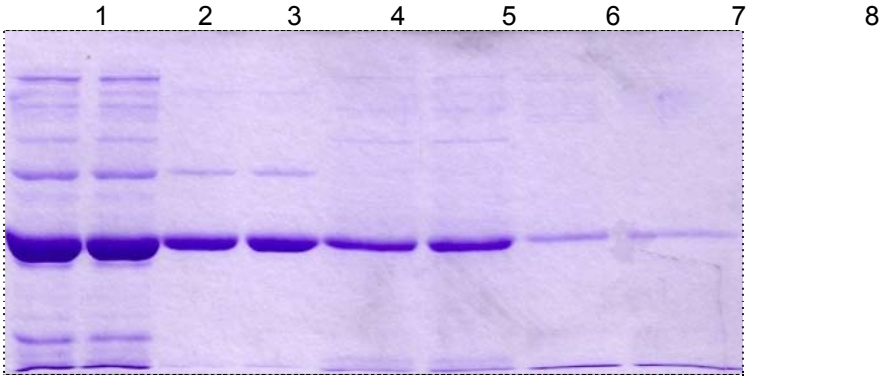


Figure 3.20: Analysis of PfHPRT purification using HPLC gel filtration. Lane 1, I M NaCl fraction from the Reactive Red 120 column; Lane 2, 1 M NaCl eluate of Reactive Red 120 filtered through a 0.22 µm filter; lanes 3-8 represent peaks collected from HPLC and concentrated as indicated in Figures 3.19 and 3.20. Two lanes were loaded for each peak to avoid cross-lane contamination and are as follows: Lanes 3-4, peak 1; Lanes 5-6, peak 2; Lanes 7- 8, peak 3.

It was of interest to determine whether Peak 1, the tetrameric species purified by HPLC gel filtration was active and exhibited the same salt-dependent tetramer-dimer equilibrium. The protein was activated with substrates and found to have an activity of 0.7 µmole/min/mg of protein, i.e. partially active. The salt-dependence of the quaternary equilibrium is shown in Figures 3.21 and 3.22, and can be seen to be similar to that in Figure 3.16, except that there is no dimer at low salt concentrations.

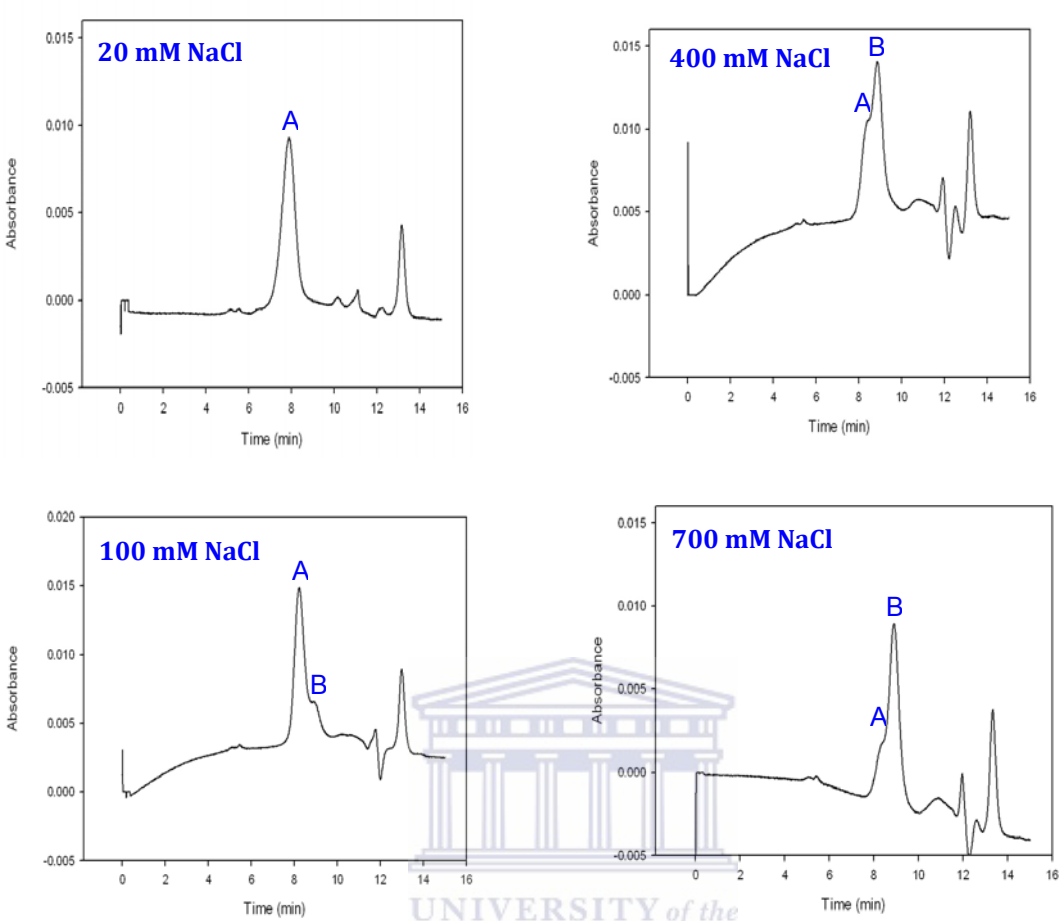


Figure 3.21: Effect of NaCl concentration on quaternary structure of HPLC gel filtration-purified HPRT. Aliquots of HPRT protein (50 μl, ~0.3 mg/ml; purified by HPLC gel filtration in 10 mM MOPS/TMAH, pH 7.0, 20 mM NaCl, dialysed against 10 mM MOPS, pH 7.0 overnight, and concentrated ~5x) were injected onto HPLC Tosohaas TSKgel G 3000 SW column (7.5 mm x 30 cm, 10 μm) at time zero with running buffer of 10 mM MOPS/TMAH, pH 7.0, without (at 20 mM NaCl) or with (at other NaCl conentrations) 0.05% NaN_3, and the concentrations of NaCl shown. A and B identify two putative quaternary forms of malarial HPRT.

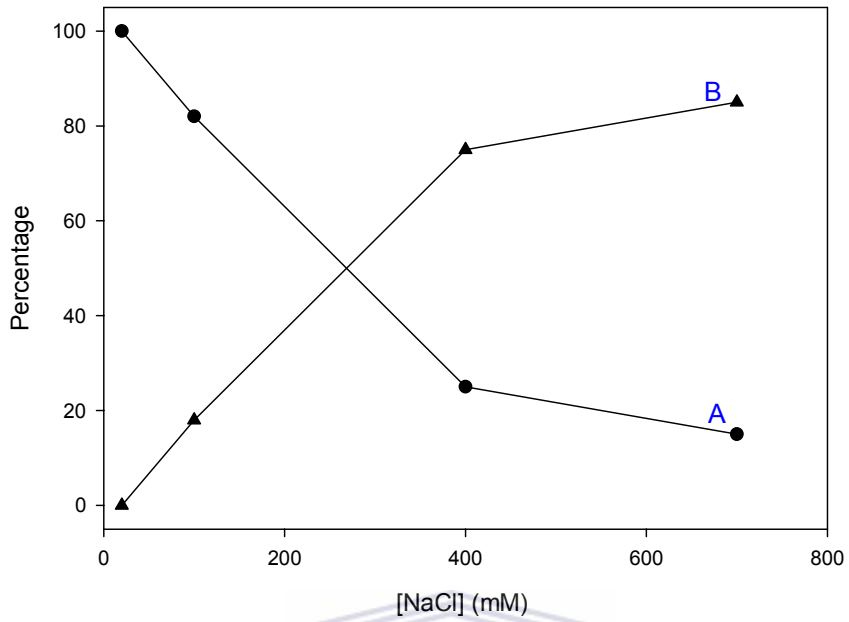


Figure 3.22: Effect of NaCl concentration on the proportion of quaternary forms of PfHPRT. The peaks shown in Figure 3.21 were quantified by gravimetric analysis of cut-outs.

Overall, gel filtration seems a very effective means of providing substantial final purification.

3.2 Comparative Modeling

In contrast to the human HPRT and HPRT's from other common disease-causing parasitic protozoa for which several crystal structures have been elucidated, the only available atomic structure for the malarial HPRT is that in complex with a transition-state analog inhibitor (PDB code 1CJB (Shi *et al.*, 1999)). The crystal structures reveal substantial differences between the holoenzyme, the substrate complexes, and product

complexes, as well as the transition-state structures. Thus, in addition to the crystallization experiments carried out, the structure of the free malarial HPRT was modeled using several solved structures of the phosphoribosyltransferase family of proteins as templates.

3.2.1 The open PfHPRT model structure

The model of the open PfHPRT was fashioned on the 1.90 Å structure of free human HPRT (PDB code 1Z7G (Keough *et al*., 2005)), the 1.90 Å structure of *Tritichomonas foetus* HPRT (PDB code 1HGX (Somoza *et al.,* 1996)) and the 2.30 Å structure of *Salmonella typhimurium* HPRT (PDB code 1J7J (Lee *et al.,* 2001)). Sequence identity between the malarial enzyme and the human, *Tritrichomonas foetus* and *Salmonella typhimurium* enzymes is 49%, 31% and 38%, respectively, indicating that the three solved structures could be useful templates. The first step in the modeling procedure was the generation of a multiple sequence –structure alignment using the advanced option of the 3DCoffee server (Poirot *et al.,* 2004). The alignment obtained was then manually refined to ensure no gaps in the core regions and that the aligned template sequences matched the corresponding PDB entries. Missing residues in the PDB templates were omitted from the alignment. A comparison of this alignment to the one obtained using the MALIGN3D routine of the MODELLER program (Sali and Blundell, 1993; Fiser *et al.,* 2000) confirmed it to be better and was, therefore, used for subsequent modeling. For modeling, an iterative approach comprising sequence-structure alignment and model building was followed. Models containing non-hydrogen atoms were then generated from

the best alignment using MODELLER which implements comparative modeling by satisfying spatial restraints extracted from the target-template(s) alignment. Out of the 50 models generated, only the one with fewest restraints violation and lowest energy was evaluated further.

Side chains for this model were optimized using SCWRL (Canutescu *et al.,* 2003). To evaluate the model quality, stereochemical analysis using PROCHECK (Laskowski *et al.,* 1993) and MOLPROBITY (Davis *et al.,* 2004) was carried out. Three-dimensional images of PfHPRT model structure were generated using PyMOL (DeLano, 2002). The sequence-structure alignment is shown in Figure 3.23.

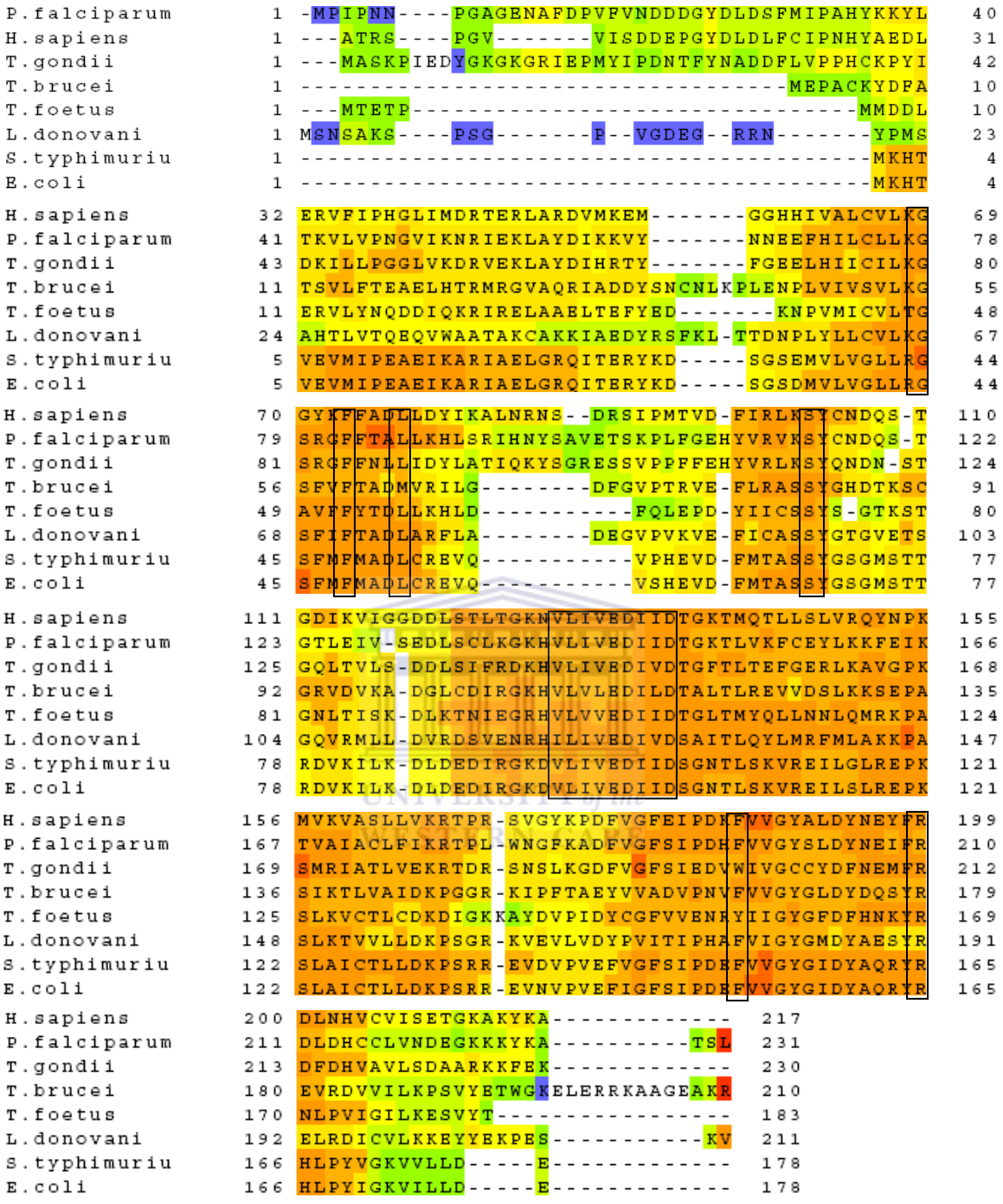


Figure 3.23: A structural-sequence alignment profile obtained using 3D-Coffee (http://igs-server.cnrs-mrs.fr/Tcoffee). Seven structures were aligned with the human HGPRT sequence. The relative reliability of the various regions is indicated by codes (blue, unreliable; green, low reliability; yellow, moderately reliable; orange, reliable; red, highly reliable portion of the alignment). Conserved

Three-dimensional models of PfHPRT dimer and monomer were generated with MODELLER (Sali and Blundell, 1993). The overall fold of the models is similar to that of previously determined phoshophoribosyltransferase structures (reviewed in Sinha and Smith, 2001) which consists of the core and hood domains as shown in Figure 3.24.

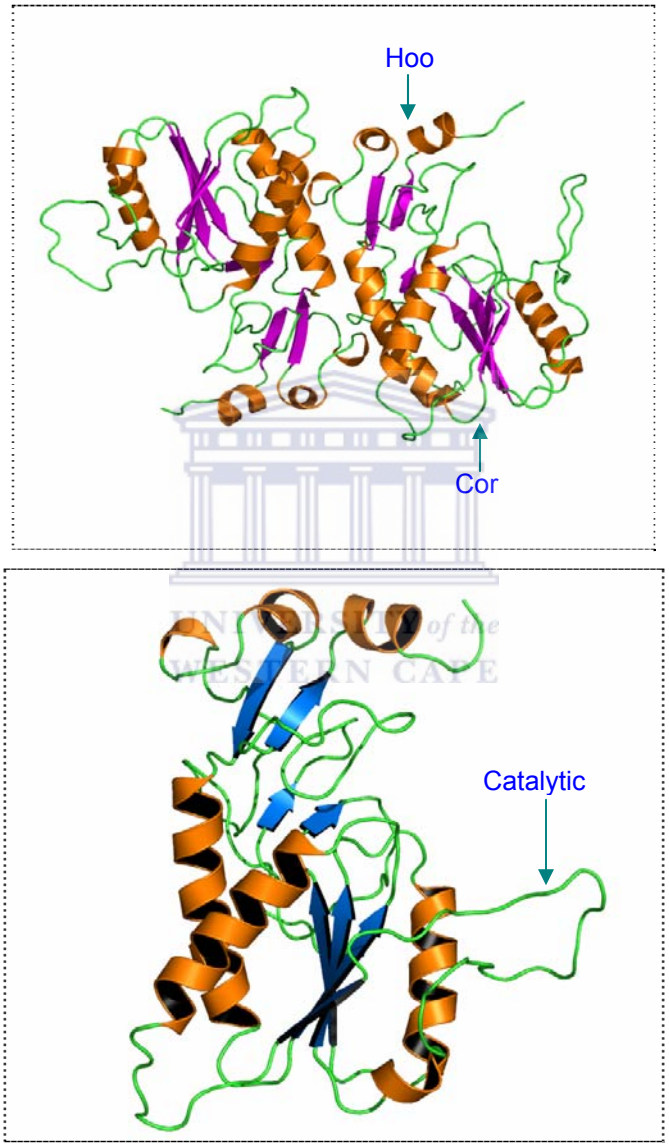


Figure 3.24: Cartoon representation of PfHPRT dimer and monomer. Top panel shows dimer with β-strands and α-helices coloured magenta and orange respectively. The tetramer can be generated by applying crystallographic symmetry to the dimer. The monomer with the catalytic loop in the open conformation is shown in the bottom panel with β-strands and α-helices coded blue and orange, respectively. Loops are

3.2.2 Subunit structure

Figure 3.24 (bottom panel) shows the three dimensional arrangement for the PfHPRT monomer. The secondary structure of the model organization consists of eight β-strands and seven α-helices linked together by loops and coils. The β-strands are formed by residues β1 (Leu40-Val45), β2 (Phe70-Leu75), β3 (Leu105-Lys113), β4 (His139-Ile146), β5 (Thr167-Lys176), β6 (Phe188-Ile193), β7 (Val198-Val199) and β8 (Cys215-Val218) while the α-helices are formed by residues α1 (Asp27-Ser29), α2 (Ala34-Tyr36), α3 (Val47-Tyr65), α4 (Arg81-Asn95), α5 (Leu132-Leu135), α6 (Lys151-Lys162) and α7 (Asp220-Tyr226). The sequence of secondary structure elements is represented in a simplified form using TOPS (Westhead *et al*., 1998) in Figure 3.25. The β-strands are organized into β-sheets with the major parallel sheet, made up of four strands (β5, β4, β2 and β3), forming part of the core region of the subunit. β4 and β5 of this sheet extends into a small parallel sheet almost perpendicular to the core sheet. The core β-sheet is braced by α-helices (α3, α4 one side and α5 on the other side). The smaller anti-parallel sheet (β1, β8) is found in the variable hood region. The arrangement seen in the model differed slightly from that predicted by PSIPRED (Jones, 1999a) using sequence information only. The secondary structure architecture predicted by PSIPRED consists of ten β-strands and seven α-helices and is schematically shown in Figure 3.26.

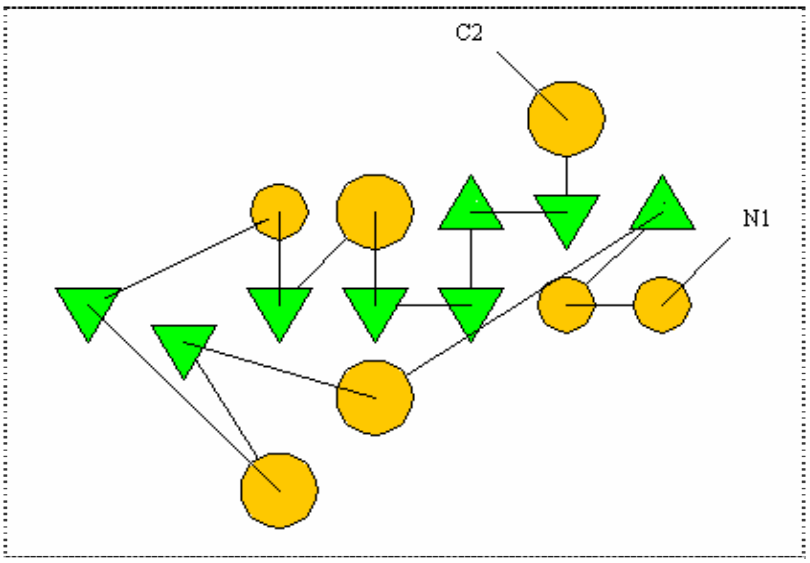


Figure 3.25: TOPS (Westhead et al., 1998) cartoon representation of the secondary structural elements of PfHPRT. The green triangles represent β-strands and orange circles helices. The N- and C-

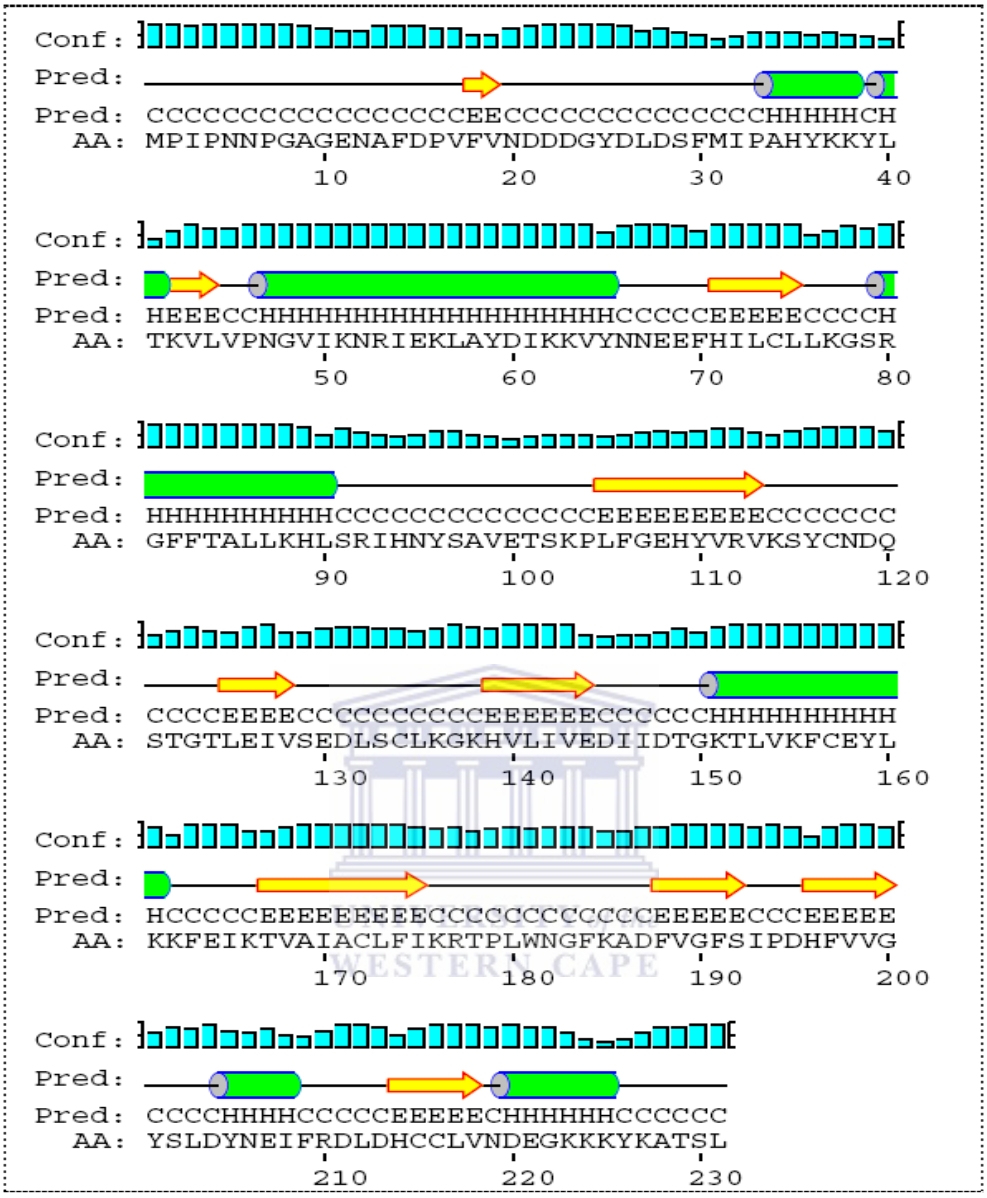


Figure 3.26: PfHPRT secondary structure prediction based on the amino acid sequence performed with PSIPRED (Jones, 1999). The top panel shows the probable secondary structure location along the sequence while the bottom panel represents the various symbolic keys used.

Pair-wise alignment of PfHPRT model structure with the solved crystal structure in complex with a transition state analog inhibitor (1CJB; Shi *et al.,* 1999) using ALIGN (Cohen, 1997) revealed that the secondary structural elements were fairly similar, but there was significant structural variability in the loop regions, particularly in the large flexible loop region as shown in Figure 3.27. This loop closes over the active site, thus, protecting the reactive transition-state from bulk solvent during catalysis in phosphoribosyltransferases (Focia *et al.,* 1998; Heroux *et al.,* 1998a; Shi *et al*., 1999b). It acts catalytically (holds a key tyrosine) as well. In closing over the active site, the loop becomes ordered and forms an antiparallel β-sheet. Overall, however, structures of free PfHPRT and 1CJB are rather similar as they superpose with a root-mean-square deviation (RMSD) of 1.29 Å for 189 C_α atoms The mobility of the catalytic loop in the unliganded enzyme could be one contributing factor as to why crystallization of free PfHPRT has been elusive thus far.

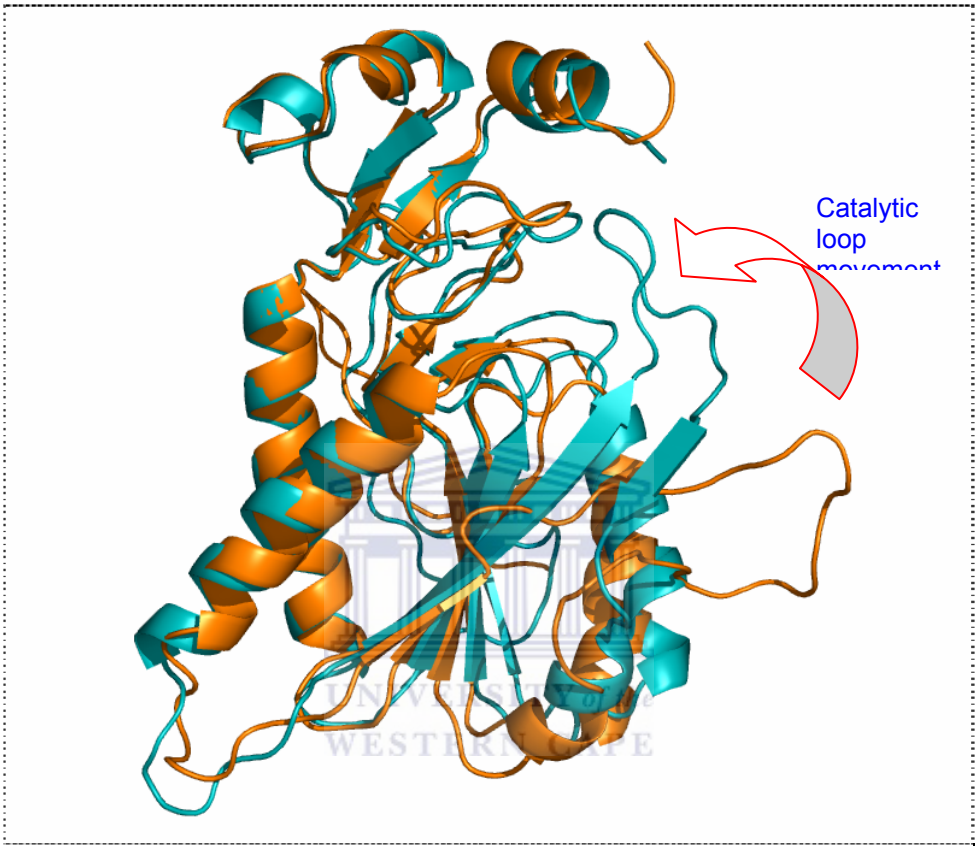


Figure 3.27: Superposition of the free PfHPRT (orange) with the solved crystal structure (PDB code 1CJB; green (Shi *et al.*, 1999)). The movement of the flexible loop which sequesters the active site from the bulk solvent is shown. The structures were superposed using ALIGN (Cohen *et al*., 1997)

3.2.3 Active site characterization

Identification and characterization of surface pockets and internal cavities in the monomer structure was performed by the CASTp server (http://cast.engr.uic.edu). The program permits identification and calculation of Connolly's molecular surface and

volume (Connolly, 1983) for all pockets and cavities in a protein structure. The cavities/pockets are then graded according to their sizes with the largest in most cases being the active site. It was found, using a probe radius of 1.4 Å, that PfHPRT monomer has a total of 38 pockets and cavities. A representation of the largest cavity/pocket, rendered in JMol is shown in Figure 3.28. This cavity has an area of 1137.8 $Å^2$ and a volume of 2913.8 $Å^3$ and is lined by 44 residues, mostly conserved residues involved in substrate binding. The large volume of this major cavity/pocket suggests that the enzyme can accommodate sizeable ligands. Possibly, therefore, both 2',3'-O-(2,4,6-trinitrophenyl)-guanosine monophosphate (TNP-GMP) and 4-iodo chalcone both of which have been shown to inhibit PfHPRT in this laboratory (Mbewe, 2005; Mbewe *et al*., 2007) bind to this site.

Several conserved residues involved in ligand binding adopt different conformations in the free and liganded PfHPRT. As shown in Figure 3.29, Leu76-Lys77 in catalytic loop I and Ser114-Tyr115 in loop II swivel by about 180 ° to be in the proper orientation for substrate binding. The effect of this translocation in the case of Leu76-Lys77 is the formation of the *cis*-peptide bond which is essential for the enzymes interaction with the pyrophosphate moiety. A similar rotation by Lys176, postulated to play a role in substrate recognition (Shi *et al*., 1999b), aids in orienting the terminal amino group towards the purine ring exocyclic oxygen (Figure 3.30, top panel). Phe197 which forms a hydrophobic interaction with the purine ring adopts a similar conformation in both the free and the liganded structures (Figure 3.30, bottom panel).

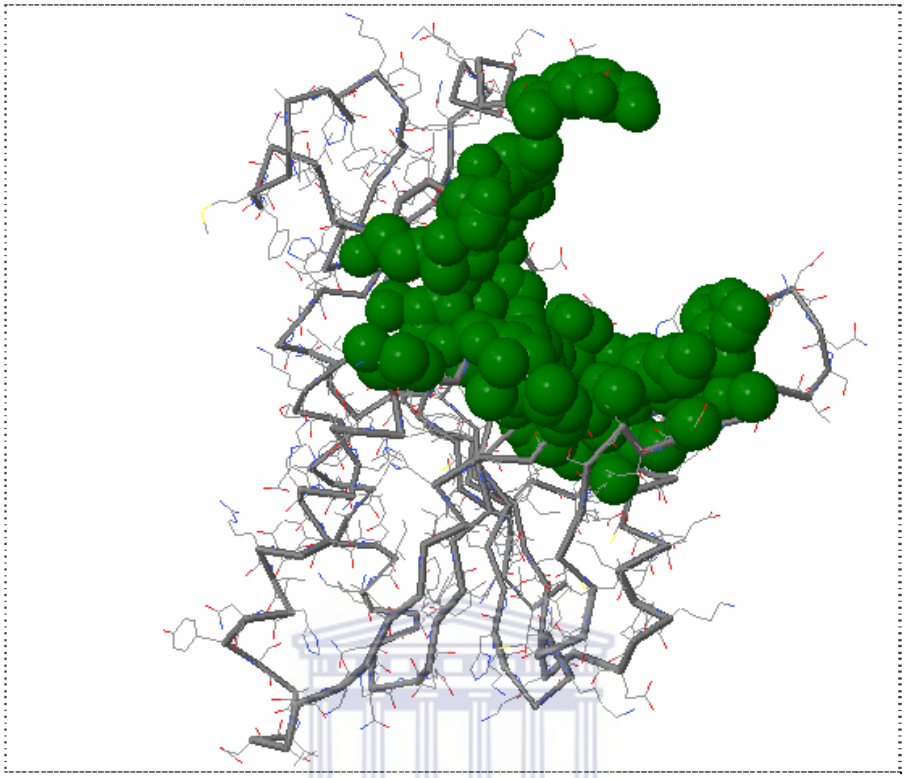


Figure 3.28: Wireframe representation of PfHPRT monomer showing the major pocket/cavity identified using a probe radius of 1.4 Å. The green spheres show the residues lining the cavity. The image was generated using CASTp (Liang *et al*., 1998) and rendered in JMol.

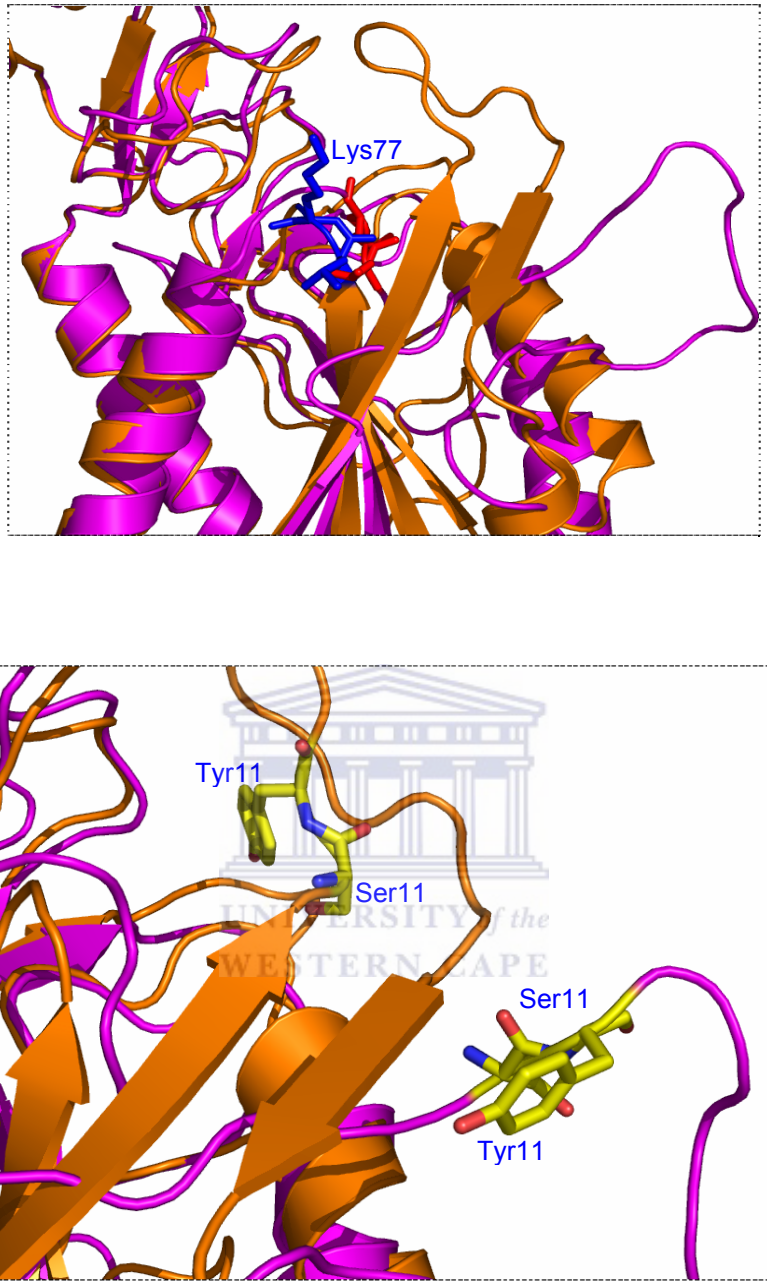


Figure 3.29: Orientations of the various conserved residues in the free (modeled) and ligand bound (1CJB) PfHPRT monomers. The free enzyme structure is coloured magenta while the liganded structure is colored orange. Top panel shows the orientations of the conserved *cis* dipeptide (Leu76-Lys77) represented as sticks and coloured blue for the liganded structure and red for the free structure. Bottom panel shows the orientation of Ser114-Tyr115 (sticks) in the two structures.

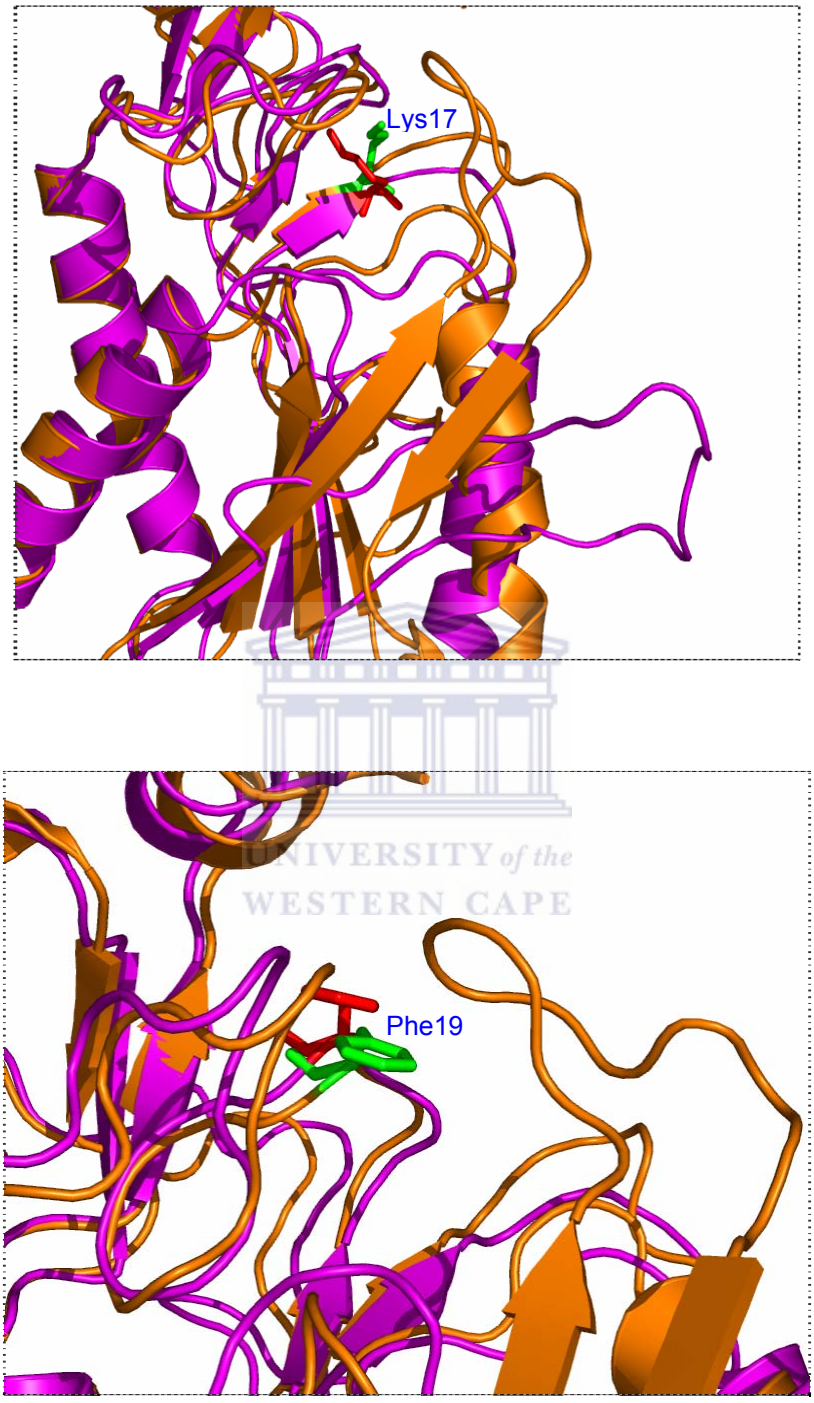


Figure 3.30: **Positioning of Lys176 and Phe197**. Top panel shows Lys176 in the free (magenta) and liganded PfHPRT (orange) structures. The residue, shown in stick representation is coloured blue and red in the free and ligand bound structures respectively. Bottom panel shows the orientation of Phe197 in both structures. Color coding is as for the top panel.

Superposition of the coordinates of the model with those of chain A of the human enzyme (1Z7G; Keough *et al.*, 2005) using ALIGN (Cohen *et al.*, 1997) showed that the two enzymes are significantly similar. The two superposed with a RMSD of only 0.67833 Å with major differences confined to the loop regions. Again the major difference was in the flexible loop region which was not built in the solved human structure because of unseen electron density due to disorder of the region.

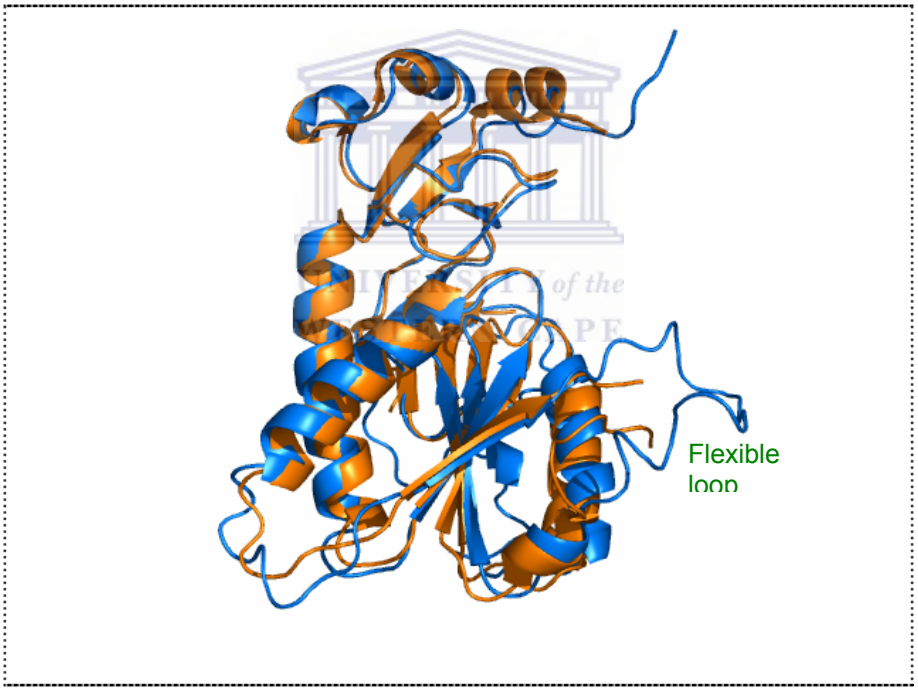


Figure 3.31: A structural superimposition of PfHPRT model (blue) and human HPRT (orange; PDB code 1Z7G; Keough *et al.*, 2005). The random mean of standard deviation (RMSD) between the two is 0.67833 Å. Structurally variable regions occur in the loops especially the flexible loop and the C-terminal end. Superposing was performed using ALIGN while the ray traced image was generated with PyMol v0.99 (DeLano, 2002).

In addition to the assessment of model quality performed by means of graphical display and C_α RMSD calculations, evaluation of the stereochemical parameters was also carried out and is shown in Figures 3.32 and 3.33.

The distribution of the *phi* and *psi* angles of apo-PfHPRT is excellent with most residues within the 'allowed' regions and only one residue (Tyr201) in the 'disallowed' region. This residue is localized in a loop where the modeling process is knowingly more difficult. The only outlier according to MolProbity was Lys77 which forms part of the unusual *cis* peptide bond in PfHPRT.

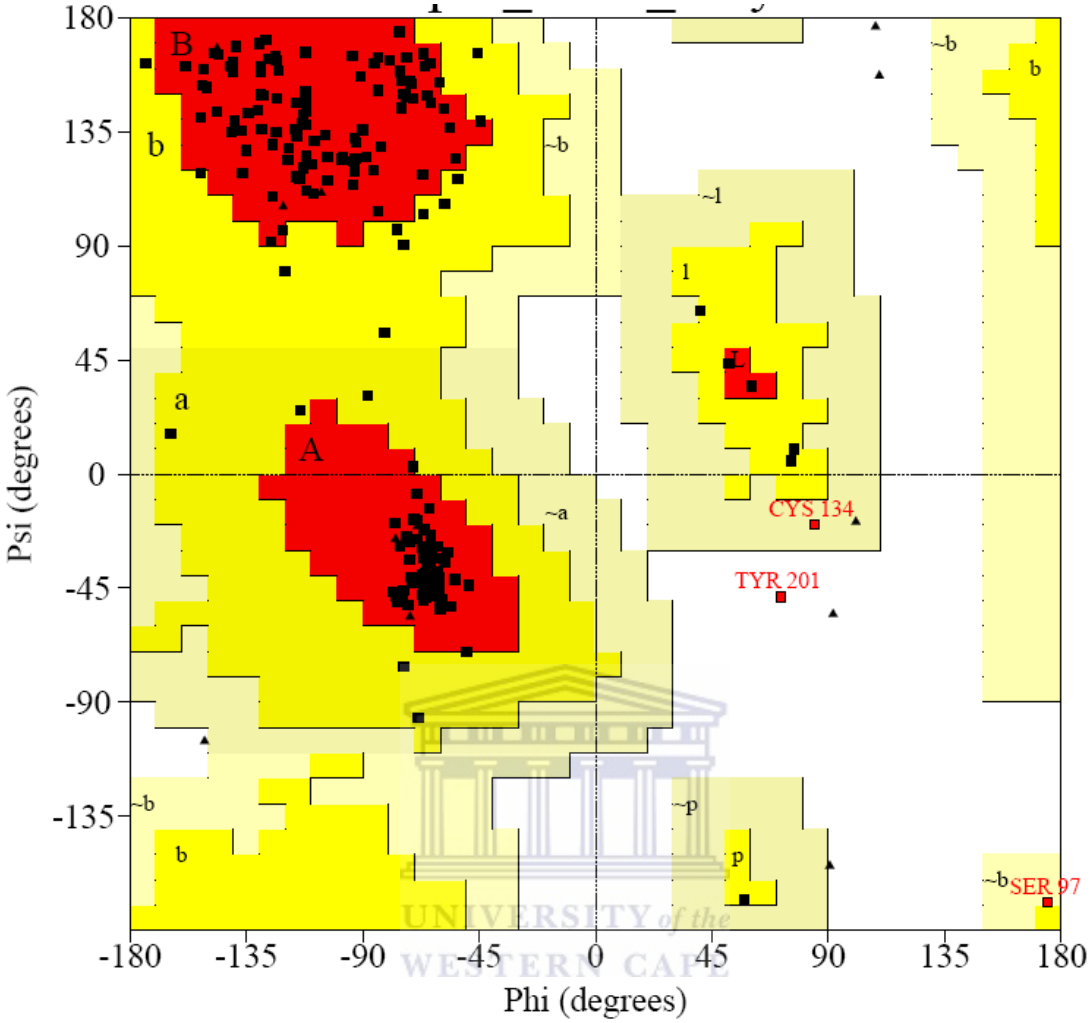


Figure 3.32: Ramachandran plot for the PfHPRT model. The percentage of residues in the favoured, additionally allowed, generously allowed and disallowed regions was 88.8% (182), 9.7% (20), 1.0% (2) and 0.5% (1) respectively. The number of residues in each category is in brackets. The plot was generated using PROCHECK.

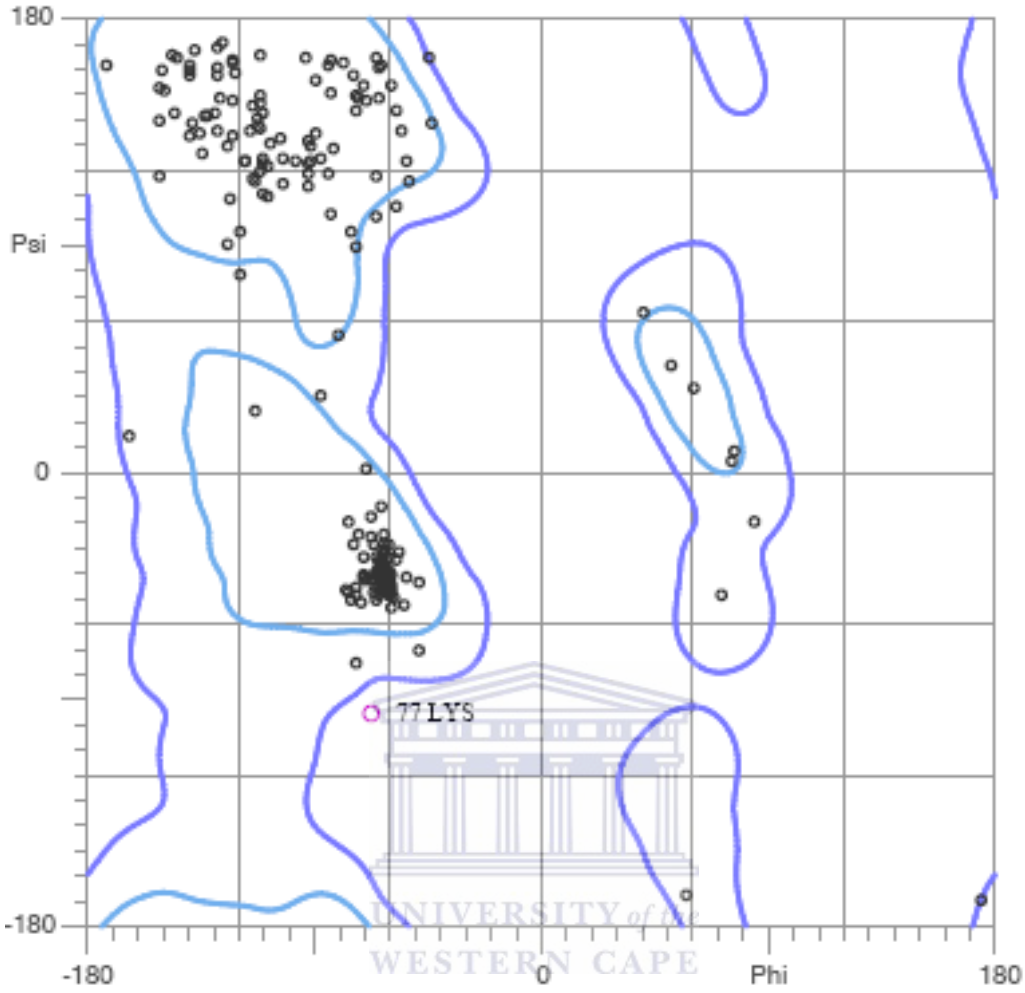


Figure 3.33: MolProbity Ramachandran plot of PfHPRT model.

.

3.3 Preliminary crystallization trials

The ultimate objective of this work was to obtain the crystal structure of PfHPRT in complex with a chalcone, which has been shown to act as accelerator of activity (in presence of Mg^{2+} and substrates), an activator of the inactive to active transition (divalent cations and substrates not necessary), and a potent inhibitor of activity (when Ca^{2+} replaces Mg^{2+}) (Phehane, 2002; Mbewe, 2005).

Numerous co-crystallization trials of PfHPRT were conducted, mostly in complex with IMP, GMP or chalcone. For the trials with IMP or GMP, an excess of the other product, Mg^{2+}-pyrophosphate, was usually included in the crystallization drops while for the chalcone trials, Mg^{2+}-pyrophosphate was usually substituted for calcium. PfHPRT concentration was kept at 15-20 mg/mL (as determined by the method of Bradford). The protein was routinely filtered through a 0.22 μm filter prior to setting up trials.

Initial screening for probable conditions was performed using Hampton screens 1 and 2 (Hampton Research) which comprises a total of 100 individual conditions. These screens are based on the sparse matrix approach (Jancarik and Kim, 1991) in which a wide range of conditions in the crystallization space are sampled. Hanging drop vapour diffusion method (McPherson, 1982), using 24-well VDX plates (Hampton Research), was employed. The 2 μL (1 μL protein plus ligand and 1 μL precipitant) hanging drops, mounted on 22 mm round siliconized cover slips, were inverted over 1 mL reservoir

solutions in the plates. Parallel screens at both 18 ºC and 4 ºC were set-up in an air-conditioned room free from vibrations and noise. The drops were observed immediately after set up and afterward daily for two weeks. Thereafter, observations were made once weekly for up to three months. Most drops, at both temperatures, were clear on set-up. After one day, however, majority of drops at 18 ºC had precipitated, while most at 4 ^{o}C remained clear.

Crystalline precipitate and phase separation, shown in Figure 3.34, generally point at potential crystallization conditions. Both occur in the supersaturation region of the phase diagram through different obscure mechanisms (Asherie, 2004). Phase separation is probably caused by the partitioning of precipitating agents present in the mother liquor as a result of immiscibility (McPherson, 1999) while precipitation is mainly a result of excessively fast nucleation due to very high supersaturation of either the protein or the precipitating agent. In phase separation, the protein sometimes concentrates and forms crystals in one phase (Kuznestov *et al.*, 2001).

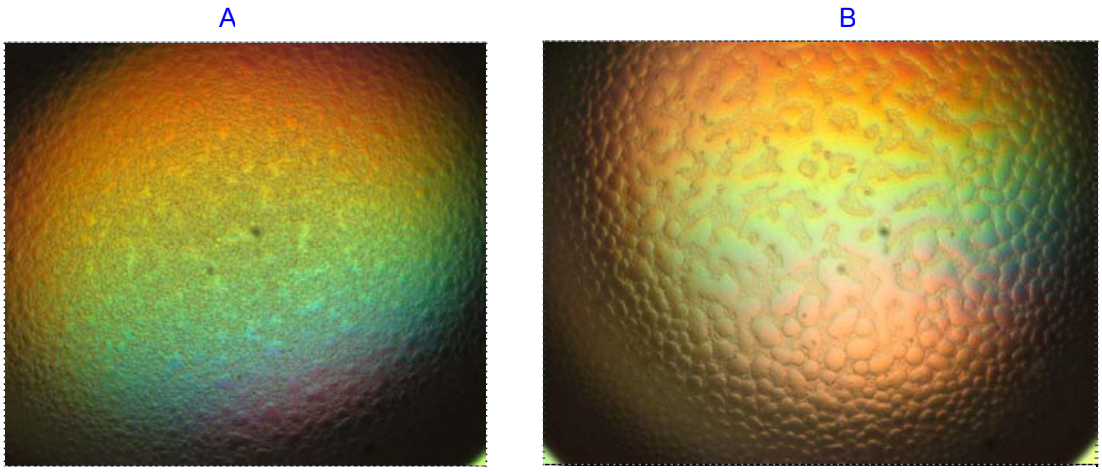


Figure 3.34: Initial crystallization results obtained using Hampton screens 1 and 2 conditions after 5 days of set-up at 18 °C. Panel A shows a crystalline precipitate of PfHPRT-IMP obtained using 30% PEG 4000, 0.1 M Tris-HCl, pH 8.5, 0.2 M $MgCl_2$. Panel B shows phase separation of PfHPRT-GMP obtained with 30% PEG 4000, 0.2 M $(NH_4)_2SO_4$.

Results obtained for some drops during the screening trials, as shown Figure 3.34 above, were encouraging. Crystalline precipitate and phase separation (hundreds of small round droplets) give an indication that the protein may crystallize with careful fine tuning of the conditions. This was explored as elaborated in section 3.3.1.

3.3.1 Refinement of initial conditions

The two conditions in Figure 3.34 were refined further using a fine systematic grid screen around some conditions as shown in Figures 3.35 and 3.36. No positive results were, however, obtained.

PEG 4000 / pH	20%	22%	24%	26%	28%	30%
7.0						
7.5						
8.0						
8.5						

Figure 3.35: A grid of pH vs. PEG concentration optimization of crystallization for PfHPRT-IMP.

PEG 4000 / $(NH_4)_2SO_4$	20%	22%	24%	26%	28%	30%
0.15 M						
0.20 M						
0.25 M						
0.30 M						

Figure 3.36: A grid of $(NH_4)_2SO_4$ concentration vs. PEG concentration optimization of crystallization for PfHPRT-GMP.

It was then decided to set up trials around the condition in which crystals of PfHPRT in complex with immucillinHP grew which was 20% PEG, 100 mM, pH 7.5 Hepes (the conditions also included 1.5 uM ImmucillinHP and an excess of Mg^{2+}– pyrophosphate). We systematically varied PEG concentration and the pH and held Hepes concentration constant. Three screens for IMP, GMP and chalcone were set-up as shown in Figure 3.37.

PEG 4000 / pH	20%	22%	24%	26%	28%	30%
7.0						
7.3						
7.5						
7.8						

Figure 3.37: **A grid screen set-up of PEG concentration vs. pH for PfHPRT co-crystallization with IMP, GMP or chalcone (each at 1 mM; with 5 mM $MgCl_2$ for IMP/GMP, and 5 mM $CaCl_2$ for the chalcone).**

This screen yielded a number of crystals with GMP and IMP but not with the chalcone, as shown in Figure 3.38.

Figure 3.38: Crystals obtained from the grid screen in Figure 3.3.7. Top panel shows showers of microcrystals with PfHPRT, 1 mM GMP, 5 mM $MgCl_2$, 30% PEG, 100 mM Hepes, pH 7.5. Middle panel, microcrystals with PfHPRT, 1 mM IMP, 5 mM $MgCl_2$, 20% PEG, 100 mM Hepes, 7.3. Bottom panel shows microcrystals with PfHPRT 1 mM IMP, mm 5 $MgCl_2$, 20% PEG, 100 mM Hepes, 7.5.

Refinement of the conditions was carried out in two ways: for the GMP-mix microcrystalline precipitate, PEG concentration lowered to 25%. After mixing the protein-ligand solution with the precipitant on a siliconized glass cover slide, the drop was seeded with material from the 30% PEG drop using the Hampton Research seeding tool. In the second approach, conditions were set up as previously (Figure 3.37) but the rate of nucleation was slowed down using oils. Three different types of oil (Al's, Paraffin and Silicon) were tested separately. After mixing the precipitant with protein on the cover slide, 400 µL of oil was layered on top of the reservoir solution and the cover slide inverted over it for equilibration.

The seeding approach resulted in microcrystals (not shown) while elegant crystals, shown in Figure 3.39, were obtained with oils. On diffraction, these crystals however turned out to be salt crystals. The salt diffraction pattern is shown in Figure 3.40.

A B

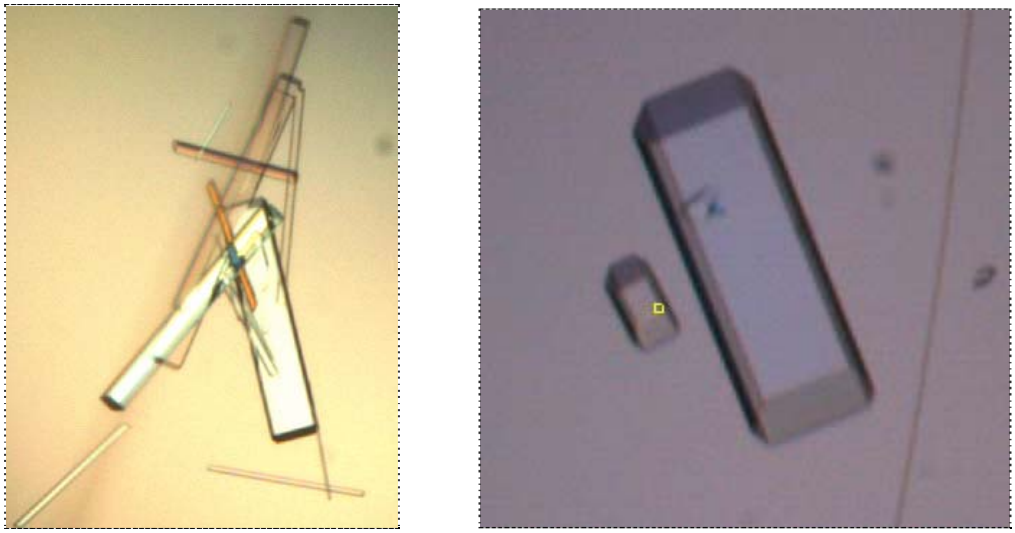


Figure 3.39: **Salt crystals.** Crystals in panel A were obtained with 20% PEG, 100 mM Hepes, 7.3 as precipitant but layered with 400 µL Al's oil. Crystals obtained with 26% PEG, 100 mM Hepes, pH 7.5 and 400 µL Al's oil are shown in panel B.

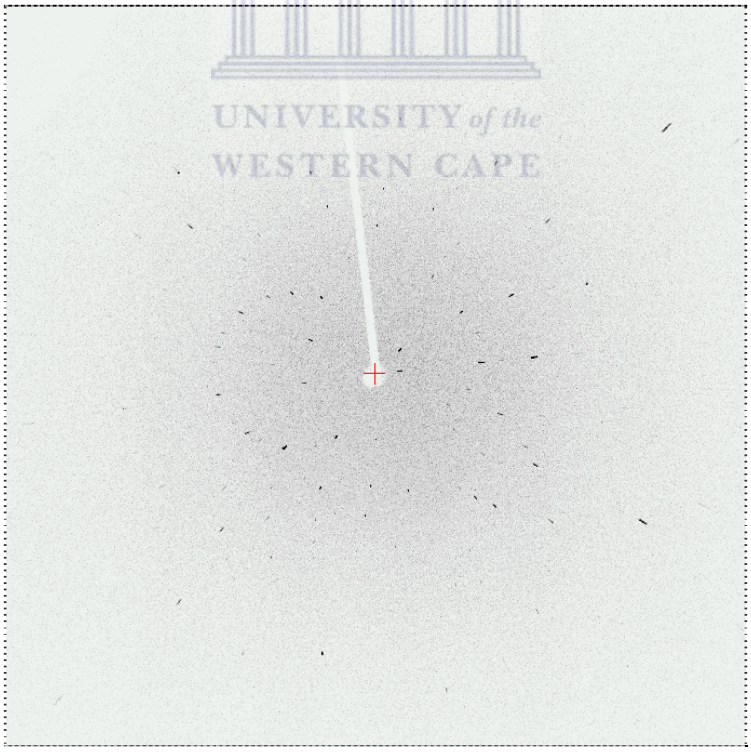


Figure 3.40: **Diffraction pattern of a salt crystal**. The crystal arose from a solution of 20% PEG, 100 mM Hepes, 7.3 as precipitant layered with 400 µL Al's oil to slow down the rate of nucleation.

Formation of salt crystals is a common phenomenon in macromolecular crystallization. Like protein crystals, salt crystals come in different forms and shape. Though it's not easy to differentiate between a protein crystal and a salt crystal by direct observation the latter are highly birefrigent. Use of methylene blue dye, which permeates (through) the solvent channels in protein crystals thus colouring the crystals blue, or poking with a fine glass rod are fast, rough ways of establishing what type of crystals one has. Protein crystals, unlike salt crystals, also crumble easily on the slightest application of pressure. Ultimately though, crystal type is confirmed by a diffraction pattern.

The diffraction pattern of salt crystals differs from that of protein crystals. Since small molecules, like salt, have shorter Bragg distances compared to protein molecules, they tend give fewer reflections with widely separated spots. Because of the longer Bragg distances, at the same crystal to detector distance, protein crystals produce more reflections and spots much closer to each other.

The fact that most of the optimized drops resulted in salt crystal formation suggested that the sodium phosphate used as the final dialysis buffer before setting up trials was possibly unfavourable. Thus, this buffer was changed to MOPS and further trials performed. Numerous salt crystals, some shown in Figure 3.41, were again obtained with this strategy.

A B

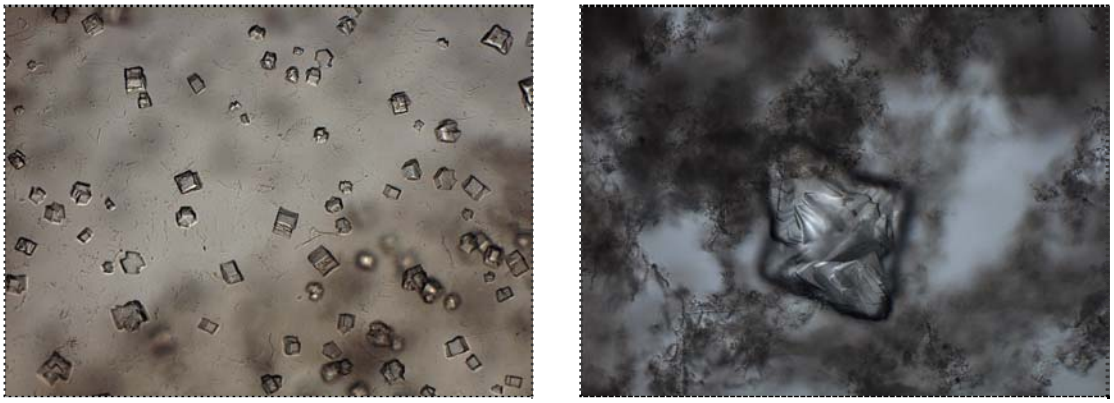
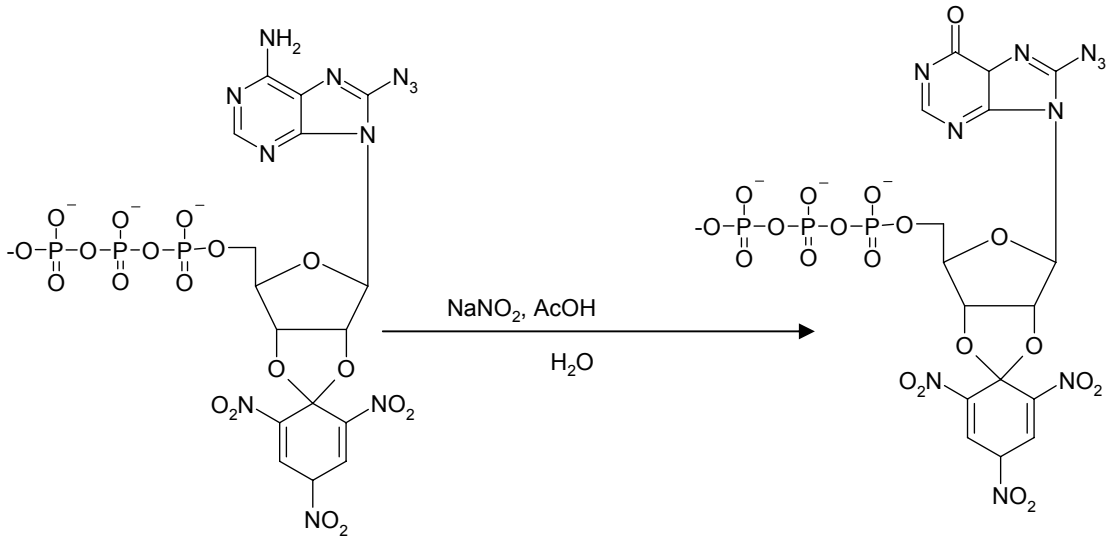


Figure 3.41: **Salt crystals obtained when PfHPRT was dialysed against 10 mM MOPS, pH 7.0, 1 mM DTT prior to setting up trials.** Panel A shows small salt crystals obtained with 20% PEG 4000, 100 mM Hepes, pH 7.3 as precipitant. The 1 mL reservoir solution was layered with 300 µL Al's oil. Panel B shows a sizeable salt crystal growing from precipitate and exhibiting twinning. This crystal was obtained with 26% PEG 4000, 100 mM Hepes, pH 7.5.

3.4 TNP-8N_3-ITP synthesis

2',3'-O-trinitrophenyl nucleotides have been widely used as fluorescent probes for nucleotide binding sites in enzymes (Hiratsuka, 1982). The 8-azido derivatives were introduced by our laboratory as photoaffinity labels of sarcoplasmic reticulum Ca^{2+}-ATPase (Seebregts and McIntosh, 1989). TNP-GMP has been found to be a potent inhibitor of PfHPRT (Phehane, 2002), and we, therefore, considered that the 8-azido derivative of TNP-GMP, and possibly TNP-IMP, could be useful photoaffinity probes of this enzyme. However, the synthesis of these compounds in a radiolabelled form presents considerable difficulties. The synthesis of 6-oxo-3',5'-cyclic phosphates from cAMP has been described (Meyer et al., 1972), and as the reaction conditions seemed fairly mild, and oxidizing (the 8-azido group is very susceptible to reduction), we considered that it may be possible to convert TNP-8N_3-ATP into TNP-8N_3-ITP directly, and possibly [γ-^{32}P]TNP-8N_3-ATP into [γ-^{32}P]TNP-8N_3-ITP. The synthesis of [γ-^{32}P]TNP-8N_3-ATP is routinely performed in our laboratory every two months for photolabelling mutants of Ca-$^{2+}$-ATPase (Clausen *et al*., 2006; McIntosh *et al*., 1999).

PfHPRT is specific for the monophosphate species, and it may be considered futile to try to use the triphosphate species as a photolabel. However, examination of the open structures (unliganded) of several HPRTs led us to believe that the extra two phosphates could easily be accommodated by extending out of the active site (see Discussion).



TNP-8N_3-ATP TNP-8N_3-ITP

Figure 3.42: Scheme for synthesis of TNP-8N_3-ITP from TNP-8N_3-ATP.

The reaction, according to Meyer *et al*. (1972), was tested first for ATP, then TNP-ATP, and finally TNP-8N_3-ATP, and in each case was found to convert these species into the inosine derivative with very good yields. The separation of TNP-8N_3-ATP and TNP-8N_3-ITP was achieved by reverse phase HPLC. Repeated injections provided a means to isolate a fairly large amount of pure TNP-8N_3-ITP. The spectrum of TNP-8N_3-ITP, as well as that following 1 min irradiation, is shown in Figure 3.43. It is evident that there is a peak at 273 nm associated with the purine ring, and another at 408 nm and a shoulder at 468 nm characteristic of the TNP Meisenheimer species. Irradiation resulted in disappearance of the 273 nm peak, demonstrating the presence of the azido group in the parent compound.

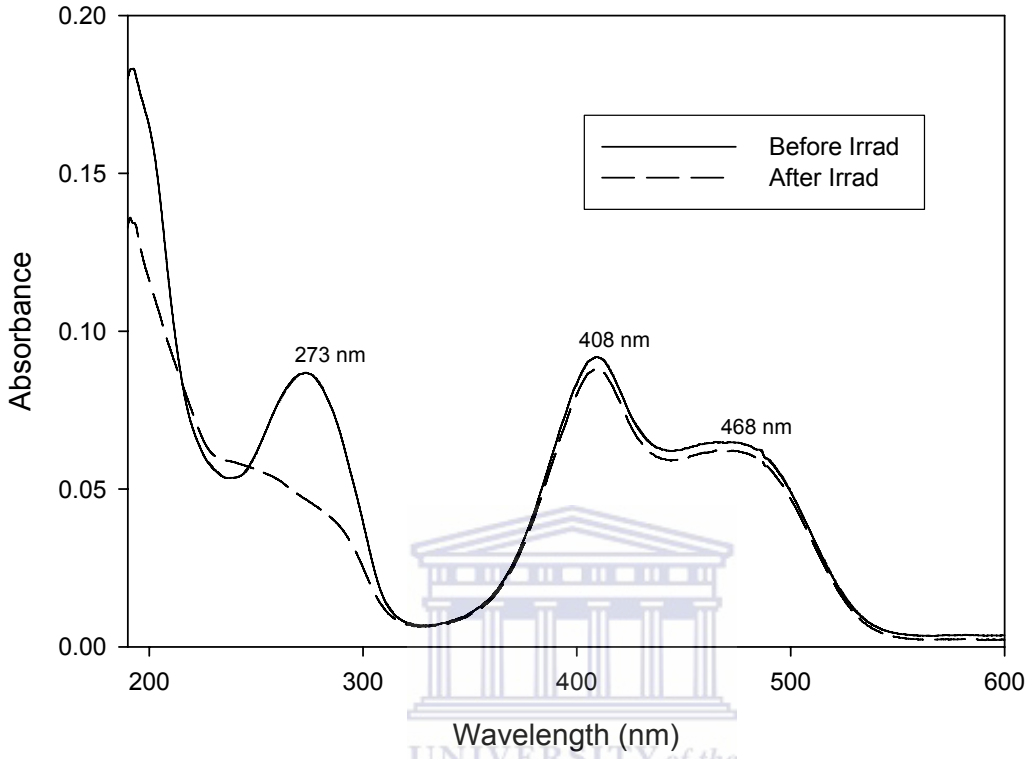


Figure 3.43: The absorption spectrum of TNP-8N_3-ITP before and after irradiation with UV light. The abolishment of the 273 nm peak upon irradiation is indicative of the presence of an azido group. The 408 nm peak with a shoulder at 468 nm is characteristic of a TNP Meisenheimer complex.

More challenging was defining conditions for conversion of [γ-^{32}P]TNP-8N_3-ATP to [γ-^{32}P]TNP-8N_3-ITP as the former species is at low concentration and very hot (typically 2 x 10^7 cpm/nmol). We could realistically only use 30 nmol [γ-^{32}P]TNP-8N_3-ATP total, and the isolation/purification needed to be performed on a C-18 Sep-Pak cartridge.

The time dependence of synthesis of a trial of cold TNP-8N_3-ITP from 30 nmol TNP-8N_3-ATP as judged by reverse-phase HPLC is shown in Figure 3.44. After 4 h reaction,

the mixture was processed through a Sep-Pak cartridge exactly as if it had been very hot, and the resulting HPLC elution at both 210 nm and 408 nm is shown in Figure 3.45. Looking at the former trace (blue), if the first peak at 2.5 min is ignored (probably acetic acid), TNP-8-N_3-ITP is the predominant peak and possibly about 50% pure. In the case of the 408 nm trace (red), it could be 95% pure. The species remaining after 3-4 h at the position of TNP-8N_3-ATP was, according to the 210/408 nm ratio, not actually TNP-8N_3-ATP, and must represent a minor contaminant of the "pure" TNP-8N_3-ATP, which is not surprising as it was purified by the same HPLC system. Taking the latter into consideration, it appears that the conversion of TNP-8N_3-ATP into TNP-8N_3-ITP was close to 100%.

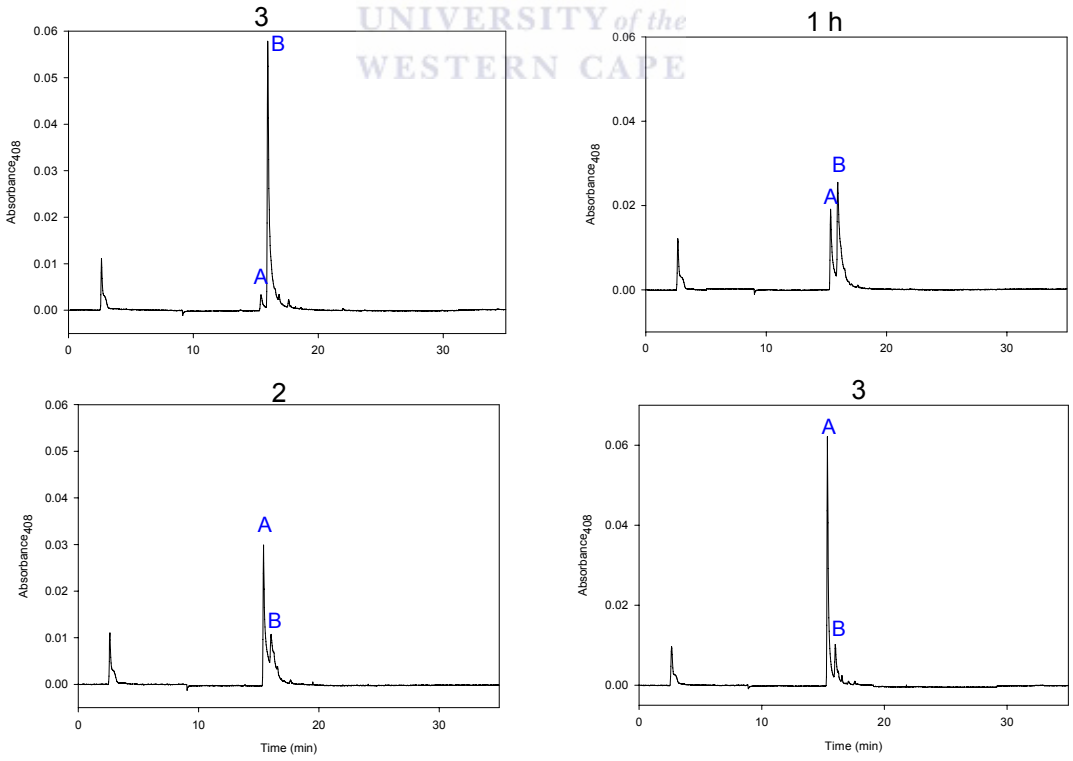


Figure 3.44: **HPLC traces for reaction mix for the synthesis of TNP-8N_3-ITP from 30 nmol TNP-8N_3-ATP**. Reaction times are as shown. A and B show the position of TNP-8N_3-ITP and TNP-8N_3-ATP respectively. The mobile phase was 10 mM KPi, pH 6, 60% (v/v) acetonitrile. Detection was at 408 nm.

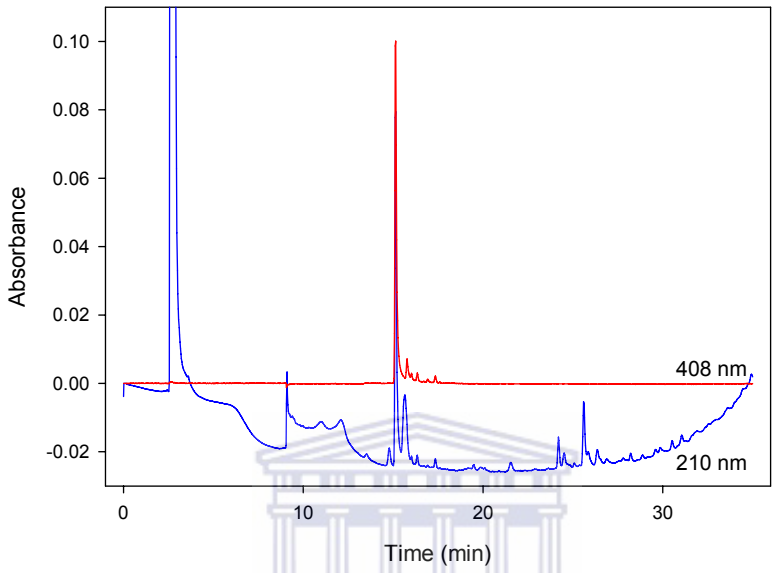


Figure 3.45: **HPLC trace of the Sep-Pak purified TNP-8N_3-ITP after 4 h reaction on ice**. The mobile phase was 10 mM KPi, pH 6, 60% (v/v) acetonitrile. The red and blue spectra represent absorption at 408 nm and 210 nm respectively.

Synthesis of [γ-^{32}P]TNP-8N_3-ITP was performed under identical conditions as described above using 30 nmol [γ-^{32}P]TNP-8N_3-ATP with a specific activity of 5.6 x 10^6 cpm/nmol. The concentrated sample after Sep-Pak purification was 130 µM, assuming the extinction is the same as that for TNP-8N_3-ATP. The volume and yield will be determined in the future after the radioactivity has declined to safer levels.

3.5 Photoaffinity labeling

Azidonucleotides, particularly the 8-azido adenine compound, have been used to probe nucleotide binding sites of proteins (McIntosh *et al*., 1992; Seebregts and McIntosh, 1989). Upon light irradiation, the azido group is converted into the highly reactive nitrene species that attacks nearby residues, particularly good nucleophiles, like tyrosine, and lysine, forming a covalent bond between the nucleotide and the protein.

Photolabeling of PfHPRT with [γ-^{32}P]TNP-8N_3-ATP under the standard conditions shown to be optimal for sarcoplasmic reticulum Ca^{2+}-ATPase (pH 8.5, 20% glycerol, 1 min irradiation) is shown in Figure 3.46, and the quantification shown in graphical form in Figure 3.47. The gels show that there is one major component of the preparation that is photolabelled that is in a position where PfHPRT is expected, and the overall banding pattern at higher concentrations of nucleotide is rather similar to that obtained with Coomassie Blue staining. In the absence of divalent cations the concentration dependence was complex, exhibiting a hyperbolic rise in labelling in the 0-5 μM range and, thereafter, a linear component, the former typifying specific binding and labelling and the latter typical of a nonspecific reaction. Fitting the curve yielded a $K_{0.5}$ of 2.06 μM and maximal derivatization of 0.018 mol/mol PfHPRT. Thus, the binding affinity was fairly tight and the efficiency rather low. In the presence of either 1 mM Mg^{2+} or Ca^{2+}, the labelling was much poorer and close to linear, suggestive of a nonspecific reaction. Either

the TNP-nucleotide does not bind specifically in the presence of divalent cations or the photolabelling reaction of bound nucleotide has very poor efficiency.

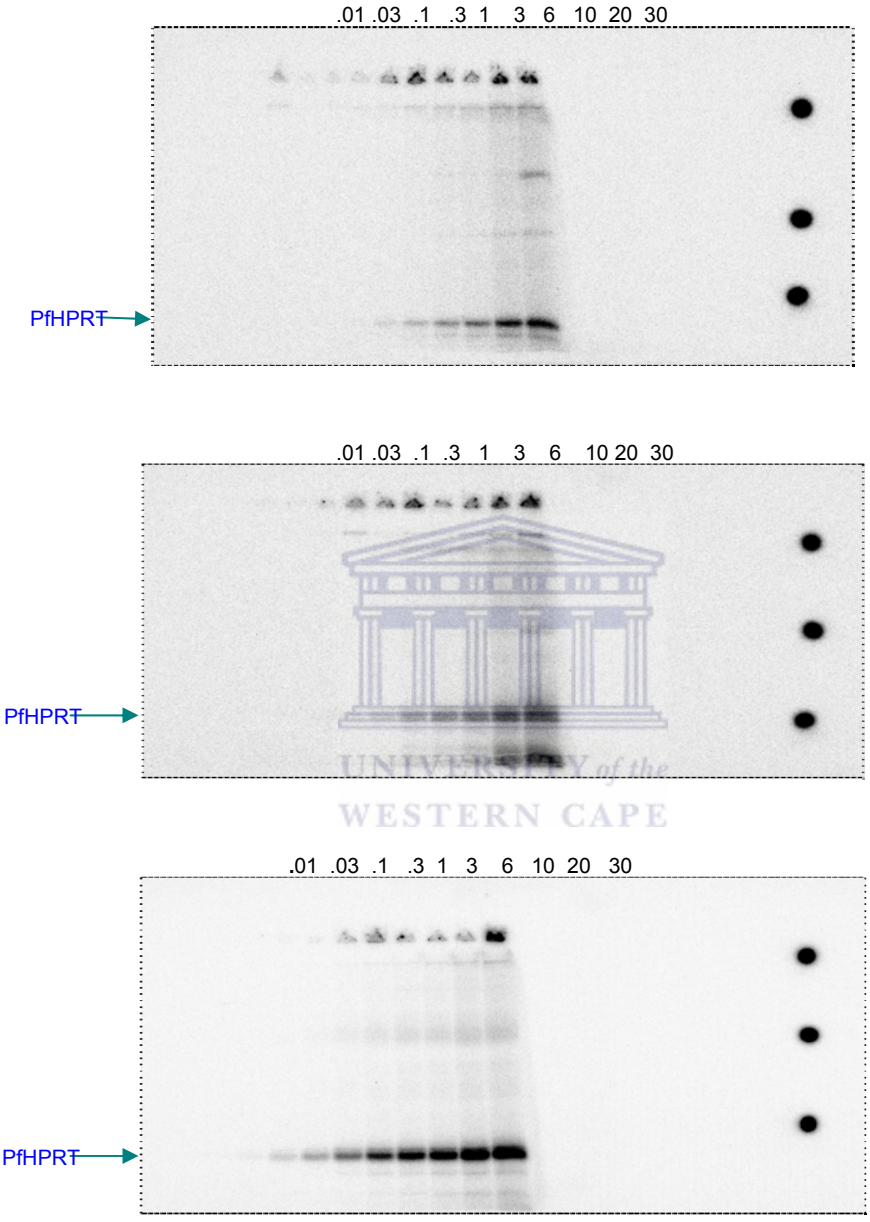


Figure 3.46: Radioactivity imaging of SDS-PAGE (12%) analysis of [γ^{32}P]TNP-8N$_3$-ATP photolabeling of PfHPRT. Top, middle, and bottom panels shows photolabeling in the presence of 1 mM Mg^{2+}, 1 mM Ca^{2+}, and 2 mM EDTA, respectively, in a standard medium of 25 mM EPPS/TMAH, pH 8.5, 20% glycerol, 0.01 mg PfHPRT/ ml, and 0.01-30 mM [γ-^{32}P]TNP-8N$_3$-ATP. The concentrations of [γ^{32}P]TNP-8N$_3$-ATP and position of PfHPRT are shown. The spots to the right of each PAGE gel represent 1 μL of 1 μM [γ-^{32}P]TNP-8N$_3$-ATP (1 pmol) applied as square millimeter filter paper patchs after electrophoresis, and are used to quantify the extent of labelling

in the bands. Note that the relative banding intensity of the gels cannot be directly compared as it depends on the specific activity of the probe and the length of exposure to the screens. The particular preparation used was inactive, but could be activated to 1 μmol/min/mg of protein on incubation with substrates.

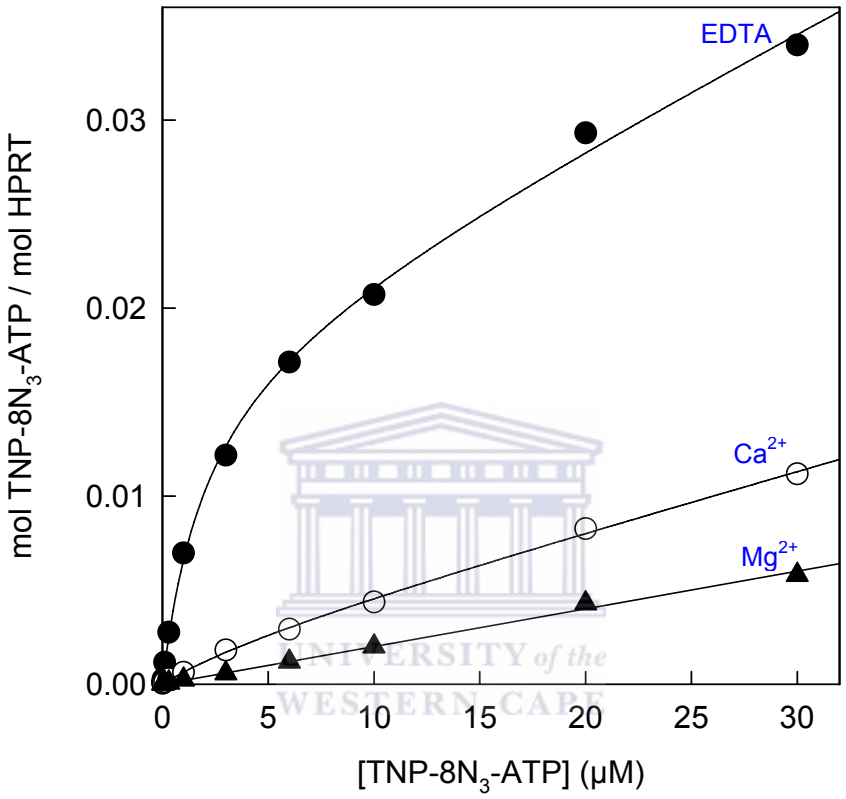


Figure 3.47: Photolabelling curves of PfHPRT with [γ-^{32}P]TNP-8N$_3$-ATP. The original gels from which the data is derived and experimental conditions are in Figure 3.46. Quantification was by exposure to radioactivity-sensitive screens for about 2 days, and analysis using a Cyclone Imager and associated software. The curves show the best fit to the data using a Hill equation with a linear component (McIntosh *et al*., 1996).

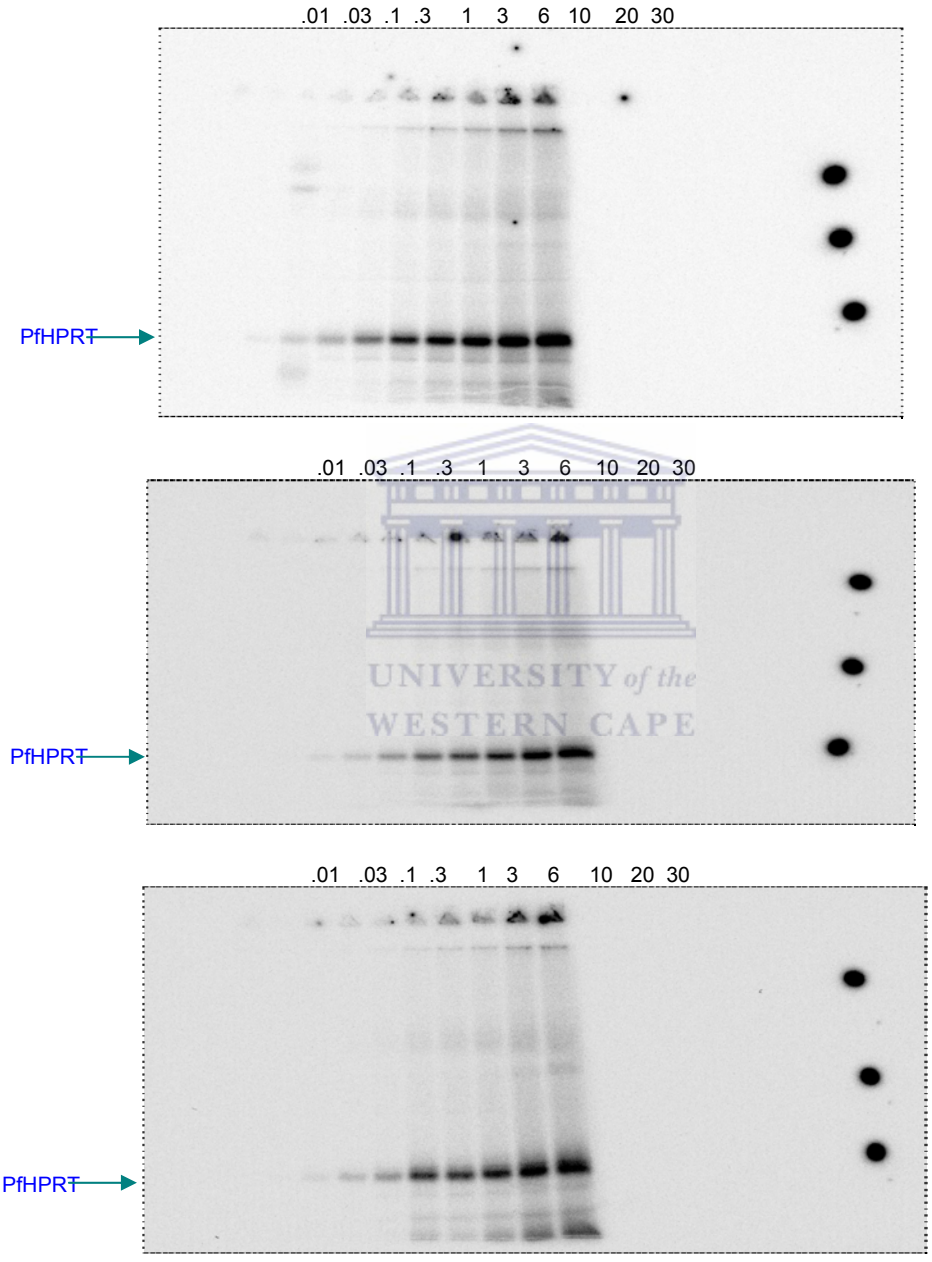


Figure 3.48: **Radioactivity imaging of SDS-PAGE (12%) analysis of [γ^{32}P]TNP-8N$_3$-ITP photolabeling of PfHPRT**. Top, middle, and bottom panels show photolabeling in the presence of 1 mM EDTA, 1 mM Mg^{2+}, and 2 mM Ca^{2+}, respectively, in a standard medium of 25 mM EPPS/TMAH, pH 8.5, 20% glycerol, 0.01 mg PfHPRT/ ml, and 0.01-30 mM [γ-^{32}P]TNP-8N$_3$-ITP. The concentrations of [γ^{32}P]TNP-8N$_3$-ITP (on top of each gel) and position of PfHPRT are shown. The spots to the right of each PAGE gel represent 1 μL of 1 μM [γ-^{32}P]TNP-8N$_3$-ATP (1 pmol)

applied as square millimeter filter paper patchs after electrophoresis, and are used to quantify the extent of labelling in the bands. Note that the relative banding intensity of the gels cannot be directly compared as it depends on the specific activity of the probe and the length of exposure to the screens. The particular preparation used was inactive, but could be activated to 1 μmol/min/mg of protein on overnight incubation with substrates.

Similar experiments using [γ^{32}P]TNP-8N_3-ITP as a photolabel are shown in Figures 3.48 and 3.49. In the absence of divalent cations (2 mM EDTA), the concentration dependence of photolabeling was similar to that obtained with the adenine nucleotide, but the amount of specific component was somewhat less ($K_{0.5}$ = 2.43 μM, and labeling maximum = 0.010 mol/mol). However, in the presence of Mg^{2+} and Ca^{2+} the inosine nucleotide exhibited specific components ($K_{0.5}$ = 4.61 and 2.85 μM respectively, and labeling maxima of 0.008 and 0.016 mol/mol, respectively), which contrasts with what was found for the adenine nucleotide. The affinity was somewhat higher and the derivatization more efficient in the presence of Ca^{2+}.

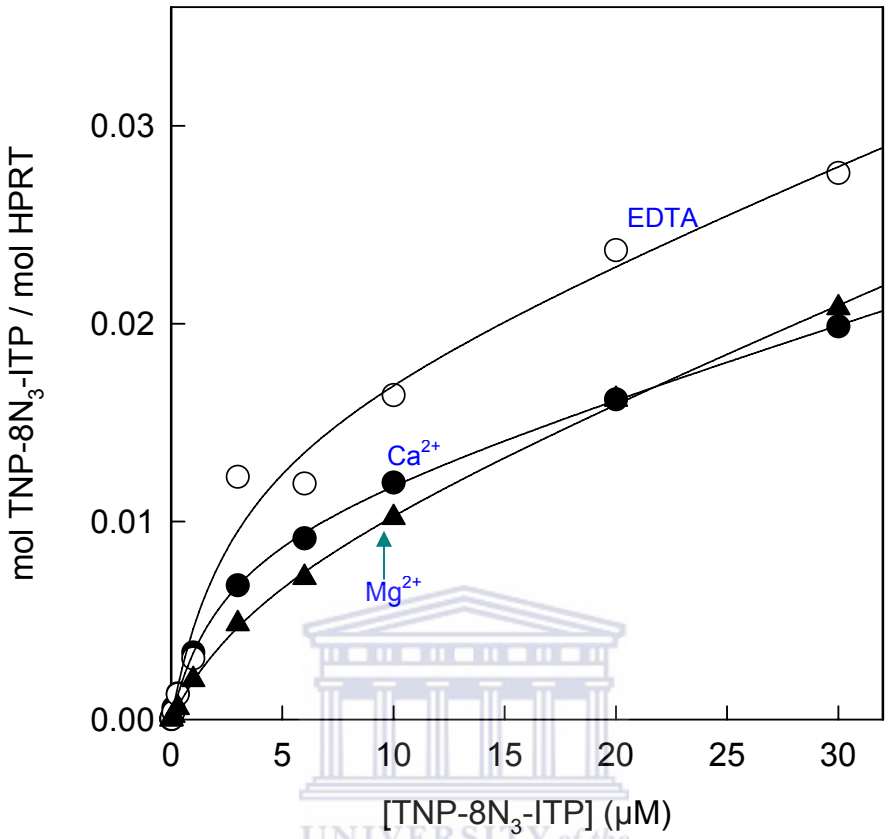


Figure 3.49: Photolabelling of PfHPRT with [γ-^{32}P]TNP-8N$_3$-ITP. The original gels from which the data is derived and experimental conditions are in Figure 3.48. Quantification was by exposure to radioactivity-sensitive screens for about 2 days, and analysis using a Cyclone Imager and OptiQuant software. The curves show the best fit to the data using a Hill equation with a linear component (McIntosh *et al*., 1996).

CHAPTER FOUR: DISCUSSION AND CONCLUSIONS

4.0 Introduction

The ultimate aim of this work was to obtain a crystal structure of PfHPRT in complex with a chalcone. To achieve this, we needed to over-express the enzyme in *E.coli* and purify as far as possible. Many crystallization trials were undertaken with partially pure enzyme. Alongside, we developed a model of the unliganded malarial enzyme based on

atomic structures of human, *Salmonella typhimurium*, and *Trypanosoma cruzi* enzymes. Methodology was also developed to determine the oligomeric status of our preparations of PfHPRT. Another aim was to develop photoaffinity labels of the active site, based on 2', 3' trinitrophenyl nucleotides, in order to determine whether the chalcone binds to the active site or not through competitive inhibition of the photolabeling.

4.1 Protein expression

Recombinant protein expression was achieved in *E.coli* BL21(DE3)pLysS cells transformed with pET-17b vector containing the PfHRPT insert. The cells were grown overnight at 37 °C to an optical density of OD_{600} of 0.4 – 0.6 at 600 nm and protein induction initiated by adding IPTG. Expression was examined by SDS-PAGE. Although PfHPRT is a soluble protein and supposed to be in the supernatant fraction after cell lysis and centrifugation, significant amounts were always present in the pellet fraction most likely as inclusion bodies. This was in spite of the rich growth media used, which slows inclusion body formation (Moore *et al*., 1993).

We tried to ameliorate this problem by lowering the temperature of the culture medium (34^{o}C) during the 3 h of induction with IPTG, but without noticeable effect. Inclusion of 10% glycerol during the entire growth period has also been found to be beneficial, but we did not obtain an improvement. It seems that primary instability of the protein, and an adverse folded/unfolded equilibrium, may be a major contributor to its deposition in inclusion bodies, rather than overproduction *per se*. The amount of soluble PfHPRT in the supernatant was however adequate for further processing.

4.2 Protein purification

Several methods exist for PfHPRT purification. For example, Shi *et al*. (1999) used a two-step procedure involving an anionic Q-Sepharose column followed by a Superdex G-75 gel filtration column to obtain a 95% pure enzyme which was subsequently crystallized and used in X-ray diffraction studies. Raman *et al*. (2005) started with ammonium sulphate precipitation of the clarified cell lysate followed sequentially by HPLC anionic-exchange chromatography on a Q-Sepharose column and cationic-exchange chromatography on a Resource S column. Keough *et al*. (1999) employed an anionic DEAE cellulose column (43 x 2.2 cm) followed by an Hg-Sepharose column (4.3 x 1.7 cm).

Our laboratory has developed a method based on a one-procedure two-step protocol of a DE-52 column (1.5 x 6 cm) placed in tandem with an agarose-linked Reactive Red 120 column (2 x 5 cm) (Mbewe *et al.,* 2006). The protein comes straight through the anion exchange column, but binds to the Reactive Red 120 column, and can be eluted with 50 mM pyrophosphate. PfHPRT has a calculated pI of 7.2 (taking the amino acids cysteine, aspartate, glutamate, histidine, lysine, arginine and tyrosine, and the amino- and carboxy-termini of the protein into account) (www.embl-heidelberg.de/cgi/pi-wrapper.pl), which

means that around this pH the protein will not bind to an anion exchange resin. Most proteins have a pK_a of between 5 and 6 and would bind well at neutral pH and higher. Experimentally, PfHPRT appears not to bind to a resin well- equilibrated at pH 8.0 (Figure 3.4) as well as at pH 8.9 (Figure 3.7), suggesting either that the actual pK_a is higher than 7.2 (surface residues are what count), or competing anions, like nucleotides (mainly derived from DNAse treatment of DNA), are loosening the interaction with the resin. Whatever the explanation, under the conditions used, PfHPRT binds very poorly to the anion exchange resin, which distinguishes it from most other proteins in the supernatant, and a significant purification is obtained. This behaviour is in agreement with Keough *et al*., 1999, but seems to disagree with Shi *et al*., 1999 and Raman *et al*., 2005, who needed salt to elute it from anion exchange resins. Reactive Red 120 is a large polysulphonated aromatic triazine dye (Figure 4.0) that mimics biological heterocyclic substrates such as nucleotides (Dean and Watson, 1979; Kaminska *et al*., 1999). It is ligated to agarose via a reactive chloro group. The polysulphonyl groups mean that it can also be considered a cation exchange resin.

However, Mbewe *et al*., 2006 considered that it behaved in the classical manner of an affinity column, due to the aromatic and sulphonated groupings mimicking GMP or IMP, PPi, or PRPP, mainly because PfHPRT was so effectively eluted with pyrophosphate, a product of the enzyme.

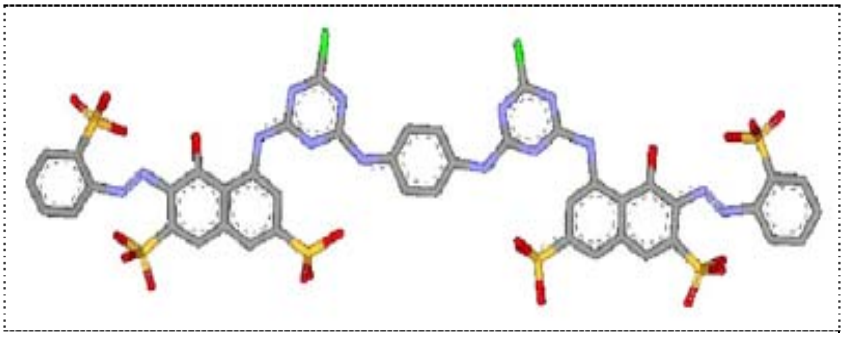


Figure 4.0: Structure of Reactive Red 120. The linkage to agarose is through the two chloro groups (coloured green). Sulpur atoms of the sulphonate groups are coloured orange while the oxygen atoms are coloured red.

In our attempt to reproduce this method, we found that some of the PfHPRT (~50%) seemed to bind tighter than usual and failed to elute with the pyrophosphate, but came off with the 1 M NaCl wash. The significance of a very tight binding fraction of PfHPRT will be addressed a bit later.

The behaviour of PfHPRT on the anion exchange resin was investigated in more detail by running the two columns separately and analysing multiple fractions for protein and pH. The majority of PfHPRT came straight through the short DE-52 column (2 x 6 cm) while the remainder was brought off by the 50 mM Tris-HCl, pH 8.0, 1 mM DDT, and 1 mM PMSF wash of the column. Compared with the flow-through fraction, the Tris-HCl wash

had more contaminants co-eluting with the malarial enzyme perhaps due to lowering of the column pH to the acidic range as the lysis buffer passed through.

The pH of the column in this case went up to 11, clearly indicating that the 10 mL of lysis buffer used for equilibration did not equilibrate the column. Proper equilibration of the

column at pH 8.9 (to induce PfHPRT binding) did not, however, result in any noticeable change in the behaviour of PfHPRT on this column.

On a longer DE-52 column (2 x 17.5 cm), washed with 30 mL of lysis buffer prior to loading of the supernatant (i.e. poorly equilibrated), PfHPRT was retarded on the column and was eluted in a very pure form by 50 mM Tris-HCl, pH 8.0, 1 mM DTT, and 1 mM PMSF. The pH of the column effluent for this preparation was extremely alkaline (up to 13) which could have aided PfHPRT interaction with the resin. While excellent purity was obtained with this approach, the resulting enzyme was incapable of being activated, most likely due to exposure to the extremely alkaline pH. When the column was well-equilibrated (pH 8.0) with the lysis buffer prior to applying the supernatant, PfHPRT did not bind or bound poorly and most came straight through the column. The purity was not as good as the poorly equilibrated long column, but nevertheless was quite pure.

It seems that the long DE-52 columns produced purer HPRT than the short columns. If one compares the concentrates (compare Figure 3.13 with those in Figure 3.1 and Figure 3.8), and bearing in mind that some of the small molecular weight species could be

proteolytic digestion products of HPRT, there appear to be relatively fewer high molecular weight contaminants with the long column. However, in our two attempts with the long column, we were not able to activate the enzyme with substrates, and the protein in both cases displayed negligible activity on assay. The reason for this lack of

activity (with the well equilibrated column) is not clear, but could be related to proteolysis.

The efficacy of the Reactive Red 120 column was not explored in detail, but its value in eliminating high molecular weight contaminants is clearly evident in Figure 3.8, where most came straight through the column, and PfHPRT was eluted in a much purer form compared with the material placed on the column. It has been shown before that PPi selectively elutes PfHPRT, whereas a salt gradient exhibits no selectivity (Mbewe *et al*., 2006), suggesting the protein interacts specifically with the resin (affinity mechanism rather than cation binding mechanism).

In some experiments, there appeared to be two populations of PfHPRT in relation to Reactive Red 120 chromatography, one that eluted with PPi and another that remained tightly bound and was eluted with 1 M NaCl. This was especially noticeable in the experiment using a poorly equilibrated, long DE-52 column, where the majority of the protein appeared in the tight binding fraction (Figure 3.11). Since this preparation had been exposed to extremely high pH, it suggests that the tight binding protein is a poorly folded form, and not desirable from the point of view of producing a homogeneous

preparation for crystallization. Thus, the Reactive Red 120 column appears to have two advantages, one in getting rid of contaminating proteins and second in eliminating an alternative structural state of PfHPRT.

One problem that was encountered during some of the purification attempts was the proteolytic degradation of PfHPRT. This could be seen in small molecular weight species in the SDS-PAGE analysis, and was proven by the co-elution of the proteolytic products with tetrameric PfHPRT on HPLC gel filtration (fraction collected and subjected to SDS-PAGE analysis).

It was apparent that the tetrameric complex consisted of full-length protein as well as nicked protein, and could explain, in some cases, why the enzyme could not be activated. Proteolysis was evident in the supernatant fraction and persisted throughout in the purification (see Figures 3.12 and 3.13). The first steps (lysis, DNAse digestion, and centrifugation) are rather lengthy and it is likely that the proteolysis is happening during this period. New, freshly prepared PMSF (which covalently modifies the active site of many serine proteases irreversibly) as well as a cocktail of other inhibitors in the lysis buffer did not help (Figure 3.12). In one experiment, the DNAse digestion was performed at 0 ^{o}C rather than the usual 25 ^{o}C, but it did not help either (results not shown). Other protease inhibitors could be explored to get around this problem.

Overall, from exploration of the elution properties of the DE-52 column, we found that a long, well-equilibrated column results in the majority of the PfHPRT coming straight through the column with better purity than with the short columns, well-equilibrated or not. This procedure seemed to eliminate the need for a large wash volume, which

generally contained several high molecular weight contaminants. We find the Reactive Red column to be very effective – likely eliminating both contaminants and poorly folded PfHPRT. HPLC gel filtration will be discussed in detail below, but it seems to offer promise to provide substantial further and final purification. Interestingly, recombinant PfHPRT used for obtaining crystals complexed with a transition-state inhibitor was purified by just ion exchange and gel filtration (Li *et al*., 1999; Shi *et al*., 1999).

4.3 Activity assay

We had difficulty in routinely obtaining active PfHPRT, even after incubation with partial substrates. It is known that PfHPRT is purified in an inactive form and requires activation by incubation with partial substrates, i.e. conditions under which there is no enzyme turnover (Keough *et al*., 1998: Keough *et al*., 1999; Sujay Subbayya *et al*., 2000; Mbewe, 2005; Mbewe *et al*., 2006). Thus the enzyme exists in at least two interconvertable forms. However, as mentioned, some preparations failed to activate, and there are evidently other species that cannot reach the activated state.

The failure of some of the protein to activate could arise from proteolytic degradation of PfHPRT, the presence of dimers, or poorly folded forms. Proteolytic cleavage is expected

to be mainly at the flexible loops, and may have affected the enzyme's ability to bind substrates. Dimer protein present in preparations after dialysis at low salt concentration appears unable to revert to tetramers. Exposure to extremely high pH results in poorly folded protein that cannot be activated.

4.4 HPLC gel filtration: PfHPRT oligomeric state analysis and further purification

Previous work has shown that that PfHPRT exists in its native, active state as a tetramer, and that high salt concentrations converts it into an inactive dimer (Keough *et al*, 1999). The possible existence of more than one oligomeric form of the enzyme could have a serious detrimental effect on crystallization. It was also of interest to find out the relationship between inactivity/activity and quaternary structure in our preparations. It was, therefore, desirable to have a method to measure the quaternary state of the enzyme. Keough *et al*. (1999) employed ultracentrifugation, but this is a lengthy procedure and unable to provide an accurate measure of the proportion of each species. We opted to try HPLC gel filtration instead. It has the advantage of speed and easy quantification.

However, we encountered difficulties in doing the analysis under low salt conditions, exactly the conditions we wanted to know the oligomeric status and the proportion of each species. Analysis of the behaviour of standard proteins on the particular HPLC gel filtration column we used, revealed that proteins migrated anomalously in low salt concentrations, probably mostly through inhibition of resin penetration through ionic

effects (i.e. most eluted earlier and together; cytochrome c was the exception and bound tightly to the resin).

Accordingly, we tested our preparations over a range of NaCl concentrations in the HPLC running buffer. There was a clear transition from a larger quaternary form to a smaller with an increase in salt concentration, with a midpoint at approximately 300 mM NaCl. The larger was attributed to the tetramer and the smaller to the dimer, even though, according to the standards, the molecular weight was about 35% lower, which suggests PfHPRT interacts with the resin in some way, rather similar to the behaviour of cytochrome c. The latter is strongly cationic, and PfHPRT exhibits similar properties in that it fails to bind to an anion exchange resin but does to a cation exchange resin (we found it bound to a cation exchange Sep-Pak), even though the overall pI is 7.2. On the other hand, the elution position of the two quaternary forms was quite insensitive to the salt concentration, the tetramer, for example, eluting at 8 min at 20 and 700 mM NaCl, with a fairly good separation of the two forms at these extremes.

Even though the analysis could not be accurately performed at concentrations below 20 mM salt, it was evident from the extrapolated curve (Figure 3.17) that the particular preparation that we analyzed consisted of 70 and 30% tetramer and dimer, respectively at zero salt concentration. This preparation had been exposed to 1 M NaCl during the purification process and then exhaustively dialyzed, and evidently had gone from being dimeric to these proportions during dialysis. The proportion is unlikely to change further

with time, as the dialysis time is already lengthy. The presence of a high proportion of the two species could inhibit crystallization.

We noticed that preparations that had been exposed to high pH through using a long, poorly equilibrated column consisted of mainly dimer, and that prolonged dialysis against 10 mM MOPS, pH 7.0, resulted in a slow, partial conversion to a tetrameric species. It was these preparations that had a high proportion of PfHPRT that bound tightly to the Reactive Red 120 column (Figure 3.11). Accordingly, our hypothesis is that the tight binding protein is the dimer, and that it is unable to be activated with substrates.

The dissociation of the tetramer into dimer as the ionic strength increases also indicates that the interacting forces between the dimers that form the tetramer are weaker than those between monomers forming the dimers, and that the former is dominated by salt-sensitive interactions (ionic), while the latter is held together by salt-independent interactions, likely hydrophobic interaction. The crystal structure indicates that the monomers interact to produce dimers via burial of a total surface area of 2812 $Å^2$, with hydrogen bonds between Lys38, Glu108, and Asp 211 in chain A and Glu100, Lys77, Arg80, and Asn95 from chain B respectively, and the involvement of many hydrophobic residues (Shi *et al*., 1999). The interaction of the dimers to form the tetramer involves a total surface area of 2070 $Å^2$, and must consist of ionic interactions, although this is not readily apparent from the PDB structure, as the interactions are not provided (dimers separated).

Our HPLC gel filtration experiments of a preparation eluted from Reactive Red 120 column with 1 M NaCl and concentrated showed that this procedure could effectively be used to purify PfHPRT as a tetramer from contaminants and other quaternary forms. In

the particular experiment shown, gel filtration failed to separate tetrameric PfHPRT from another protein of approximately 30 000 daltons (as judged by SDS-PAGE, Figure 3.20, and presumed to be trimer or tetramer in the absence of SDS). Normally PfHPRT would be eluted with PPi from the Reactive Red 120 column and this protein is eliminated in the 200 mM NaCl wash (Figure 3.6).

We show that it is possible to obtain reasonable amount of protein by repeated 200 μL injections of a concentrate of 0.3 mg/mL. The limitation is the volume rather than the concentration of protein. Preparations from DE-52 and Reactive Red 120 purification routinely produced approximately 5 mg. Thus, by concentrating such preparations to 2 mg/mL or so, it would be possible to purify the entire preparation in 4 h. Overall, one could complete an entire purification, from cells to pure protein in 24 h.

It is interesting that the preparations analyzed at low salt concentrations consist of predominantly the tetrameric species, and yet are not active, i.e. the preparation still needs to be activated by incubation with partial substrates. This means that the activation process is not the conversion from dimer to tetramer, but rather a conformational change of the tetramer, which may involve rearrangement of the flexible loops in each monomer

around the substrates. In fact, in one experiment, activation (incubation overnight with PRPP + hypoxanthine) resulted in a small, partial conversion to dimer. This agrees with the finding of Raman *et al*. (2005) that substrates actually destabilize the protein.

4.5 Crystallization

For the numerous crystallization trials performed, a myriad of results, including precipitates, phase separations, and salt crystals, were obtained but no protein crystals, as judged by the prick test, dye binding, and diffraction patterns. Several factors could have contributed to the failure to obtain PfHPRT crystals. Among these, the presence of a few contaminants in the enzyme preparation, particularly persistent high molecular impurities, could have played a major role. While the purity of a protein is a critical factor which influences the outcome of macromolecular crystallization experiments, our rationale for proceeding with crystallization attempts while the enzyme still contained contaminants was that crystallization is often used as a purification step to recover proteins from impure mixtures due to the ability of the crystal surfaces' to recognise cognate molecules for adsorption (Bergfors, 2003).

Another factor that may have disfavoured crystallization was the heterogenous oligomeric state of PfHPRT. Like purity, the existence of the protein to be crystallized in a singular structural species is an important factor. HPLC gel filtration experiments pointed to the existence of significant dimer among the tetramer species of PfHPRT at low salt concentration, which mitigates against formation of ordered crystals.

The apparent presence of proteolytic degradation products in several of our preparations would also hamper crystal formation. The most likely sites for attack by proteolytic enzymes are at the flexible loops that close over the substates, and nicks here could

prevent ligand binding and closure of the loops. A compact shape may be needed for crystallization.

Due to the instability of PfHPRT, it is also possible that the protein aggregated over time.

One of the main problems encountered during crystallization trials was the precipitation of chalcone, which was usually added to a final concentration of 1 mM (concentrate in DMSO) in order to saturate the ligand binding site (active site concentration 0.4-0.6 mM at 10-15 mg/ml). It may be possible to increase the DMSO concentration in the crystallization medium to lessen precipitation. Alternatively, it may be advantageous to dry the chalcone onto the cover slip and let the chalcone dissolve with time into the crystallization droplet.

4.6 Homology modeling

PfHPRT holoenzyme was modeled based on the structures of the human, *T. cruzi* and *S. typhimurium* HGPRTs using MODELLER (Sali and Blundell, 1993). Superposition of this open, unliganded structure of PfHPRT with the solved closed crystal structure of PfHPRT (PDB code 1CJB; Shi *et al*., 1999) revealed a good overall agreement of

secondary structural features as reflected by the low RMSD values. A major difference was the less extensive core β-sheet attributed to the opening and flexibility of the catalytic loop. The terminal β-strand was thus separated from the sheet and the

penultimate β-strand shortened by four residues. The other significant difference was the opening of all the loops that close over the substrates (Figure 4.1). The side chain conformations of most residues lining the active site remained unaltered. However, the opening of the loops caused several side chains to be displaced.

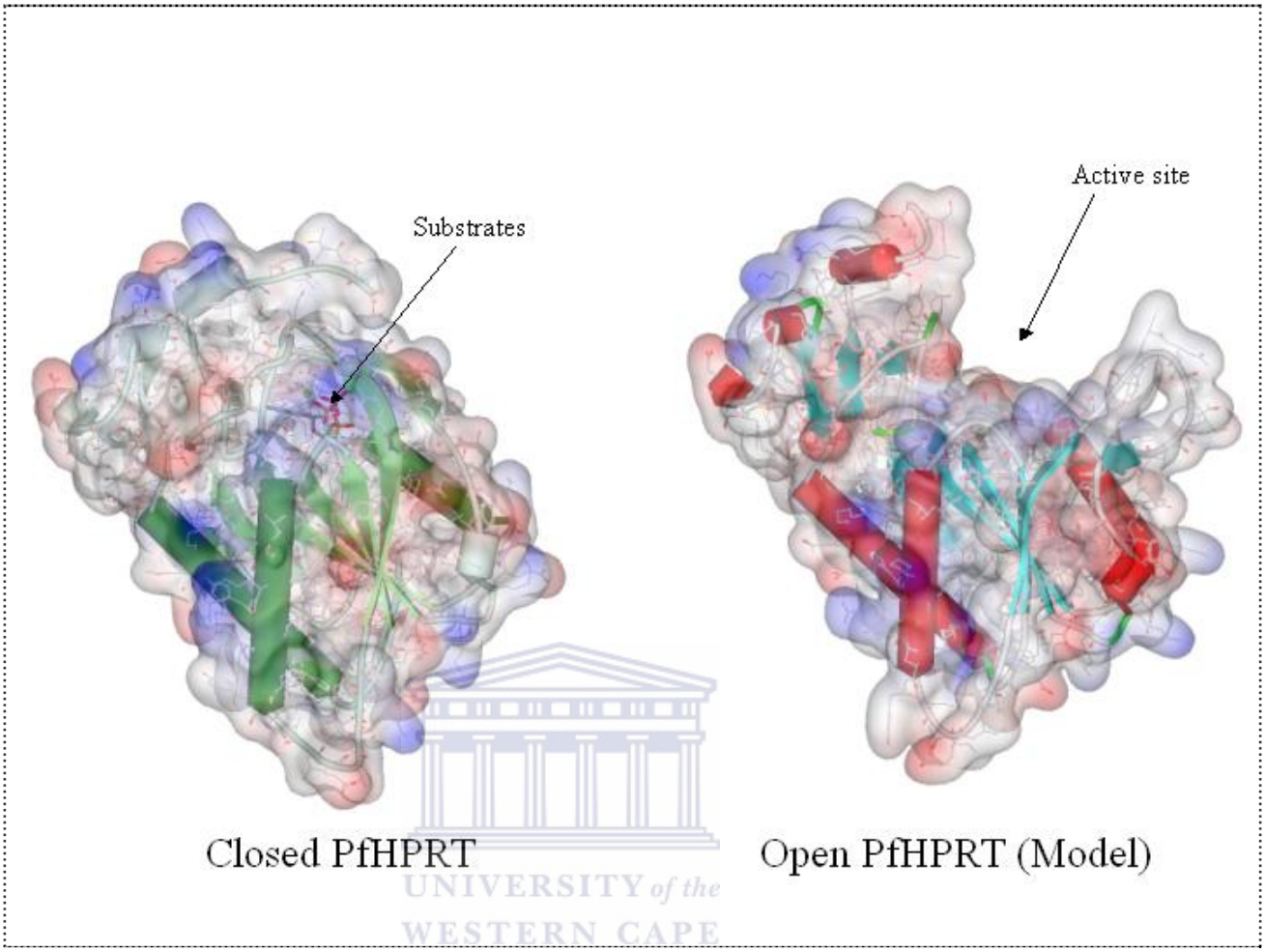


Figure 4.1: **A semitransparent surface representation showing a comparison of the closed atomic structure of PfHPRT with bound ImmHP, PPi and two Mg^{2+} ions at the active site with a model of the open unliganded form of PfHPRT. Blue and red indicate basic and acidic charges, respectively.**

The homology model emphasized the large size of the open active site cavity. This cavity, 2913.8 $Å^3$ in size, is large enough to accommodate sizeable ligands, even larger than the substrates. Note the loop movements in the open structure (Figure 4.2).

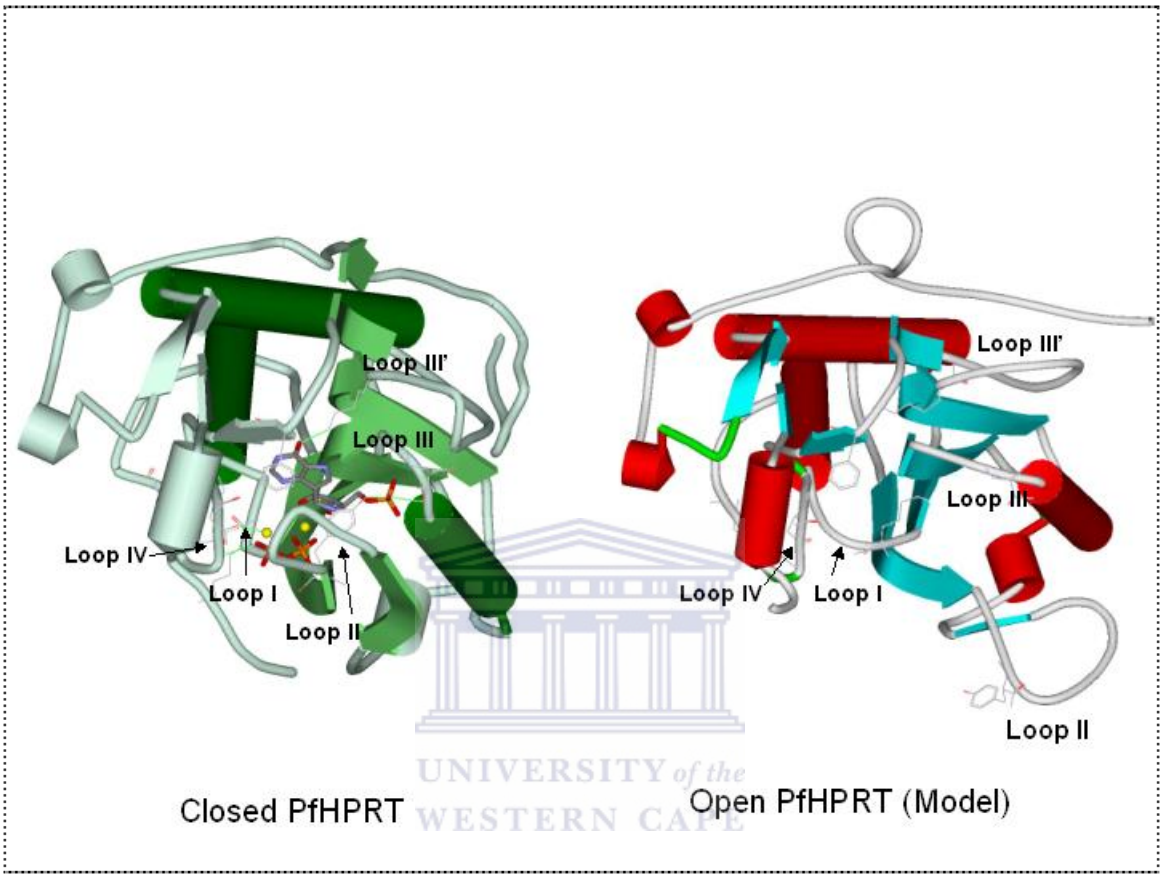


Figure 4.2: Comparison of closed and open PfHPRT (model). This top view shows some key substrate binding residues in stick form.

The value in having an open PfHPRT structure, albeit a model, is being able to dock ligands in an appropriate structure, i.e. one that is likely to bind a competitive inhibitor.

4.7 Photoaffinity labeling

TNP nucleotides have been widely used as probes of nucleotide binding sites because of their fluorescent properties and also on account of their generally higher affinity toward these sites (Ward and Cavieres, 1998). The 8-azido derivatives were introduced by our laboratory as photoaffinity probes. Photoaffinity probes are useful because they can identify residues at the active site and be used to investigate whether a ligand of interest binds to this site.

In this study, we developed a method for synthesizing [γ-^{32}P]TNP-8N_3-ITP (2',3'-O-(2,4,6-trinitrophenyl)-8-azidoinosine triphosphate) and explored its efficacy as a photolabel of the IMP site of PfHPRT. We compared its properties to that of [γ-^{32}P]TNP-8N_3-ATP (2',3'-O-(2,4,6-trinitrophenyl)-8-azidoadenosine triphosphate).

The synthesis entailed directly converting the exocyclic amino group TNP-8N3-ATP to an oxo group. The reaction in nitrous acid was mild and quantitative enough to consider modification of highly radiolabeled [γ-^{32}P]TNP-8N_3-ATP and use the product directly following simple purification through a C_{18} Sep-Pak cartridge.

Labeling PfHPRT with [γ-^{32}P]TNP-8N_3-ITP and [γ-^{32}P]TNP-8N_3-ATP was rather inefficient – derivatization was only to the extent of 0.005 to 0.01 mol probe/mol protein. This should be compared with [γ-^{32}P]TNP-8N_3-ATP photolabeling of Ca^{2+}-ATPase, which proceeds to close to 1 mol probe/mol protein, because of the derivatization of a

lysine (K492) – a powerful nucleophile (McIntosh *et al*., 1992). Presumably there is no nucleophile in the vicinity of the azido group in PfHPRT.

The two probes had rather different properties in their interaction with PfHPRT. [γ-^{32}P]TNP-8N_3-ATP bound and photolabeled with high affinity in the absence of divalent cations, whereas no specific labeling could be detected in the presence of Mg^{2+} or Ca^{2+}. In contrast, [^{32}P]TNP-8N_3-ITP bound and photolabeled with high affinity in the absence and in the presence of divalent cations. The poor photolabeling efficiency meant that there was a rather high background labeling relative to the specific derivatization.

It is difficult to understand the different effects of divalent cations with the two nucleotides, but somehow the Mg- or Ca-complex with the inosine nucleotide makes for better interaction with the protein than does the adenine species. One rather remote possibility is that it is known that a large substituent at the 8-position of adenine nucleotides, such as a bromo or azido group causes the nucleotide to adopt a *syn* configuration rather than the more usual *anti* configuration about the glycosidic bond, and this might not be the case for the inosine nucleotide, and interaction of divalent cations with the phosphates could cause other perturbations which alter the *syn/anti* equilibrium.

It may be argued that the triphospho-nucleotides are not good analogues of IMP. However, the structures indicate that in the closed structure and with the flexible catalytic

loop II pulled back, there should be nothing preventing the extra phosphates from sticking out into the medium or making favorable interactions with residues in the vicinity (Figure 4.3, top). In the open model structure, the phosphates are even more free to extend into the medium (Figure 4.3, bottom).

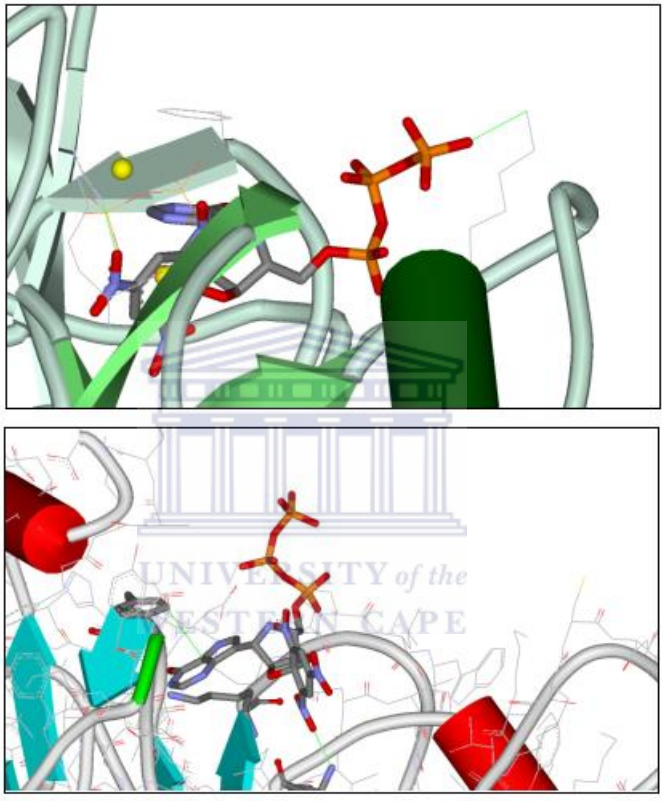


Figure 4.3: Docking of TNP-ImmHP into closed PfHPRT structure with flexible loop II removed (top frame) and of TNP-ATP into open PfHPRT (model, bottom frame). Some favorable interactions are shown with green lines.

4.8 Future directions

This study has suggested that a good future purification strategy will be to use a long, properly equilibrated DE-52 column (2 x 17.5 cm) as the first step of the purification. Most of the enzyme will come off the column in the flow-through fraction. This will avoid contaminants that are dislodged from the column by the Tris-HCl wash. Reactive Red 120 column chromatography will be retained as the second step, which will eliminate most high molecular weight contaminants co-eluting with PfHPRT from the DE-52 column and a tight binding PfHPRT fraction, which cannot be activated. HPLC gel filtration will be employed as the final step to ensure a homogeneous tetrameric preparation and to remove remaining contaminants. We are hopeful that such a preparation will lead to protein crystals.

A His-tagged PfHPRT has been expressed previously in this laboratory, and it purifies to homogeneity using nickel-affinity chromatography (Phehane, 2002). The histidine tag can be cleaved off with the retention of 3 extra amino acid residues. We would like to perform crystallization trials on such a protein as well.

The chalcones which we have been identified as activators, accelerators, and inhibitors of PfHPRT activity under various conditions are unique and only crystal structures of the

chalcone-PfHPRT complexes, under the different conditions, will provide an explanation for how these ligands perform such a diversity of functions in a single protein.

BIBLIOGRAPHY

Allen, T.E., Ullman, B. (1993) Cloning and expression of the hypoxanthine-guanine phosphoribosyltransferase gene from *Trypanosoma brucei*. *Nucleic Acids Res*. **21**: 5431-8.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**: 403-10.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**: 3389-402.

Arav-Boger, R., Shapiro, T.A. (2005) Molecular mechanisms of resistance in antimalarial chemotherapy: the unmet challenge. *Annu. Rev. Pharmacol. Toxicol.* **45**: 565-85.

Asherie, N. (2004) Protein crystallization and phase diagrams. *Methods* **34**: 266-72.

Baker, D., Sali, A. (2001) Protein structure prediction and structural genomics. *Science* **294**: 93-6.

Bayley, H., Knowles, J.R. (1977) Photoaffinity labeling. *Methods Enzymol.* **46**: 69-114.

Bayley, H.; Staros, J. V. (1984). In *Azides and Nitrenes - Reactivity and Utility*; Scriven, E. F. V. Ed., Academic Press Inc.(New York), Chapter 9.

Berendt, A.R., Simmons, D.L., Tansey, J., Newbold, C.I., Marsh, K. (1989) Intercellular adhesion molecule-1 is an endothelial cell adhesion receptor for *Plasmodium falciparum*. *Nature* **341**: 57-9.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E. (2001) The Protein Data Bank. *Nucleic Acids Res*. **28:** 235-42.

Berman, J.D., Nielsen, R., Chulay, J.D., Dowler, M., Kain, K.C., Kester, K.E., Williams, J., Whelen, A.C., Shmuklarsky, M.J. (2001) Causal prophylactic efficacy of atovaquone-proguanil (Malarone) in a human challenge model. *Trans. R. Soc. Trop. Med. Hyg.* **95**: 429-32.

Blundell, T.L., Patel, S. (2004) High-throughput x-ray crystallography for drug design. *Curr. Opin. Pharmacol*. **4**: 490-6.

Blundell, T.L., Sibanda, B.L., Sternberg, M.J., Thornton, J.M. (1987) Knowledge-based prediction of protein structures and the design of novel molecules. *Nature* **326**: 347-52.

Bonneau, R., Strauss, C.E., Rohl, C.A., Chivian, D., Bradley, P., Malmstrom, L., Robertson, T., Baker, D. (2002) De novo prediction of three-dimensional structures for major protein families. *J Mol. Biol.* **322**: 65-78.

Bourne, P.E., Weissig, H. (2003) Structural Bioinformatics. Wiley-Liss, Inc., Hoboken, New Jersey.

Bower, M.J., Cohen, F.E., Dunbrack, R.L. (1997) Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modelling tool. *J. Mol. Biol.* **267**: 1268-82.

Bowie, J.U., Clarke, N.D., Pabo, C.O., Sauer, R.T. (1990) Identification of protein folds: matching hydrophobicity patterns of sequence sets with solvent accessibility patterns of known structures. *Proteins* **7**: 257-64.

Bray, P.G., Deed, S., Fox, E., Kalkanidis, M., Mungthin, M., Deady, L.W., Tilley, L. (2005) Primaquine synergises the activity of chloroquine against chloroquine-resistant *P. falciparum*. *Biochem. Pharmacol.* **70**: 1158-66.

Bray, P.G., Hawley, S.R., Ward, S.A. (1996) 4-Aminoquinoline resistance of *Plasmodium falciparum*: insights from the study of amodiaquine uptake. *Mol. Pharmacol.* **50**: 1551-8.

Breman, J.G. (2001) The ears of the hippopotamus: manifestations, determinants and estimates of the malaria burden. *Am. J. Trop. Med. Hyg.* **64**: 1-11.

Breman, J.G., Alilio, M.S., Mills, A. (2004) Conquering the intolerable burden of malaria: what's new, what's needed: a summary. *Am. J. Trop. Med. Hyg.* **71**: 1-15.

Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., Karplus, M. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**: 187.

Browne, W.J., North, A.C., Phillips, D.C., Brew, K., Vanaman, T.C., Hill, R.L. (1969) A possible three-dimensional structure of bovine alpha-lactalbumin based on that of hen's egg-white lysozyme. *J. Mol. Biol.* **42**: 65-86.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., Warren, G.L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Cryst. D.* **54**: 905-21.

Brunner, J. (1993). New photolabeling and crosslinking methods. *Annu. Rev. Biochem.* **62:**483-514.

Canutescu, A.A., Shelenkov, A.A., Dunbrack, R.L. (2003) A graph-theory algorithm for rapid protein side-chain prediction. *Protein Sci.* **12**: 2001-14.

Canyuk, B., Medrano, F.J., Wenck, M.A., Focia, P.J., Eakin, A.E., Craig, S.P. (2004) Interactions at the dimer interface influence the relative efficiencies for purine nucleotide synthesis and pyrophosphorolysis in a phosphoribosyltransferase. *J.Mol.Biol.* **335**: 905-21.

Chakrabarti, S., John, J., Sowdhamini, R. (2003) Improvement of comparative modeling by the application of conserved motifs amongst distantly related proteins as additional restraints. *J. Mol. Model.* **10**: 69-75.

Chong, C.R., Sullivan, D.J. (2003) Inhibition of heme crystal growth by antimalarials and other compounds: implications for drug discovery. *Biochem. Pharmacol*. **66**: 2201-12.

Chothia, C., Lesk, A.M. (1986) The relation between the divergence of sequence and structure in proteins. *EMBO J.* **5**: 823-6.

Chothia, C., Lesk, A.M. (1987) Canonical structures for the hypervariable regions of immunoglobulins*. J. Mol. Biol.* **196**: 901-17.

Clark, I.A., Alleva, L.M., Mills, A.C., Cowden, W.B. (2004) Pathogenesis of malaria and clinically similar conditions. *Clin. Microbiol. Rev.* **17**: 509-39.

Clausen, J.D., McIntosh, D.B., Vilsen, B., Woolley, D.G., Andersen, J.P. (2006) Importance of conserved N-domain residues Thr441, Glu442, Lys515, Arg560, and Leu562 of sarcoplasmic reticulum Ca2+-ATPase for MgATP binding and subsequent catalytic steps. Plasticity of the nucleotide-binding site. *J. Biol. Chem*. **278**: 20245-58.

Congreve, M., Murray, C.W., Blundell, T.L. (2005) Structural biology and drug discovery. *Drug Discov. Today*. **10**: 895-907.

Connolly, M. L. (1983) Solvent-accessible surfaces of proteins and nucleic acids. *Science* **221**: 709-13.

Cowman, A.F. (1991) The P-glycoprotein homologues of Plasmodium falciparum: Are they involved in chloroquine resistance? *Parasitol. Today*. **7**: 70-6.

Craig, S.P., Eakin, A.E. (2000) Purine phosphoribosyltransferases. *J. Biol Chem.* **275**: 20231-4.

Crehuet, R., Thomas, A., Field, M.J. (2005) An implementation of the nudged elastic band algorithm and application to the reaction mechanism of HGXPRTase from *Plasmodium falciparum*. *J. Mol. Graph. Model.* **24**: 102-10.

Dayhoff, M. O., Schwartz, R.M., Orcutt, B.C. (1978) A model of evolutionary change in proteins. *Atlas of Protein Sequence and Structure 5*, (National Biomedical Research Foundation, Washington, D.C) Dayhoff, M.O., ed. pp. 345–352.

Dean, P. D., Watson, D. H. (1979) Protein purification using immobilised triazine dyes. *J. Chromatogr*. **165**: 301-319.

Deanne, C.M., Blundell, T.L. (2000) A novel exhaustive search algorithm for predicting the conformation of polypeptide segments in proteins. *Proteins* **40**: 135-44.

Deanne, C.M., Blundell, T.L. (2001) CODA: a combined algorithm for predicting the structurally variable regions of protein models. *Protein Sci.* **10**: 599-612.

DeTitta, G.T., Luft, J.R. (1995) Rate of water equilibration in vapor-diffusion crystallization: dependence on the residual pressure of air in the vapor space. *Acta Crysta. D.* **51**: 786-91.

De Wet, H., McIntosh, D.B., Conseil, G., Baubichon-Cortay, H., Krell, J., Jault, J.M., Daskiewicz, J.B., Barron, D., DiPietro, A. (2001) Sequence requirements of the ATP-binding site within the C-terminal nucleotide-binding domain of mouse P-glycoprotein: structure-activity relationships for flavonoid binding. *Biochemistry* **40**: 10382-91.

Dorman, G., Prestwich, G.D. (2000) Using photolabile ligands in drug discovery and development. *Trends Biotechnol.* **18**: 64-77.

Dorsey, G, Kamya, M.R., Singh, A., Rosenthal, P.J. (2001) Polymorphisms in the *Plasmodium falciparum* pfcrt and pfmdr-1 genes and clinical response to chloroquine in Kampala, Uganda. *J. Infect. Dis*. **183**: 1417-20.

Dunbrack, J.R. (2006) Sequence comparison and protein structure prediction. *Curr. Opin. Struct. Biol.* **16**: 374-84.

Eads, J.C., Scapin, G., Xu, Y., Grubmeyer, C., Sacchettini, J.C. (1996) The crystal structure of human hypoxanthine-guanine phosphoribosyltransferase with bound GMP. *Cell* **78:** 325-34.

Eckstein-Ludwig, U., Webb, R,J., Van Goethem, I.D., East, J.M., Lee, A.G., Kimura, M., O'Neill, P.M., Bray, P.G., Ward, S.A., Krishna, S. (2003) Artemisinins target the SERCA of *Plasmodium falciparum*. *Nature* **424**: 957-61.

El Kouni, M.H. (2003) Potential chemotherapeutic targets in the purine metabolism of parasites. *Pharmacol. Ther.* **99**: 283-309.

Famin, O., and Ginsburg, H. (2002) Differential effects of 4-aminoquinoline-containing antimalarial drugs on hemoglobin digestion in *Plasmodium falciparum*-infected erythrocytes. *Biochem. Pharmacol*. **63**: 393-8.

Famin, O., Krugliak, M., Ginsburg, H. (1999) Kinetics of inhibition of glutathione-mediated degradation of ferriprotoporphyrin IX by antimalarial drugs. *Biochem. Pharmacol*. **58**: 59-68.

Fedorov A, Shi W, Kicska G, Fedorov E, Tyler PC, Furneaux RH, Hanson JC, Gainsford GJ, Larese JZ, Schramm VL, Almo SC. (2001) Transition state structure of purine nucleoside phosphorylase and principles of atomic motion in enzymatic catalysis. *Biochemistry* **40**: 853-60.

Feng, D.F., Doolittle, R.F. (1987) Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.* **25**: 351-60.

Fernandez-Fuentes, N., Zhai, J., Fiser, A. (2006) ArchPRED: a template based loop structure prediction server. *Nucleic Acids Res.* **34**: 173-6.

Fidock, D.A., Nomura, T., Talley, A.K., Cooper, R.A., Dzekunov, S.M., Ferdig, M.T., Ursos, L.M., Sidhu, A.B., Naude, B., Deitsch, K.W., Su, X.Z., Wootton, J.C., Roepe, P.D., Wellems, T.E. (2000) Mutations in the *P. falciparum* digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. *Mol. Cell* **6**: 861-71.

Fischer, D., Eisenberg, D. (2003) Predicting structures for genome proteins. *Curr. Opin. Struct. Biol.* **9**: 208-11.

Fiser, A. (2006) Protein structure modeling in the proteomics era. *Expert Rev. Proteomics* **1**: 97-110.

Fiser, A., Feig, M., Brooks, C.L., Sali, A. (2002) Evolution and physics in comparative protein structure modeling. *Acc. Chem. Res.* **35**: 413-21.

Focia, P.J., Craig, S.P., Eakin, A.E. (1998) Approaching the transition state in the crystal structure of a phosphoribosyltransferase. *Biochemistry* **37**: 17120-7.

Focia, P.J., Craig, S.P., Nieves-Alicea, R., Fletterick, R.J., Eakin, A.E. (1998) A 1.4 Å crystal structure for the hypoxanthine phosphoribosyltransferase of *Trypanosoma cruzi*. *Biochemistry* **37**: 15066-75.

Freymann, D.M., Wenck, M.A., Engel, J.C., Feng, J., Focia, P.J., Eakin, A.E., Craig, S.P. (2000) Efficient identification of inhibitors targeting the closed active site conformation of the HPRT from *Trypanosoma cruzi*. *Chem. Biol.* **7**: 957-68.

Fry, M., Pudney, M. (1992) Site of action of the antimalarial hydroxynaphthoquinone, 2-[trans-4-(4'-chlorophenyl) cyclohexyl]-3-hydroxy-1,4-naphthoquinone (566C80). *Biochem. Pharmacol.* **43**: 1545-53.

Gallup, J.L., Sachs, J.D. (2001) The economic burden of malaria. *Am. J. Trop. Med. Hyg.* **64**: 85-96.

Garg, R.K. (2000) Cerebral malaria. *J. Assoc. Physicians India* **48**: 1004-13.

Gaur, D., Mayer, D.C., Miller, L.H. (2004) Parasite ligand-host receptor interactions during invasion of erythrocytes by *Plasmodium* merozoites. *Int. J. Parasitol.* **34**: 1413-29.

Ginalski, K. (2006) Comparative modeling for protein structure prediction. *Curr. Opin. Struct. Biol.* **16**: 172-7.

Ginsburg, H., Krugliak, M. (1999) Chloroquine – some open questions on its antimalarial mode of action and resistance. *Drug Resist. Updat.* **2**: 180-7.

Go, M., Go, N., Scheraga, H.A. (1970) Molecular theory of the helix-coli transition in polyamino acids. II. Numerical evaluation of s and sigma for polyglycine and poly-L-alaine in the absence (for s and sigma) and presence (for sigma) of solvent. *J. Chem. Phys.* **52**: 2060-79.

Greenwood, B., Mutabingwa, T. (2002) Malaria in 2002. *Nature* **415**: 670-72.

Greer, J. (1990) Comparative modeling methods: application to the family of the mammalian serine proteases. *Proteins* **7**: 317-34.

Gregson, A., Plowe, C.V. (2005) Mechanisms of resistance of malaria parasites to antifolates. *Pharmacol. Rev*. **57**: 117-45.

Guex, N., Peitsch, M.C. (2000) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**: 2714-23.

Hanna, M. M., Bentsen, L., Lucido, M., and Sapre, A. (1999) RNA-protein cross-linking with photoreactive nucleotide analog. *Methods Mol. Biol.* **118**: 21-33.

Haque, T.S., Skillman, A.G., Lee, C.E., Habashita, H., Gluzman, I.Y., Ewing, T.J., Goldberg, D.E., Kuntz, I.D., Ellman, J.A. (1999) Potent, low-molecular-weight non-peptide inhibitors of malarial aspartyl protease plasmepsin II. *J. Med. Chem.* **42**: 1428-40.

Hardy, L.W., Malikayil, A. (2003) The impact of structure-guided drug design on clinical agents. *Curr. Drug Discov*. **4**: 15–20.

Hastings, I.M., Bray, P.G., Ward, S.A. (2002) A requiem for chloroquine. *Science* **298**: 74-5.

Hendlich, M., Lackner, P., Weitckus, S., Floeckner, H., Froschauer, R., Gottsbacher, K., Casari, G., Sippl, M.J. (1990) Identification of native protein folds amongst a large number of incorrect models. The calculation of low energy conformations from potentials of mean force. *J. Mol. Biol.* **216**: 167-80.

Henikoff, S., Henikoff, J.G. (1993) Performance evaluation of amino acid substitution matrices. *Proteins* **17**: 49-61.

Henry, M., Alibert., S., Orlandi-Pradines, E., Bogreau, H., Fusai, T., Rogier, C., Barbe, J., Pradines, B. (2006) Chloroquine resistance reversal agents as promising antimalarial drugs. *Curr. Drug Targets* **7**: 935-48.

Heroux, A., White, E.L., Ross, L.J., Davis, R.L., Borhani, D.W. (1999) Crystal structure of *Toxoplasma gondii* hypoxanthine-guanine phosphoribosyltransferase with XMP, pyrophosphate, and two Mg(2+) ions bound: insights into the catalytic mechanism. *Biochemistry* **38**: 14495-506.

Heroux, A., White, E.L., Ross, L.J., Borhani, D.W. (1999) Crystal structures of the *Toxoplasma gondii* hypoxanthine-guanine phosphoribosyltransferase-GMP and -IMP complexes: comparison of purine binding interactions with the XMP complex. *Biochemistry* **38**: 14485-94.

Hohenester, E., Maurer, P., Hohenadl, C., Timpl, R., Jansonius, J.N., Engel, J. (1996) Structure of a novel extracellular Ca(2+)-binding module in BM-40. *Nat. Struct. Biol.* **3**: 67-73.

Holm, L., Sander, C. (1998) Dictionary of recurrent domains in protein structures. *Proteins* **33**: 88-96.

Huy, N.T., Kamei, K., Kondo, Y., Serada, S., Kanaori, K., Takano, R., Tajima, K., Hara, S. (2002) Effect of antifungal azoles on the heme detoxification system of malarial parasite. *J. Biochem*. **131**: 437-44.

Ismaili, J., Van der Sande, M., Holland, M.J., Sambou, I., Keita, S., Allsopp, C., Ota, M.O., McAdam, K.P.W.J., Pinder, M. (2003) *Plasmodium falciparum* infection of the placenta affects newborn immune responses. *Clin. Exp. Immunol.* **133**: 414-21.

Jones, D.T. (1994) De novo protein design using pairwise potentials and a genetic algorithm. *Protein Sci.* **3**: 567-74.

Jones, D.T. (1999) GenTHREADER: An efficient and reliable protein fold recognition method for genomic sequences. *J. Mol. Biol.* **287**: 797-815.

Jones, D.T. (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292**: 195-202.

Jones, D.T., Taylor, W.R., Thornton, J.M. (1992) A new approach to protein fold recognition. *Nature* **358**: 86-9.

Jones, T.A., Thirup, S. (1986) Using known substructures in protein model building and crystallography. *EMBO J.* **5**: 819-22.

Khan, S.M., and Walters, A.P. (2004) Malaria parasite transmission stages: an update. *Trends Parasitol.* **20**: 575-80.

Keana, J.F.W., Cai, S.X. (1990) New reagents for photoaffinity labeling: synthesis and photolysis of functionalized perfluorophenyl azides. *J. Org. Chem.* **55**: 3640-47.

Kelley, L.A., MacCallum, R.M., Sternberg, M.J. (2003) Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J. Mol. Biol.* **299**: 499-520.

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

Kirk K. (2004) Channels and transporters as drug targets in the *Plasmodium*-infected erythrocyte. *Acta Trop.* **89**: 285-98.

Kline, P.C., Schramm, V.L. (1995) Pre-steady-state transition-state analysis of the hydrolytic reaction catalyzed by purine nucleoside phosphorylase. *Biochemistry* **34**: 1153-62.

Koehl, P., Delarue, M. (1995) A self consistent mean field approach to simultaneous gap closure and side-chain positioning in homology modelling. *Nat. Struct. Biol.* **2**: 163-70.

Koehl, P., Levitt, M. (1999) Structure-based conformational preferences of amino acids. *PNAS* **96**: 12524-9.

Kohler, S., Delwiche, C.F., Denny, P.W., Tilney, L.G., Webster, P., Wilson, R.J., Palmer, J.D., Roos, D.S. (1997) A plastid of probable green algal origin in Apicomplexan parasites. *Science* **275**: 1485-9.

Koppisch, A.T., Fox, D.T., Blagg, B.S., Poulter, C.D. (2002) *E. coli* MEP synthase: steady-state kinetic analysis and substrate binding. *Biochemistry* **41**: 236-43.

Korsinczky, M., Chen, N., Kotecka, B., Saul, A., Rieckmann, K., Cheng, Q. (2000) Mutations in *Plasmodium falciparum* cytochrome b that are associated with atovaquone resistance are located at a putative drug-binding site. *Antimicrob. Agents Chemother.* **44**: 2100-8.

Kuhn, P., Wilson, K., Patch, M.G., Stevens, R.C. (2002) The genesis of high-throughput structure-based drug discovery using protein crystallography. *Curr. Opin. Chem. Biol.* **6**: 704-10.

Kurreck, J. (2003) Antisense technologies: Improvement through novel chemical modifications. *Eur. J. Biochem.* **270**: 1628–44.

Kuznestov, Y.G., Larson, S.B., Day, J., Greenwood, A., McPherson, A. (2001) Structural transitions of satellite tobacco mosaic virus particles. *Virology* **284**: 223-34.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**: 283-91.

Lattman, E.E., Rose, G.D. (1993) Protein folding-what's the question? *PNAS* **90**: 439-41.

Leavitt, S., Freire, E. (2001) Direct measurement of protein binding energetics by isothermal titration calorimetry. *Curr. Opin. Struct. Biol.* **11**: 560-66.

Lee, C.C., Craig, S.P., Eakin, A.E. (1998) A single amino acid substitution in the human and a bacterial hypoxanthine phosphoribosyltransferase modulates specificity for the binding of guanine. *Biochemistry* **37**: 3491-8.

Lesk, A.M., Chothia, C. (1980) How different amino acid sequences determine similar protein structures: the structure and evolutionary dynamics of the globins. *J. Mol. Biol.* **136**: 225-70.

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

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., Kenyon, G.L. (1996) Structure-based design of parasitic protease inhibitors. *Bioorg. Med. Chem*.**4**: 1421-7.

Li, Y.Z., Kirby, J.P., George, M.W., Poliakoff, M., Schuster, G.B. (1988) 1,2-Didehydroazepines from the photolysis of substituted aryl azides: analysis of their chemical and physical properties by time-resolved spectroscopic methods. *J. Am. Chem. Soc.* **110**: 8092-8.

Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J. (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*. **23:** 3-25.

Lipman, D.J., Altschul, S.F., Kececioglu, J.D. (1989) A tool for multiple sequence alignment. *PNAS* **86**: 4412-5.

Loria, P., Miller, S., Foley, M., Tilley, L. (1999) Inhibition of the peroxidative degradation of haem as the basis of action of chloroquine and other quinoline antimalarials. *Biochem. J.* **339**: 363-370.

Lovell, S.C., Word, J.M., Richardson, J.S., Richardson, D.C. (2000) The penultimate rotamer library. *Proteins* **40**: 389-408.

Maitland, K., Makanga, M., Williams, T.N. (2004) *Falciparum* malaria: current therapeutic challenges. *Curr. Opin. Infect. Dis*. **17**: 405-12.

Maitland, K., Marsh, K. (2004) Pathophysiology of severe malaria in children. *Acta Trop.* **90**: 131-40.

Marti-Renom, M.A., Stuart, A.C., Fiser, A., Sanchez, R., Melo, F., Sali, A. (2000) Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys .Biomol. Struct*. **29**: 291-325.

Mbewe, B. (2005) Cloning, expression, purification and drug targeting of *Plasmodium falciparum* hypoxanthine guanine xanthine phosphoribosyltransferase, Ph.D. Thesis, University of Cape Town, South Africa.

Mbewe, B., Chibale, K., McIntosh, D.B. (2007) Purification of human malaria parasite hypoxanthine guanine xanthine phosphoribosyltransferase (HGXPRT) using immobilized Reactive Red 120. *Protein Expr. Purif*. **52**: 153-58.

McIntosh, D.B., Parrish, J.C., Wallace, C.J. (1996) Definition of a nucleotide binding site on cytochrome c by photoaffinity labeling. *J. Biol. Chem.* **271**: 18379-86.

McGready, R., Cho, T., Villegas, L., Brockman, A., van Vugt, M., Looareesuwan, S., White, N.J., Nosten, F. (2001) Artemisinin antimalarials in pregnancy: a prospective treatment study of 539 episodes of multidrug-resistant Plasmodium falciparum. *Clin. Infect. Dis*. **33**: 2009-16.

McGuffin, L.J., Jones, D.T. (2003) Improvement of the GenTHREADER method for genomic fold recognition. *Bioinformatics* **19**: 874-81.

McPherson, A. (1999) Crystallization of biological macromolecules. Cold Spring Harbour Laboratory Press, New York, USA.

Meshnick, S.R., Taylor, T.E., Kamchonwongpaisan, S., (1996) Artemisinin and the antimalarial endoperoxides: from herbal remedy to targeted chemotherapy. *Microbiol. Rev*. **60**: 301-15.

Miller, H.L., Baruch, D.I., Marsh, K., Doumbo, O.K. (2002) The pathogenic basis of Malaria. *Nature* **415**: 673-79.

Mizuguchi, K., Deane, C.M., Blundell, T.L., Johnson, M.S., Overington, J.P. (1998) JOY: protein sequence-structure representation and analysis. *Bioinformatics* **14**: 617-23.

Moore, J. T., Uppal, A., Maley, F., Maley, G. F. **(**1993**)** Overcoming Inclusion Body Formation in a High-Level Expression System. *Protein Expression and Purification* **4:** 160-168.

Mota, M.M., Hafalla, J.C., Rodriguez, A. (2002) Migration through host cells activates Plasmodium sporozoites for infection. *Nat. Med.* **8**: 1318-22.

Mota, M.M., Pradel, G., Vanderberg, J.P., Hafalla, J.C., Frevert, U., Nussenzweig, R.S., Nussenzweig, V., Rodriguez, A. (2001) Migration of Plasmodium sporozoites through cells before infection. *Science* **291**: 24-5.

Mota, M.M., Rodriguez, A. (2001) Migration through host cells by apicomplexian parasites. *Microbes Infect*. **3**: 1123-8.

Munagala, N., Basus, V.J., Wang, C.C. (2001) Role of the flexible loop of hypoxanthine-guanine-xanthine phosphoribosyltransferase from *Tritrichomonas foetus* in enzyme catalysis. *Biochemistry* **14:** 4303-11.

Murzin, A.G., Brenner, S.E., Hubbard, T., Chothia, C. (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247:** 536-40.

Musick, W.D. (1981) Structural features of the phosphoribosyltransferases and their relationship to the human deficiency disorders of purine and pyrimidine metabolism. *CRC Crit. Rev. Biochem.* **11**: 1-34.

Needleman, S.B., Wunsch, C.D. (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48**: 443-53.

Nissani, E., Ginsburg H. (1989) Protonophoric effects of antimalarial drugs and alkylamines in Escherichia coli membranes. *Biochim. Biophys. Acta* **978**: 293-8.

Notredame, C., Higgins, D.G., Heringa, J. (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**: 205-17.

Ockenhouse, C.F., Tegoshi, T., Maeno, Y., Benjamin, C., Ho, M., Kan, K.E., Thway, Y., Win, K., Aikawa, M., Lobb, R.R. (1992) Human vascular endothelial cell adhesion receptors for *Plasmodium falciparum*-infected erythrocytes: roles for endothelial leukocyte adhesion molecule 1 and vascular cell adhesion molecule 1. *J. Exp. Med.* 176: 1183-9.

Olliaro, P. (2001) Mode of action and mechanisms of resistance for antimalarial drugs. *Pharmacol.Ther*. **89**: 207-19.

Orengo, C.A., Michie, A.D., Jones, S., Jones, D.T., Swindells, M.B., Thornton, J.M. (1997) CATH-a hierarchic classification of protein domain structures. *Structure* **5**: 1093-108.

Panchenko, A.R., Wolf, Y.I., Panchenko, L.A., Madej, T. (2005) Evolutionary plasticity of protein families: coupling between sequence and structure variation. *Proteins* **61**: 535-44.

Pasvol G. (2003) How many pathways for invasion of the red blood cell by the malaria parasite? *Trends Parasitol*. **19**: 430-2.

Pearson, W.R., Lipman, D.J. (1988) Improved tools for biological sequence comparison. *PNAS* **85**: 2444-8.

Pedersen, J.T., Moult, J. (1995) *Ab initio* structure prediction for small polypeptides and protein fragments using genetic algorithms. *Proteins* **23**: 454-60.

Peterson, R.W., Dutton, P.L., Wand, A.J. (2004) Improved side-chain prediction accuracy using an *ab initio* potential energy function and a very large rotamer library. *Protein Sci.* **13**: 735-51.

Petrella, R.J., Karplus, M. (2001) The energetics of off-rotamer protein side-chain conformations. *J. Mol. Biol.* **312**: 1161-75.

Phehane, V. N. (2002) Expression and drug targeting of parasite hypoxanthine-guanine phosphoribosyltransferase (HGXPRT), Ph.D. Thesis, University of Cape Town, South Africa.

Polshakov, D., Rai, S., Wilson, R.M., Mack, E.T., Vogel, M., Krause, J.A., Burdzinski, G., Platz, M.S. (2005) Photoaffinity labeling with 8-azidoadenosine and its derivatives: chemistry of closed and opened adenosine diazaquinodimethanes. *Biochemistry* **44**: 11241-53.

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

Ralph, S.A., D'Ombrain, M.C., McFadden, G.I. (2001) The apicoplast as an antimalarial drug target. *Drug Resist. Updat.* **4**: 145-51.

Ramachandran, G.N., Ramakrishnan, C., Sasisekharan, V. (1963) Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**: 95-9.

Raman, J., Ashok, C.S., Subbayya, S.I., Anand, R.P., Selvi, S.T., Balaram, H. (2005) *Plasmodium falciparum* hypoxanthine guanine phosphoribosyltransferase. Stability studies on the product-activated enzyme. *FEBS J.* **272:** 1900-11.

Rapp, C.S., Friesner, R.A. (1999) Prediction of loop geometries using a generalized born model of solvation effects. *Proteins* **35**: 173-83.

Reed, M.B., Saliba, K.J., Caruana, S.R., Kirk, K., Cowman, A.F. (2000) Pgh1 modulates sensitivity and resistance to multiple antimalarials in *Plasmodium falciparum*. *Nature* **403**: 906-9.

Reeder, J.C., Cowman, A.F., Davern, K.M., Beeson, J.G., Thompson, J.K., Rogerson, S.J., Brown, G.V. (1999) The adhesion of *Plasmodium falciparum*-infected erythrocytes to chondroitin sulfate A is mediated by *P. falciparum* erythrocyte membrane protein 1. *PNAS* **96**: 5198-202.

Retief, J.D. (2000) Phylogenetic analysis using PHYLIP. *Methods Mol. Biol.* **132**: 243-58.

Ridley R.G., (2002) Medical need, scientific opportunity and the drive for antimalarial drugs. *Nature* **415**: 686-93.

Ridley, R.G., Dorn, A., Vippagunta, S.R., Vennerstrom, J.L. (1997) Haematin (haem) polymerization and its inhibition by quinoline antimalarials. *Ann. Trop. Med. Parasitol*. **91**: 559-66.

Rizk, M.S., Shi, X., Platz, M.S. (2006) Lifetimes and reactivities of some 1,2-didehydroazepines commonly used in photoaffinity labeling experiments in aqueous solutions. *Biochemistry* **45**: 543-51.

Roberts, D.D., Haverstick, D.M., Dixit, V.M., Frazier, W.A., Santoro, S.A., Ginsburg, V. (1985) The platelet glycoprotein thrombospondin binds specifically to sulfated glycolipids. *J. Biol. Chem.* **260**: 9405-11.

Rosenthal PJ. (2003) Antimalarial drug discovery: old and new approaches. *J. Exp. Biol.* **206**: 3735-44.

Rost, B. (1997) Protein structures sustain evolutionary drift. *Fold. Des.* **2**: 19-24.

Russell, R.B., Saqi, M.A., Sayle, R.A., Bates, P.A., Sternberg, M.J. (1997) Recognition of analogous and homologous protein folds: analysis of sequence and structure conservation. *J. Mol. Biol.* **269**: 423-39.

Sachs, J., Malaney, P. (2002) The economic and social burden of malaria. *Nature* **415**: 680-85.

Sali, A. (1995) Modeling mutations and homologous proteins. *Curr. Opin. Biotechnol.* **6**: 437-51.

Sanchez, R., Sali, A. (1998) Large-scale protein structure modeling of the *Saccharomyces cerevisiae* genome. *PNAS* **95**: 13597-602.

Sanchez, C.P., Stein, W., Lanzer, M. (2003) Trans stimulation provides evidence for a drug efflux carrier as the mechanism of chloroquine resistance in *Plasmodium falciparum*. *Biochemistry* **42**: 9383-94.

Sauder, J.M., Arthur, J.W., Dunbrack, R.L. (2000) Large-scale comparison of protein sequence alignment algorithms with structure alignments. *Proteins* **40**: 6-22.

Scapin, G., Grubmeyer, C., Sacchettini, J.C. (1994) Crystal structure of orotate phosphoribosyltransferase. *Biochemistry* **33**: 1287-94.

Schramm, V.L., Shi, W. (2001) Atomic motion in enzymatic reaction coordinates. *Curr. Opin. Struct. Biol*. **11**: 657-65.

Sculley, D.G., Dawson, P.A., Emmerson, B.T., Gordon, R.B. (1992) A review of the molecular basis of hypoxanthine-guanine phosphoribosyltransferase (HPRT) deficiency. *Hum .Genet.* **90**: 195-207.

Seebregts, C.J., McIntosh, D.B. (1989) 2',3'-O-(2,4,6-trinitrophenyl)-8-azido-adenosine mono-, di-, and triphosphates as photoaffinity probes of the Ca2+-ATPase of sarcoplasmic reticulum. Regulatory/superfluorescent nucleotides label the catalytic site with high efficiency. *J. Biol. Chem.* **264**: 2043-52.

Sidhu,A.B., Verdier-Pinard, D., Fidock, D.A. (2002) Chloroquine resistance in *Plasmodium falciparum* malaria parasites conferred by pfcrt mutations. *Science* **298**: 210-3.

Singh, A., Thornton, E.R., Westheimer, F.H. (1962) The photolysis of diazoacetylchymotrypsin. *J. Biol. Chem.* **237:** 3006-8.

Sinha, S.C., Smith, J.L. (2001) The PRT protein family. *Curr. Opin. Struct. Biol.* **11**: 733-9.

Sippl, M.J. (1995) Knowledge-based potentials for proteins. *Curr. Opin. Struct. Biol.* **5**: 229-35.

Shi, J., Blundell, T.L., Mizuguchi, K. (2001) FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. *J. Mol. Biol.* **310**: 243-57.

Shi, W., Li, C.M., Tyler, P.C., Furneaux, R.H., Grubmeyer, C., Schramm, V.L., Almo, S.C. (1999) The 2.0 Å structure of human hypoxanthine-guanine phosphoribosyltransferase in complex with a transition-state analog inhibitor. *Nat. Struct. Biol.* **6**: 588-93.

Shi, W., Li, C.M., Tyler, P.C., Furneaux, R.H., Cahill, S.M., Girvin, M.E., Grubmeyer, C., Schramm, V.L., Almo, S.C. (1999) The 2.0 Å structure of malarial purine phosphoribosyltransferase in complex with a transition-state analogue inhibitor. *Biochemistry* **38**: 9872-80.

Smith, T.F., Waterman, M.S. (1981) Identification of common molecular subsequences. *J. Mol. Biol.* **147**: 195-7.

Snow, R.W., Guerra, C.A., Noor, A.M., Myint, H.Y., Hay, S.I. (2005) The global distribution of clinical episodes of *Plasmodium falciparum* malaria. *Nature* **434**: 214-217.

Somoza, J.R., Chin, M.S., Focia, P.J., Wang, C.C., Fletterick, R.J. (1996) Crystal structure of the hypoxanthine-guanine-xanthine phosphoribosyltransferase from the protozoan parasite *Tritrichomonas foetus*. *Biochemistry* **35**: 7032-40.

Srinivasan, N., Blundell, T.L. (1993) An evaluation of the performance of an automated procedure for comparative modelling of protein tertiary structure. *Protein Eng.* **6**: 501-12.

Srivastava, I.K., Vaidya, A.B. (1999) A mechanism for the synergistic antimalarial action of atovaquone and proguanil. *Antimicrob. Agents Chemother*. **43**: 1334-9.

Staines, H.M., Ellory, J.C., Chibale K. (2005) The new permeability pathways: targets and selective routes for the development of new antimalarial agents. *Comb. Chem. High Throughput Screen.* **8**: 81-8.

Stoye, J., Evers, D., Meyer, F. (1997) Generating benchmarks for multiple sequence alignments and phylogenetic reconstructions. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **5**: 303-6.

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

Sujjay Subbayya, I.N., Sukumaran, S., Shivashankar, K., Balaram, H. (2000) Unusual substrate specificity of a chimeric hypoxanthine-guanine phosphoribosyltransferase containing segments from the Plasmodium falciparum and human enzymes. *Biochem. Biophys. Res. Commun.* **272:** 596-602.

Surolia, N., Surolia, A. (2001) Triclosan offers protection against blood stages of malaria by inhibiting enoyl-ACP reductase of *Plasmodium falciparum*. *Nat. Med*. **7**: 167-73.

Sutcliffe, M.J., Hayes, F.R., Blundell, T.L. (1987) Knowledge based modelling of homologous proteins, part II: rules for the conformations of substituted side chains. *Protein Eng.* **1**: 385-92.

Swain, M.T., Kemp, G.J. (2001) Modelling of protein side-chain conformations using constraint logic programming. *Comput. Chem.* **26**: 85-95.

Tanimura, R., Kidera, A., Nakamura, H. (1994) Determinants of protein side-chain packing. *Protein Sci.* **3**: 2358-65.

Tao, W., Grubmeyer, C., Blanchard, J.S. (1996) Transition state structure of *Salmonella typhimurium* orotate phosphoribosyltransferase. *Biochemistry* **35**: 14-21.

Taylor, W.R. (1988) A flexible method to align large numbers of biological sequences. *J. Mol. Evol.* **28**: 161-9.

Ten Eyck, L.F. (1985) Fast Fourier transform calculation of electron density maps. *Methods Enzymol.* **115:** 324-37.

Tickle, A., Sharff, A., Vinkovic, M., Yon, J., Jhoti, H. (2004) High-throughput protein crystallography and drug discovery. *Chem. Soc. Rev.* **33**: 558-65.

Thomas A, Field MJ. (2002) Reaction mechanism of the HGXPRTase from *Plasmodium falciparum*: a hybrid potential quantum mechanical/molecular mechanical study. *J. Am. Chem. Soc*. **124**: 12432-8.

Thomas, P.J., Qu, B.H., Pedersen, B.L. (1995) Defective protein folding as a basis of human disease. *Trends Biochem. Sci.* **20**: 456-9.

Thompson, J.D., Higgins, D.G., 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**: 4673-80.

Todd, A.E., Orengo, C.A., Thornton, J.M. (2001) Evolution of function in protein superfamilies, from a structural perspective. *J. Mol. Biol.* **307**: 1113-43.

Trampuz, A., Jereb, M., Muzlovic, I., Prabhu, R.M. (2003) Clinical review: severe malaria. *Crit. Care* **7**: 315-23.

Udomsangpetch, R., Webster, H.K., Pattanapanyasat, K., Pitchayangkul, S., Thaithong, S. (1992) Cytoadherence characteristics of rosette-forming *Plasmodium falciparum*. *Infect. Immun*. **60**: 4483-90.

Varghese, J.N. (1999) Development of neuraminidase inhibitors as anti-influenza virus drugs. *Drug Dev. Res.* **46**: 176–96.

Vasquez, M (1996) Modeling side-chain conformation. *Curr. Opin. Struct. Biol.* **6**: 217-21.

Vlijmen, H.W., Karplus, M. (1997) PDB-based protein loop prediction: parameters for selection and methods for optimization. *J. Mol. Biol.* **267**: 975-1001.

Voigt, C.A., Gordon, D.B., Mayo, S.L. (2000) Trading accuracy for speed: A quantitative comparison of search algorithms in protein sequence design. *J. Mol. Biol.* **299**: 789-803.

Vos, S., de Jersey, J., Martin, J.L. (1997) Crystal structure of *Escherichia coli* xanthine phosphoribosyltransferase. *Biochemistry* **36**: 4125-34.

Ward, W.H., Holdgate, G.A. (2001) Isothermal titration calorimetry in drug discovery. *Prog. Med. Chem.* **38**: 309-76.

Warhurst, D.C. (2003) Polymorphism in the *Plasmodium falciparum* chloroquine-resistance transporter protein links verapamil enhancement of chloroquine sensitivity with the clinical efficacy of amodiaquine. (http://www.malariajournal.com/content/2/1/31)

Wellems, T.E., Plowe, C.V. (2001) Chloroquine-resistant malaria. *J. Infect. Dis.* **184**: 770-6.

White, N.J. (2004) Antimalarial drug resistance. *J. Clin. Invest.* **113**: 1084-92.

Wiener, J., Hintz, M., Altincicek, B., Sanderbrand, S., Weidemeyer, C., Beck, E., Jomaa, H. (2000) *Plasmodium falciparum*: detection of the deoxyxylulose 5-phosphate reductoisomerase activity. *Exp. Parasitol*. **96**: 182-6.

Wiesner, J., Borrmann, S., Jomaa, H. (2003) Fosmidomycin for the treatment of malaria. *Parasitol. Res.* **90**: 71-6.

Wilson, C.M., Serrano, A.E., Wasley, A., Bogenschutz, M.P., Shankar, A.H., Wirth, D.F. (1989) Amplification of a gene related to mammalian mdr genes in drug-resistant *Plasmodium falciparum*. *Science* **244**: 1184-6.

Wilson, R.J. (2002) Progress with parasite plastids. *J. Mol. Biol*. **319**: 257-74.

Winstanley, P.A. (2000) Chemotherapy for falciparum malaria: the armoury, the problems and the prospects. *Parasitol. Today* **16**: 146-53.

Winstanley, P.A., Coleman, J.W., Maggs, J.L., Breckenridge, A.M., Park, B.K. (1990) The toxicity of amodiaquine and its principal metabolites towards mononuclear leucocytes and granulocyte/monocyte colony forming units. *Br. J. Clin. Pharmacol.* **29**: 479-85.

Wood, T.C., Pearson, W.R. (1999) Evolution of protein sequences and structures. *J. Mol. Biol.* **291**: 977-95.

Wright, C.W., Addae-Kyereme, J., Breen, A.G., Brown, J.E., Cox, M.F., Croft, S.L., Gokcek, Y., Kendrick, H., Phillips, R.M., Pollet, P.L .(2001) Synthesis and evaluation of cryptolepine analogues for potential as new antimalarial agents. *J. Med. Chem.* **44**: 3187-94.

Xu, Y., Eads, J., Sacchettini, J.C., Grubmeyer, C. (1997) Kinetic mechanism of human hypoxanthine-guanine phosphoribosyltransferase: rapid phosphoribosyl transfer chemistry. *Biochemistry* **36**: 3700-12.

Xu, Y., Grubmeyer, C. (1998) Catalysis in human hypoxanthine-guanine phosphoribosyltransferase: Asp 137 acts as a general acid/base. *Biochemistry* **37**: 4114-24.

Yeung, S., Pongtavornpinyo, W., Hastings, I.M., Mills, A.J., White, N.J. (2004) Antimalarial drug resistance, artemisinin-based combination therapy, and the contribution of modeling to elucidating policy choices. *Am. J. Trop. Med. Hyg.* **71**: 179-86.

You, D., Chen, Q., Liang, Y., An, J., Li, R., Gu, X., Luo, M., Su, X.D. (2003) Protein preparation, crystallization and preliminary X-ray crystallographic studies of a thermostable hypoxanthine-guanine phosphoribosyltransferase from *Thermoanaerobacter tengcongensis*. *Acta Crysta. D* **59**: 1863-5.

Young, M.J.T., Platz, M.S. (1991) Mechanistic analysis of the reactions of (pentafluorophenyl)nitrene in alkanes. *J. Org. Chem.* **56**: 6403-06.

Yuvaniyama, J., Chitnumsub, P., Kamchonwongpaisan, S., Vanichtanankul, J., Sirawaraporn, W., Taylor, P., Walkinshaw, M.D., Yuthavong, Y. (2003) Insights into antifolate resistance from malarial DHFR-TS structures. *Nat. Struct. Biol.* **10**: 357-65.

Zhang, H., Howard, E.M., Roepe, P.D. (2002) Analysis of the antimalarial drug resistance protein Pfcrt expressed in yeast. *J. Biol. Chem.* **277**: 49767-75.

Zhang, K., Rathod, P.K. (2002) Divergent regulation of dihydrofolate reductase between malaria parasite and human host. *Science* **296**: 545-7.