THE UNIVERSITY *of* EDINBURGH

Biochemical and Structural Studies on Trypanosomatid Pyruvate Kinases

A Thesis

Submitted for the Degree of

Doctor of Philosophy

by

Wenhe Zhong, B.Sc

Structural Biochemistry Group

Institute of Structural and Molecular Biology

School of Biological Sciences

University of Edinburgh

Scotland

United Kingdom

February 2013

Abstract

Glycolytic enzymes have been indicated as potential drug targets in trypanosomatid parasites such as *Trypanosoma brucei* (*T. brucei*), *Trypanosoma cruzi* (*T. cruzi*) and *Leishmania* spp. Pyruvate kinase (PYK) catalyses the final reaction in the glycolytic pathway to produce ATP and pyruvate from ADP and phospho*enol*pyruvate (PEP), and has been validated by RNAi experiments as a suitable drug target in *T. brucei*. This thesis describes biochemical and structural studies of PYKs from *T. cruzi* (*Tc*PYK) and *T. brucei* (*Tb*PYK), providing not only a foundation but also new clues for PYK-specific inhibitor screening and structure-based drug design.

Soluble *Tc*PYK and *Tb*PYK (81% sequence identity) have been expressed and purified from *E. coli*, and their kinetics have been fully characterised. X-ray crystal structures of apoenzyme *Tc*PYK (apo *Tc*PYK), and of *Tb*PYK in complex with fructose 2,6-bisphosphate (F26BP) (*Tb*PYK/F26BP/Mg) have been determined, and each possesses a tetrameric architecture composed of four identical protein chains. Each chain contains four domains which are A-domain, B-domain, C-domain and N-terminal domain. The active site is located in the cleft between the A- and B-domains, while the F26BP-bound effector site is within the C-domain. The conformational transition between inactive T-state and active R-state for both enzymes requires a concerted 8^{o} rigid-body rotation of each of the four AC-cores (A- and C-domains) in the tetramer. During the T- to R-state transition induced by F26BP binding, the side chain of Arg311 is re-orientated to stabilise the short Aα6′ helix at the active site, and the flexible loop at the effector site is stabilised by F26BP. In this active conformation additional salt bridges form across the C-C interface to lock the enzyme in a more stable R-state.

*Tb*PYK/F26BP/Mg is the first ‘effector only’ PYK structure and identifies a third

Mg^{2+} binding site (Mg-3) which is distinct from the two canonical Mg^{2+} binding sites. The substrate PEP was soaked into crystals of *Tb*PYK/F26BP/Mg resulting in an '*in crystallo*' 23° B-domain rotation forming a partially closed active site. This is accompanied by active site side-chain reorientations, and the movement of Mg^{2+} from its 'priming' position Mg-3 to its canonical position Mg-1. It is plausible that Mg^{2+} is retained in its 'priming' position after product release to act as a co-activator with F26BP to maintain the enzyme in its R-state conformation, as long as F26BP is present.

The inherent oxaloacetate decarboxylase activity of PYK was reported over 30 years ago and has been further characterised by ^{1}H NMR studies in this thesis. In addition, a series of *Tb*PYK structures in complex with product (pyruvate), with analogues of the decarboxylase substrate oxaloacetate (D-malate and α-ketoglutarate), or with the competitive inhibitor oxalate have been determined by crystal soaking, and indicate that both decarboxylase activity and kinase activity share a common active site. A proposed mechanism explains the conserved decarboxylase activity of PYK where the active-site Mg^{2+} and Lys239 in *Tb*PYK (which is conserved between species) play essential roles in the decarboxylation reaction.

Three strategies for designing novel inhibitors against trypanosomatid PYKs have been proposed in this thesis. (1) Develop selective modulators to increase the binding affinity of inhibitors. As an example, F16BP has been shown to regulate the inhibitory effect of PEP analogues (oxalate, D-malate, α-ketoglutarate, malonate and L-tartrate) on *Tb*PYK activity. (2) Develop allosteric inhibitors in order to lock trypanosomatid PYKs in an inactive state where the enzyme has low affinity for substrate binding. (3) A third strategy is to combine multiple modulators and inhibitors to increase the inhibition efficiency and selectivity.

Declaration

I hereby declare that the research presented in this thesis and the thesis itself has been composed and originated completely by myself at The University of Edinburgh, United Kingdom, except when otherwise stated. This thesis has not been submitted in whole or partially for any other degree.

Wenhe Zhong

Edinburgh, February 2013

Acknowledgements

First I would like to express my gratitude to my first supervisor Prof. Malcolm Walkinshaw for his kind support, encouragement and extremely professional mentoring throughout my PhD study. Another gratitude will be given to my second supervisor Dr. Linda Gilmore for her endless support, comments and care. Linda also inspired me as a scientist because she loves her work and has been spending most of her time on science even though she is retired.

I would like to thank Dr. Hugh Morgan for his guidance and teaching in parasitic enzyme research. Hugh encouraged me every time when there was no smile on my face. I could not have this many crystal structures presented in the thesis without his advice of screening dye molecules in protein crystallisation.

I would like to thank all the members on the 3rd floor. Thank Dr. Paul Taylor and Dr. Julia Richardson for being in my PhD committee and giving ideas and supports. Thank Dr. Iain McNae for teaching me X-ray crystallography and thank Dr. Matthew Nowicki for offering his professional answers on Chemistry questions. Not forgetting Dr. Jacqueline Dornan, many thanks for 'guiding' me in lab S3.4. All the protein production and analysis work were done in PPF and BCF, so I would like to thank all the staffs in PPF and BCF (Dr. Martin Wear, Sandra Bruce, Dr. Janice Bramham, Dr. Liz Blackburn, Dr. Matthew Nowicki) for their tireless helps and technical assistance throughout the period of study. I must thank Prof. Paul Michels in Brussels for kindly sharing his knowledge in trypanosomatid parasites. Also thank all the PhD students on the 3rd floor for sharing happiness and sadness. Additionally, thank Mr. Juraj Bella in the School of Chemistry for running proton NMR and analysing the raw data. I am grateful to Dr David A. Okar at the Veterans Administration Medical Centre (USA) as well for the gift of F26BP.

Last but not least, I would like to give my very huge and warm thank to my dear parents. Without their love, care, support, understanding and encouragement throughout my life, I would never be so healthy, happy, fearless, and have this opportunity to be studying in one of the best universities in the world. Also thank to University of Edinburgh and China Scholarships Council for offering scholarship in my PhD study.

Abbreviations

AC_{50}	Half-maximum activating concentration
AC_{90}	90% maximum activating concentration
ADP	Adenosine diphosphate
AKG	α-ketoglutarate
ALD	Aldolase
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
CCP4	Collaborative Computational Project Number 4
CIT	Citrate
CL	Cutaneous Leishmaniasis
CNS	Central Nervous System
DALYs	Disability-Adjusted Life Years
DMSO	Dimethyl sulfoxide
Ec	*E. coli, Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
ENO	Enolase
F16BP	Fructose 1,6-bisphosphate
F26BP	Fructose 2,6-bisphosphate
GAPDH	Glyceraldehydes-3-phophate dehydrogenase
GPDH	Glycerol-3-phophate dehydrogenase
Gs	*G. stearothermophilus, Geobacillus stearothermophilus*
h	Hill coefficient
HAT	Human African Trypanosomiasis
HK	Hexokinase
Hs	*Homo sapiens*
IC_{50}	Half maximal inhibitory concentration
IMAC	Immobilised Metal Affinity Chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
k_{cat}	Turnover number
kDa	Kilodalton
K_i	Binding affinity of the inhibitor
*K*m	Michaelis constant
LDH	Lactate dehydrogenase
Lm	*L. mexicana, Leishmania mexicana*
MCL	Muco-Cutaneous (mucosal) Leishmaniasis
MLA	Malonate
MLT	D-malate
NAD^+	Nicotinamide adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide
NCS	Non-crystallographic symmetry

NIH	National Institutes of Health
NTD(s)	Neglected tropical disease(s)
OAD	Oxaloacetate decarboxylase
Oc	*Oryctolagus cuniculus*
OD_{600}	Optical density 600 nm
OX	Oxalate
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
PEP	Phosphoenolpyruvate
PFK	Phosphofructokinase
PGAM	Phosphoglycerate mutase
PGI	Glucose-6-phosphate isomerase
PGK	Phosphoglycerate kinase
pI	Isoelectric point
PKDL	Post-kala-azar Dermal Leishmaniasis
PYK	Pyruvate kinase
PYR	Pyruvate
R5P	Ribose 5-phosphate
RMS	Root mean square
RMSD	Root mean square deviation
RNAi	RNA interference
rpm	Revolutions per minute
$S_{0.5}$	Substrate concentration giving half-maximal reaction rate
Sa	*S. aureus, Staphylococcus aureus*
Sc	*S. cerevisiae, Saccharomyces cerevisiae*
SDS-PAGE	Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis
Tb	*T. brucei, Trypanosoma brucei*
Tc	*T. cruzi, Trypanosoma cruzi*
TEA	Triethanolamine
TIM	Triosephosphate isomerase
TLA	L-tartrate
TLS	Translation/Libration/Screw
Tm	Transition temperature
UV	Ultraviolet
VL	Visceral Leishmaniasis
WHO	World Health Organisation

Contents

List of Figures

List of Tables

CHAPTER 1: Introduction

1.1. Neglected tropical diseases

Neglected tropical diseases (NTDs) are infectious diseases which predominate in the tropical belt and are poorly managed (Chatelain and Ioset 2011; Feasey et al. 2010; Hotez and Kamath 2009; Hotez and Pecoul 2010). The World Health Organization (WHO) delineates a total of 17 main NTDs which are covered by the NTD Department, and are highlighted on the WHO website (http://www.who.int/neglected_diseases/diseases/en/): Chagas disease (American trypanosomiasis), cysticercosis/taeniasis, echinococcosis, foodborne trematode infections, human African trypanosomiasis (HAT, sleeping sickness) and leishamaniases are caused by protozoa; dracunculiasis, lymphatic filariasis, onchocerciasis, schistosomiasis and soil transmitted helminthiases are caused by helminthes; Buruli ulcer, leprosy, trachoma and yaws are caused by bacteria; dengue and rabies are caused by viruses.

NTDs are a group of chronic and disabling conditions leading to acute illness, long-term disability and early death (Feasey et al. 2010; Hotez et al. 2007; Hotez and Kamath 2009). However, NTDs have been 'neglected' for decades, and have primarily become the diseases of the 'bottom billion' populations who are the world's poorest people mostly living in low- and lower middle- income countries in tropical and subtropical climates (Collier 2007; Manderson 2012).

NTDs are responsible for substantial morbidity and mortality in developing countries. According to the report from WHO in 2006, 100% of low-income countries are affected by at least five NTDs concurrently, and many individuals who live in those counties are simultaneously afflicted with more than one pathogen (WHO 2006b). In this case, these diseases have a clear economic burden on the

affected countries (Conteh et al. 2010; Mathers et al. 2007). Nevertheless, the progress of drug discovery and development for the treatment of NTDs is relatively slow in history partially due to the lack of links between profitable market and public health policies in low-income countries (Chatelain and Ioset 2011; Hotez and Kamath 2009; World Health Organization/Industry Drug Development Working Group 2001). During 1975-2004, only 21 out of 1556 approved drugs (~1.3%) were specifically developed to address NTDs, even if NTDs account for 11.4% of the global disease burden (Chatelain and Ioset 2011; Chirac and Torreele 2006).

In recent years an increasing number of not-for-profit organisations and academia are filling the gap between research and industry to greatly enhance new drug development to address NTDs (Chatelain and loset 2011; Croft 2005; Moran 2005; Nwaka and Ridley 2003; Tanne 2012). Public funding remains the top global source of investment in neglected diseases including HIV/AIDS, malaria and tuberculosis (TB) (**Figure 1.1**). However, the immediate needs of the most neglected patients affected by NTDs are still largely unmet (Moran et al. 2009). A 5-year review (2007-2011) of global neglected disease research and development funding recently released through G-FINDER (Moran et al. 2012) reveals a low percentage share of worldwide funding for NTDs compared to the top-tier diseases (HIV/AIDS, malaria and tuberculosis (TB)). For example, the share of worldwide funding for the diseases caused by kinetoplastids, including human African trypanosomiasis (HAT), Chagas disease and leishmaniases, did not exceed 5% in 2011 compared to ~70% share for the three top-tier diseases. Therefore, more high-level attention and investments should be given towards NTDs to develop safe and effective new drugs.

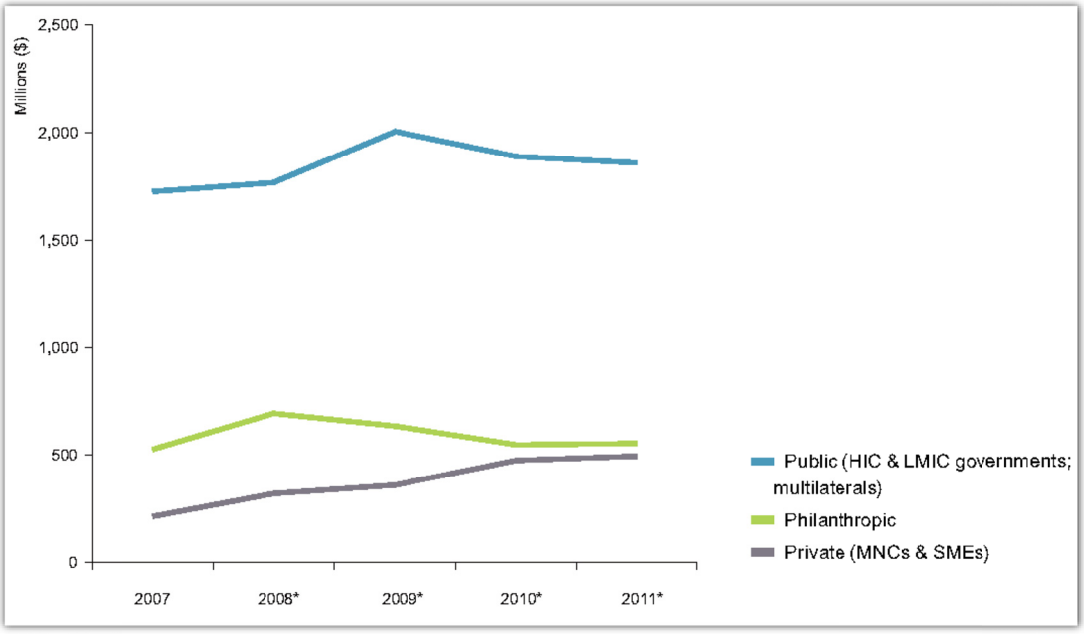


Figure 1.1 Total funding for neglected diseases by funder type 2007-2011
Figures are adjusted for inflation and reported in 2007 US dollars.
HIC, High-income country; LMIC, Low- and middle-income country; MNC, Multinational pharmaceutical company; SMEs, Small pharmaceutical and biotechnology firms.
This figure is obtained from Figure 26 in Moran et al. 2012 (G-FINDER 2012).

1.2. NTDs caused by trypanosomatid parasites

Human African trypanosomiasis (HAT), Chagas disease and leishmaniases are three wide-spread, debilitating and often fatal NTDs which are caused by *Trypanosoma brucei* (*T. brucei*), *T. cruzi* and *Leishmania* spp., respectively (Steverding 2008; Stuart et al. 2008; von Geldem et al. 2011). These three flagellated parasitic protozoa belong to the order of Kinetoplastida. *T. brucei* and *T. cruzi* are two species of the genus *Trypanosoma*, and *Leishmania* spp are species of the genus *Leishmania*.

1.2.1. Human African trypanosomiasis and T. brucei

Human African trypanosomiasis (HAT), also known as sleeping sickness, is generally fatal if left untreated (Burri and Brun 2003) and mainly affects rural

populations in 36 countries in sub-Saharan Africa (**Figure 1.2**; von Geldern et al. 2011). Tens of thousands of people are infected each year and millions are at risk of contracting HAT (Fèvre et al. 2008). It was estimated by WHO in 2002 that ~1.5 million disability-adjusted life years (DALYs) were lost due to this disease (WHO 2004).

HAT is caused by *T. brucei* parasites which are transmitted to humans through the bite of the *Glossina* tsetse fly (Simarro et al. 2008; von Geldern et al. 2011; WHO 2006a). Two stages of HAT have been identified. In the first (early) stage of the disease (the hemolymphatic stage), the parasites (trypomastigote form) proliferate in the blood and lymphatic systems and remain extracellularly in the blood stream (**Figure 1.3**). Although rapid parasite growth is countered by host immune responses, parasite antigenic variation enables immune evasion and the parasites continuously remain one step ahead of the host (Morrison et al. 2009; Rudenko et al. 2011; Taylor and Rudenko 2006). This stage involves adenopathy, fever, headache, joint pain, and pruritus (Stuart et al. 2008). In the second (late) stage of HAT (the central nervous system or CNS stage), the parasites migrate to the brain tissue across the blood-brain barrier resulting in parasitic invasion of the central nervous system and eventual neuronal death (Grab and Kennedy 2008; Kennedy 2008a; Kennedy 2008b; Simarro et al. 2008; von Geldem et al. 2011). This later stage is accompanied by a multitude of neurological symptoms including hallucinations, sleep disorders, coma and ultimately death (Grab and Kennedy 2008).

Two subspecies of *T. brucei* parasites have been observed to infect humans, namely *T. brucei gambiense* and *T. brucei rhodesiense* (Barrett et al. 2003; Brun et al. 2010): (i) *T. brucei gambiense*causes a chronic HAT in central and west Africa (**Figure 1.2a**) representing ~97% of reported HAT cases (von Geldern et al. 2011); (ii) *T. brucei rhodesiense* causes an acute form of HAT in east and southern Arica (**Figure 1.2b**). Additionally, *T. brucei brucei*, which is a third subspecies of *T. brucei*, usually only infects domestic and wild animals but not humans.

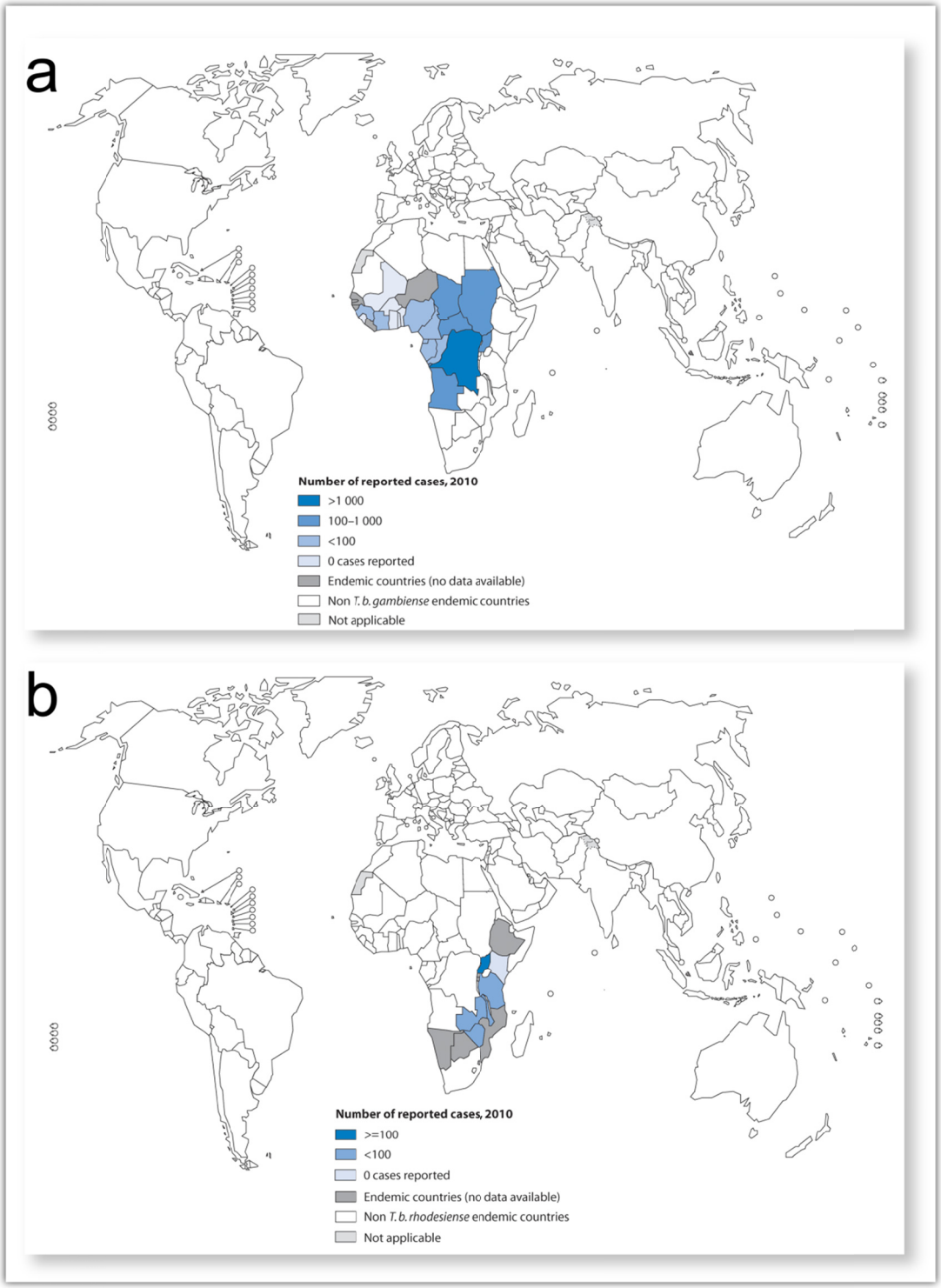


Figure 1.2 Worldwide distribution of human African trypanosomiasis, 2010
(a) Distribution of HAT caused by *T. brucei gambiense*
(b) Distribution of HAT caused by *T. brucei rhodesiense*
This modified figure was originally obtained from the WHO website:
http://gamapserver.who.int/mapLibrary/app/searchResults.aspx

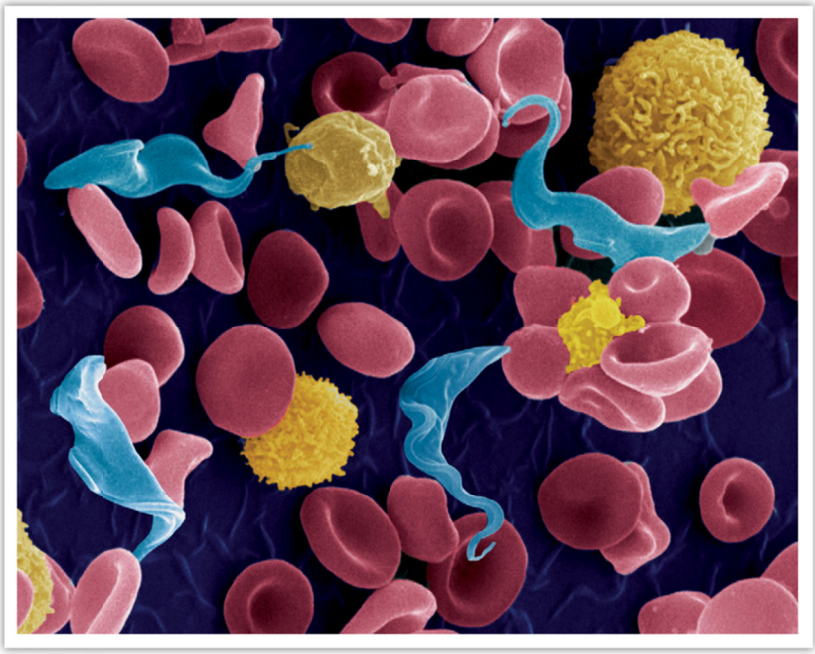


Figure 1.3 *T. brucei* parasites among blood cells.
This figure is from Figure 4 in Brun et al. 2010.

HAT has been observed since the 19th century and is often fatal without proper treatment.Currently more than 60 million people are at risk (Brun et al. 2010; von Geldern et al. 2011), and unfortunately, the treatment options for HAT are still far from satisfactory mainly because of the ineffectiveness, toxicity and complexity of treatment regimens (**Table 1.1**; **Figure 1.4**; Croft et al. 2005; Jacobs et al. 2011; Wilkinson and Kelly 2009). Additionally, the development of a vaccine seems to be unfeasible for HAT due to the high degree of antigenic variation of *T. brucei* parasites in the host (Stuart et al. 2008).

Table 1.1 Drugs in use for HAT, Chagas disease and Leishmaniases

Drug	Date first used	Limitations
Human African trypanosomiasis (HAT)		
Suramin (First stage treatment)	1921	Toxicity, injected; used for *T. brucei rhodesiense* disease
Pentamidine (First stage treatment)	1941	Toxicity, resistance and required injection; used for *T. brucei gambiense* disease
Melarsoprol (Second stage treatment)	1949	Toxicity, resistance and required injection; available for *T. brucei rhodesiense* disease
Eflornithine (Second stage treatment)	1990	Costly, injected and only effective against *T. brucei gambiense*
Combination of eflornithine and nifurtimox (Second stage treatment)	2009	used for *T. brucei gambiense* disease
Chagas disease		
Nifurtimox	1970	Toxicity, long-treatment compliance, activity limited to acute stage of disease
Benznidazole	1974	Toxicity, activity limited to acute stage of disease
Leishmaniases		
Pentamidine	1939	Toxicity, resistance and required injection
Pentavalent antimonials	1950	Toxicity, resistance and required injection
Liposomal amphotericin B	1990	Costly and required injection
Miltefosine	2002	Contraindicated in pregnancy
Paromomycin	Available in 1960 but abandoned until 2005	Toxicity and required injection

This modified table was originally obtained from Table 1.1 in Fuad 2012 (PhD thesis) and Table 1 in Barrett and Croft 2012.

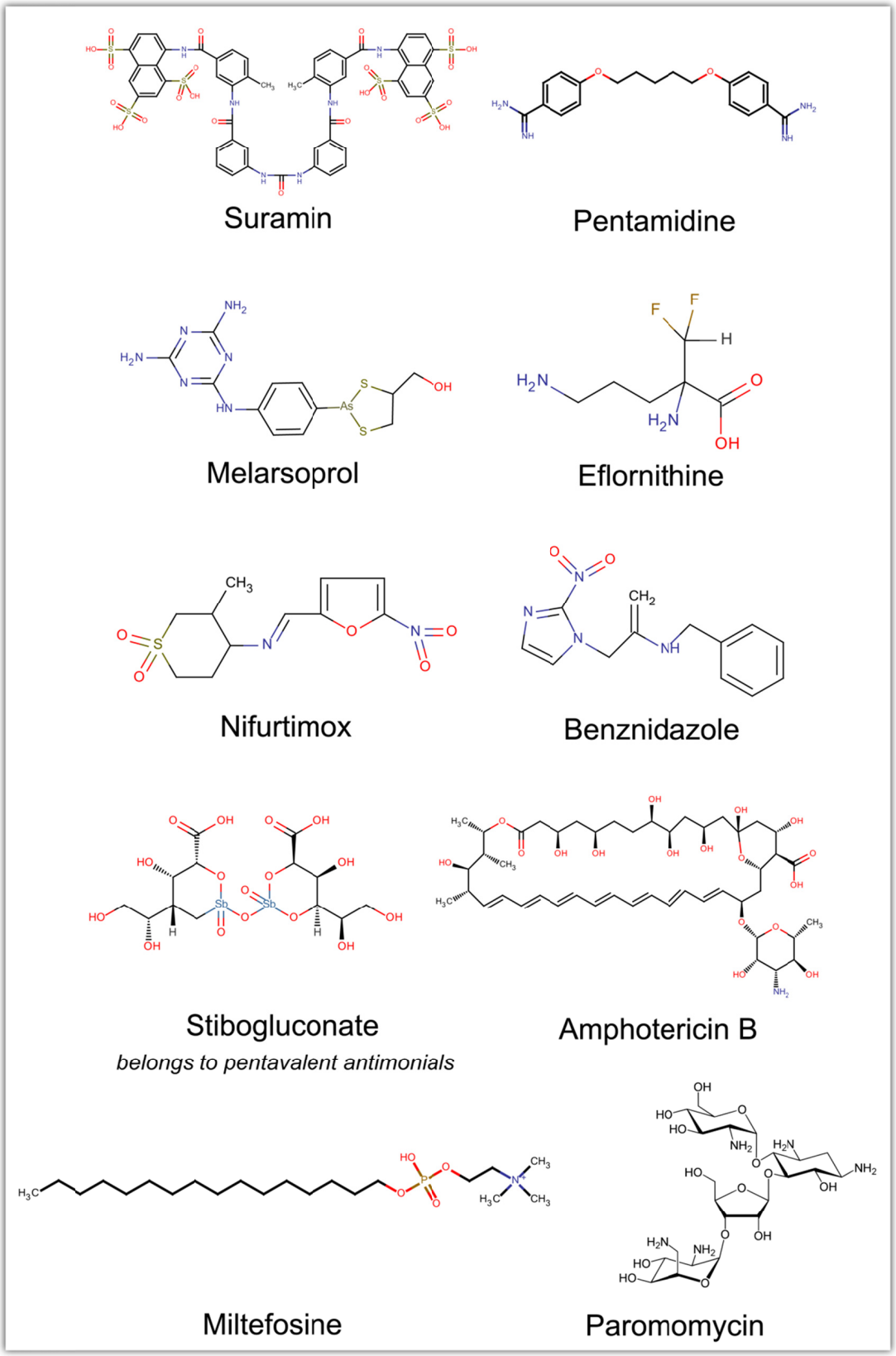


Figure 1.4 Two-dimensional representations of current drugs for the treatment of HAT, Chagas disease and Leishmaniases, corresponding to Table 1.1.

1.2.2. Chagas disease and T. cruzi

Chagas disease (American trypanosomiasis), which is caused by infection by the protozoan parasite *T. cruzi* and generally transmitted to humans by blood-sucking reduviid bugs (kissing bugs) of the subfamily Triatominae (Buscaglia et al. 2006; Deane 1964; Zeledón and Rabinovich 1981), has been recognised by WHO as one of the world's most neglected tropical diseases (Hotez et al. 2007; Rassi et al. 2010; WHO 2002). Apart from the vector-borne transmission, other routes of Chagas disease transmission such as blood and organ transplant transmissions, and foodborne transmission all exist currently (Benchimol Barbosa 2006; CDC 2006; Gurtler et al. 2003; Pereira et al. 2009; Quijano-Hernandez and Dumonteil 2011; Rassi et al. 2009; Ríos et al. 2011; Young 2007). Although Chagas disease was discovered in 1909 by the Brazilian physician Carlos Chagas (1879-1934), it has been proved that this disease afflicted humans as early as 9000 years ago by the findings from paleoparasitology studies on human mummies who were infected by *T. cruzi*.(Aufderheide et al. 2004).

Chagas disease is the major disease in endemic areas such as Central and South America where ~8-15 million people in 18 countries are infected and ~30 million people at risk (**Figure 1.5**; Nagajyothi et al. 2012; Rassi et al. 2012). Furthermore, Chagas disease is becoming a global public health problem as a result of *T. cruzi*-infected immigrants in non-endemic countries (**Figure 1.5**; Pérez-Molina et al. 2012). Therefore, this international immigration has gradually expanded the disease to affect richer suburban and urban areas in many countries instead of being limited to rural areas (**Figure 1.5**), and has brought new challenges for researchers to develop new vaccines and drugs against this disease (Nagajyothi et al. 2012).

The parasite *T. cruzi* possesses four distinct stages for its life cycle which are involved in three phases of Chagas disease. The four stages of *T. cruzi* are: epimastigotes (proliferative stage in bugs), metacyclic trypomastigotes (nonproliferative stage both in bugs and hosts), intra-cellular amastigotes

(proliferative stage in human cells), and bloodstream-form trypomastigotes (nonproliferative stage in hosts) (Nagajyothi et al. 2012).

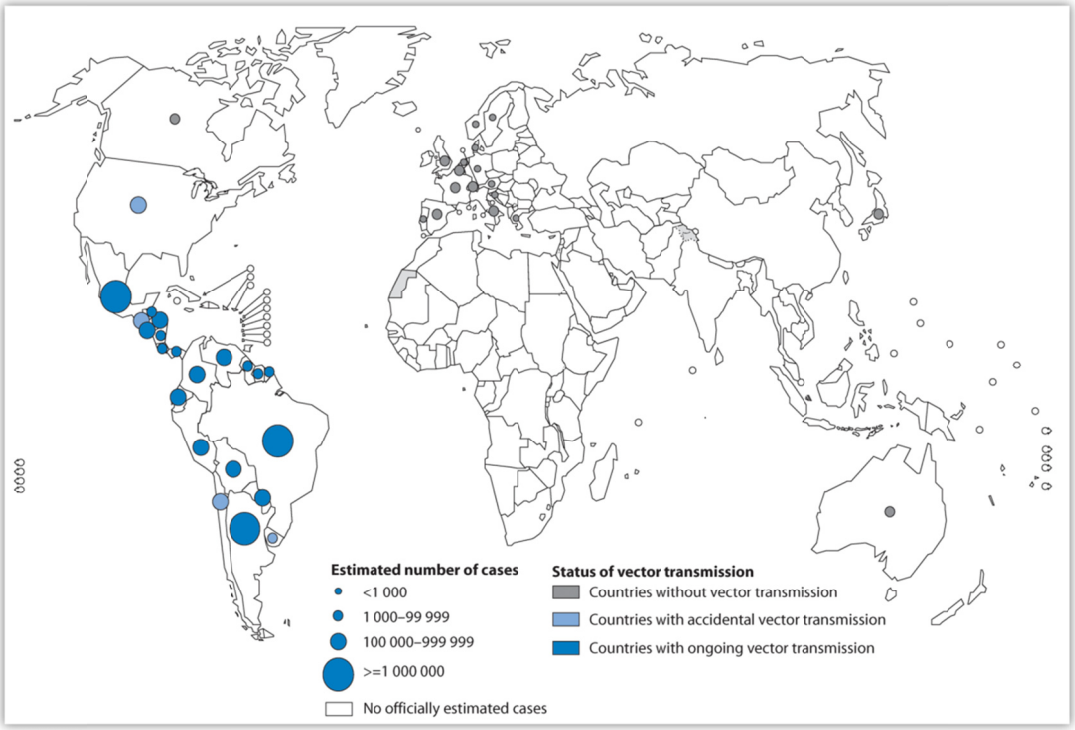


Figure 1.5 Worldwide distribution of cases of *T. cruzi* infection (Chagas disease), 2006-2009
This modified figure was originally obtained from the WHO website based on official estimates and status of vector transmission:
http://gamapserver.who.int/mapLibrary/app/searchResults.aspx

Briefly, *T. cruzi* parasites (epimastigotes) amplify in Triatominae bugs and invade many cell types of humans including macrophages and muscle cells in the form of metacyclic trypomastigotes (Andrade and Andrews 2005). The parasites in the form of amastigotes enter the parasite-amplification process and the amplified parasites can be released into the bloodstream in the form of trypomastigotes. This stage allows parasitological diagnosis and is known as the acute phase of Chagas disease which is associated with fever, headache and nausea (Barrett et al. 2003). Without proper treatment, the disease transitions into an indeterminate chronic phase,

where parasite numbers drop significantly. Then 10-30 years after the initial infection, ~30% of infected people will experience the chronic stage characterised by chronic heart failure, digestive lesions and peripheral nervous damage (Moncayo and Ortiz 2006; Rassi et al. 2010).

Chagas disease is a lifelong infection and has become a social and economic problem in many endemic countries (Mathers et al. 2007; Moncayo and Silveira 2009; Rassi et al. 2010). Unfortunately, similar to HAT, there is not a vaccine or appropriate drugs available nowadays, and the current treatment options for Chagas disease are far from satisfactory mostly due to unwanted effects, poor clinical efficacy and the requirement for parenteral administration (**Table 1.1**; **Figure 1.4**; Quijano-Hernandez and Dumonteil 2011; Rassi et al. 2010).

1.2.3. Leishmaniases and Leishmania

It has been highlighted by WHO (http://www.who.int/leishmaniasis/burden/en/) that leishmaniases occur in four continents currently (including Africa, Southern America, Asia and Southern Europe) and is considered to be endemic in 88 counties, 72 of which are developing counties, indicating that leishmaniasis is a poverty-related disease (**Figure 1.6**; Chappuis et al. 2007; von Geldern et al. 2011). This family of diseases is caused by the infection by several species of the protozoan kinetoplastid parasite *Leishmania* and is transmitted by phlebotomine sandflies (Mubayi 2009). Furthermore, it has been confirmed by WHO that the co-infection of leishmaniasis and HIV exists which significantly hampers the control of *Leishmania* (http://www.who.int/leishmaniasis/burden/hiv_coinfection/burden_hiv_coinfection/en/index.html).

There are a total of ~21 leishmanial species transmitted by ~30 species of phlebotomine sandflies (Shaw 1994). Similar to the vectors in trypanosomiasis, *Leishmania* spp also exists in more than one morphological form. Two

morphologically distinct forms have been observed involved in infection by *Leishmania*: a motile flagellated form (promastigotes) which can multiply in the sandfly; an intracellular non-flagellaed form (amastigotes) which are transformed from promastigotes and multiply in human cells (Molyneux 2007; Mubayi 2009).

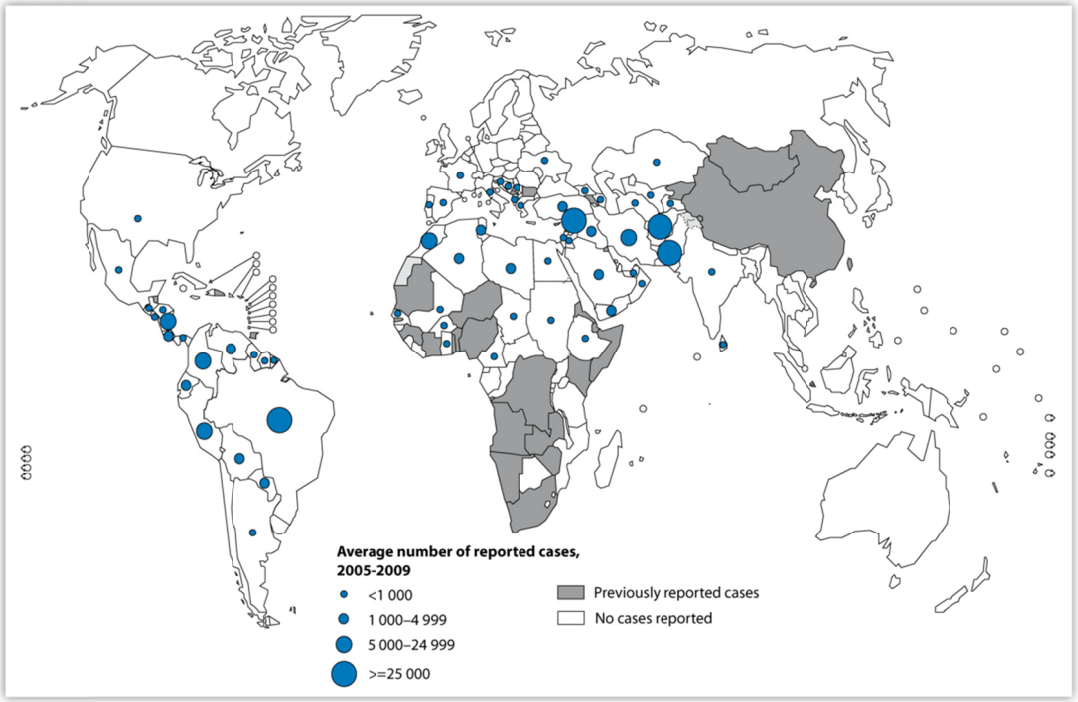


Figure 1.6 Worldwide distribution of cutaneous leishmaniasis, 2009
This modified figure was originally obtained from the WHO website: http://gamapserver.who.int/mapLibrary/app/searchResults.aspx

Currently, four main clinical syndromes of leishmaniasis are clearly identified: (i) cutaneous leishmaniasis (CL) (Minodier and Parola 2007), (ii) muco-cutaneous (mucosal) leishmaniasis (MCL) (Amato 2008), (iii) visceral leishmaniasis (VL, also known as kala-azar) (Sundar and Chatterje 2006), (iv) post-kala-azar dermal leishmaniasis (PKDL) (Zijlstra 2003). The first three clinical syndromes of leishmaniases are most prominent, and are caused by the replication of the parasite in macrophages in the dermis, the naso-oropharyngeal mucosa and the mononuclear phagocyte system, respectively (David and Craft 2009).

About 1.5-2.0 million new cases of leishmaniases are estimated yearly, and of these, 1-1.5 million are due to cutaneous leishmaniasis and 0.5 million due to visceral leishmaniasis (Murray 2005). However, these numbers are thought to be significantly under-estimated (Bern et al. 2008; Desjeux 2004; Mubayi 2009). Currently, the available drugs for the treatment for leishmaniases are far from satisfactory and are summarised in **Table 1.1** and **Figure 1.4**.

1.3. Glycolysis as a target for the design of novel anti-trypanosomatid drugs

Glycolysis is one of the most ancient known metabolic pathways and occurs in nearly all organisms. The pathway contains ten enzymes to convert glucose into pyruvate by a series of biochemical reactions, producing high-energy compounds ATP and NADH (Fothergill-Gilmore and Michels 1993; Romano and Conway 1996; **Figure 1.7**). The glycolytic pathway occurs in the cytosol of most organisms, however, an unusual organisation of glycolysis in trypanosomatid parasites has been identified (Gualdrón-López et al. 1982; Opperdoes 1990).

Compartmentation of glycolysis is the essential difference in glycolysis between trypanosomatids and most other organisms (Gualdrón-López et al. 1982). In 1977, the first seven glycolytic enzymes in the glycolytic pathway were for the first time described in *T. brucei* to be sequestered in a microbody-like organelle called the glycosome which is similar to a peroxisome (Opperdoes and Borst 1977; **Figure 1.7**). Later, glycosomes were found in other members of *Trypanosomatida* (Opperdoes 1990), including *T. cruzi* and *Leishmania* (Cannata et al. 1982; Hart and Operdoes 1984; Taylor et al. 1980).

These seven glycolytic enzymes located in glycosomes are represented in **Figure 1.7**: HK (hexokinase), PGI (glucose-6-phosphate isomerase), PFK (phosphofructokinase), ALD (aldolase), TIM (triosephosphate isomerase), GAPDH (glyceraldehydes-3-phophate dehydrogenase) and PGK (phosphoglycerate kinase)

(Gualdrón-López et al. 1982; Opperdoes and Borst 1977). The remaining last three glycolytic enzymes were located in cytosol (PGAM: phosphoglycerate mutase; ENO: enolase; PYK: pyruvate kinase) (Gualdrón-López et al. 1982; Opperdoes and Borst 1977; Verlinde et al. 2001; **Figure 1.7**).

This high degree of specialization in glycolysis may, in some case, make the glycolytic enzymes different from the enzymes from other organisms. In particular, it has been found that mechanisms of allosteric regulation are significantly different for trypanosomatids which have many of the glycolytic enzymes sequestered in the glycosome. For example, unlike the mammalian enzymes, HK and PFK from trypanosomatid parasites are unregulated by glucose-6-phosphate (G-6-P; product of HK reaction) and fructose 2,6-bisphosphate (F26BP), respectively (Berens et al. 1980; Cáceres et al. 2003; Cronin et al. 1985; Gualdrón-López et al. 1982; López et al. 2002; Opperdoes 1987; Racagni and Machado de Domenech 1983; Sols et al. 1981; Urbina and Crespo 1984). Although PYK is present in the cytosol of the parasite instead of glycosome, it also shows obviously different enzymatic behaviour compared to that from mammals. PYK is highly regulated in trypanosomatids by F26BP instead of fructose 1,6-bisphosphate (F16BP) which is the natural effector for PYKs (M2PYK, LPYK, RPYK) in mammals (Callens et al.1991; Cazzulo et al.1989; Ernest et al.1994; Juan et al. 1976; van Schaftingen et al.1985; van Schaftingen et al.1987).

When living in the host, trypanosomatids rely heavily on glycolysis to generate net ATP. Indeed bloodstream form *T. brucei* uses ATP solely from glycolysis (Albert et al. 2005; Cazzulo 1992; Hannaert et al. 2003; Opperdoes et al. 1987; Tielens and Van Hellemond, 1998). Therefore, glycolysis is a potential drug target for the design of new drugs to control the glycolytic flux by inhibiting glycolytic enzyme activity in order to kill the parasites (Verlinde et al. 2001). Additionally, the glycolytic enzymes in trypanosomatids and their mammal hosts have distinct properties due to their long evolutionary distance (Fernandes et al. 1993) and unusual organisation of glycolysis

in trypanosomatid parasites (Opperdoes 1990; Gualdrón-López et al. 1982). These differences provide potential opportunities to design highly selective inhibitors against parasitic enzymes.

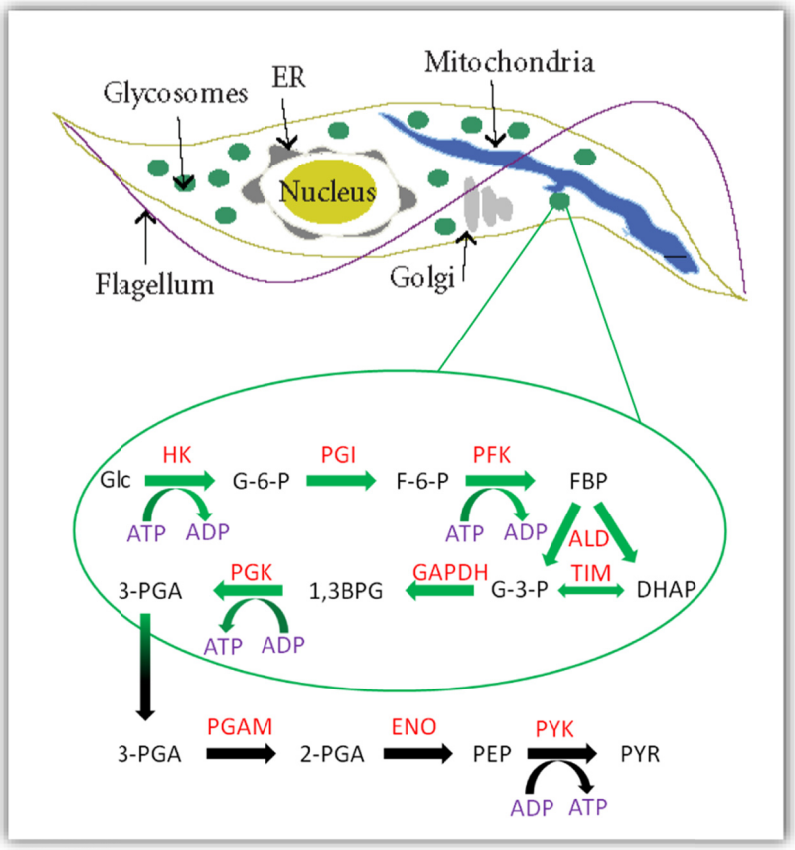


Figure 1.7 Glycolysis in the bloodstream form of *T. brucei*.
Abbreviations: HK: hexokinase; PGI: glucose 6-phosphate isomerase; PFK: phosphofructokinase; ALD: aldolase; TIM: triosephosphate isomerase; GAPDH: glyceraldehyde 3-phosphate dehydrogenase; PGK: phosphoglycerate kinase; PGAM: phosphoglycerate mutase; ENO: enolase; PYK: pyruvate kinase; Glc: glucose; G-6-P: glucose 6-phosphate; F-6-P: fructose 6-phosphate; FBP: fructose 1,6-bisphosphate; G-3-P: glyceraldehyde 3-phosphate; DHAP: dihydroxyacetone phosphate; 1,3BPGA: 1,3-bisphosphoglycerate; 3-PGA: 3-phosphoglycerate; 2-PGA: 2-phosphoglycerate; PEP: phosphoenolpyruvate; PYR: pyruvate
This modified figure was originally from Figure 1 in Coley et al. 2011.

1.4. Pyruvate kinase

1.4.1. Pyruvate kinase is a validated drug target

Glycolytic enzymes have been identified as potential drug targets in trypanosomatid parasites such as *Trypanosoma brucei*, *T. cruzi* and *Leishmania* spp. (Verlinde et al. 2001). Pyruvate kinase (PYK) (E.C. 2.7.1.40) catalyses the last reaction in the glycolytic pathway to yield adenosine triphosphate (ATP) and pyruvate from adenosine diphosphate (ADP) and phospho*enol*pyruvate (PEP) (**Figure 1.7**). Among the glycolytic enzymes, PYK has been validated by RNAi experiments as a suitable drug target in bloodstream-form *T. brucei* (Albert et al. 2005).

1.4.2. Similar homotetrameric architectures of PYKs

PYK is typically a tetrameric enzyme (**Figure 1.8a**), with each identical subunit made up of four domains (**Figure 1.8b**). The active site is located between the A- and B-domains, and is 40 Å distant from the effector site that is contained within the C-domain. Adjacent C-domains form the C-C or 'small' interface, as distinguished from the 'large' interface between adjacent A-domains. The B-domain has been observed in a variety of positions in different PYK crystal structures (Larsen et al. 1997; Morgan et al. 2010b; Tulloch et al. 2008), and forms a mobile lid over the active site. The relationship between B-domain position and active-site ligand binding will be discussed in **Chapter 5** and **Chapter 6**. The A- and C-domains comprise the relatively static AC-core of the enzyme.

An extra C-terminal domain (indicated by green box in the sequence alignment in **Figure 1.9**) has been observed in PYKs from bacteria *Geobacillus stearothermophilus* (Lovell et al. 1998; Suzuki et al. 2008) and *Staphylococcus aureus* (Axerio-Cilies et al. 2012; Zoraghi et al. 2011b). This extra C-terminal domain is suggested to weakly interact with the neighbouring monomer (A-domain

and C-domain) (Lovell et al. 1998; Suzuki et al. 2008) and the role of this extra domain will be further discussed in **Chapter 8**.

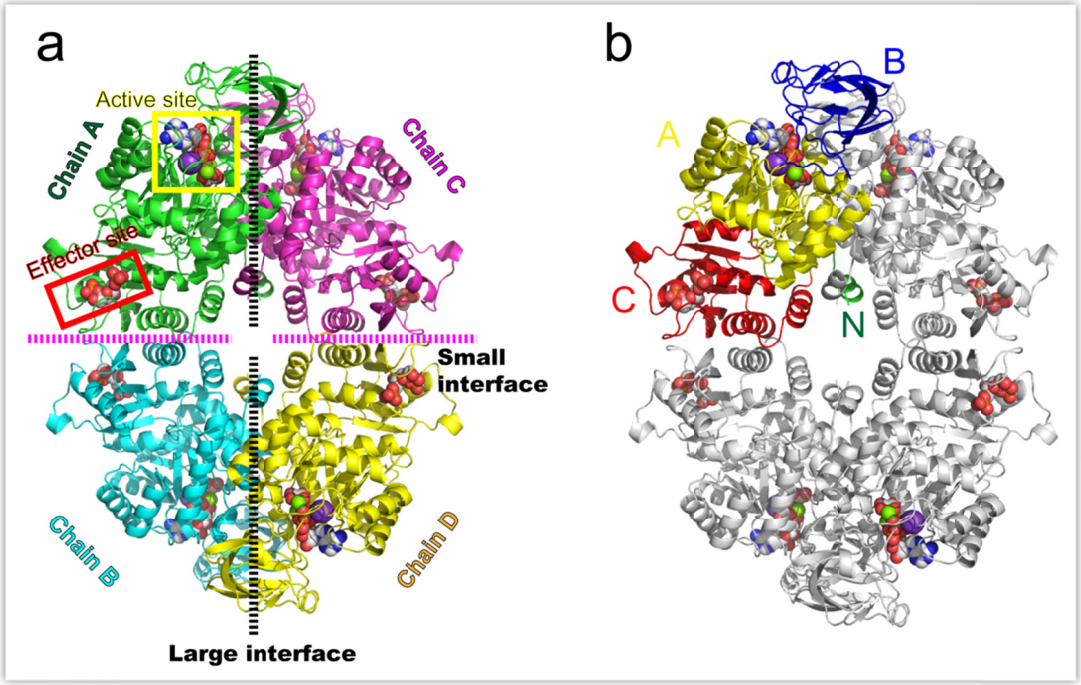


Figure 1.8 Tetramer architecture and domain boundaries of *Lm*PYK.
(a) The crystal structure of the *Lm*PYK/F26BP/OX/ATP/Mg complex (chains A, C, B, D; PDB code: 3HQP) shows the tetramer architecture, active site and effector site. Metals and ligands are represented by spheres. Each monomer has been coloured to aid the identification of subunit interfaces. The large (A-A) and small (C-C) interfaces between subunits are shown as dashed lines. (b) The *Lm*PYK/F26BP/OX/ATP/Mg tetramer where one subunit is coloured to show domains: N-terminal (green = residues 2-18), A-domain (yellow = residues 19-89, 188-358), B-domain (blue = residues 90-187), C-domain (red = residues 359-499). The residue numbering of *Lm*PYK includes the Met coded by the initiation codon AUG.

1.4.3. Sequence identity of PYKs from diverse species

PYKs are generally well conserved (particularly at the active site), and have retained similar sequences throughout evolution, with enzymes from even quite evolutionarily distant organisms showing 34-48% sequence identity (**Table 1.2**; **Figure 1.9**). The PYKs from trypanosomatids (*T. brucei*, *T. cruzi* and *Leishmania* spp.) are 74-81% identical in sequence. However, the effector sites of PYKs are varied among species compared to the highly conserved active sites (**Figure 1.9**). In particular, kinetic properties and the mechanism of activity regulation are rather distinct between trypanosomatid PYKs and human PYKs. These differences potentially become interesting drug targets to design selective inhibitors against trypanosomatid PYKs. Detailed discussion of this topic can be found in **Chapter 8**.

Table 1.2 Pairwise protein sequence comparisons of PYKs

	*Tb*PYK	*Tc*PYK	*Lm*PYK	*Hs*M2PYK	*Hs*M1PYK	*Hs*RPYK	*Hs*LPYK	*Oc*M1PYK	*Sc*PYK	*Ec*PYK	*Gs*PYK	*Sa*PYK
*Tb*PYK	100	81	74	48	47	45	43	48	48	42	35	34
*Tc*PYK		100	76	47	47	45	43	47	48	42	35	35
*Lm*PYK			100	47	48	45	42	47	48	42	34	34
*Hs*M2PYK				100	96	70	66	93	49	44	36	34
*Hs*M1 PYK					100	68	65	97	48	44	36	34
*Hs*RPYK						100	94	100	47	43	35	34
*Hs*LPYK							100	100	47	43	35	34
*Oc*M1PYK								100	47	43	35	34
*Sc*PYK									100	43	34	35
*Ec*PYK										100	42	39
*Gs*PYK											100	60
*Sa*PYK												100

The pairwise sequence analysis was obtained from the EMBL-EBI web server (EMBOSS Stretcher):
http://www.ebi.ac.uk/Tools/psa/emboss_stretcher/
Values are overall percent sequence identities.
Tb, T. brucei; Tc, T. cruzi; Lm, L. mexicana; Hs, Homo sapiens; Oc, Oryctolagus cuniculus, (rabbit); Sc, Saccharomyces cerevisiae; Ec, Escherichia coli; Gs, Geobacillus stearothermophilus; Sa, Staphylococcus aureus subsp. aureu

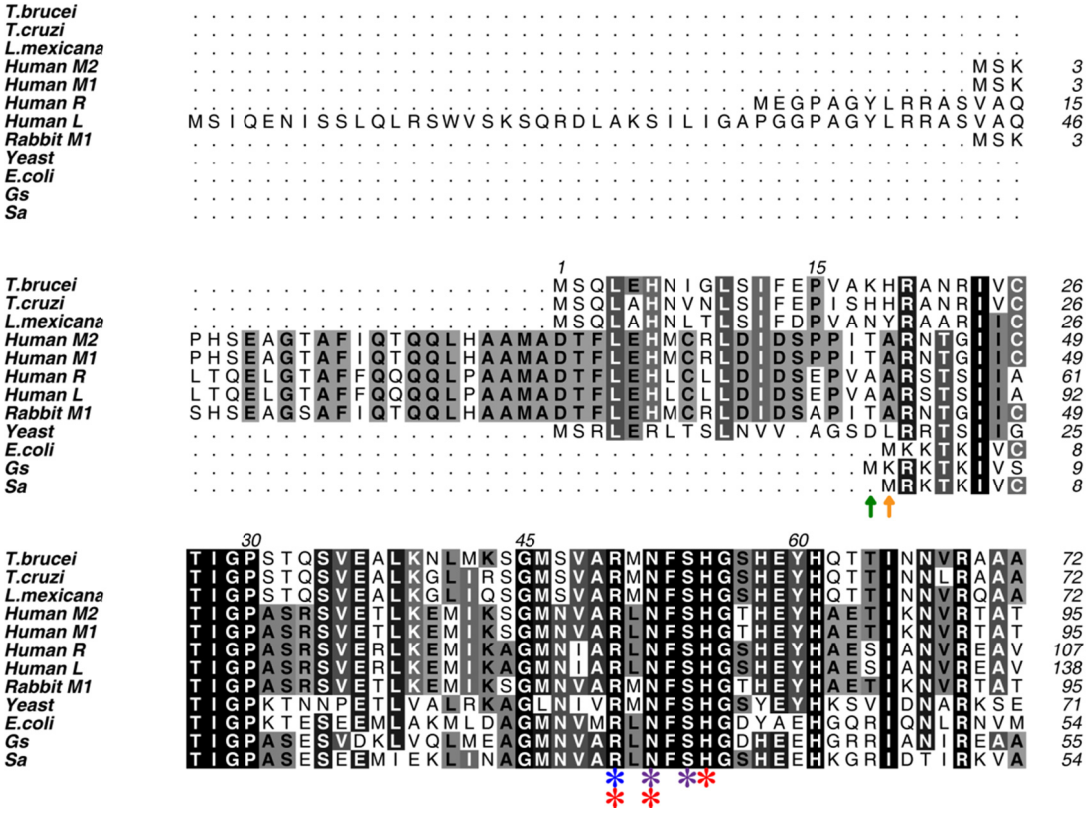


Figure 1.9 Sequence alignment of pyruvate kinases.

The alignment includes PYKs from *T. brucei, T. cruzi, L. mexicana,* human M2, human M1, human R, human L, rabbit M1, yeast, *E. coli, Geobacillus stearothermophilus (Gs) and* methicillin-resistant *Staphylococcus aureus (Sa).*The sequence alignment was performed using the program Clustal Omega at the European Bioinformatics Institute (Sievers et al 2011; Goujon et al. 2010). Domain boundaries are indicated by vertical arrows in domain-specific colours. The conservation of the residues is indicated by shading from black (identical in twelve or eleven sequences) to grey (conserved from five to ten) to white (low or no conservation). Residue numbers corresponding to each PYK are listed after the sequences. The residue numbers for *Tb*PYK are shown above the sequence alignment as well. At the active site, the amino acids involved in divalent metal ion binding (*, in green), potassium metal ion binding (*, in purple), substrate PEP binding (*, in blue) and nucleotide binding (*, in red) are indicated by asterisks. In the poorly conserved effector site, ligand binding residues of trypanosomatid PYKs are indicated by pink asterisks (*). The small α-helix Aα6′ which are involved in allosteric regulation and in binding divalent metal and the substrate PEP are indicated by a dashed box (cyan). The effector loop residues are indicated by a pink dashed box. The extra C-terminal sequence (ECTS) domain (C'-domain) which potentially plays a role in tetramer stabilisation are shown by a green box. The figure was generated using the program Aline (Bond and Schüttelkopf 2009).

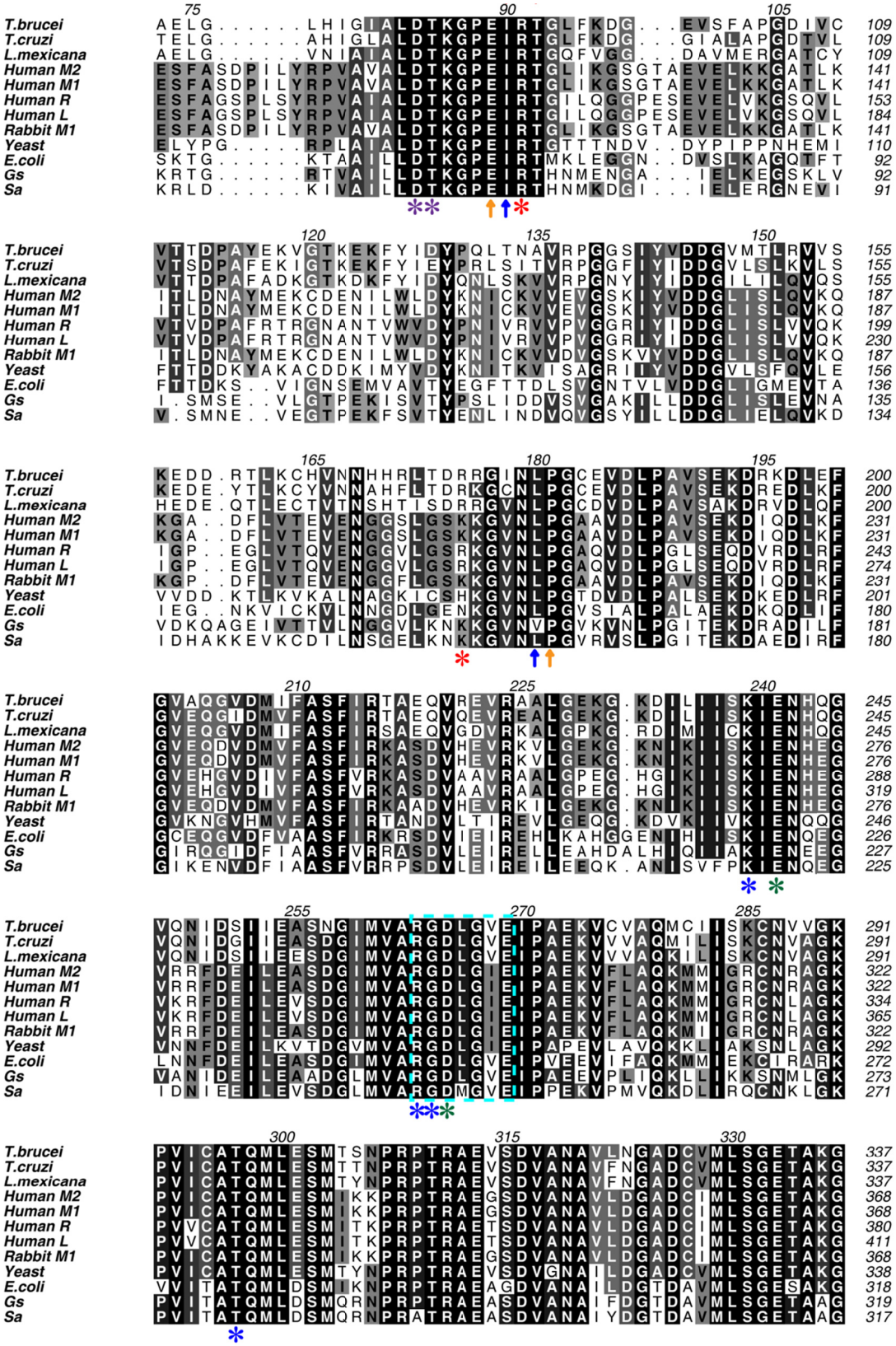


Figure 1.9 (*continued)*

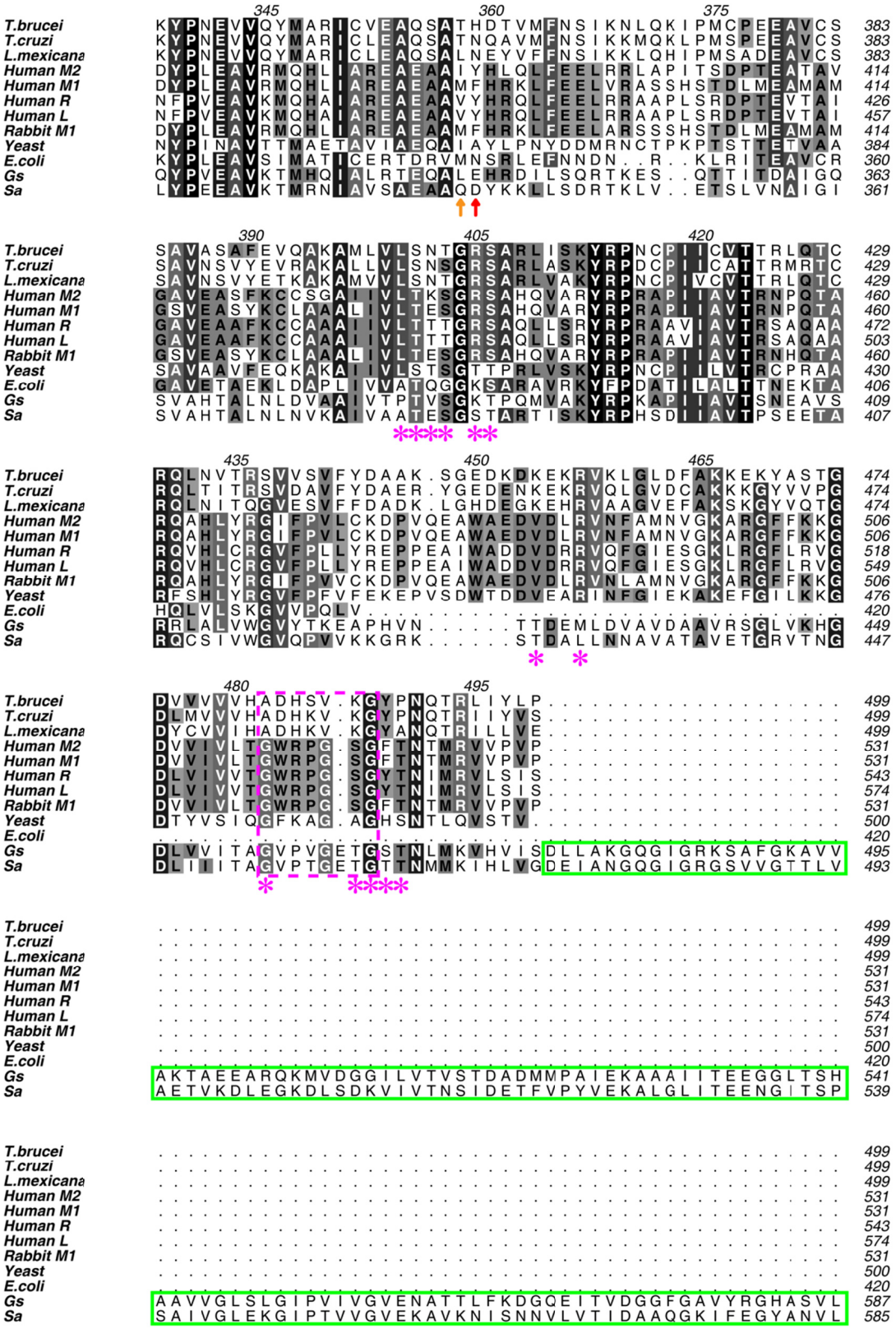


Figure 1.9 (*continued*)

1.4.4. Previous studies on PYKs

Enzyme kinetic properties of PYKs from trypanosomatids (*T. cruzi*, *T. brucei* and *L. mexicana*), human [R isoform (erythrocyte), L isoform (liver), M2 isoform (embryonic or tumour), and M1 isoform (muscle)], yeast, and bacteria (*E. coli* and *G. stearothermophilus*) are summerised in **Table 1.3** (the data for *T. cruzi* and *T. brucei* are obtained from the work in this thesis). The PYK structures deposited in the Protein Data Bank (PDB, website link: http://www.rcsb.org) are summarised in **Table 1.4**. A total of 15 X-ray crystal structures of *T.cruzi* PYK (*Tc*PYK) and *T. brucei* PYK (*Tb*PYK) determined from the work in this thesis will be summarised and presented in **Table 2.1** of **Chapter 2**.

PYKs are either constitutively active (as exemplified by the muscle isoenzyme M1 in mammals) or are allosterically regulated by effectors (essentially all other forms of PYK). For example, human isoenzymes RPYK, LPYK and M2PYK are allosterically regulated by F16BP (**Table 1.3**) but the isoenzyme M1PYK is unregulated by F16BP (Fothergill-Gilmore and Michels 1993). Another example is *Geobacillus stearothermophilus* PYK (*Gs*PYK) which is allosterically regulated by adenosine monophosphate (AMP) (**Table 1.3**) or ribose 5-phosphate (R5P) but not F16BP (Sakai and Ohta 1993).

Trypanosomatid PYKs differ from human PYKs and *Gs*PYK by being allosterically activated by F26BP with nanomolar affinity (**Table 1.3**). The detailed allosteric mechanism for *L. mexicana* PYK (*Lm*PYK) has been elucidated by Morgan et al. 2010b (**Figure 1.10**), from four crystal structures of *Lm*PYK (**Table 1.4**) including apoenzyme *Lm*PYK (3HQN) and *Lm*PYK complexed with active site ligands (oxalate and ATP) and/or effector F26BP (3HQO, 3HQP, 3HQQ). In this 'rock-and-lock' mechanism, the conformational transition of *Lm*PYK from T- to R-state initiated by F26BP and/or oxalate binding requires a 6^{o} concerted rigid body rotation of each of the four AC-cores in the tetramer, followed by the formation of eight salt-bridges across the C-C interface to lock the enzyme in a

thermodynamically more stable conformation.

Table 1.3 Comparison of kinetic properties of PYKs [#]

Ligand	Modulator	Kinetic properties Parameter	*Tc*PYK[†] 25°C	*Tb*PYK[†] 25°C	*Lm*PYK 37°C	*Hs*RPYK 37°C (Wang et al. 2001)	*Hs*LPYK 25°C (Imamura and Tanaka 1982)	*Hs*M2PYK 37°C	*Hs*M1PYK 37°C	*Sc*PYK 25°C (Collins et al. 1995)	*Ec*PYK 25°C (Valentini et al. 2000)	*Gs*PYK*
PEP	None	$S_{0.5}$(mM)	1.2	1.03	1.06	1.1	0.96	0.86	0.05	2.76	3.63	0.96
		h	2.1	1.88	1.5	1.6	2.0	1.2	1.0	2.93	3.2	2.2
		k_{cat}	169	145	296	355	494	117	347	188	123	187
		$k_{cat}/S_{0.5}$	138	141	279	323	515	136	6939	68	40	85
	F26BP	$S_{0.5}$(mM)	0.14	0.12	0.11	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
		h	1.13	1.2	1.0	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
		k_{cat}	212	231	300	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
		$k_{cat}/S_{0.5}$	1548	1942	2727	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
	F16BP	$S_{0.5}$(mM)	0.35	0.26	0.84[‡]	0.18	0.3	0.10	0.05	0.22	0.08	0.28 (AMP)
		h	2.17	1.31	2.68[‡]	1.05	1.6	1.0	1.0	1.7	1.0	1.0 (AMP)
		k_{cat}	226	186	417[‡]	355	494	222	357	195	160	187 (AMP)
		$k_{cat}/S_{0.5}$	651	711	496[‡]	1972	1647	2218	7140	886	2000	668 (AMP)
ADP	None	K_m(mM)	0.40	0.14	0.10[‡]	0.17	0.4	0.34	0.37 (Ikeda et al. 1997)	0.25	0.30	n.d.
F26BP	None	$Ka_{0.5}$(μM)	0.03	0.01 (Ernest et al. 1998)	0.29[‡]	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
F16BP	None	$Ka_{0.5}$(mM)	1.24	0.05	0.52[‡]	0.0004	0.0001	0.006	n.d.	n.d.	n.d.	n.d.

k_{cat} values in s^{-1}; $k_{cat}/S_{0.5}$ values in $s^{-1}\cdot mM^{-1}$; h, Hill coefficient

Tc, T. cruzi; Tb, T. brucei; Lm, L. mexicana; Hs, Homo sapiens; Sc, Saccharomyces cerevisiae; Ec, Escherichia coli; Gs, Geobacillus stearothermophilus

[#] This modified table is originally obtained from my colleague Dr. Hugh Morgan; unless stated otherwise, values are provided by Dr. Hugh Morgan and unpublished.

[†] Values are obtained from the work in this thesis unless stated otherwise.

[‡] Values are obtained from Ernest et al. 1994.

* V_{max} (average) is 200 $\mu mol^{-1}\cdot min^{-1}\cdot mg^{-1}$ (Suzuki et al. 2008) and converted to k_{cat} assuming a monomer with mass of 56,000; all other values are obtained from Sakai 2004. The modulator for *Gs*PYK is AMP instead of F16BP. The temperature for the assay is not stated in the references.

n.d., not done

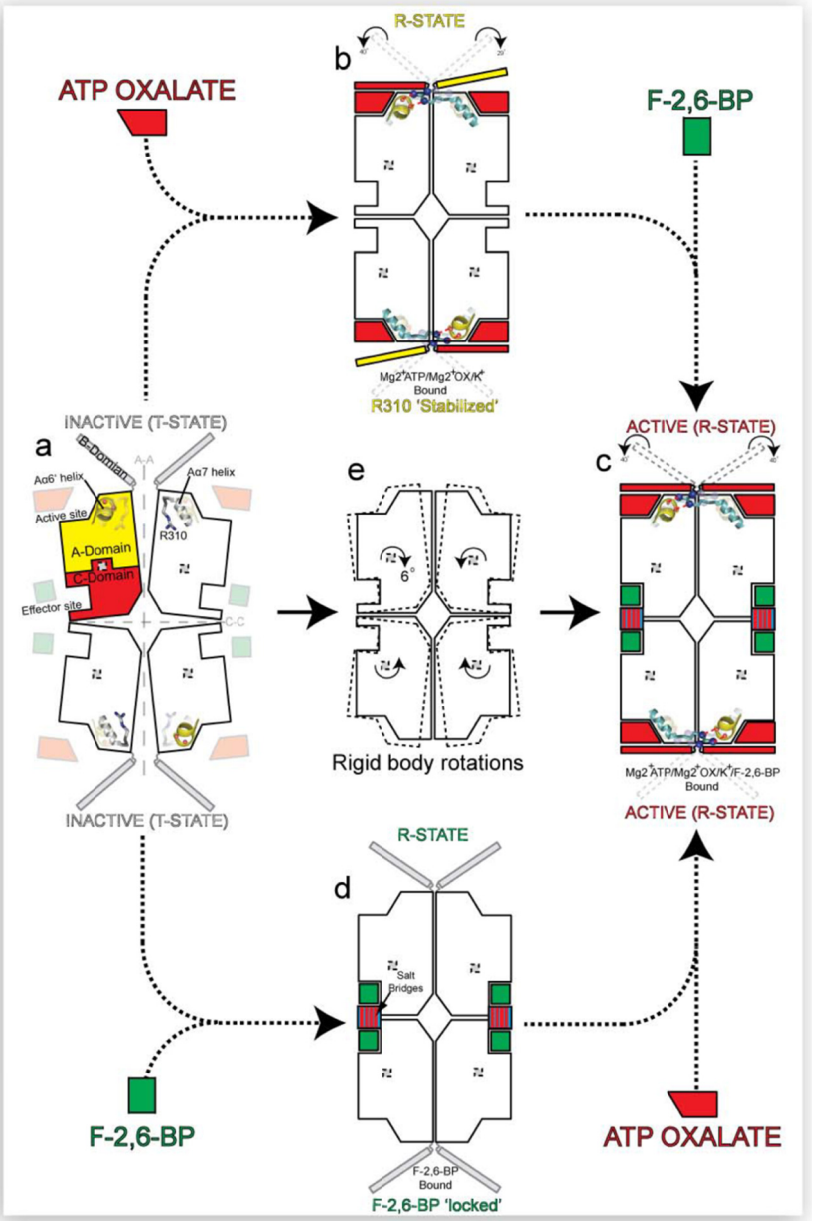


Figure 1.10 Schematic representations reveal the allosteric mechanism of *Lm*PYK.

This figure and modified legend were obtained from Figure 2 in Morgan 2010b. (a) Schematic representation of the inactive T-state *Lm*PYK structure (corresponding to the crystal structure of apo *Lm*PYK with PDB code 3HQN). The relative positions of the AC core (A- and C-domains), pivot point (grey circle), active site, effector site, Arg311 (R311), Aα6' and Aα7 helices are shown for the *Lm*PYK tetramer. (b) The binding of ATP and oxalate (red lozenges) to the apoenzyme (corresponding to the crystal structure of *Lm*PYK/OX/ATP/Mg with PDB code 3HQO) initiates Arg311 to move into the R-state position, forming hydrogen bonds (shown as dashed red lines) with the active-site Aα6' helix of the adjacent chain and a stabilising bridge at the A-A interface, stabilising the AC core in an R conformation. (c) The active R-state

*Lm*PYK/ATP/OX/F26BP/Mg structure (corresponding to PDB code 3HQP). All monomers have rotated AC cores, fully closed B domains, F-2,6-BP (F26BP) induced salt bridges across the C-C interface and stabilising Arg311…Aα6' hydrogen bonds across the A-A interface. (d) The binding of F26BP to the apoenzyme (corresponding to the crystal structure of *Lm*PYK/F26BP with PDB code 3HQQ) results in the formation of four pairs of stabilising salt bridges (shown as red lines) formed across each C-C interface, which lock the AC core in an R-state conformation. The *Lm*PYK/F26BP structure has fully open B-domains. (e) Superposition of the AC cores from T- (dashed lines) and R-state tetramers.

Table 1.4 X-ray crystal structures of PYK[†]

Species	PDB Code	Resolution (Å)	Active site ligands	Effector site ligands	Other ligands	Reference
L. mexicana	1PKL	2.35	none	SO_4^{2-}	none	Rigden et al. 1999
L. mexicana	3E0V	3.30	SO_4^{2-}	SO_4^{2-}	none	Tulloch et al. 2008
L. mexicana	3E0W	3.10	none	none	none	Tulloch et al. 2008
L. mexicana	3KTX	2.10	none	none	PTK	Morgan et al. 2010a
L. mexicana	3IS4	2.10	none	none	PTK	Morgan et al. 2010a
L. mexicana	3HQN	2.00	K^+	SO_4^{2-}	none	Morgan et al. 2010b
L. mexicana	3HQO	3.40	K^+, Mg^{2+} oxalate, ATP	none	none	Morgan et al. 2010b
L. mexicana	3HQP	2.30	K^+, Mg^{2+} oxalate, ATP	F26BP	none	Morgan et al. 2010b
L. mexicana	3HQQ	5.00	none	F26BP	none	Morgan et al. 2010b
L. mexicana	3PP7	2.35	K^+, suramin	none	PTK	Morgan et al. 2011
L. mexicana	3QV6	2.85	K^+, AB80	PO_4^{3-}	none	Morgan et al. 2011
L. mexicana	3QV7	2.70	K^+, ponceau S	SO_4^{2-}	QV8	Morgan et al. 2011
L. mexicana	3QV8	2.45	K^+, BDS	PO_4^{3-}	none	Morgan et al. 2011
L. mexicana	3SRK	2.65	K^+, DBS	none	PTK	Morgan et al. 2012a
E. coli	1PKY	2.50	none	none	none	Mattevi et al. 1995
E. coli R292D	1E0T	1.80	none	SO_4^{2-}	none	Valentini et al. 2000
E. coli R271L	1E0U	2.00	none	SO_4^{2-}	none	Valentini et al. 2000
Geobacillus stearothermophilus	2E28	2.40	none	SO_4^{2-}	none	Suzuki et al. 2008
Staphylococcus aureus	3T05	3.05	none	PO_4^{3-}	none	Zoraghi et al. 2011b
Staphylococcus aureus	3T07	3.30	none	PO_4^{3-}	09C	Zoraghi et al. 2011b
Staphylococcus aureus	3T0T	3.10	none	PO_4^{3-}	IS-130	Axerio-Cilies et al. 2012
Pyrobaculum aerophilum	3QTG	2.20	none	SO_4^{2-}	none	Solomons et al. 2012
Toxoplasma gondii	3EOE	2.30	none	none	none	Bakszt et al. 2010
Cryptosporidium parvum	4DRS	2.50	none	SO_4^{2-}	acetate	Cook et al. 2012
Cryptosporidium parvum	3MA8	2.64	citrate	SO_4^{2-}	none	Wernimont et al. 2012
S. cerevisiae	1A3W	3.00	K^+, Mn^{2+}, PG	F16BP	none	Jurica et al. 1998
S. cerevisiae	1A3X	3.00	K^+, Mn^{2+}, PG	none	none	Jurica et al. 1998
Cat (M1)	1PYK	2.60	none	none	none	Stuart et al. 1979
Cat (M1)	1PKM	2.60	none	none	none	Allen and Muirhead 1996

Table 1.4 ***(continued)***

Rabbit (M1)	1PKN	2.90	K^+, Mn^{2+} pyruvate	none	none	Larsen et al. 1994
Rabbit (M1)	1AQF	2.70	K^+, Mg^{2+} phospholactate	none	none	Larsen et al. 1997
Rabbit (M1)	1A49	2.10	K^+, Mg^{2+} oxalate, ATP	none	none	Larsen et al. 1998
Rabbit (M1)	1A5U	2.35	K^+, Mg^{2+} oxalate	none	none	Larsen et al. 1998
Rabbit (M1)	1F3W	3.00	K^+, Mn^{2+} pyruvate	none	none	Wooll et al. 2001
Rabbit (M1) S402P	1F3X	2.80	K^+, Mn^{2+} pyruvate	none	none	Wooll et al. 2001
Rabbit (M1)	2G50	1.65	Na^+, Mn^{2+} pyruvate	alanine	none	Williams et al. 2006
Human erythrocyte	2VGB	2.70	K^+, Mn^{2+}, PG	F16BP	none	Valentini et al. 2002
Human erythrocyte T384M	2VGF	2.75	K^+, Mn^{2+}, PG	F16BP	none	Valentini et al. 2002
Human erythrocyte R479H	2VGG	2.74	K^+, Mn^{2+}, PG	F16BP	none	Valentini et al. 2002
Human erythrocyte R486W	2VGI	2.87	K^+, Mn^{2+}, PG	F16BP	none	Valentini et al. 2002
Human M1	3N25	2.41	K^+, Mn^{2+} pyruvate	none	proline, Na^+	Fenton et al. 2010
Human M1	3SRF	2.85	K^+, Mg^{2+} pyruvate	PO_4^{3-}	none	Morgan et al. 2012b
Human M2	1ZJH	2.20	none	none	none	Atanassova et al. 2005
Human M2	1T5A	2.80	K^+, Mg^{2+} oxalate, PO4	F16BP	none	Dombrauckas et al. 2005
Human M2	3BJT	2.50	Mg^{2+}, oxalate	none	none	Christofk et al. 2008
Human M2	3BJF	2.03	K^+, Mg^{2+} oxalate	F16BP	none	Christofk et al. 2008
Human M2	3H6O	2.00	none	F16BP	D8G	Hong et al. 2009
Human M2	3GQY	1.85	tartrate	F16BP	DZG	Hong et al. 2009
Human M2	3GR4	1.60	tartrate, ADP	F16BP	DYY	Hong et al. 2009
Human M2 S437Y	3G2G	2.00	none	SO_4^{2-}	SO_4^{2-}	Hong et al. 2009
Human M2	4B2D	2.30	Mg^{2+}	F16BP	serine	Chaneton et al. 2012
Human M2	4G1N	2.30	Mg^{2+}, oxalate	none	NZT	Kung et al. 2012
Human M2	3U2Z	2.10	none	F16BP	TEPP-46, SO4	Anastasiou et al. 2012
Human M2	3ME3	1.95	none	F16BP	DASA-58	Anastasiou et al. 2012
Human M2	3SRD	2.90	K^+, Mg^{2+} oxalate	F16BP	none	Morgan et al. 2012c
Human M2	3SRH	2.60	none	PO_4^{3-}	none	Morgan et al. 2012d

Abbreviations:
PTK, pyrene-1,3,6,8-tetrasulfonic acid;
PG, phosphoglycollate;
AB80, acid blue 80, 3-({4-[(2,4-dimethyl-5-sulfophenyl)amino]-9,10-dioxo-9,10-dihydroanthracen-1-yl}amino)-2,4,6-trimethylbenzenesulfonic

acid;
QV8, acid blue 25, 9,10-dioxo-4-(phenylamino)-9,10-dihydroanthracene-2-sulfonic acid;
BDS, 1,3-benzothiazole-2,5-disulfonic acid;
DBS, 4-[(1,1-dioxo-1,2-benzothiazol-3-yl)sulfanyl]benzoic acid;
09C, *cis*-3,4-dihydrohamacanthin B (Bao et al. 2005; Casapullo et al. 2000);
IS-130, N'-[(1E)-1-(1H-benzimidazol-2-yl)ethylidene]-
5-bromo-2-hydroxybenzohydrazide;
D8G, 6-(2-fluorobenzyl)-2,4-dimethyl-4,6-dihydro-
5H-thieno[2',3':4,5]pyrrolo[2,3-d]pyridazin-5-one;
DZG, 1-(2,3-dihydro-1,4-benzodioxin-6-ylsulfonyl)-
4-[(4-methoxyphenyl)sulfonyl]piperazine;
DYY, 1-[(2,6-difluorophenyl)sulfonyl]-4-(2,3-dihydro-
1,4-benzodioxin-6-ylsulfonyl)piperazine;
NZT, N-(4-{[4-(pyrazin-2-yl)piperazin-1-yl]carbonyl}phenyl)quinoline-
8-sulfonamide;
TEPP-46, thieno-[3,2-b]pyrrole[3,2-d]pyridazinone (Jiang et al. 2010);
DASA-58, substituted N,N'-diarylsulfonamide (Boxer et al. 2010);
† Glycerol which is commonly used to enhance crystal growth is not included in this table.

1.4.5. Discovery of inhibitors or activators against PYKs

PYKs not only play roles in some diseases but also become interesting drug targets for the treatment of several diseases, such as cancer (Christofk et al. 2008a), and infectious diseases caused by trypanosomatid parasites (Nowicki et al. 2008), bacteria methicillin-resistant *S. aureus* (Zoraghi et al. 2011b), and the malaria parasites *Plasmodium* spp. (Ayi et al. 2008). Especially for cancer which is a proliferative disease, tumour-specific PYK (M2PYK) is allosterically regulated by F16BP and the properties of this enzyme may facilitate the accumulation of phosphometabolites which are required for the cancer cell to proliferate. Drug discovery targeting of M2PYK to develop M2PYK-specific inhibitors or activators could be of potential interest in elucidating the etiology of cancer and designing novel treatment tools.

PYK has been implicated as an interesting target for the treatment of many

diseases, however, high-affinity and selective inhibitors or activators against PYK are still limited. **Table 1.5** and **Figure 1.11** summarise some inhibitors or activators against parasite PYKs, bacterial PYKs, or human PYKs. The details of interactions between protein and ligand have been revealed by X-ray crystal structures. Inhibitors for methicillin-resistant *S. aureus* PYK (MRSA PYK) show nanomolar inhibition and some of the inhibitors have been proved in crystal structures to bind at the C-C (small) interface of the enzyme resulting in high selectivity of the inhibitors (Axerio-Cilies et al. 2012; Zoraghi et al. 2011a; Zoraghi et al. 2011b). Nanomolar activators for human M2PYK show inhibition of cell proliferation *in vivo*, and some of the activators are confirmed to bind at the A-A (large) interface in crystal structures (Anastasiou et al. 2012; Hong et al. 2009; Kung et al. 2012). Inhibitors of human M2PYK which are not included in **Table 1.5** have been reported as well, but have relatively low affinity (IC50 = 10-20 μM) and show cross inhibition of human LPYK (Vander Heiden et al. 2010).

Unfortunately, efficient and selective inhibitors for trypanosomatid PYKs are still not available. Oxalate is a structural analogue of the enolate form of pyruvate and is a general inhibitor for PYKs, but with poor specificity and weak affinity (K_i = 220 μM for human M2PYK) (Dombrauckas et al. 2005). Recently, the old drug suramin (**Table 1.1**) is reported to mimic ADP/ATP and binds competitively with ADP to the active site of trypanosomatid PYKs at micromolar affinity (Morgan et al. 2011b; **Table 1.5**; **Figure 1.11**). However, this compound inhibits human PYKs as well due to the highly conserved active site between species. Another micromolar inhibitor has been identified that selectively reacts with Lys336 at the active site of *Lm*PYK, to form a covalent bond and sterically hinder the binding of ADP/ATP (Morgan et al. 2012a; **Table 1.5**; **Figure 1.11**). Nevertheless, the active lysine residue is conserved in human PYKs, resulting in cross inhibition of human PYKs (**Table 1.5**). Therefore, obtaining selective inhibitors against trypanosomatid PYKs is urgent for drug discovery of trypanosomiasis and leishmaniases.

Table 1.5 Inhibitors/Activators against PYKs

Compound	Type	PYK target	*in vitro* assay		Validated binding site	PYK complex (PDB ID)	Reference
			IC_{50} (μM)	AC_{50}/AC_{90} (μM)			
Suramin	Inhibitor	*Lm*PYK	160 80†	n.a.	ATP/ADP binding site‡ (active site)	3PP7	Morgan et al. 2011
		*Tc*PYK	150	n.a.		n.d.	
		*Hs*M2PYK	33 17†	n.a.		n.d.	
		*Hs*M1PYK	20 18†	n.a.		n.d.	
		*Hs*LPYK	2.2 1.8†	n.a.		n.d.	
DBS	Covalent inhibitor	*Lm*PYK	2.9	n.a.		3SRK	Morgan et al. 2012a
		*Hs*M2PYK	8	n.a.		n.d.	
		*Hs*RPYK	16.3	n.a.		n.d.	
NSK5-15*	Inhibitor	MRSA PYK	0.185	n.a.	Speculate to be at C-C interface	n.d.	Zoraghi et al. 2011a
09A			1-10	n.a.		n.d.	Zoraghi et al. 2011b
09D			0.06	n.a.		n.d.	
09C			0.016	n.a.	C-C interface	3T07	
IS-130*			0.091	n.a.		3T0T	Axerio-Cilies et al. 2012
D8G	Activator	*Hs*M2PYK	n.a.	n.d.	A-A interface	3H6O	Hong et al. 2009
DZG			n.a.	n.d.		3GQY	
DYY			n.a.	n.d.		3GR4	
NZT*			n.a.	0.07/n.d.		4G1N	Kung et al. 2012
TEPP-46			n.a.	0.092/0.47		3U2Z	Anastasiou et al. 2012
DASA-58			n.a.	0.038/0.68		3ME3	

n.a., not applicable; n.d., not done

IC_{50}, half maximal inhibitory concentration; K_i, binding affinity of the inhibitor

AC_{50}, half-maximum activating concentration; AC_{90}, 90% maximum activating concentration

† The assay was done in the presence of activator (F26BP for *Lm*PYK, F16BP for the other PYKs).

‡ The compound binding sites on PYKs are considered to be the same due to the highly conserved active site of PYKs (**Figure 1.9**).

* Refer to the corresponding reference for more similar inhibitors or activators which are not described in the table.

Abbreviations:

MRSA, methicillin-resistant *Staphylococcus aureus (S. aureus)*;

DBS, 4-[(1,1-dioxo-1,2-benzothiazol-3-yl)sulfanyl]benzoic acid;

09A, spongotine A (Bao et al. 2007; Murai et al. 2007);

09D, bromodeoxytopsentin (Bao et al. 2005; Shin et al. 1999);

09C, *cis*-3,4-dihydrohamacanthin B (Bao et al. 2005; Casapullo et al. 2000);
IS-130, N'-[(1E)-1-(1H-benzimidazol-2-yl)ethylidene]-
5-bromo-2-hydroxybenzohydrazide;
D8G, 6-(2-fluorobenzyl)-2,4-dimethyl-4,6-dihydro-
5H-thieno[2',3':4,5]pyrrolo[2,3-d]pyridazin-5-one;
DZG, 1-(2,3-dihydro-1,4-benzodioxin-6-ylsulfonyl)-
4-[(4-methoxyphenyl)sulfonyl]piperazine;
DYY, 1-[(2,6-difluorophenyl)sulfonyl]-4-(2,3-dihydro-
1,4-benzodioxin-6-ylsulfonyl)piperazine;
NZT, N-(4-{[4-(pyrazin-2-yl)piperazin-1-yl]carbonyl}phenyl)quinoline-
8-sulfonamide;
TEPP-46, thieno-[3,2-b]pyrrole[3,2-d]pyridazinone (Jiang et al. 2010);
DASA-58, substituted N,N'-diarylsulfonamide (Boxer et al. 2010)

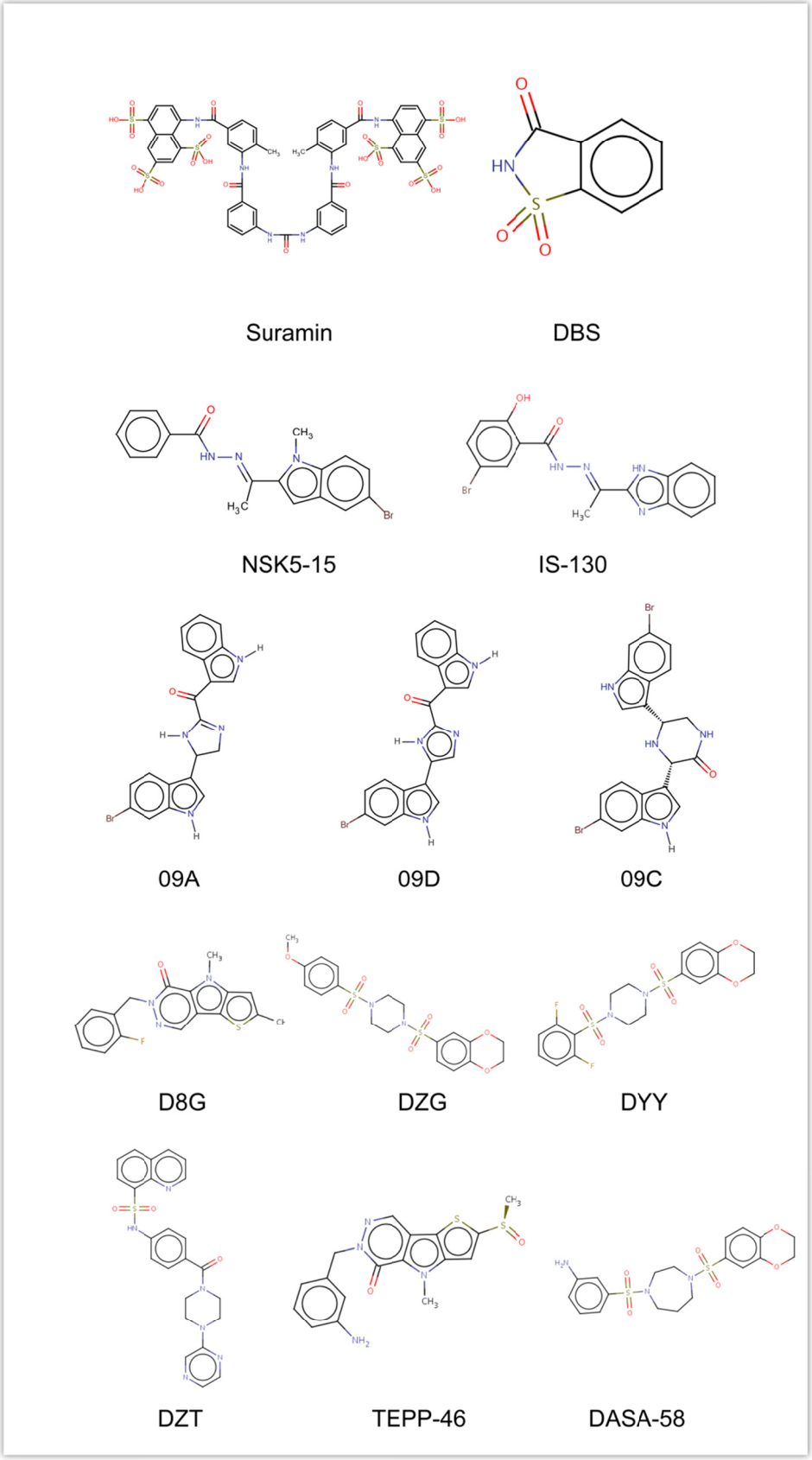


Figure 1.11 Two-dimensional representations of PYK inhibitors or activators, corresponding to Table 1.4.

1.5. The aims and general methods in this thesis

Although *Lm*PYK has been extensively studied in the past few years, relatively little work has been done on *Trypanosoma* PYKs (*Tc*PYK and *Tb*PYK are described in this thesis), partially due to the difficulty of obtaining pure enzyme of high quality.

To understand the unique structural properties and allosteric mechanism of the enzymes, many methods have been developed in the work of this thesis (**Figure 1.12**): the methods of cloning, protein expression and purification were developed and optimised to obtain large quantities of pure and active *Tc*PYK and *Tb*PYK expressed in *E.coli*; enzyme kinetic assay, thermal shift assay and enzyme inhibition/activation assay were set up (mainly in plate-reader format) to characterise the purified enzyme and to screen potential inhibitors and activators; X-ray crystallography was used to understand the structural details of the enzyme and its allosteric mechanism. Additionally, a crystal soaking method was introduced into the study of ligand-protein interactions for mapping interesting binding sites which could be hotspots to design selective inhibitors. Thereafter the knowledge of enzyme structures and allosteric mechanism could be used on structure-based and allostery-based drug design.

The project outlines and detailed methods for this thesis will be presented in the next chapter (**Chapter 2**).

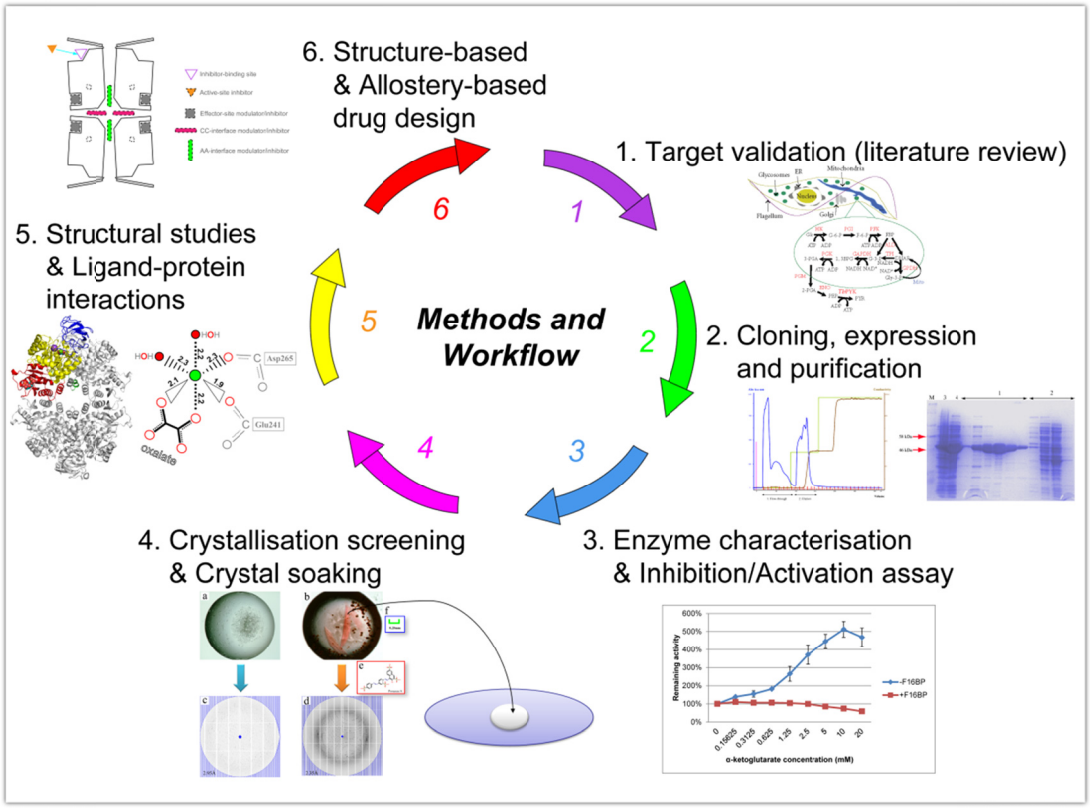


Figure 1.12 A representation of workflow for this thesis

CHAPTER 2: Project Outlines and Methods

2.1. Outline of X-ray crystal structures of PYK determined in this thesis

X-ray crystal structures of pyruvate kinases (PYKs) from *T. cruzi* and *T. brucei* have been determined for the first time and are discussed in this thesis. A total of 15 novel PYK structures including *Tc*PYK and *Tb*PYK with and without bound ligand(s) are summarised in **Table 2.1**. The corresponding protein crystals were obtained by either co-crystallisation or crystal soaking.

2.2. Project outlines

2.2.1. Purification and characterisation of TcPYK and TbPYK (Chapter 3)

Very few *Tc*PYK publications were reported previously, and the enzyme kinetic studies were performed on partially pure enzyme which was extracted and purified from parasites (Cazzulo et al. 1989). *Tb*PYK has been studied more extensively; however, the expression and purification procedures to obtain pure and active *Tb*PYK were still not efficient (Flynn and Bowman 1980; Flynn and Bowman 1981; Barnard and Pedersen 1988; Callens et al. 1991; Callens and Opperdoes 1992; Ernest et al. 1998). To solve these problems, a novel strategy to express and purify bacterially expressed *Tc*PYK and *Tb*PYK from *E. coli* has been developed, and the kinetic properties of both enzymes have been fully characterised.

Publications on *Tc*PYK from this project:

Morgan, H. P., McNae, I. W., Nowicki, M. W., **Zhong, W.**, Michels, P. A., Auld, D. S., Fothergill-Gilmore, L. A., Walkinshaw, M. D. (2011) The trypanocidal drug suramin and other trypan blue mimetics are inhibitors of pyruvate kinases and bind to the adenosine site. *J. Biol. Chem.* **286**, 31232-31240.

Morgan, H. P. [*], **Zhong, W.** [*], *et al*. Evolutionary divergence of pyruvate kinase allostery. (Manuscript in preparation) [*] Co-first authors

Publication on *Tb*PYK from this project:

Zhong, W.[*], Morgan, H. P.[*], McNae, I. W., Michels, P. A., Fothergill-Gilmore, L. A., Walkinshaw, M. D. *'In crystallo'* substrate binding triggers major domain movements and reveals magnesium as a co-activator of *Trypanosoma brucei* pyruvate kinase. (Manuscript submitted to *Acta Crystallogr. D*) [*] Co-first authors

2.2.2. Optimisation of crystallisation for both TcPYK and TbPYK to obtain reproducible and high-quality diffracting crystals (Chapter 4)

Neither the crystal structure of *Tc*PYK nor *Tb*PYK were reported before this work. In particular for *Tb*PYK [which is a validated drug target against *T. brucei* (Drew et al. 2003; Albert et al. 2005)], many attempts have been made since 1998 to try to obtain high-quality diffracting crystals (Ernest et al. 1998) but without success. To solve these problems, large-scale crystallisation trials were set up in combination with the addition of various dye molecules. Ponceau S was eventually selected as a 'silver bullet' to crystallise *Tc*PYK and *Tb*PYK to high quality.

The structure of *Tc*PYK in complex with ponceau S has been deposited in the PDB (*Tc*PYK/ponceau S: 3QV9) and published in:

Morgan, H. P., McNae, I. W., Nowicki, M. W., **Zhong, W.**, Michels, P. A., Auld, D. S., Fothergill-Gilmore, L. A., Walkinshaw, M. D. (2011) The trypanocidal drug suramin and other trypan blue mimetics are inhibitors of pyruvate kinases and bind to the adenosine site. *J. Biol. Chem.* **286**, 31232-31240.

The structure of *Tb*PYK in complex with F26BP determined from a crystal grown with ponceau S has been deposited in the PDB (*Tb*PYK/F26BP/Mg: 4HYW) and will be submitted for publication in: ([*]Co-first authors)

Zhong, W.[*], Morgan, H. P.[*], McNae, I. W., Michels, P. A., Fothergill-Gilmore, L. A.,

Walkinshaw, M. D. *'In crystallo'* substrate binding triggers major domain movements and reveals magnesium as a co-activator of *Trypanosoma brucei* pyruvate kinase. (Manuscript submitted to *Acta Crystallogr. D*)

Table 2.1 PYK crystal structures determined in this thesis

Structure	Crystallisation method	Resolution (Å)	T- or R-state[†]	Active site ligands	Effector site ligand	B-domain rotation	Described in Chapter
apo *Tc*PYK	co-crystallisation	2.50	T-state	K^+	none	n.a.	Chapter 4, 5
***Tc*PYK/ponceau S**	co-crystallisation	2.10	T-state	K^+	none	n.a.	Chapter 4
*Tc*PYK/F26BP/OX/Mg[‡]	co-crystallisation	2.80	R-state	Mg^{2+}, K^+ oxalate	F26BP	22°-35°	Chapter 5
*Tb*PYK/F26BP/Mg	co-crystallisation	2.95	R-state	Mg^{2+}, K^+	F26BP	n.a.	Chapter 4
***Tb*PYK/F26BP/Mg**[a]	co-crystallisation	2.35	R-state	Mg^{2+}, K^+	F26BP	n.a.	Chapter 4, 6
***Tb*PYK/F26BP /PEP/Mg**	crystal soaking	2.30	R-state	Mg^{2+}, K^+ PEP	F26BP	23°	Chapter 6
*Tb*PYK/F26BP/EDTA	crystal soaking	2.40	R-state	K^+ EDTA	F26BP	12°	Chapter 6
*Tb*PYK/F26BP/PYR/Mg	crystal soaking	2.15	R-state	Mg^{2+}, K^+ pyruvate	F26BP	22°	Chapter 7
*Tb*PYK/F26BP/PYR/Mg[b]	crystal soaking	1.95	R-state	Mg^{2+}, K^+ pyruvate	F26BP	22°	Chapter 7
*Tb*PYK/F26BP/MLT/Mg	crystal soaking	2.35	R-state	Mg^{2+}, K^+ D-malate	F26BP	22°	Chapter 7, 8
*Tb*PYK/F26BP/AKG/Mg	crystal soaking	2.18	R-state	Mg^{2+}, K^+ α-ketoglutarate	F26BP	22°	Chapter 7, 8
*Tb*PYK/F26BP/OX/Mg	crystal soaking	2.50	R-state	Mg^{2+}, K^+ oxalate	F26BP	22°	Chapter 7, 8
*Tb*PYK/F26BP/MLA/Mg	crystal soaking	2.30	R-state	Mg^{2+}, K^+ malonate	F26BP	23°	Chapter 8
*Tb*PYK/F26BP/TLA	crystal soaking	2.45	R-state	K^+ L-tartrate	F26BP	22°	Chapter 8
*Tb*PYK/F26BP/CIT	crystal soaking	2.30	R-state	K^+ citrate	F26BP	n.a.	Chapter 8

Abbreviations: *Tc*PYK, *T. cruzi* PYK. *Tb*PYK, *T. brucei* PYK; F26BP, fructose 2,6-bisphosphate; OX, oxalate; PEP, phosphoenolpyruvate; PYR, pyruvate; MLT, D-malate; AKG, α-ketoglutarate; MLA, malonate; TLA, L-tartrate; CIT, citrate.
The structures in bold have been deposited in the PDB: *Tc*PYK/ponceau S (3QV9), *Tb*PYK/F26BP/Mg (4HYW), *Tb*PYK/F26BP/PEP/Mg (4HYV).
[†] T-state: inactive state; R-state: active state.
[‡] This structure was determined by my colleague Dr. Hugh Morgan.
[a] This structure was determined from a crystal grown with ponceau S. These *Tb*PYK/F26BP/Mg crystals grown with ponceau S were used in the crystal soaking

experiments.
[b] This structure was determined from *Tb*PYK/F26BP/Mg crystals soaked with oxaloacetate. But the product pyruvate instead of oxaloacetate was observed in the structure.
n.a., not applicable

2.2.3. Structural studies of TcPYK (Chapter 5)

*Tc*PYK in the absence of active-site and effector-site ligands is in a relatively unstable and inactive state, and is therefore difficult to crystallise. The structure of this inactive form (T-state) of *Tc*PYK (apo *Tc*PYK) is described in Chapter 5, and allows detailed comparisons with the active R-state *Tc*PYK (*Tc*PYK/F26BP/OX/Mg) to study its allosteric mechanism at a molecular level. Furthermore, *Tc*PYK structures in both states could be used in structure-based inhibitor design. The work from this project forms the basis of a manuscript to be submitted for publication:
Morgan, H. P. [*], **Zhong, W.** [*], *et al*. Evolutionary divergence of pyruvate kinase allostery. (Manuscript in preparation) [*] Co-first authors

2.2.4. Structural studies of TbPYK (Chapter 6)

*Tb*PYK is a validated drug target against *T. brucei* (Drew et al. 2003; Albert et al. 2005). Therefore, it is of particular interest to obtain its structure for structure-based drug design. The results from this project would offer the opportunity for the first time to explore the structural properties of *Tb*PYK and the role of divalent metal in PYK activity. The work from this project forms the basis of one publication:
Zhong, W.[*], Morgan, H. P.[*], McNae, I. W., Michels, P. A., Fothergill-Gilmore, L. A., Walkinshaw, M. D. *'In crystallo'* substrate binding triggers major domain movements and reveals magnesium as a co-activator of *Trypanosoma brucei* pyruvate kinase. [Manuscript submitted to *Acta Crystallogr. D*; Structures deposited in PDB: *Tb*PYK/F26BP/Mg (4HYW) and *Tb*PYK/F26BP/PEP/Mg (4HYV)] [*] Co-first authors

2.2.5. Inherent and conserved decarboxylase activity of TbPYK (Chapter 7)

Oxaloacetate decarboxylases have not been extensively studied yet compared to many other enzymes in central metabolic pathways. Only a limited number of decarboxylases have been identified so far, and are all from bacteria. In addition to specialized oxaloacetate decarboxylases, the inherent oxaloacetate decarboxylase activity of PYKs was reported for the first time 30 years ago for codfish muscle PYK (Creighton and Rose 1976a; Creighton and Rose 1976b) and rabbit muscle PYK (Creighton and Rose 1976a; Jursinic and Robinson 1978). Both ^{1}H NMR enzyme assays and X-ray crystallography were used in this thesis to 1. reveal the conservation of decarboxylase activity of PYKs from different species, 2. elucidate the overlapping active sites for decarboxylase activity and kinase activity, and 3. propose a reaction mechanism for decarboxylase activity. The work from this project forms the basis of a publication:

Zhong, W., *et al.* Pyruvate kinases have an inherent and conserved decarboxylase activity. (Manuscript in preparation)

2.2.6. Proposed strategies for designing selective inhibitors against Trypanosomatid PYKs (Chapter 8)

The crystal structures of trypanosomatid PYKs and the regulation of enzyme have been extensively studied in this thesis. The in-depth structural information provides essential knowledge to design selective inhibitors for trypanosomatid PYKs. In particular, differences at the effector site, in the allosteric mechanism, and in interface residues between trypanosomatid PYKs and human PYKs are suggested to be potential targets to design allosteric inhibitors. Additionally, the combination of multiple compounds to inhibit enzyme activity has been proved to be a feasible strategy due to the communication between effector site and active site, where ligand inhibition at the active site can be regulated by the binding of a modulator at a different site such as the effector site.

2.3. Materials

ADP, ATP, PEP, pyruvate, oxalate, oxaloacetate, D-malate, α-ketoglutarate, malonate, L-tartrate, citrate, ethylenediaminetetraacetic acid (EDTA), ponceau S, dimethyl sulfoxide (DMSO), F16BP, F26BP, lactate dehydrogenase (LDH), polyethylene glycol (PEG) 8,000, antibiotics and buffers were obtained from Sigma-Aldrich. Reduced nicotinamide adenine dinucleotide (NADH) and EDTA-free protease inhibitor mixture tablets were from Roche, glycerol was from BDH Prolabo, isopropyl β-D-1-thiogalactopyranoside (IPTG) was from Melford, and salts were from Fisher Scientific. Restriction enzymes and other enzymes used for cloning were purchased from New England BioLabs. Vector and *E. coli* competent cells were from Novagen.

2.4. Methods

Section 1: Cloning, PYK expression and purification

2.4.1. Cloning

2.4.1.1. Cloning His_6-*Tc*PYK and His_6-*Tb*PYK

Two constructs of the vector pET28a containing the genes encoding *Tc*PYK (pET28a_*Tc*PYK) and *Tb*PYK (pET28a_*Tb*PYK), are available in the laboratory. 6xHistidine coding regions (His_6-) are present in the upstream portions of the *NdeI-BamHI*-digested insert for each construct, generating N-terminal His_6-*Tc*PYK and N-terminal His_6-*Tb*PYK, respectively. The extra N-terminal His_6-tag sequence is MGSSHHHHHHSSGLVPRGSH, where sequence LVPRGS is the thrombin cleavage site. The protein sequences (excluding His_6-tag sequence) of *Tc*PYK and *Tb*PYK encoded from the genes can be found in **Figure 3.1** from **Chapter 3**.

2.4.1.2. Cloning untagged *Tc*PYK and untagged *Tb*PYK

The vector pET30a containing the gene encoding *Tc*PYK (pET30a_*Tc*PYK) and the vector pET24a containing the gene encoding *Tb*PYK (pET24a_*Tb*PYK) for expression of the untagged PYKs are available in the laboratory as well. For a codon optimized gene encoding *Tb*PYK (opt-*Tb*PYK), the digested insert which contained opt-*Tb*PYK was engineered into the *NdeI*-*BamHI*-digested cloning vector pUC57. This cloning construct pUC57_opt-*Tb*PYK was synthesized and identified by GenScript (http://www.genscript.com/index.html). To obtain the expression construct, the insert was amplified and digested with *NdeI* and *BamHI* (New England Biolabs). The digested insert was then ligated into an *NdeI*-*BamHI*-digested pET30a vector (Novagen). The resulting construct is referred to as pET30a_opt-*Tb*PYK. The DNA sequence of codon optimized *Tb*PYK will be shown in **Figure 3.2** from **Chapter 3**. Soluble and active protein could be expressed in large amount using the original construct of *Tc*PYK so there was no need to do codon optimization. Refer to **Figure 3.1** from **Chapter 3** for the protein sequences of *Tc*PYK and *Tb*PYK encoded from the genes.

2.4.2. Expression of PYKs in E. coli

2.4.2.1. Expression of His_6-*Tc*PYK

Chemically competent *E. coli* Rosetta(DE3)pLysS cells were transformed with the sequence-verified construct pET28a_*Tc*PYK. Single colonies of transformed *E. coli* harbouring the *Tc*PYK gene were picked from LB (Bertani 2004) kanamycin (50 μg ml^{-1}) and chloramphenicol (30 μg ml^{-1}) plates and used to inoculate 50 ml of LB medium containing 50 μg ml^{-1} of kanamycin and 30 μg ml^{-1} chloramphenicol. Cultures were grown overnight at 37°C with agitation. 10 ml of the overnight culture was used to inoculate 500 ml of 2xTY medium containing 50 μg ml^{-1} of kanamycin and 30 μg ml^{-1} of chloramphenicol. 500 ml cultures were grown to an OD_{600} of

0.6-0.8 and incubated in the cold room (4°C) for 30 min. Expression was induced by adding IPTG in the cold room to a final concentration of 1 mM. The incubation was then continued for 22 h at 18°C with agitation (250 rpm) and the cells were harvested by centrifugation at 8,000 rpm in a JLA-9.1000 rotor for 15 min at 4°C. Cell pellets from 500 ml culture were then frozen in liquid nitrogen and stored at -80°C.

2.4.2.2. Expression of untagged *Tb*PYK (codon optimized)

Chemically competent *E. coli* BL21(DE3) cells were transformed with the sequence-verified construct pET30a_opt-*Tb*PYK. Single colonies of transformed *E. coli* harbouring the opt-*Tb*PYK gene were picked from LB kanamycin (50 μg ml^{-1}) plates and used to inoculate 50 ml of LB medium containing 50 μg ml^{-1} of kanamycin. Cultures were grown overnight at 37°C with agitation. 10 ml of the overnight culture was used to inoculate 500 ml of LB medium containing 50 μg ml^{-1} of kanamycin. 500 ml cultures were grown to an OD_{600} of 0.6-0.8, cooled and incubated in the cold room (4°C) for 30 min. In the cold room the expression was induced by adding IPTG to a final concentration of 1 mM. The incubation was then continued for 22 h at 18°C with agitation (250 rpm) and the cells were harvested by centrifugation at 8,000 rpm in a JLA-9.1000 rotor for 15 min at 4°C. Cell pellets from 500 ml culture were then frozen in liquid nitrogen and stored at -80°C.

2.4.3. Preparation of cell lysates

2.4.3.1. Preparation of cell lysates for His_6-*Tc*PYK

Cell pellets from 500 ml of *E. coli* culture were thawed on ice and resuspended in 20 ml ice-cold cell-lysis buffer (Buffer A: 50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl, 20 mM imidazole, and 10% glycerol) supplemented with one tablet of EDTA-free protease inhibitors (Roche). Lysis was performed at 6°C by a single passage through a Constant System Cell Disruptor TS Series Benchtop instrument set to 152 MPa (equivalent to 22 KPSI instrument reading). Cellular debris was removed by

centrifugation at 23000 rpm in a JA-25.50 rotor for 45 min at 4°C. The supernatant was filtered through a 0.22 μm syringe filter and kept on ice for the purification procedure.

2.4.3.2. Preparation of cell lysates for untagged *Tb*PYK

Cell pellets from 500 ml of *E. coli* culture were thawed on ice and resuspended in 20 ml ice-cold cell-lysis buffer [Buffer X: 50 mM triethanolamine/HCl (TEA), pH 7.9, 20 mM KCl, and 10% glycerol] supplemented with one tablet of EDTA-free protease inhibitors (Roche). Lysis was performed at 6°C by a single passage through a Constant System Cell Disruptor TS Series Benchtop instrument set to 152 MPa (equivalent to 22 KPSI instrument reading). Cellular debris was removed by centrifugation at 23000 rpm in a JA-25.50 rotor for 45 min at 4°C. The supernatant was filtered through a 0.22 μm syringe filter and kept on ice for the purification procedure.

2.4.4. Purification of bacterially expressed PYKs

All purification steps were performed on an ÄKTApurifier (100 ml/min) system (GE Healthcare) at 6°C unless stated otherwise. The columns for protein purification used in this thesis are summarised in **Table 2.2**.

2.4.4.1. Purification of His_6-*Tc*PYK

The supernatant containing His_6-*Tc*PYK was loaded onto a 5 ml HiTrap IMAC HP Sepharose column (pre-charged with nickel and pre-equilibrated in buffer A) (**Table 2.2**) at 2 ml min^{-1}. The column was maintained at a constant flow rate of 2 ml min^{-1} throughout the purification process. The column was washed with twenty column volumes of buffer A. A linear gradient (0-100% over thirty column volumes) of buffer B (50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl, 500 mM imidazole, and 10%

glycerol) was applied to elute bound proteins. Over five column volumes of buffer B was used to achieve elution. Fractions containing His_6-*Tc*PYK (validated by running SDS-PAGE for each eluted fraction) were pooled and concentrated to ≤5 ml using a Vivaspin column [Molecular Weight Cut Off (MWCO) = 100 kDa] (Sartorius). The concentrated proteins were loaded onto a Superdex 200pg XK 26/60 gel-filtration column (319 ml column volume) (**Table 2.2**), pre-equilibrated in buffer C (20 mM TEA, pH 7.2, 50 mM KCl, 10 mM $MgCl_2$, 20% glycerol). The column was maintained at a constant flow rate of 2 ml min^{-1} throughout the purification process. The purified protein was concentrated to 10 mg ml^{-1} using a Vivaspin column (MWCO = 100 kDa) and stored at -80°C.

2.4.4.2. Purification of untagged *Tb*PYK

The supernatant containing *Tb*PYK was loaded onto two tandem ion-exchange columns (Hiprep DEAE FF 16/10 and Hiprep SP FF 16/10 – fitted sequentially, pre-equilibrated in buffer X) (**Table 2.2**). The columns were maintained at a constant flow rate of 2 ml min^{-1} throughout the purification process. The columns were washed with three column volumes of buffer X after the sample loading process. Bound proteins were eluted by one single step of 40% elution buffer (Buffer Y: 50 mM TEA, pH 7.9, 2 M KCl, and 10% glycerol) over four column volumes. After that, over seven column volumes of buffer Y were applied to further wash the columns. Fractions containing *Tb*PYK (validated by running SDS-PAGE for each eluted fraction) were pooled and concentrated to ≤5 ml using a Vivaspin column (MWCO = 100 kDa). The concentrated proteins were loaded onto a Superdex 200pg XK 26/60 gel-filtration column (319 ml of column volume) (**Table 2.2**), pre-equilibrated in buffer C. The column was maintained at a constant flow rate of 2 ml min^{-1} throughout the purification process. The purified protein was concentrated to 10 mg ml^{-1} using a Vivaspin column (MWCO = 100 kDa) and stored at -80°C.

Table 2.2 Columns used for protein purification in this thesis†

Column	Column volume	Maximum pressure	Void volume	Recommended sample volume	Recommended flow rate
HiTrap IMAC HP	5 ml	0.3 MPa	n.a.	n.a.	5 ml/min
HiPrep DEAE FF 16/10	20 ml	0.15 MPa	n.a.	n.a.	2-10 ml/min
HiPrep SP FF 16/10	20 ml	0.15 MPa	n.a.	n.a.	2-10 ml/min
Superdex 200pg XK26/60	319 ml	0.5 MPa	105 ml	≤5 ml; ≤50 mg	2-2.5 ml/min

n.a., not applicable

† All the columns were bought from GE Healthcare. The information in this table is summarised from the manufacturer's instructions.

2.4.5. Measurement of protein concentration

Protein concentration was determined by the measurement of the absorbance at 280 nm, using the molar absorption coefficients shown in **Table 3.1** of **Chapter 3** and the Beer-Lambert law: $A=\varepsilon \times c \times l$, where A is the absorbance at a particular wavelength, ε is the absorption coefficient, and l is the path length of the cuvette (or cell) which holds the solution of concentration c.

Section 2: Enzyme assays

2.4.6. Enzyme kinetics assays

All enzyme assay experiments were performed at least three times on a SpectraMax® M5 Multi-Mode Microplate Reader in 96-well plate format at 25°C unless stated otherwise.

PYK activity was measured by monitoring the decrease of NADH absorbance at 340 nm in a lactate dehydrogenase (LDH)-coupled system **(Figure 2.1)**. Kinetic curves were plotted by KaleidaGraph software (Synergy Software, Reading, PA). One activity unit is defined as the conversion of 1 μmol substrate min^{-1} under the assay conditions. In general, the PYK enzyme activity assay was performed in 100-μl reaction mixtures containing 50 mM TEA buffer, pH 7.2, 50 mM KCl, 10

mM $MgCl_2$, 0.5 mM NADH, and 3.2 units/ml of LDH from rabbit muscle (Sigma, 61309). The substrates were added depending on the purpose of the study, which will be described later. The reaction was initiated by the addition of PYK solution, which had been diluted in 1x assay buffer (50 mM TEA, pH 7.2, 50 mM KCl, 10 mM $MgCl_2$) beforehand, to a final concentration of 1 μg ml^{-1}. The PYK solution should be added to the reaction mixture immediately after dilution. The plate was then gently agitated automatically for 5 sec and the decrease in absorbance at 340 nm was measured for 5 min. The specific activity of the enzyme was obtained under saturating conditions (4 mM of ADP and 10 mM of PEP).

Enzyme kinetics with regard to ADP was studied under a saturated concentration of PEP (10 mM) and variable concentrations of ADP (from 0 to 4 mM). Enzyme kinetics with regard to PEP were studied under a saturated concentration of ADP (4mM) and variable concentrations of PEP (from 0 to 10 mM) in the absence and presence of 1 μM F26BP, or in the absence and presence of 4.5 mM F16BP. When hyperbolic kinetics was obtained, the Michaelis-Menten equation was fitted to the initial rates: $v_o = V_{max} \times [S] / (K_m + [S])$, where v_o = initial reaction rate, V_{max} = maximum rate, [S] = substrate concentration, K_m = Michaelis constant of investigated enzyme. The Michaelis-Menten plot related the reaction rate v to the substrate concentration [S]. When sigmoidal kinetics was obtained, an extended Michaelis-Menten equation was fitted to the initial rates: $v_o = V_{max} \times [S]^n / (S_{0.5}^n + [S]^n)$, where v_o = initial reaction rate, V_{max} = maximum rate, [S] = substrate concentration, $S_{0.5}$ = substrate concentration giving half-maximal reaction rate, n = Hill coefficient (h). The extended Michaelis-Menten plot related the reaction rate v to the substrate concentration [S]. The turnover number (k_{cat}) of the enzyme was calculated based on the enzyme specific activity and the subunit molecular mass of PYK. The ratio k_{cat}/K_m or $k_{cat}/S_{0.5}$ which can indicate the catalytic efficiency of an enzyme was calculated as well.

When activators F26BP or F16BP were studied, the activation constants were

calculated from v-v_o versus the activator concentration plots using the equation: v-v_o = (V_{max}-v_o) x $[S]^n$ / ($K_{a0.5}{}^n$+$[S]^n$), where v_o is the initial rate in the absence of effector. Enzyme kinetics with regard to effector F26BP were studied at 4 mM ADP, 0.4 mM PEP and variable concentrations of F26BP. Enzyme kinetics with regard to effector F16BP were studied at 4 mM ADP, 0.7 mM PEP and variable concentrations of F16BP.

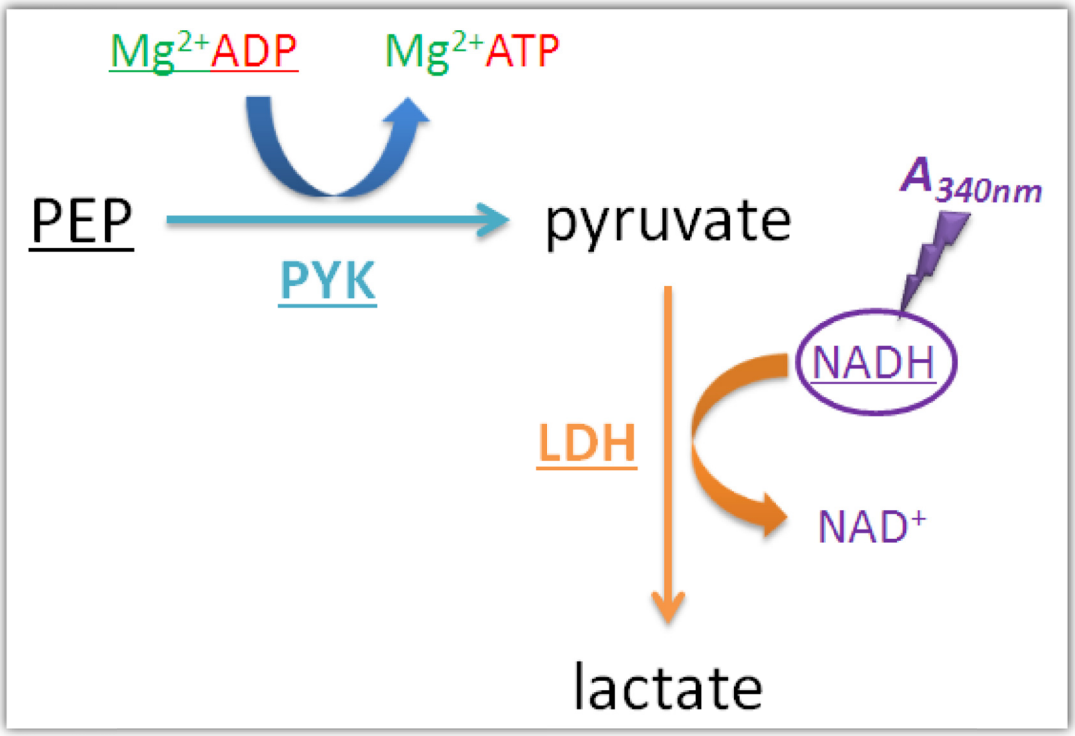


Figure 2.1 LDH-coupled enzyme reactions for PYK activity assay
The reagents added to the plate are indicated by underlines. NADH absorbs light at 340 nm.

2.4.7. PYK activity assay in the presence of selected ligands

All the enzyme assay experiments were performed based on the LDH-coupled enzyme system at 25°C by monitoring the decrease of NADH absorbance at 340 nm.

2.4.7.1. DMSO activation assay

DMSO activation of His_6-*Tc*PYK was monitored by the LDH-coupled assay at 25°C which is similar to the method applied in the enzyme kinetic assay in **section 2.4.6** but on a 1-ml reaction system using a UV-VIS Spectrophotometer (JASCO V-550).

To manage the assay, the following reagents were added to a 1-ml cuvette: 770 µl of assay mix (50 mM TEA buffer, pH 7.2, 50 mM KCl, 10 mM $MgCl_2$, 0.2 mM NADH, and 3.2 units/ml of LDH from rabbit muscle), 10 µl of His_6-*Tc*PYK (2 µg ml^{-1}), and 200 µl of DMSO solution. All the concentrations stated above were the final concentrations in the 1-ml reaction system. The mixture was incubated for 5 min before the reaction was started by adding 20 µl of substrate mixture of PEP and ADP (final concentrations were 0.4 mM and 0.2 mM, respectively). The mixture was then gently agitated and the decrease in absorbance at 340 nm was monitored for 2 min. A concentration range of 1-10% (v/v) for DMSO was tested in this experiment. The control assay with no DMSO present was also determined. The rate for each activation assay with a series of DMSO titrations was expressed as a percentage of the control assay. The assay was performed in triplicate.

2.4.7.2. *Tb*PYK activity assay in the presence of selected substrate analogues

Six analogues of the substrate PEP were selected to test their effects on *Tb*PYK activity in the presence and absence of F16BP (an analogue of the natural activator F26BP): oxalate, D-malate, α-ketoglutarate, malonate, L-tartrate and citrate. A series of concentration titrations (0-20 mM) for each analogue was performed in the assay to monitor their effects on *Tb*PYK activity.

The assay conditions were similar to that in the DMSO activation assay (**section 2.4.7.1)** but were performed on a SpectraMax® M5 Multi-Mode Microplate Reader using a 96-well plate (100 µl reaction system) instead of a 1-ml cuvette. All the reagents were dissolved or diluted in the assay buffer which contains 50 mM TEA, pH 7.2, 100 mM KCl, 10 mM $MgCl_2$.

For the assay without adding F16BP, 10 µl of analogue with varied concentrations were added to each well. Then the analogue solution in each well was mixed by adding 80 µl of assay mix containing 50 mM TEA, pH 7.2, 100 mM KCl, 10 mM $MgCl_2$, 0.5 mM NADH, 3.2 units/ml of LDH, 0.4 mM PEP and 0.75 µg/ml of *Tb*PYK, The concentrations described for the assay mix were the final concentrations in the 100 µl reaction system. 10 µl of 2 mM ADP (final concentration = 0.2 mM) was added to each well to start the reaction. The plate was then gently agitated automatically for 5 sec and the decrease in absorbance at 340 nm was measured for 5 min. For the assay in the presence of 1 mM F16BP, 0.2 mM of PEP (final concentration) was used instead of 0.4 mM. The substrate concentrations used for the assay were subsaturating. The control assay with no analogue present was also determined.

The rate for each activity assay with a series of analogue titrations was expressed as a specific activity (U/mg) of *Tb*PYK or a percentage of the control assay. The assay was performed at least in triplicate.

2.4.8. Decarboxylase activity assay: ^{1}H NMR analysis of the oxaloacetate decarboxylation reaction catalyzed by PYK

2.4.8.1. Oxaloacetate decarboxylation

The decarboxylation of oxaloacetate converts oxaloacetate to pyruvate and CO_2. Metals ions including divalent metals have been shown to enhance this reaction (Refer to section **7.3.6** of **Chapter 7** for more details). A schematic representation for the reaction is shown in **Figure 2.2**. The process of oxaloacetate decarboxylation can occur spontaneously, and can also be accelerated by some enzymes which possess oxaloacetate decarboxylase activity. The enzyme activity is defined as µmol/min/mg (U/mg). Here we applied ^{1}H NMR spectroscopy to monitor the protons of the methyl group ($–CH_3$) from the product pyruvate to determine the oxaloacetate decarboxylase

activity of PYKs.

2.4.8.2. ^{1}H NMR assay

NMR spectra (^{1}H) were recorded at 25^{o} at 800 MHz with a Bruker Avance 800 NMR spectrometer (School of Chemistry, The University of Edinburgh) immediately after dissolving and diluting oxaloacetate into the sample solution. The spectra were recorded every 1 min for up to 30 min. An example showing the change in ^{1}H NMR spectra over time can be found in **Figure 2.3**. The rates of oxaloacetate decarboxylation were measured by converting the rates of increasing spectral signal (peak areas) of protons from the methyl group (–CH_3) over time into the rates of increasing amount of the product pyruvate (µmol/min). 5 mM DMSO was used as an internal standard to calculate the reaction rate (µmol/min). The spectral data were analysed by the NMR software TopSpin (Bruker).

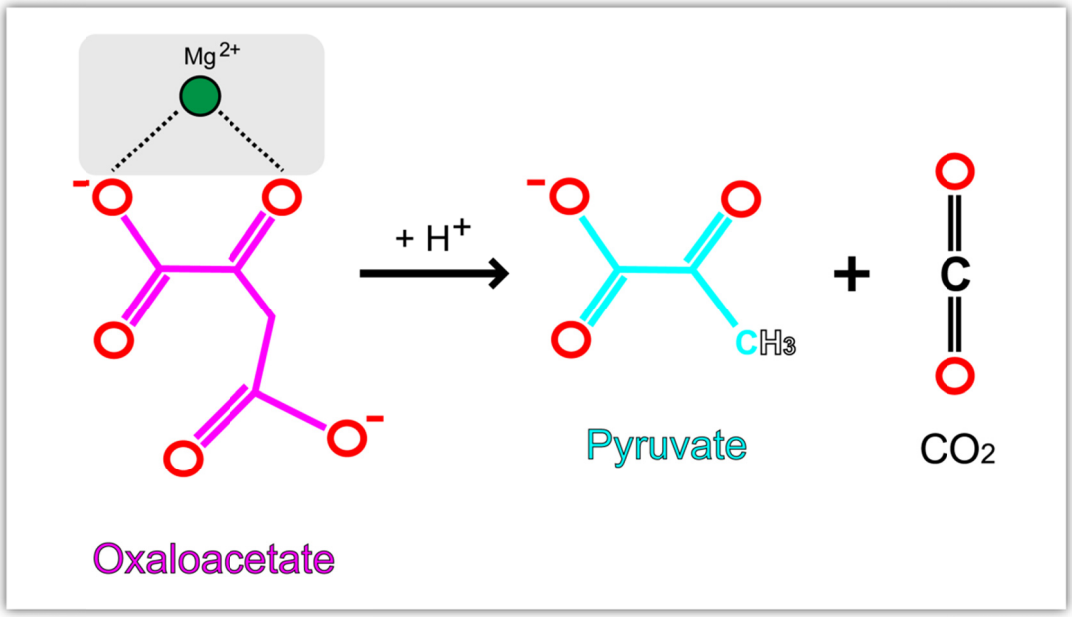


Figure 2.2 A schematic representation to show oxaloacetate decarboxylation. The divalent metal ion Mg^{2+}, which is one of the most common metal ions facilitating the decarboxylation of oxaloacetate is shown as a green sphere. The coordination between Mg^{2+} and oxaloacetate is indicated as dashed lines.

The sample solution in a total volume of 0.5 ml contained: 50 mM TEA (pH 7.2), 10 mM $MgCl_2$, 50 mM KCl, 50 mM oxaloacetate, 0.72 mg/ml enzyme [*Tb*PYK, Human M1 PYK; or rabbit muscle LDH (Sigma, L1254) as a negative control enzyme], 10% D_2O, 5 mM DMSO as the internal standard, with or without 18 μM F26BP when *Tb*PYK was present. The assay containing 0.36 mg/ml *Tb*PYK was carried out as well. In the presence of 18 μM F26BP, only 0.36 mg/ml *Tb*PYK was used because the reaction rate with 0.72 mg/ml *Tb*PYK was too quick to measure. Assay buffer instead of the enzyme was added to the solution to measure the spontaneous reaction rate of oxaloacetate decarboxylation. The rates of oxaloacetate decarboxylation catalysed by enzymes were obtained by subtracting the spontaneous rates. The substrate analogue oxalate (20 mM in final concentration) was tested as well to study its effect on decarboxylase activity. The following molecular mass of each enzyme was used to calculate k_{cat} (s^{-1}) for its oxaloacetate decarboxylase activity: 54466.6 for *Tb*PYK, 60225.3 for Human M1 PYK [Values were obtained from the online server of program ProtParam (Wilkins et al. 1999): http://web.expasy.org/protparam/].

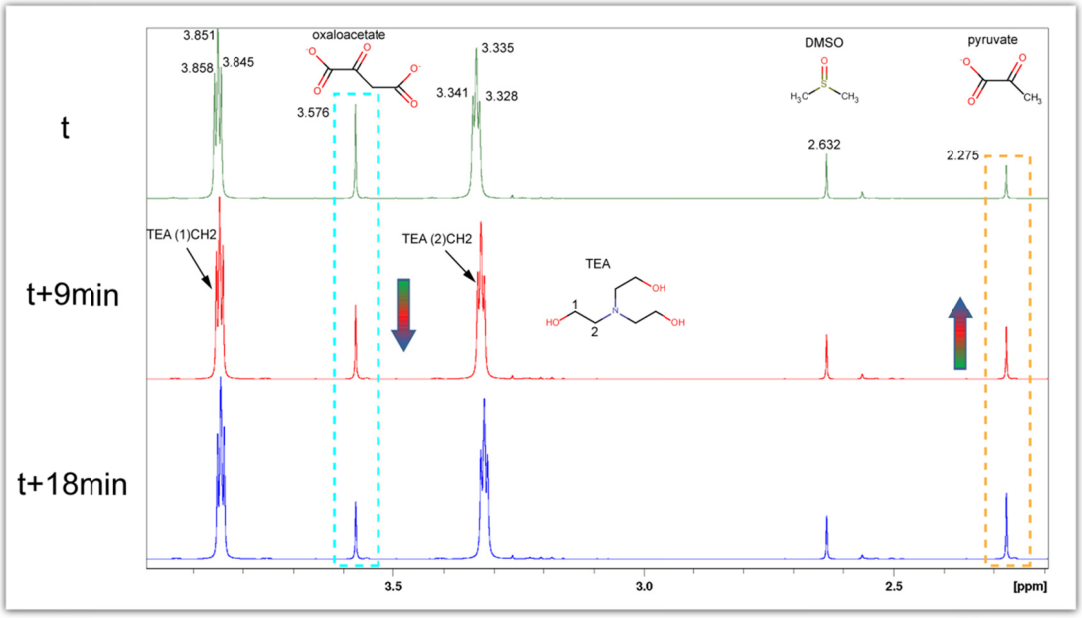


Figure 2.3 Comparison of peak intensities in the ^{1}H NMR spectra of an oxaloacetate solution recorded at three different time points to show the conversion from oxaloacetate to pyruvate.

This oxaloacetate solution contained 50 mM TEA (pH 7.2), 10 mM $MgCl_2$, 50 mM KCl, 50 mM oxaloacetate, and 5 mM DMSO. ^{1}H chemical shifts for each compound have been labelled. The ^{1}H peak intensity for oxaloacetate (-CH2-) (3.576 ppm) decreased (represented by arrow and cyan dashed box) over time (t, t+9min, t+18min) indicating the conversion of oxaloacetate. The ^{1}H peak intensity for pyruvate (-CH3) (2.275 ppm) increased (represented by arrow and yellow dashed box) over time (t, t+9min, t+18min) indicating the production of pyruvate. The ^{1}H peak intensities for the buffer TEA and the internal standard DMSO remained unchanged.

Section 3: Biophysical assay

2.4.9. Thermal shift assay

The assay involves monitoring changes in the fluorescence signal of SYPRO Orange dye as it interacts with a protein undergoing thermal unfolding. The SYPRO Orange dye is supplied by Invitrogen (catalogue number S6650) at 5000x concentration in DMSO, and diluted in assay buffer for use. Each of 50 µl of test samples in a 96-well PCR plate (BioRad) contained 50 mM TEA (pH 7.2), 100 mM KCl, 10 mM

$MgCl_2$, 10% glycerol, 10x SYPRO Orange dye, 4 μM enzyme, or/and 10 mM of test ligand(s) except 1.5 μM F26BP. Buffer was added instead of test ligand(s) in negative control samples. The temperature midpoint for the protein unfolding transition, T_m, was calculated using the BioRad iQ5 software. The plate was sealed with optical quality sealing tape (BioRad) and heated in an i-Cycler iQ5 real-time PCR detection system (BioRad) from 20 to 80 ^oC in increments of 1 ^oC. Fluorescence changes in the wells were monitored simultaneously with a charge-coupled (CCD) camera. The wavelengths for excitation and emission were 485 and 575 nm, respectively. Due to the lack of commercially available F26BP, the assay on F26BP could only be carried out once. The results for the other ligands were duplicated.

Section 4: X-ray crystallography

2.4.10. Crystallisation of TcPYK and TbPYK

2.4.10.1. Dye molecules used in crystallisation screening

All the dyes used in the crystallisation trials for optimising the crystal growth of *Tc*PYK and *Tb*PYK were purchased from Sigma. Detailed information about the dyes is summarised in **Chapter 4** (**Table 4.1** and **Figure 4.1)**. The dyes were normally dissolved to 100 mM using 200 mM TEA buffer (pH 7.2), otherwise solutions of saturating concentration were made up. The solution stocks were either diluted to working concentration using 200 mM TEA buffer (pH 7.2) when the experiment was performed, or frozen at -20 ^oC for storage. To prepare the working solution from each dye stock, the stock was diluted 20x-50x depending on its solubility when mixing with well solution.

2.4.10.2. Crystallisation screening procedures for *Tc*PYK and *Tb*PYK

Purified His_6-*Tc*PYK and untagged *Tb*PYK samples (10 mg ml^{-1}) were used for the crystallisation experiments. Frozen proteins samples were thawed on ice before using. Both crystallisation and co-crystallisation experiments were carried out by the vapour diffusion method using the hanging-drop technique at 4 ^{o}C. The drops were equilibrated against a reservoir filled with 1 ml of well solution.

To obtain T-state *Tc*PYK or T-state *Tb*PYK crystals, the drops were set up by mixing 1.5 µl of a well solution with 1 µl of 10 mg ml^{-1} protein sample. The well solution contained 6-20% PEG 8,000, 10-20% glycerol, 50 mM TEA buffer, pH 7.2, 100 mM KCl, and 50 mM $MgCl_2$. The procedure for the preparation of well solutions can be seen in **Table 2.3**. Dye molecules were included in the drops for crystallisation trials by mixing 0.5 µl of a working solution of the dye, 1.5 µl of a well solution with 1 µl of 10 mg ml^{-1} protein sample. The final concentration of dyes was around 50 - 800 µM depending on the solubility of the dye.

The allosteric activator F26BP was added to the crystallization conditions for co-crystallisation in order to obtain R-state *Tc*PYK or R-state *Tb*PYK crystals. In these cases, the drops were set up by mixing 0.5 µl of F26BP solution (500 µM), 1.5 µl of a well solution with 1 µl of 10 mg ml^{-1} protein. The final concentration of F26BP was ~80 µM. To include dye molecules in the crystallisation trials, the drops were set up by mixing 0.5 µl of a solution containing a mixture of dye (in working concentration) and F26BP (500 µM), 1.5 µl of a well solution with 1 µl of 10 mg ml^{-1} protein sample. The final concentration of dyes was around 50 - 800 µM. In addition, the substrate analogue oxalate which is a competitive inhibitor of PYKs, and the product ATP (both at a final concentration of 3 mM) were also used (in the presence or absence of F26BP) to crystallise *Tc*PYK or *Tb*PYK in its R-state.

The pH of each diluted dye solution was checked before using and adjusted to pH 7.2 when necessary.

Table 2.3 Preparation of well solutions used for crystallisation

10% glycerol															
PEG conc. (%)	**6**[†]	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
40%PEG8000	150	175	200	225	250	275	300	325	350	375	400	425	450	475	500
50% glycerol	200	200	200	200	200	200	200	200	200	200	200	200	200	200	200
10× Buffer	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
H2O	550	525	500	475	450	425	400	375	350	325	300	275	250	225	200
20% glycerol															
PEG conc. (%)	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
40%PEG8000	150	175	200	225	250	275	300	325	350	375	400	425	450	475	500
50%glycerol	400	400	400	400	400	400	400	400	400	400	400	400	400	400	400
10× Buffer	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
H2O	350	325	300	275	250	225	200	175	150	125	100	75	50	25	0

Stock solutions:
40% (w/v) PEG 8,000
50% glycerol
10x Buffer: 500 mM TEA, pH 7.25, 1 M KCl, 500 mM $MgCl_2$
[†] Example: 1 ml of well solution shaded by brown colour shown in the table contains 6% PEG 8,000, 10% glycerol, 50 mM TEA (pH 7.2), 100 mM KCl and 50 mM $MgCl_2$. It was made up by mixing 150 µl 40% PEG 8,000, 200 µl 50% glycerol, 100 µl 10x buffer and 550 µl distilled water.

2.4.10.3. Crystal soaking

For crystal soaking experiments, a crystal of *Tb*PYK/F26BP/Mg was placed into a 2 µl drop of soaking solution, composed of 5 mM soaking ligand (PEP, pyruvate, oxaloacetate, oxalate, D-malate, α-ketoglutarate, malonate, L-tartrate or citrate), 80 µM F26BP, 800 µM ponceau S, 20% PEG 8,000, 20% glycerol, 50 mM TEA buffer, pH 7.2, 100 mM KCl, and 50 mM $MgCl_2$. When an attempt was made to soak EDTA into the crystal, the soaking solution contained 25 mM EDTA, 80 µM F26BP, 800 µM ponceau S, 20% PEG 8,000, 20% glycerol, 50 mM TEA buffer, pH 7.2 and 100 mM KCl. The drop harbouring the crystal was equilibrated overnight (10 min for the oxaloacetate sample) against a reservoir of 1 ml of well solution, containing 20% PEG 8,000, 20% glycerol, 50 mM TEA buffer (pH 7.2), 100 mM KCl, and 50 mM $MgCl_2$.

2.4.11. Data collection and processing

X-ray intensity data from crystals were collected at the Diamond synchrotron radiation facility in Oxfordshire, United Kingdom. Data from crystals of apo *Tc*PYK, *Tb*PYK/F26BP/PEP/Mg, *Tb*PYK/F26BP/pyruvate/Mg, and *Tb*PYK/F26BP/malonate/Mg were collected on beamline I04; data from crystals of *Tb*PYK/F26BP/Mg (2.95 Å) and *Tb*PYK/F26BP/L-tartrate were collected on beamline I04-1; data from crystals of *Tb*PYK/F26BP/EDTA, *Tb*PYK/F26BP/pyruvate/Mg (oxaloacetate as soaking ligand), *Tb*PYK/F26BP/D-malate/Mg and *Tb*PYK/F26BP/citrate were collected on beamline I03; data from crystals of *Tb*PYK/F26BP/Mg, *Tb*PYK/F26BP/oxalate/Mg and *Tb*PYK/F26BP/α-ketoglutarate/Mg were collected on beamline I24; data from crystals of *Tc*PYK/ponceau S was collected on beamline I02. The intensity data were collected from single crystals flash frozen in liquid nitrogen at 100K. Data were then processed with MOSFLM (Potterton et al. 2003) and scaled with SCALA (Evans 2005).

2.4.12. Structure determination

2.4.12.1. Phasing

Structures of *Tc*PYK and *Tb*PYK were all solved by molecular replacement using the program PHASER (McCoy et al. 2007). To determine the *Tc*PYK structures, a monomer from the previously determined structure of *Tc*PYK/F26BP/OX/Mg (unpublished; determined by my colleague Dr. Hugh Morgan) was used as a search domain in the molecular replacement experiment. To determine the first *Tb*PYK structure (*Tb*PYK/F26BP/Mg), a monomer from the determined structure of apo *Tc*PYK (81% sequence identity compared to *Tb*PYK) in this thesis was applied as a search domain in the molecular replacement experiment. A monomer (chain B) of *Tb*PYK/F26BP/Mg was used as a search domain for solving other *Tb*PYK structures

in PHASER.

2.4.12.2. Model building and refinement

For the resulting PYK models after the phasing procedure described above, the B-domains with poor electron density were removed from the models at this early stage. For the first *Tb*PYK structure, residues were manually mutated from the *Tc*PYK sequence to the *Tb*PYK sequence and fitted into the density maps using COOT (Emsley and Cowtan 2004). The initial models were subjected to five cycles of rigid-body refinement followed by five cycles of restrained refinement using the program REFMAC (Murshudov et al. 1997). The side chains in the models were then manually adjusted using COOT, followed by several cycles of restrained refinement using REFMAC. When appropriate, relevant ligands (e.g. K^+, Mg^{2+}, F26BP, ponceau S, and/or other ligands) and water molecules were manually added to the structure using COOT. After additional cycles of restrained refinement and manual adjustments to side chains, ligands and water molecules, the overall quality of the map improved. For some models, the electron density signals for the missing B-domains also became stronger. The missing B-domains were then built up manually using COOT followed by cycles of TLS and restrained refinement in REFMAC and COOT adjustments. TLS groups were generated according to the domain region of PYK for each chain: (1) residues 1-14; (2) residues 15-87; (3) residues 88-189; (4) residues 190-357; (5) residues 358-499. NCS restrains were not used in all refinements.

2.4.12.3. Deposition of coordinates

Coordinates for crystal structures of *Tc*PYK/Ponceau S, *Tb*PYK/F26BP/Mg (2.35 Å), and *Tb*PYK/F26BP/PEP/Mg have been deposited in the Protein Data Bank (PDB) with the codes 3QV9, 4HYW and 4HYV, respectively.

Section 5: Structure analysis

2.4.13. Analysis of rigid-body rotation function

The method described here to analyse the rigid-body motion of trypanosomatid PYKs has been published by Dr. Hugh Morgan to describe the rigid-body rotation of *Lm*PYK (Morgan et al. 2010b).

A completed tetramer of T-state PYK or R-state PYK was generated from the symmetry monomers using PyMol (Delano 2002). Superposition of these two tetramers was performed using CCP4 superpose (C-α atoms fit) (Potterton et al. 2002). The RMS difference numbers and the rotation matrix were calculated automatically by the program during the superposition process. The allosteric rigid-body rotation was calculated by two steps of superposition: 1. The A- and C-domains (AC cores, residues 19-87 and 190-499) of two tetramers were superposed (C-α atoms fit) simultaneously to each other; 2. Using the new coordinates generated from the first step, each individual chain of R-state PYK tetramer was superposed onto the T-state PYK tetramer by CCP4 superpose (C-α atoms fit); the rotation matrix can be obtained from the program output file. The rotation angle (θ) was calculated by converting the rotation matrices to degrees using the following equation: $\theta = \arccos[0.5x(a1+b2+c3-1)]$. The numbers a1, b2 and c3 in this equation are from rotation matrix (**Figure 5.13** in **Chapter 5**). **Figure 5.13** explains the procedures for analysing this rigid-body rotation.

2.4.14. Analysis of B-domain rotation

Two PYK monomers were superposed onto each other by CCP4 superpose (Potterton et al. 2002) to visualise and analyse the movement of the B-domains of these two chains. The comparison was performed by superposing the AC-domains

(residues 19-89, 188-499) of the two monomers (C-α atoms fit). Then the B-domains (residues 90-187) from the superposed coordinates generated from the previous step were superposed onto each other by CCP4 superpose. The rotation angle (θ) of the B-domain was determined by converting the rotation matrices to degrees using the following equation: θ = arccos[0.5x(a1+b2+c3-1)]. This B-domain rotation can be automatically analysed by the online program DynDom (Hayward et al. 1997; Hayward and Berendsen 1998; Website link: http://fizz.cmp.uea.ac.uk/dyndom/).

2.4.15. Analysis of interface interactions

A completed tetramer of PYK was generated from the symmetry monomers using PyMol (Delano 2002). The residue interactions across the large (A-A) and small (C-C) interfaces of the structure were analysed by the online service PDBePISA (Krissinel and Henrick 2007). The website link for this service is: http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver.

2.4.16. Analysis of solvent accessible residues

The solvent-accessible residues were obtained by the online service PDBePISA program (Krissinel and Henrick 2007; Website link: http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html).

CHAPTER 3: Cloning, Expression, Purification and Characterisation of Pyruvate Kinases from *Trypanosoma*

3.1. Introduction

The purification and kinetic characterization of pyruvate kinase from blood-stream form *T. brucei* was published for the first time by Flynn and Bowman in 1980 (Flynn and Bowman 1980). However, the enzyme could only be partially purified at that time. Subsequently, enzyme purification and characterization protocols were developed over several years to attain relatively high purity and quality of *Tb*PYK from parasite cells (Flynn and Bowman 1981; Barnard and Pedersen 1988; Callens et al. 1991; Callens and Opperdoes 1992). In 1998, the first protocol to obtain bacterially expressed *Tb*PYK from *E. coli* was achieved by Ernest et al. (Ernest et al. 1998). The kinetic properties of bacterially expressed *Tb*PYK were characterized, and proved to be similar to those of the enzyme purified from *T. brucei*. However, only a modest yield of ~7 mg of pure *Tb*PYK could be obtained from 1-liter of cell culture.

For *Tc*PYK, enzyme kinetics were reported for the first time by Cazzulo et al. in 1989 (Cazzulo et al. 1989), but the assays were carried out on partially purified samples from parasite cell extracts, not on pure enzyme. No subsequent improved purification strategy for *Tc*PYK has been reported, and so far, kinetic data from pure high quality *Tc*PYK is still unavailable.

In this chapter, a complete and efficient strategy to express and purify bacterially expressed *Tc*PYK and *Tb*PYK from *E. coli* has been successfully developed. A series of enzyme kinetic characterizations have been carried out as well. The pure and active enzymes obtained in these experiments will be used for further studies later on, such as protein crystallization for structural studies of the

enzyme, and inhibitor screenings on the enzyme for anti-parasite drug design.

The work on *Tc*PYK from this chapter is published in:
Morgan, H. P., McNae, I. W., Nowicki, M. W., **Zhong, W.**, Michels, P. A., Auld, D. S., Fothergill-Gilmore, L. A., Walkinshaw, M. D. (2011) The trypanocidal drug suramin and other trypan blue mimetics are inhibitors of pyruvate kinases and bind to the adenosine site. *J. Biol. Chem.* **286**, 31232-31240.

Morgan, H. P.[*], **Zhong, W.**[*], *et al*. Evolutionary divergence of pyruvate kinase allostery. (Manuscript in preparation) [*] Co-first authors

The work on *Tb*PYK from this chapter is published in:
Zhong, W.[*], Morgan, H. P.[*], McNae, I. W., Michels, P. A., Fothergill-Gilmore, L. A., Walkinshaw, M. D. *'In crystallo'* substrate binding triggers major domain movements and reveals magnesium as a co-activator of *Trypanosoma brucei* pyruvate kinase. (Manuscript submitted to *Acta Crystallogr. D*) [*] Co-first authors

3.2. Materials and Methods

The materials and the general methods used in this chapter have been described in **Chapter 2**. In short, the work of 'cloning, expression and purification' is described in **Section 1**; the enzyme characterisation was carried out by the methods in **Section 2**.

3.3. Results and Discussion

3.3.1. Sequence analysis of TcPYK and TbPYK

The amino acid sequences of *Tc*PYK and *Tb*PYK which have been used to develop the expression constructs are shown in the sequence alignment in **Figure 3.1**. The coding region of the N-terminal His_6-tag (MGSSHHHHHHSSGLVPRGSH) had been engineered in the upstream region of native PYK, resulting in a total of 519 amino acids for His_6-tagged PYK, while untagged PYK is composed of 499 amino acids. However, we do not actually know whether the first residue methionine corresponding to the start codon (ATG) of the sequence has been translated or not, although we cannot see the ordered electron density for this residue in both crystal structures of *Tc*PYK and *Tb*PYK. These two *Trypanosoma* PYKs share a high degree of similarities shown in **Figure 3.1** with 80% sequence identity. General physical and chemical parameters (theoretical pI, molecular weight and extinction coefficient) of His_6-*Tc*PYK and untagged *Tb*PYK are summarised in **Table 3.1**. The protein expression systems shown in **Table 3.1** are the chosen strategies in the experiments which will be discussed later. The solvent accessible residues of *Tc*PYK, *Tb*PYK and *Lm*PYK which may contribute to the overall surface charge of proteins are summarised in **Table 3.2**.

Figure 3.1 Alignment of the amino acid sequences of *Tc*PYK and *Tb*PYK showed 81% sequence identity, with residues coloured according to type.
Tc, *Tc*PYK; *Tb*, *Tb*PYK. The red colour corresponds to small and hydrophobic amino acids including phenylalanine (F), the magenta corresponds to basic amino acids but not including histidine (H), the blue corresponds to acidic amino acids, and the green corresponds to hydroxyl, sulfhydryl and amide amino acids including glycine (G). The symbols under the residues indicate '*' as identical residues; ':' as

strongly similar residues; and '.' as weakly similar residues. This sequence alignment was performed using CLUSTALW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) with parameters set at default values.

Table 3.1 Expression systems and physical and chemical parameters of expressed PYKs

Protein	Expression vector	Expression cell strain (*E. coli*)	pI[a]	MW[b]	Extinction coefficient (Ex.co.) [b]
*Tc*PYK (His_6-)	pET-28a	Rosetta(DE3)pLysS	7.46	56769.1	22350 M^{-1} cm^{-1} 0.39370 $(mg/ml)^{-1}$ cm^{-1}
*Tb*PYK[c] (notag)	pET-30a	BL21 (DE3)	7.74	54466.6	17880 M^{-1} cm^{-1} 0.32827$(mg/ml)^{-1}$ cm^{-1}

[a] Theoretical pI (isoelectric point) of protein was obtained from EMBL Isoelectric Point Service (http://www3.embl.de/cgi/pi-wrapper.pl).

[b] Molecular weight (MW) and extinction coefficient (Ex.co.) were obtained from ExPASy Proteomics Server (http://www.expasy.ch/tools/protparam.html). All values include the initiator Met. The extinction coefficient of which unit was $(mg/ml)^{-1}$ cm^{-1} was calculated by the equation:

Ex.co. ($(mg/ml)^{-1}$ cm^{-1}) = Ex.co. (M^{-1} cm^{-1}) / Molar mass (g mol^{-1}).

[c] The untagged *Tb*PYK was expressed from codon optimized gene opt-*Tb*PYK.

Table 3.2 Solvent accessible residues of trypanosomatid PYKs with positive or negative charges under physiological pH [a]

	*Tb*PYK				*Tc*PYK				*Lm*PYK			
Residue	Lys	Arg	Asp	Glu	Lys	Arg	Asp	Glu	Lys	Arg	Asp	Glu
Residue Number [b]	18	20	84	36	39	20	84	36	38	19	83	58
	39	23	106	59	86	23	97	59	85	22	98	73
	43	50	113	74	96	43	106	74	118	49	112	88
	86	69	128	89	118	50	113	89	121	68	117	102
	118	91	145	99	122	69	145	117	123	90	122	156
	122	137	158	117	124	91	158	123	133	136	127	158
	124	152	159	123	152	131	174	128	192	174	144	162
	156	160	174	157	156	137	186	157	224	175	157	202
	163	171	186	184	163	175	194	159	229	194	173	217
	193	175	194	192	176	195	197	184	238	214	183	240
	196	176	197	199	193	215	207	192	290	223	185	253
	230	195	207	218	199	224	233	196	335	231	193	254
	232	215	233	222	230	405	250	203	394	404	196	268
	239	221	250	229	232	408	257	218	411	407	206	332
	291	224	265	241	239	414	265	222	446	413	221	341
	336	405	326	254	291	425	326	225	453	424	232	359
	338	408	360	269	336	427	416	229	465	429	249	378
	372	414	444	333	369	430	439	241	467	456	256	438
	395	425	451	342	372	436	444	254	486		264	451
	412	430	453	378	395	447	451	269			325	462
	447	436	463	379	412	457	463	333			377	
	452	457	475	450	454		475	342			443	
	454			468	456			378			445	
	456				466			379			450	
	459				467			446			474	
	466				468			450				
	467				487			452				
	469											
	487											
Total residues	***29***	***22***	***22***	***23***	***27***	***21***	***22***	***27***	***19***	***18***	***25***	***21***
	51		***45***		***48***		***49***		***37***		***46***	

[a] The solvent accessible residues were obtained by the online service PDBePISA program (Krissinel and Henrick 2007). The X-ray crystal structures of trypanosomatid PYKs applied in the analysis are as follows: *Tb*PYK/F26BP/Mg (this work), *Tc*PYK/F26BP/OX/Mg (determined by colleague Dr. Hugh Morgan)

and *Lm*PYK/F26BP/OX/ATP (3HQP). Only Lys, Arg, Asp and Glu are considered in the analysis. Residues Lys and Arg have p*K*a values of 12.1 and 10.67, respectively, and contain positive charges at physiological pH. Residues Asp and Glu have p*K*a values of 3.71 and 4.15, respectively, and contain negative charges at physiological pH.

[b] Residue numbers correspond to the numbers in the structure.

3.3.2. Cloning of TcPYK and TbPYK

3.3.2.1. Amplification and cloning of the *Tc*PYK genes

Plasmid pET28a_*Tc*PYK for N-terminal His_6-tagged *Tc*PYK expression and plasmid pET30a_*Tc*PYK for untagged *Tc*PYK expression have been constructed and sequence verified, and are available in the laboratory.

3.3.2.2. Codon optimization and cloning of the *Tb*PYK genes

Plasmid pET28a_*Tb*PYK for N-terminal His_6-tagged *Tb*PYK expression and plasmid pET24a_*Tb*PYK for untagged *Tb*PYK expression have been constructed and sequence verified, and are available in the laboratory as well.

To improve the soluble expression of *Tb*PYK, a series of codon optimization procedures were managed by GenScript to improve the efficiency of gene expression in *E. coli*. The DNA alignment highlighting the optimized regions within the sequence is shown in **Figure 3.2**. The optimized DNA sequence was originally inserted into a pUC57 vector, which was referred to as pUC5_opt-*Tb*PYK. The results from the double-digested *NdeI-BamHI* pUC5_opt-*Tb*PYK can be seen in **Figure 3.3**. The construct was amplified in *E. coli* DH5α cell culture, and digested by *NdeI* and *BamHI* restriction enzymes. The *NdeI-BamHI* digested inserts (~1.51 kb) were ligated into *NdeI-BamHI* digested pET30a vector (~5.26 kb) resulting in an expression construct pET30a_opt-*Tb*PYK for untagged *Tb*PYK expression. The new construct was transformed into chemically competent *E. coli* DH5α cells, and the gel electrophoresis showing the restriction digests of pET30a_opt-*Tb*PYK isolated from the transformation can be seen in **Figure 3.4**.

```
Optimized   ATGAGCCAACTGGAACATAACATCGGCCTGAGCATCTTTGAACCGGTCGCAAAACATCGT 60
Original    ATGTCGCAGCTAGAGCACAACATTGGGCTCTCCATTTTTGAGCCGGTGGCCAAACACCGT 60

Optimized   GCCAATCGTATCGTCTGCACCATCGGCCCGAGTACCCAGTCCGTTGAAGCTCTGAAAAAC 120
Original    GCCAACCGCATTGTGTGCACCATCGGACCCAGCACACAGAGCGTTGAAGCCCTGAAAAAC 120

Optimized   CTGATGAAATCAGGCATGTCGGTGGCGCGTATGAATTTTAGCCACGGCAGCCACGAATAT 180
Original    CTAATGAAGAGTGGTATGTCTGTTGCCCGTATGAACTTTTCCCATGGGTCCCATGAATAC 180

Optimized   CATCAGACCACGATTAACAATGTTCGTGCGGCGGCGGCAGAACTGGGTCTGCACATTGGT 240
Original    CACCAGACGACAATAAACAACGTTCGCGCAGCTGCTGCTGAACTTGGCCTTCATATTGGT 240

Optimized   ATCGCCCTGGATACCAAAGGCCCGGAAATTCGTACGGGCCTGTTTAAAGACGGTGAAGTG 300
Original    ATTGCGCTGGACACGAAAGGTCCTGAAATCCGTACAGGCCTCTTCAAGGATGGTGAGGTG 300

Optimized   TCTTTCGCCCCGGGCGATATCGTTTGTGTCACCACGGACCCGGCGTATGAAAAAGTTGGC 360
Original    AGCTTCGCCCCTGGTGACATCGTGTGTGTCACTACCGATCCCGCCTATGAGAAGGTTGGC 360

Optimized   ACCAAAGAAAAATTCTACATCGATTATCCGCAGCTGACCAATGCCGTGCGTCCGGGCGGT 420
Original    ACAAAGGAAAAGTTCTATATCGACTACCCACAACTGACGAACGCGGTACGCCCGGGCGGT 420

Optimized   TCTATCTACGTGGATGACGGTGTTATGACCCTGCGTGTTGTTAGTAAAGAAGATGACCGC 480
Original    TCCATTTATGTTGACGACGGTGTGATGACTCTCCGCGTGGTGAGCAAGGAAGACGATCGG 480

Optimized   ACGCTGAAATGTCACGTGAACAATCATCACCGCCTGACCGATCGTCGCGGCATTAATCTG 540
Original    ACGCTGAAGTGCCACGTGAACAATCACCACCGCCTGACCGACCGCAGGGGTATCAACCTC 540

Optimized   CCGGGTTGCGAAGTTGACCTGCCGGCTGTCTCGGAAAAAGATCGTAAAGACCTGGAATTT 600
Original    CCAGGTTGTGAGGTGGATCTGCCCGCCGTTTCCGAGAAGGACCGCAAGGACTTGGAGTTC 600

Optimized   GGCGTCGCACAGGGTGTGGATATGATTTTTGCCAGCTTCATCCGCACCGCAGAACAGGTC 660
Original    GGTGTGGCACAAGGTGTTGATATGATTTTTGCCTCGTTCATCCGCACGGCTGAGCAGGTT 660

Optimized   CGTGAAGTGCGTGCCGCACTGGGTGAAAAGGGTAAAGATATCCTGATCATCAGCAAAATC 720
Original    CGTGAGGTGCGCGCCGCCCTCGGTGAGAAGGGAAAAGACATCCTTATCATTTCCAAGATT 720

Optimized   GAAAACCATCAGGGCGTGCAAAATATTGATAGCATTATCGAAGCGAGCAATGGTATTATG 780
Original    GAAAACCACCAGGGTGTGCAGAACATTGACTCCATCATTGAAGCGAGCAATGGCATTATG 780
```

Figure 3.2 Alignment of the DNA sequences of the optimized *Tb*PYK gene and the original *Tb*PYK gene.
The optimized regions are highlighted by red colour.

```
Optimized   GTTGCGCGCGGCGACCTGGGTGTGGAAATTCCGGCTGAAAAAGTGTGCGTTGCGCAGATG 840
Original    GTTGCTCGTGGTGATCTTGGCGTCGAGATCCCCGCCGAGAAGGTGTGTGTGGCGCAGATG 840

Optimized   TGTATTATCTCCAAATGCAACGTCGTGGGCAAACCGGTGATCTGTGCGACCCAAATGCTG 900
Original    TGCATCATCAGCAAGTGCAACGTGGTGGGCAAACCTGTCATCTGTGCCACGCAAATGCTC 900

Optimized   GAAAGTATGACCAGCAATCCGCGTCCGACGCGCGCCGAAGTTTCCGATGTCGCGAACGCC 960
Original    GAGAGCATGACGTCAAACCCCCGCCCCACACGTGCCGAAGTATCCGATGTTGCCAATGCT 960

Optimized   GTTCTGAATGGTGCCGACTGTGTCATGCTGTCAGGCGAAACGGCAAAGGGTAAATATCCG 1020
Original    GTTCTCAACGGAGCCGACTGTGTCATGCTTTCCGGTGAAACGGCCAAAGGGAAGTACCCC 1020

Optimized   AATGAAGTTGTCCAGTACATGGCCCGTATTTGCGTTGAAGCTCAATCGGCGACCCATGAT 1080
Original    AACGAAGTGGTACAGTACATGGCTCGCATCTGTGTCGAGGCGCAGAGCGCAACGCACGAT
1080

Optimized   ACGGTCATGTTCAACAGCATCAAAAACCTGCAGAAAATCCCGATGTGTCCGGAAGAAGCA 1140
Original    ACTGTTATGTTCAACAGCATTAAGAACCTGCAGAAAATCCCCATGTGCCCCGAAGAGGCC 1140

Optimized   GTGTGCAGCTCTGCAGTTGCGAGCGCGTTCGAAGTGCAAGCCAAAGCAATGCTGGTTCTG 1200
Original    GTTTGCAGCAGCGCCGTCGCCTCTGCATTTGAGGTACAAGCCAAGGCAATGCTCGTGCTA 1200

Optimized   AGTAACACCGGCCGTTCCGCGCGCCTGATCTCAAAATACCGCCCGAATTGCCCGATTATC 1260
Original    AGCAACACGGGCCGAAGTGCTCGCCTCATCAGCAAGTACCGCCCGAACTGCCCCATCATT 1260

Optimized   TGTGTTACCACGCGTCTGCAGACCTGCCGCCAACTGAACGTCACGCGTTCAGTGGTTTCG 1320
Original    TGTGTCACCACACGTCTGCAAACGTGCCGTCAACTCAACGTGACGCGCTCCGTCGTGAGT 1320

Optimized   GTGTTTTATGATGCAGCTAAATCTGGCGAAGATAAAGACAAAGAAAAACGCGTGAAACTG 1380
Original    GTTTTCTACGACGCCGCAAAGAGTGGTGAGGACAAGGATAAGGAAAAGCGCGTGAAACTC 1380

Optimized   GGTCTGGATTTCGCCAAAAAAGAAAAATACGCAAGTACCGGTGACGTCGTGGTTGTCGTG 1440
Original    GGTTTGGATTTTGCGAAGAAGGAGAAATATGCCAGCACTGGAGACGTGGTGGTTGTCGTG 1440

Optimized   CACGCTGACCACTCTGTGAAAGGCTACCCGAACCAAACCCGTCTGATTTATCTGCCGTAA 1500
Original    CATGCCGACCACTCTGTGAAGGGGTACCCGAACCAAACTCGCCTCATCTATCTTCCTTAA
1500
```

Figure 3.2 (*continued*)

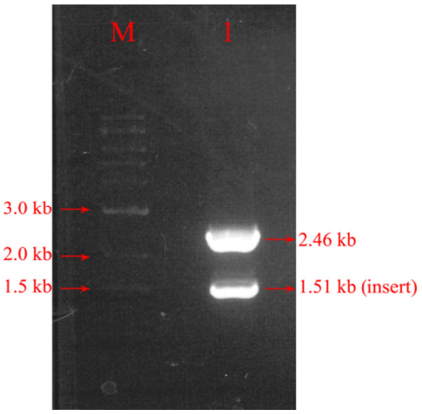


Figure 3.3 Agarose gel electrophoresis (0.8% w/v) of *NdeI-BamHI*-digested plasmid pUC57_opt-*Tb*PYK.

Lanes:

(1) *NdeI-BamHI*-digested plasmid construct of pUC57_opt-*Tb*PYK;

(M) GeneRulerTM 1kb DNA ladder is shown as M.

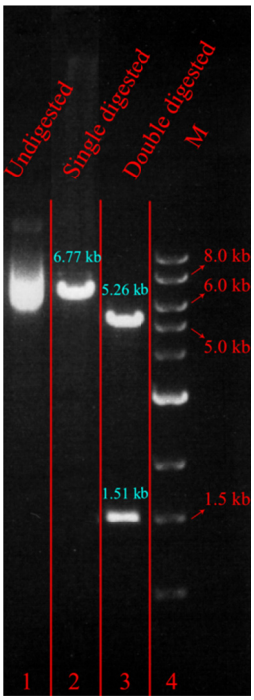


Figure 3.4 Agarose gel electrophoresis (0.8% w/v) of digested and undigested pET30a_opt-*Tb*PYK. †

Lanes:

(1) Undigested pET30a_opt-*Tb*PYK;

(2) *NdeI*-digested (single digested) pET30a_opt-*Tb*PYK;

(3) *NdeI-BamHI*-digested (double digested) pET30a_opt-*Tb*PYK;

(4) GeneRulerTM 1kb DNA ladder is shown as M.

† The non-relevant lanes were removed from the figure.

3.3.3. Expression of TcPYK and TbPYK in E. coli

3.3.3.1. Expression of His_6-*Tc*PYK

E. coli Rosetta(DE3)pLysS cells were transformed with pET28a_*Tc*PYK, and the expression of the His_6-*Tc*PYK gene was achieved by the T7 promoter-driven system in *E. coli*. After 30 min incubation at 4°C, expression was induced by adding IPTG to a final concentration of 1 mM. Maximum protein expression could be achieved by a further 22 h growth at 18°C. ~3 grams (wet weight) of cell pellets were harvested from 500 ml cell culture, where the yield of soluble His_6-*Tc*PYK was ~30 mg.

Untagged *Tc*PYK expression was carried out as well using the same expression method to that of His_6-*Tc*PYK, but with the plasmid pET30a_*Tc*PYK. The yield of active untagged *Tc*PYK can also achieve ~30 mg from 500 ml cell culture.

The similar final yields of His_6-*Tc*PYK and untagged *Tc*PYK indicate that the His_6-tag does not have a significant effect on protein expression in this case.

3.3.3.2. Expression of untagged *Tb*PYK (codon optimized)

Various *E. coli* expression strains, growth conditions (e.g. media, temperatures) and IPTG induction levels was carried out to express both untagged and His_6-tagged soluble *Tb*PYK. Unfortunately, almost all the expressed *Tb*PYK was inactive and was contained within inclusion bodies. The yield of the inactive *Tb*PYK can achieve up to 30 mg from 500 ml cell culture. All attempts at protein refolding were unsuccessful.

Codon optimization as another attempt was performed to obtain expression of pET30a_opt-*Tb*PYK, which eventually helped to achieve a high yield of soluble and active *Tb*PYK from *E. coli*. Six *E. coli* strains and three expression temperatures were applied in the expression trials trying to obtain the most efficient expression strategy. All the expression trials were under the T7 promoter-driven system and the expression was induced by IPTG (1 mM in final concentration). The results of

expression trials were analysed by SDS-PAGE and summarized in **Figure 3.5**. The expression conditions of 22 h of *E. coli* BL21(DE3) cell growth at 18°C in LB medium after 1 mM IPTG induction, were chosen as the optimum condition to express soluble untagged *Tb*PYK in *E. coli*. ~3 grams (wet weight) of cell pellets were harvested from 500 ml cell culture, where the yield of the soluble untagged *Tb*PYK was ~30 mg.

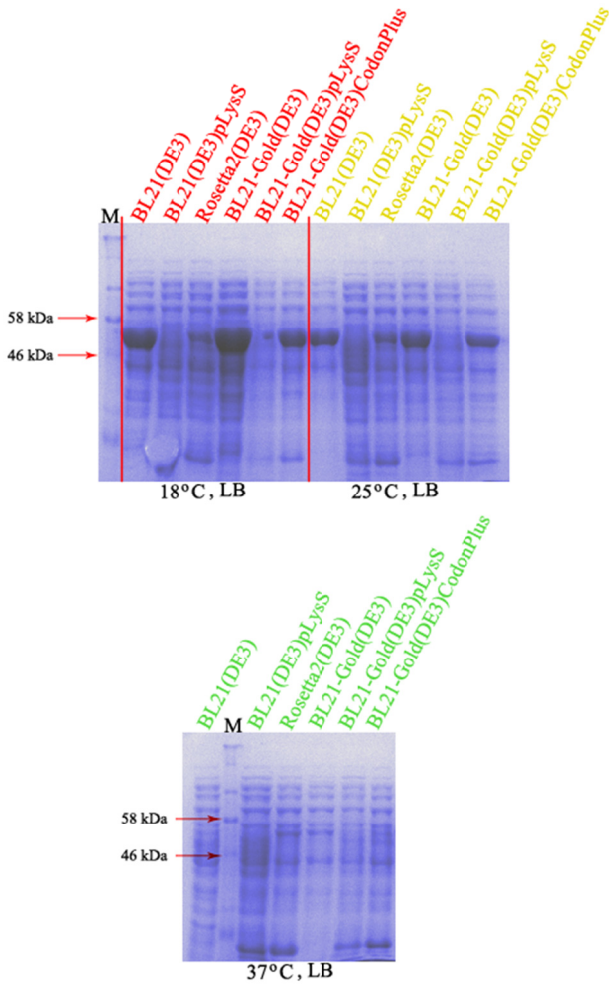


Figure 3.5 Expression trials for untagged *Tb*PYK .
Cell extracts from six different *E. coli* cell strains harbouring pET30a_opt-*Tb*PYK for 22 h growth in LB medium at three different temperatures after adding IPTG (1mM). Cells were lysed by a few cycles of sonication on ice and the soluble fractions were obtained by 13,000 rpm of centrifugation for 20 min at 4 °C. Soluble fractions from each induction sample were analysed by 12% SDS-PAGE and visualized by Coomassie staining. Protein markers are shown as M.

3.3.4. Purification of bacterially expressed PYKs

3.3.4.1. Purification of His_6-*Tc*PYK

i) Step One: IMAC (Immobilised Metal-Affinity Chromatography) (**Figure 3.6**)

The supernatants of cell lysates containing soluble His_6-*Tc*PYK were loaded onto a Ni-Sepharose column, and unbound proteins were washed away to the flow-through fractions. 20 mM of imidazole was maintained in the initial loading buffer (buffer A) to remove any weakly-bound proteins. Unbound proteins in the flow-through can be seen in SDS-PAGE in **Figure 3.6b**. His_6-*Tc*PYK was eluted gradually from the column by increasing the imidazole concentration using a linear gradient method at a constant pH, where imidazole displaced the His_6- tag from the nickel ion (**Figure 3.6a**). The purified proteins were analysed on a 12% SDS-PAGE where His_6-*Tc*PYK migrated as a single protein band (~56 kDa) with a few other protein bands indicating the presence of small amounts of contaminants (**Figure 3.6b**). A second purification procedure (gel-filtration chromatography in this experiment) was required for a purer quality of His_6-*Tc*PYK.

ii) Step Two: gel-filtration chromatography (**Figure 3.7**)

Purified samples from step one were concentrated and loaded onto a gel-filtration column (319 ml) to separate His_6-*Tc*PYK from other contaminants, making use of the difference in molecular size. His_6-*Tc*PYK eluted predominantly as a single peak (peak 2 in **Figure 3.7a**) with a retention volume of 164.97 ml which indicated His_6-*Tc*PYK was in a tetrameric form (~220 kDa) in solution. More than 95% purity was judged by 12% SDS-PAGE (**Figure 3.7b**). Peaks 1 were composed of multiple peaks eluted before the elution of His_6-*Tc*PYK, which possibly represented aggregated His_6-*Tc*PYK judged by SDS-PAGE (**Figure 3.7b**). Protein from peaks 1 migrated as a single protein band at the same position with the migration of His_6-*Tc*PYK from peak 2. Furthermore, the smaller retention volume indicated it is a larger oligomeric form of His_6-*Tc*PYK than its tetrameric form. A few minor peaks

eluted after the elution of peak 2 were suspected to be protein contaminants. Only the protein eluted from peak 2 containing pure His_6-*Tc*PYK was stored for further study. At least 20 mg of pure His_6-*Tc*PYK can be purified from 500-ml cell culture.

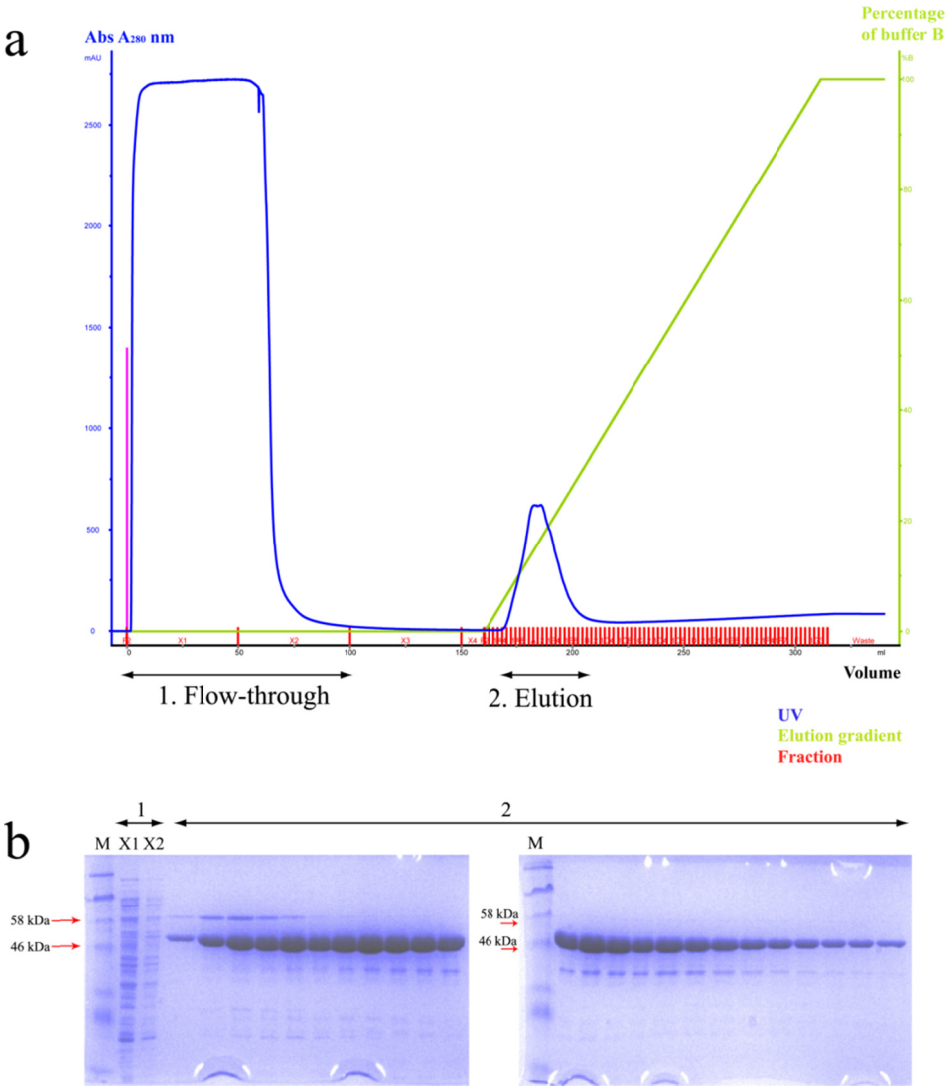


Figure 3.6 Purification of His_6-*Tc*PYK - Step One: IMAC

(a) The elution profile of His_6-*Tc*PYK from a 5 ml Hitrap HP Sepharose column using an imidazole linear gradient method. Fractions from the flow-through and the elution are indicated, and labelled as 1 and 2, respectively.

(b) A 12% SDS-PAGE with Coomassie staining is shown to analyse the corresponding fractions indicated in (a). Protein markers are shown as M.

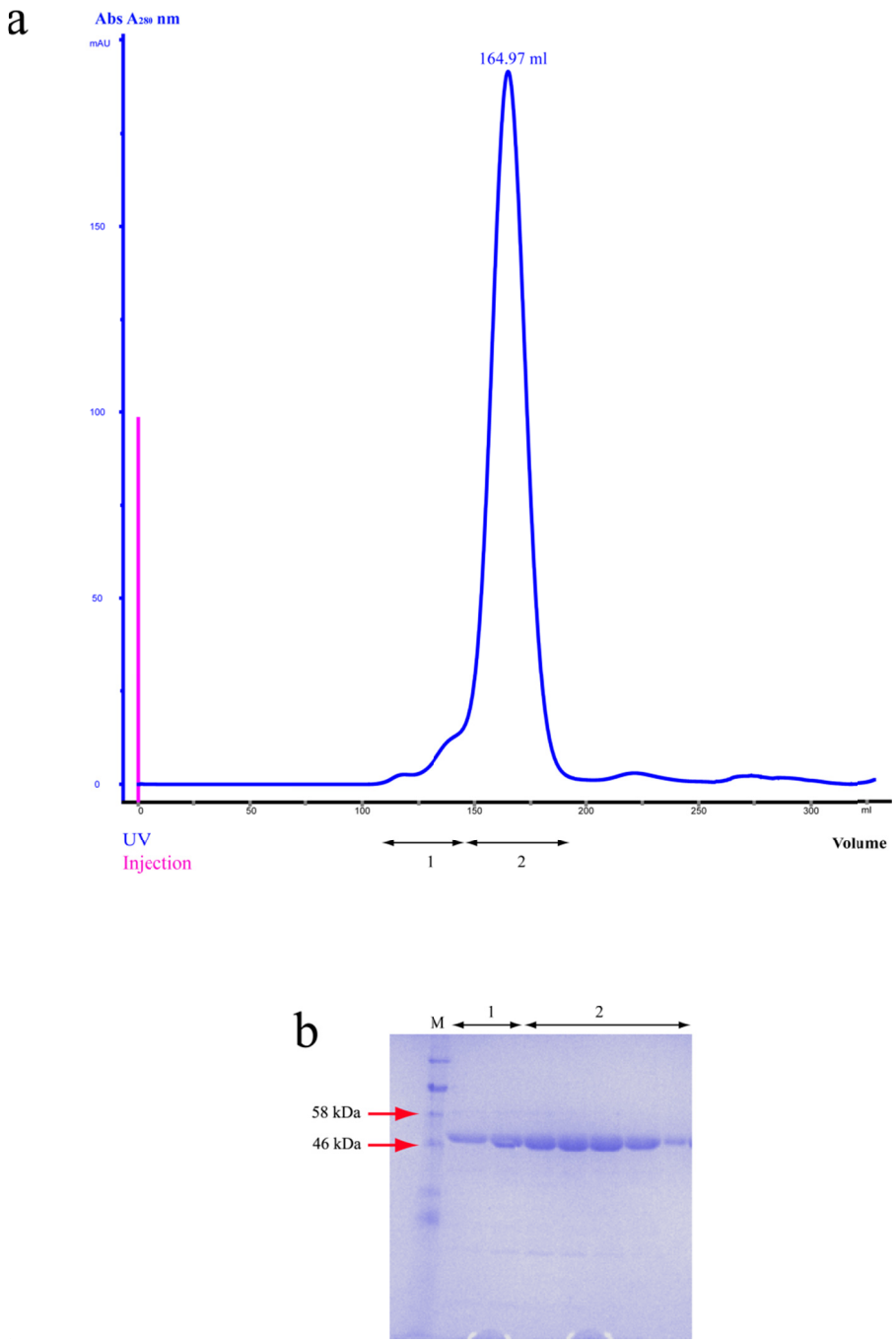


Figure 3.7 Purification of His_6-*Tc*PYK - Step Two: gel-filtration chromatography

(a) The elution profile of His_6-*Tc*PYK from a Superdex 200pg XK 26/60 gel-filtration column (column colume = 319 ml; void volume = 105 ml). The small peak shoulders and the major peak (retention volume = 164.97 ml) are indicated, and labelled as 1 and 2, respectively.

(b) A 12% SDS-PAGE with Coomassie staining is shown to analyse the corresponding peak fractions indicated in (a). Protein markers are shown as M.

3.3.4.2. Purification of untagged *Tb*PYK

i) Step One: ion-exchange chromatography (**Figure 3.8**)

The supernatants of cell lysates containing soluble untagged *Tb*PYK were loaded onto two tandem ion-exchange columns (Hiprep DEAE FF 16/10 and Hiprep SP FF 16/10 – fitted sequentially). Under the conditions of the buffer pH 7.9, column Hiprep DEAE FF 16/10 contained positively-charged resin while column Hiprep SP FF 16/10 contained negatively-charged resin; untagged *Tb*PYK had minimal net charge due to its theoretical pI = 7.74 which was close to buffer pH = 7.9; and contaminants were expected to have enough negative or positive net charge, resulting in binding to either of the ion-exchange columns. Therefore, untagged *Tb*PYK with minimal net charge had rather weak interactions with both of the ion-exchange columns and would be found in the flow-through. The bound proteins were eluted by increasing the salt concentration. Fractions from the flow-through and salt elution were analysed on a 12% SDS-PAGE (**Figure 3.8b**). Flow-through fractions were found to contain abundant untagged *Tb*PYK which migrated predominantly as a single protein band (~55 kDa) on SDS-PAGE but was still contaminated by a few other proteins judged by the protein bands on the gel. Salt elution fractions were composed of contaminants judged from the SDS-PAGE. A second purification procedure (gel-filtration chromatography in this experiment) was required for a purer quality of untagged *Tb*PYK.

ii) Step Two: gel-filtration chromatography (**Figure 3.9**)

Purified samples from step one were concentrated and loaded onto a gel-filtration column (319 ml) to separate untagged *Tb*PYK from contaminants making use of the difference in molecular size. Untagged *Tb*PYK eluted as a single peak (peak 2 in **Figure 3.9a**) with a retention volume of 171.69 ml which indicated untagged *Tb*PYK was in a tetrameric form (~220 kDa) in solution. More than 95% purity was judged by 12% SDS-PAGE (**Figure 3.9b**). Similar to the gel-filtration elution profile

for His_6-*Tc*PYK, peaks 1 which were composed of multiple and broad peaks eluted before the elution of untagged *Tb*PYK, contained aggregated untagged *Tb*PYK (data not shown). Multiple peaks belonging to peak region 3 represented the contaminating proteins (data not shown). Only the protein eluted from peak 2 containing pure untagged *Tb*PYK was stored for further study. At least 20 mg of pure *Tb*PYK can be purified from 500-ml cell culture.

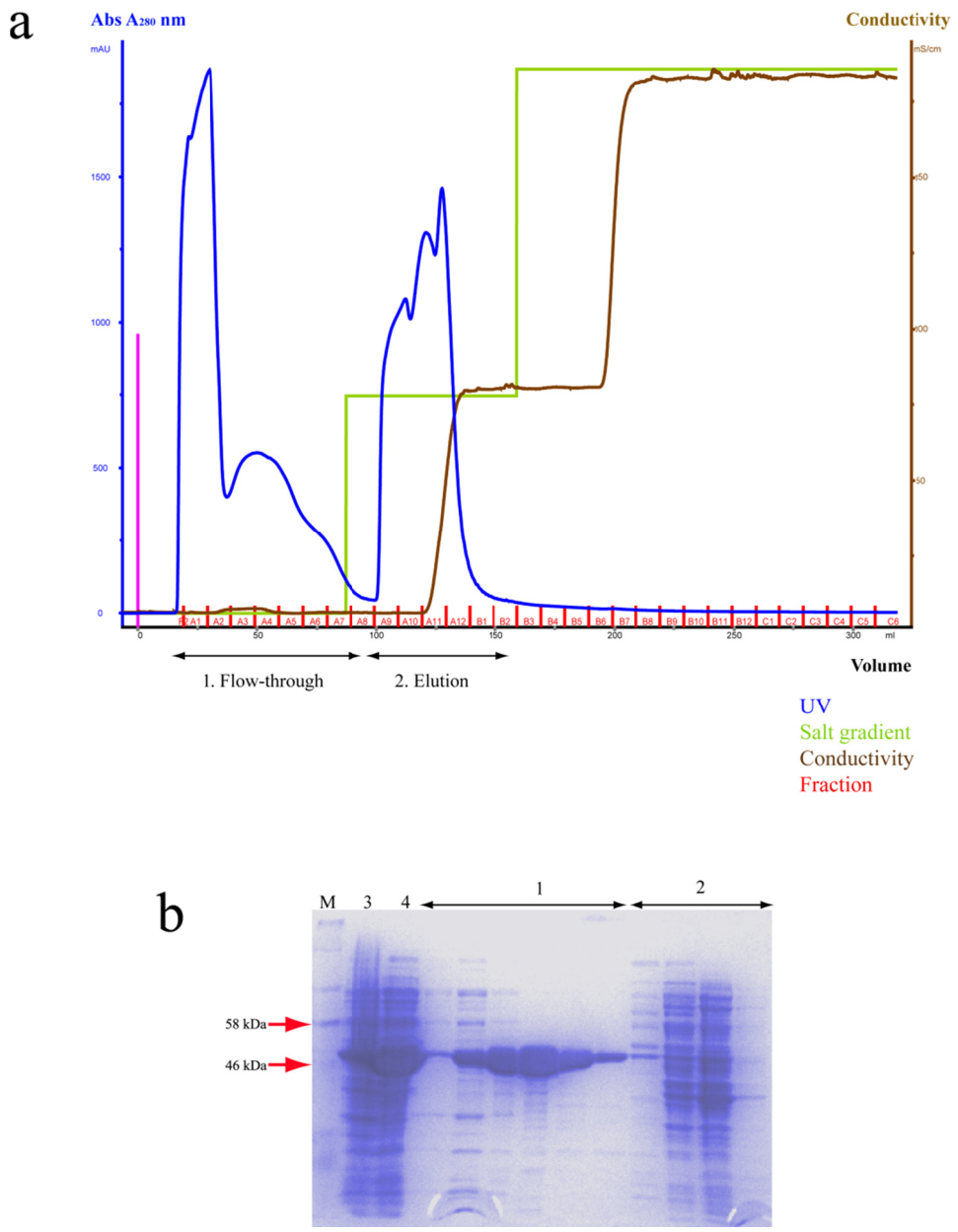


Figure 3.8 Purification of untagged *Tb*PYK Step One: ion-exchange chromatography

(a) The elution profile of untagged *Tb*PYK from two tandem ion-exchange columns (Hiprep DEAE FF 16/10 and Hiprep SP FF 16/10 – fitted sequentially). Fractions from the flow-through and elution are indicated, and labelled as 1 and 2, respectively.

(b) A 12% SDS-PAGE by Coomassie staining is shown to analyse the corresponding fractions indicated in (a). Untagged *Tb*PYK was contained in the flow-through fractions, while most of the other proteins were eluted by increasing salt concentration and can be seen in the elution fractions. Lane 3 is the total extract from a cell lysate; lane 4 is the soluble fraction. Protein markers are shown as M.

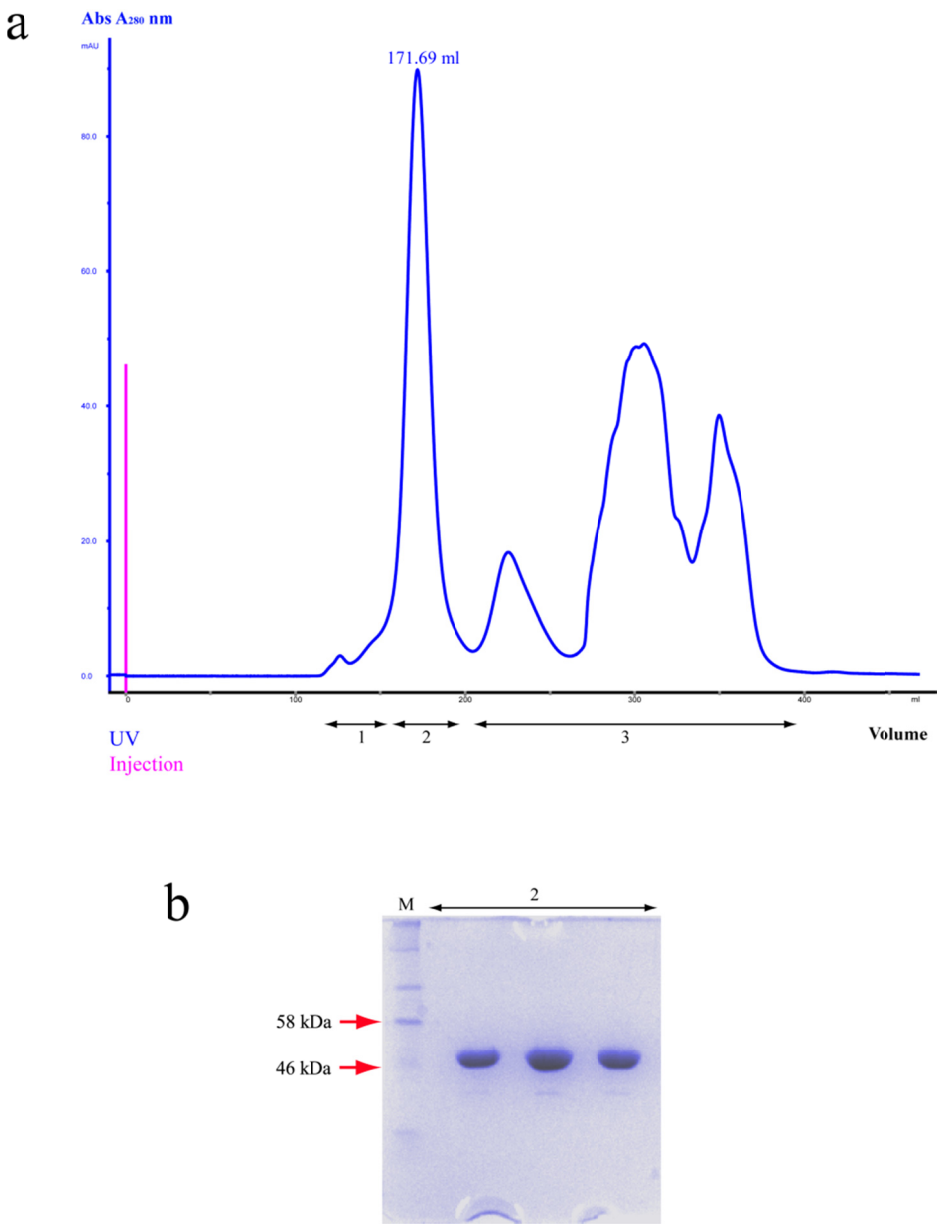


Figure 3.9 Purification of untagged *Tb*PYK Step Two: gel-filtration chromatography

(a) The elution profile of untagged *Tb*PYK from a Superdex 200pg XK 26/60 gel-filtration column (void volume = 105 ml). The first major peak (retention volume = 171.69 ml) corresponding to untagged *Tb*PYK is indicated, and labelled as 2. Peaks corresponding to aggregates of *Tb*PYK and other contaminants are labelled as 1 and 3, respectively.

(b) A 12% SDS-PAGE by Coomassie staining is shown to analyse the elution from peak 2 which was indicated in (a). Protein markers are shown as M.

3.3.5. TcPYK and TbPYK have similar kinetic parameters

*Tc*PYK and *Tb*PYK purified from *E. coli* in this experiment have many kinetic features in common, and their kinetic properties are also comparable with data previously reported for the enzymes from various trypanosomatids (Barnard and Pedersen 1988; Callens et al. 1991; Callens and Opperdoes 1992; Ernest et al. 1998; Cazzulo et al. 1989; Ernest et al. 1994). The kinetic comparisons for pyruvate kinases from different trypanosomatid species are summarized in **Table 3.3**. The kinetic parameters shown in **Table 3.3** were determined under similar assay conditions.

3.3.5.1. Effector F26BP abolishes sigmoidal kinetics with respect to substrate PEP

In the absence of effector, *Tc*PYK and *Tb*PYK exhibit sigmoidal kinetics with respect to the substrate PEP with an apparent affinity $S_{0.5}$ of 1.23 mM (±0.06) and 1.03 mM (±0.08), respectively. Additionally, the related Hill coefficients (*h*) of 2.06 (±0.06) and 1.88 (±0.12) for *Tc*PYK and *Tb*PYK respectively, indicate the positive cooperation of substrate PEP binding. The presence of 1 μM of the allosteric activator F26BP increases the apparent affinity for PEP to 0.137 mM (±0.017) and 0.119 mM (±0.026), for *Tc*PYK and *Tb*PYK respectively, abolishing the sigmoidal response to hyperbolic (lowering *h* to ~1.0). Moreover, the presence of F16BP at saturating concentration (4.5 mM), which is the allosteric activator for PYKs from many other species (e.g. human M2, *E. coli* PYKs), increases the apparent affinity for PEP as well (**Table 3.3**). Interestingly, unlike F26BP, the modulator F16BP fails to abolish or completely abolish the enzyme sigmoidal kinetics with respect to PEP, where *Tc*PYK has a Hill coefficient of 2.17 (±0.12) and *Tb*PYK has a Hill coefficient of 1.31 (±0.09) in the presence of F16BP. This is not an uncommon feature for F16BP on trypanosomatid PYKs which has already been shown in

previously reported results (**Table 3.3**). However, *Tb*PYK seems to behave quite differently compared to *Tc*PYK and *Lm*PYK, as *Tb*PYK is the only one to show a significant decrease in positive cooperation to PEP when adding F16BP, which can be seen from the decrease of Hill coefficient.

3.3.5.2. Turnover number is increased for both enzymes in the presence of activators F26BP or F16BP

At saturating concentrations of substrates, the turnover numbers (k_{cat}) for both *Tc*PYK and *Tb*PYK are slightly increased in the presence of F26BP or F16BP, compared to k_{cat} in the absence of effector. The k_{cat} for His_6-*Tc*PYK increased from ~10.15 ($min^{-1}x10^{-3}$) to ~12.72 ($min^{-1}x10^{-3}$), while the k_{cat} for *Tb*PYK increased from ~8.72 ($min^{-1}x10^{-3}$) to ~13.87 ($min^{-1}x10^{-3}$). The increased k_{cat} values indicate the increase of V_{max} for the same enzyme after adding F26BP or F16BP. The results are reproducible from batch to batch.

The increase of V_{max} at saturating concentrations of substrates on addition of effectors has previously been reported for *Tc*PYK in 1989 (Cazzulo et al. 1989). In the case of *Tb*PYK is not entirely clear why an increase in k_{cat} on addition of F26BP or F16BP has not previously been reported (**Table 3.3**). It other words, the addition of F26BP or F16BP did not affect the V_{max} of the enzyme from previously reported data. However, the different purification methods might be relevant to this observation. Previously, 0.1-0.5 mM PEP or/and 5-10 μM F26BP were used in the purification procedure to elute the enzyme from the column (Barnard and Pedersen 1988; Callens et al. 1991; Callens and Opperdoes 1992; Ernest et al. 1998). In this case, the small amount of substrate PEP and allosteric activator F26BP remaining in the purified enzyme might possibly have influenced the kinetic results.

The other factor which might be relevant to the observation is the presence of DMSO in the protein purification procedure. DMSO was sometimes used for *Tb*PYK purification and stabilization previously (Barnard and Pedersen 1988; Ernest

et al. 1998). Nevertheless, DMSO shows considerable activation of *Trypanosoma* PYKs even at 1% (v/v) concentration in our experiments. **Figure 3.10** shows the activation effects of DMSO on His_6-*Tc*PYK, where the enzyme doubles its activity in the presence of 2% DMSO compared to the activity in the absence of DMSO. DMSO has similar activation effects on *Tb*PYK as well (data not shown). A large amount of DMSO (up to 20% v/v) was used to purify or/and stabilize the enzyme *Tb*PYK previously (Barnard and Pedersen 1988; Ernest et al. 1998) and no evidence was shown whether DMSO had been completely removed from the purified proteins or not. Therefore, for previously reported data, the potential presence of DMSO in the enzyme will also affect the kinetic parameters, although the activation mechanism by DMSO is still unknown. In addition, the purity and quality of the purified enzymes should be considered.

Additionally, sulfate (SO_4^{2-}) which has been proved to bind to the effector site of *Lm*PYK and affect the enzyme activity (Tulloch et al. 2008) were present in the enzyme assay previously (Barnard and Pedersen 1988; Callens et al. 1991; Callens and Opperdoes 1992; Ernest et al. 1998; Ernest et al. 1994). By contrast, none of the substrates, effector F26BP or DMSO was used in the purification procedures and sulfate was not used in the enzyme assay as well in this work.

3.3.5.3. Inconsistent turnover numbers from different laboratories

For *Tb*PYK, the kinetic experiments were all performed under similar assay conditions, including pH, temperature, salt concentrations, and substrate or effector concentrations. Many kinetic parameters reported by different authors including the results from this experiment are quite consistent or similar, including the affinity for substrates PEP or ADP and the affinity for effector F26BP or F16BP. Interestingly, differences in turnover numbers (k_{cat}) for *Tb*PYK reported by different authors were found. At saturating concentrations of substrates, in the absence or presence of effector F26BP or F16BP, k_{cat} for *Tb*PYK has a range of ~9-14 ($min^{-1}x10^{-3}$)

determined in our experiments, while the number of k_{cat} was around 30 ($min^{-1}x10^{-3}$) as reported by Callens M. and Opperdoes F.R. in 1992 (Callen and Opperdoes 1992), and was around 20 ($min^{-1}x10^{-3}$) as reported by Ernest I. *et al* in 1998 (Ernest et al. 1998). These differences were discussed in the latter paper with the observation that the quantitative results for the purified enzyme were largely dependent on the method used for measurement of protein concentration. The measured quantity of the enzyme would directly affect the calculation of k_{cat} number. Additionally, as had been stated in **section 3.3.5.2**, sulfate was maintained in the enzyme assay and the enzyme purified by different groups may have been contaminated with PEP, F26BP and possibly DMSO as well. All those contaminants may affect the activity of the enzyme and then influence the determination of k_{cat} number.

3.3.5.4. Both PYKs display hyperbolic kinetics with respect to substrate ADP

*Tc*PYK and *Tb*PYK display hyperbolic kinetics with respect to their substrate ADP, with apparent affinity K_m of 0.398 mM (±0.036) and 0.136 mM (±0.028), respectively. These values are comparable to those from previously reported data in **Table 3.3**.

3.3.5.5. Both PYKs show nanomolar affinity for allosteric activator F26BP

As stated in **section 3.3.5.1**, F26BP at micromolar concentration (1 μM) converted the kinetic behaviors of *Tc*PYK and *Tb*PYK with respect to the substrate PEP to a hyperbolic response, and considerably increased the enzymes' apparent affinities for PEP. The apparent affinity of His_6-*Tc*PYK ($S_{0.5}$) for F26BP exhibits a hyperbolic saturation curve with $S_{0.5}$ of 31 nM (±1). The $S_{0.5}$ value is closer to that of *Tb*PYK from previously reported data shown in **Table 3.3** than that for *Lm*PYK, which has a ten-fold lower affinity for F26BP compared to *Tb*PYK and *Tc*PYK. Surprisingly, *Lm*PYK displays sigmoidal behavior for F26BP with h=1.96 (±0.31). In short, *Tc*PYK and *Tb*PYK have more similarities with respect to F26BP binding when

compared to *Lm*PYK. However, we should be aware that *Lm*PYK was purified in the presence of phosphate ions and assayed in the presence of sulphate (SO_4^{2-}) ions (Ernest et al. 1994). Both ligands competitively bind to the phospho group binding sites of effector F26BP, substrate ADP or product ATP on the enzymes (Tulloch et al. 2008).

3.3.5.6. *Tb*PYK shows significantly higher affinity (μM range) for F16BP compared to *Tc*PYK and *Lm*PYK

Although F26BP is generally considered to be the principal allosteric activator for trypanosomatid PYKs, F16BP also shows considerable activation. It increases the enzymes' apparent affinity for PEP without completely abolishing the sigmoidal response (**Table 3.3**). The apparent affinities ($S_{0.5}$) of His_6-*Tc*PYK and *Tb*PYK for F16BP are given in **Table 3.3**. All three trypanosomatid PYKs display almost hyperbolic responses to F16BP with *h* close to 1.0. For *Tb*PYK, both our results and previously reported data indicate that it has μM range affinity for F16BP ($K_{a0.5}$≈47 μM in our experiments). Interestingly, both *Tc*PYK and *Lm*PYK have mM range affinities for F16BP, namely 1.24 mM (±0.24) and 0.520 mM (±0.211) respectively.

Table 3.3 Comparison of kinetic properties of trypanosomatid PYKs

Ligand	Modulator	Kinetic properties Parameter	*Tc*PYK[a]	*Tb*PYK[b]	*Tb*PYK[c]	*Tb*PYK[d]	*Lm*PYK[e]
PEP	None	$S_{0.5}$(mM)	1.23±0.06	1.03±0.08	1.48	1.28	1.44±0.28
		h	2.06±0.16	1.88±0.12	2.4	3.2	2.81±0.52
		k_{cat}($min^{-1}x10^{-3}$)	10.15±0.32	8.72±0.16	33.6	20.28	21±1.5
		$k_{cat}/S_{0.5}$ ($mM^{-1}min^{-1}x10^{-3}$)	8.25	8.49	22.70	15.84	14.58
	F26BP	$S_{0.5}$(mM)	0.137±0.017	0.119±0.026	0.15	0.055	0.195±0.004
		h	1.13±0.13	1.19±0.10	1.0	1.0	2.72±0.70
		k_{cat}($min^{-1}x10^{-3}$)	12.72±0.45	13.87±0.26	28.6±0.7	18.96	11.0±0.6
		$k_{cat}/S_{0.5}$ ($mM^{-1}min^{-1}x10^{-3}$)	92.85	116.5	190.67	344.73	56.41
	F16BP	$S_{0.5}$(mM)	0.347±0.010	0.261±0.040	0.46	n.d.	0.840±0.003
		h	2.17±0.12	1.31±0.09	1.4	n.d.	2.68±0.37
		k_{cat}($min^{-1}x10^{-3}$)	13.55±0.17	11.14±0.23	31	n.d.	25.0±0.1
		$k_{cat}/S_{0.5}$ ($mM^{-1}min^{-1}x10^{-3}$)	39.06	42.68	67.39	n.d.	29.76
ADP	None	K_m(mM)	0.398±0.036	0.136±0.028	0.114±0.013	0.109±0.048	0.264±0.007
		k_{cat}($min^{-1}x10^{-3}$)	11.74±0.36	8.92±0.22	n.d.	16.872±0.502	8.95±0.31
		k_{cat}/K_m ($mM^{-1}min^{-1}x10^{-3}$)	29.50	65.76	n.d.	154.79	33.90
F26BP	None	$K_{a0.5}$(μM)	0.031±0.001	n.d.	0.056±0.002	0.013±0.006	0.290±0.001
		h	0.94±0.04	n.d.	1.03±0.10	1	1.96±0.31
		k_{cat}($min^{-1}x10^{-3}$)	6.27±0.04	n.d.	27.3±0.46	14.088±1.080	12.9±0.5
F16BP	None	$Ka_{0.5}$(mM)	1.24±0.24	0.047±0.029	0.121±0.003	0.026±0.009	0.520±0.211
		h	0.92±0.09	1.25±0.19	1.2±0.06	1	0.92
		k_{cat}($min^{-1}x10^{-3}$)	7.44±0.48	5.64±0.19	23.5±0.25	13.872±1.368	8.6±0.9

[a] The *Tc*PYK is the N-terminal His_6-tagged version and purified from *E. coli*. (this work); assayed in the presence of $MgCl_2$ instead of $MgSO_4$.
[b] The *Tb*PYK is the untagged version and purified from *E. coli*. (this work); assayed in the presence of $MgCl_2$ instead of $MgSO_4$.
[c] The *Tb*PYK is from bloodstream form *T. brucei*. (Barnard and Pedersen 1988; Callens et al. 1991; Callens and Opperdoes 1991)
[d] The *Tb*PYK is the untagged version and purified from *E. coli* (Ernest et al. 1998)
[e] The *Lm*PYK is the untagged version and purified from *E. coli*. (Ernest et al. 1994)
The values given are means ± SD.
h: Hill coefficient
n.d.: not done

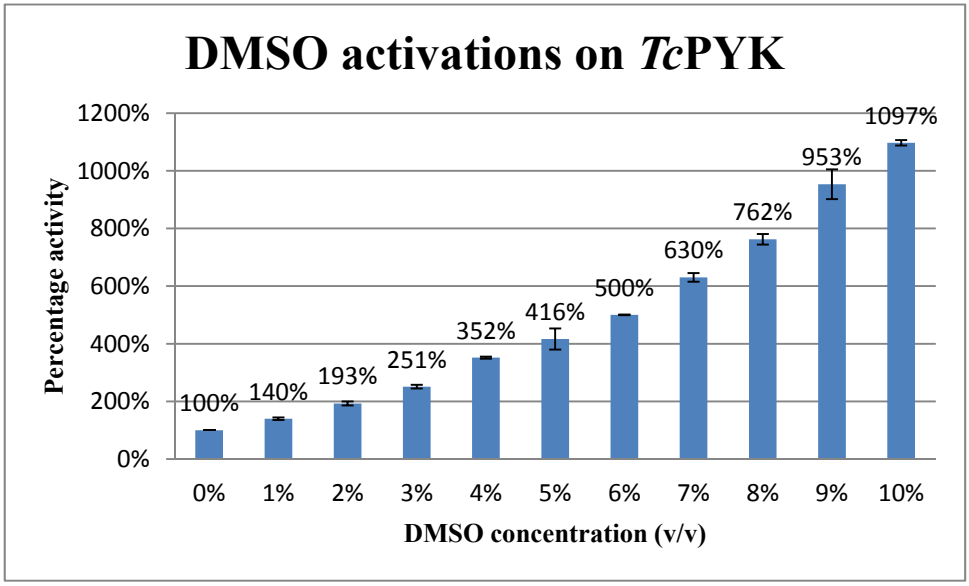


Figure 3.10 DMSO activation of His_6-*Tc*PYK.
Concentration-response bar chart observed for the titration of DMSO against His_6-*Tc*PYK activity. The assay was performed in triplicate at 25°C in standard buffer condition (50 mM TEA, pH7.2, 50 mM KCl, 10 mM $MgCl_2$), supplemented with 0.2 mM ADP, 0.4 mM PEP and variable concentrations (v/v) of DMSO. Each error bar represents the standard deviation in the experiment.

CHAPTER 4: Dye Molecule Additives: an Improved Strategy for the Crystallisation of Pyruvate Kinases from *Trypanosoma*

4.1. Introduction

The production and purification of active *Tc*PYK and *Tb*PYK from *E. coli* cells with high efficiency has been described in **Chapter 3**. The development of a repeatable crystallisation protocol to obtain high-quality diffracting crystals of *Tc*PYK and *Tb*PYK therefore became a priority for the purposes of structural studies and structure-based inhibitor design. Previously, collections of small molecules have been screened by McPherson and his colleagues (McPherson and Cudney 2006; Larson et al. 2007) on quite a broad range of proteins for crystallisation trials, and have demonstrated that small molecules or cocktails of small molecules are capable of promoting the formation of crystal lattice. These molecules act by providing crosslinks and by stabilising intermolecular interactions, thereby favouring the growth of crystals and increasing their diffracting quality. However, these proven additives in crystallisation conditions are unlikely to be 'silver bullets' for every protein.

A series of sulfonic acid dye-like compounds have been identified in *L. mexicana* PYK (*Lm*PYK, ~74% sequence identity to *Tc*PYK and *Tb*PYK) structures where several compounds are clearly shown to act as intermolecular bridges to stabilise lattice formation in crystals (Morgan et al. 2010a; Morgan et al. 2011). In the work reported in this chapter, a collection of dye molecules, most of which contain sulfonic acid groups such as ponceau S and trypan blue, were used as additives in crystallisation trails with the goal of finding customised 'silver bullets' to crystallise *Tc*PYK and *Tb*PYK. Additionally, it can be mentioned that quite a few

candidate additives (e.g. ponceau S, trypan blue, acid blue 80) are chemically related to suramin which is the drug for the treatment of human African trypanosomiasis (sleeping sickness) caused by *T. brucei* (Morgan et al. 2011; Ehrlich and Shiga 1904; Hawking 1978). The structural basis of enzyme inhibition effects on *Lm*PYK by these dye-like compounds including suramin have been characterised by colleague Dr. Hugh Morgan (Morgan et al. 2011). Therefore, it will also be of interest to explore the interactions between the additive molecules and the protein macromolecule from a structural viewpoint.

The structure of *Tc*PYK in complex with ponceau S has been deposited in PDB (*Tc*PYK/ponceau S: 3QV9) and published in:

Morgan, H.P., McNae, I.W., Nowicki, M.W., **Zhong, W.**, Michels, P.A., Auld, D.S., Fothergill-Gilmore, L.A., Walkinshaw, M.D. (2011) The trypanocidal drug suramin and other trypan blue mimetics are inhibitors of pyruvate kinases and bind to the adenosine site. *J. Biol. Chem.* **286**, 31232-31240.

The structure of *Tb*PYK in complex with F26BP determined from the crystal growing with ponceau S has been deposited in PDB (*Tb*PYK/F26BP/Mg: 4HYW) and published in:

Zhong, W.[*], Morgan, H. P.[*], McNae, I. W., Michels, P. A., Fothergill-Gilmore, L. A., Walkinshaw, M. D. *'In crystallo'* substrate binding triggers major domain movements and reveals magnesium as a co-activator of *Trypanosoma brucei* pyruvate kinase. (Manuscript submitted to *Acta Crystallogr. D*) [*] Co-first authors

4.2. Materials and Methods

The materials and the general methods used in this chapter have been described in **Chapter 2**. In brief, the structures of apoenzyme *Tc*PYK, *Tc*PYK/Ponceau S, *Tb*PYK/F26BP/Mg (2.95 Å) and *Tb*PYK/F26BP/Mg (2.35 Å) were solved by the methods described in **Section 4** of **Chapter 2**.

Table 4.1 Dyes used in crystallisation trials for *Tc*PYK and *Tb*PYK[†]

	Dye molecule	Supplier& Product number	Formula	Molecular Weight
1	Ponceau S	Sigma, P3504	$C_{22}H_{12}N_4Na_4O_{13}S_4$	760.57
2	1,3,6,8-Pyrenetetrasulfonic acid	Sigma, 82658	$C_{16}H_6Na_4O_{12}S_4$	610.43
3	Acid blue 80	Sigma, 210323	$C_{32}H_{28}N_2Na_2O_8S_2$	678.68
4	Acid blue 25	Sigma, 210684	$C_{20}H_{13}N_2NaO_5S$	416.38
5	Benzothiazole-2,5-disulfonic acid	Sigma, S628492	$C_7H_3NNa_2O_6S_3$	339.277
6	Suramin	Sigma, S2671	$C_{51}H_{34}N_6Na_6O_{23}S_6$	1429.17
7	Trypan blue	Sigma, T6146	$C_{34}H_{24}N_6O_{14}S_4Na_4$	960.81
8	Reactive blue 2	Sigma, R115	$C_{29}H_{17}ClN_7Na_3O_{11}S_3$	840.10
9	Reactive blue 4	Sigma, 244813	$C_{23}H_{14}Cl_2N_6O_8S_2$	637.43
10	Dimethyl 5-sulfoisophthalate	Sigma, 150010	$NaO_3SC_6H_3$-1,3-$(CO_2CH_3)_2$	296.23
11	4-Amino-5-hydroxy-2,7-naphthalenedisulfonic acid	Sigma, 528765	$H_2NC_{10}H_4(OH)$-$(SO_3H)SO_3Na$	341.29
12	Morin hydrate	Sigma, M4008	$C_{15}H_{10}O_7$	302.24
13	Lucifer yellow CH	Sigma, L0144	$C_{13}H_9K_2N_5O_9S_2$	521.57
14	Hydroxy naphthol blue	Sigma, 219916	$C_{20}H_{12}N_2O_{11}S_3Na_2$	598.49
15	Pyridoxal phosphate-6-azo(benzene-2,4-disulfonic acid)	Sigma, P178	$C_{14}H_{10}N_3Na_4O_{12}PS_2$	599.31
16	2,2′-Dihydroxy-1,1′-azonaphthalene-3,3′,6,6′-tetrasulfonic acid	Sigma, D5021	$C_{20}H_{14}N_2O_{14}S_4$	634.59
17	Naphthol blue black	Sigma, N9002	$C_{22}H_{14}N_6Na_2O_9S_2$	616.49
18	Sulfonazo III	Sigma, 86170	$C_{22}H_{12}N_4Na_4O_{14}S_4$	776.57

[†] Dyes 1-6 which have been reported in published *Lm*PYK structures (Morgan et al. 2011) are shaded in blue. Ponceau S (dye No.1), which has been shown in the experiments reported here to successfully improve the crystallisation of *Tc*PYK and *Tb*PYK, is highlighted in red.

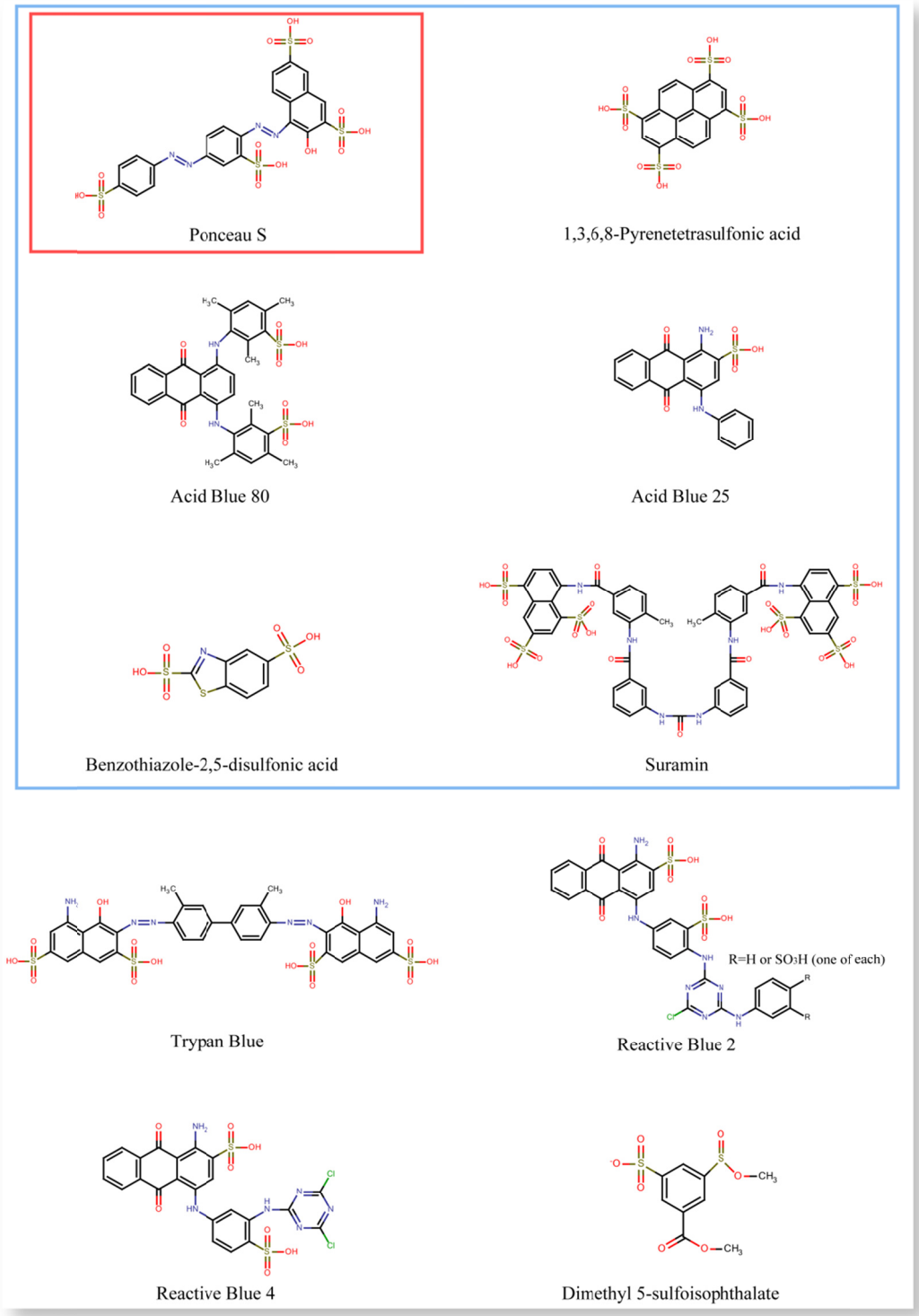


Figure 4.1 Two-dimensional representations of selected dyes used in crystallisation trials for *Tc*PYK and *Tb*PYK, corresponding to Table 4.1.
The dyes which have been reported in published *Lm*PYK structures (Morgan et al. 2011) are enclosed within the blue box. The dye ponceau S which is proved to successfully improve the crystallisation of *Tc*PYK and *Tb*PYK in the experiments reported here is enclosed within the red box.

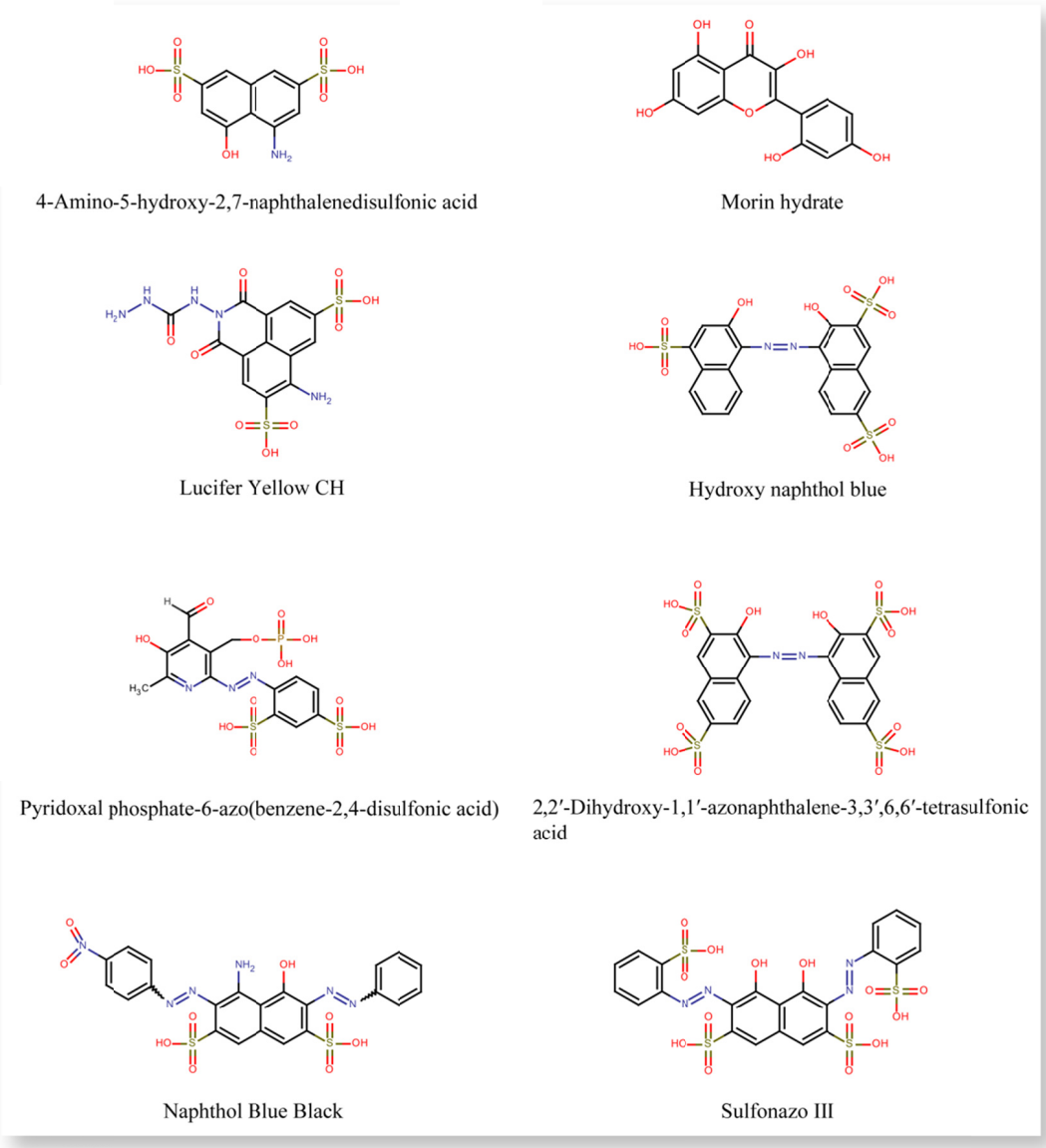


Figure 4.1 ***(continued)***

Table 4.2 X-ray data-collection parameters and refinement statistics

	apo *Tc*PYK	*Tc*PYK/ponceau S	*Tb*PYK/F26BP/Mg	*Tb*PYK/F26BP/Mg†
Data collection				
Space group	*C*2	*C*2	$P2_1$	*I*222
Cell dimensions				
a, *b*, *c* (Å)	113.78,121.42,97.16	113.20,121.83,96.54	177.86,110.07,185.47	103.68,109.00,268.38
α, *β*, *γ* (°)	90.00,115.73,90.00	90.0,109.81,90.0	90.00,105.57,90.00	90.00,90.00,90.00
Solvent content (%)	53	55	54	65
Wavelength (Å)	0.98	0.98	0.98	0.98
Resolution (Å)	43.76-2.50	90.82-2.10	30.00-2.95	54.50-2.35
No. reflections	252349(31324)	440370(36069)	375823(54438)	411424(60956)
No. unique reflections	38483(5416)	70435(9217)	131115(19481)	63644(9232)
Wilson B-factor	37.4	25.9	48.2	42.2
R_{merge} (%)	9.0(25.7)	11.8(47.9)	15.8(40.2)	8.8(54.6)
R_{meas} (%)	10.5(30.6)	13.8(61.1)	19.1(49.0)	10.5(64.8)
$R_{p.i.m.}$(%)	5.5(16.4)	6.8(37.4)	10.5(27.5)	5.6(34.4)
$\langle I/\sigma I \rangle$	13.7(5.8)	9.4(2.4)	5.4(2.5)	12.4(3.1)
% completeness	93.7(90.5)	98.2(88.6)	90.7(92.6)	100.0(100.0)
Multiplicity	6.6(5.8)	6.3(3.9)	2.9(2.8)	6.5(6.6)
Refinement statistics				
Monomers in ASU	2	2	12	2
No. reflections	36536	66852	124292	60417
R_{work} / R_{free}	16.15/22.12	19.8/24.9	22.49/27.29	16.38/19.87
RMS deviations				
Bond lengths (Å)	0.01	0.02	0.01	0.01
Bond angles (degree)	0.99	1.92	1.05	0.99

Values in parentheses are for the highest resolution shell.
† This structure is determined from the crystal growing with ponceau S.

4.3. Results and Discussion

4.3.1. Ponceau S is the 'silver bullet' for obtaining high-quality diffracting crystals of TcPYK and TbPYK

Six of the dye molecules studied in the work reported here (dyes 1-6 in **Table 4.1** and **Figure 4.1**) have previously been co-crystallised with *Lm*PYK (Morgan et al. 2010a; Morgan et al. 2011), and some of them (e.g. 1,3,6,8-pyrenetetrasulfonic acid and acid blue 80) have been shown to stabilise the crystal lattice of *Lm*PYK. However, their effects on the crystal growth of *Tc*PYK and *Tb*PYK (both have ~74% sequence identity to *Lm*PYK) are varied. In the absence of dyes, crystals of apo *Tc*PYK and crystals of *Tb*PYK/F26BP/Mg were obtained. However, it took rather a long time to grow apo *Tc*PYK crystals, and the growth conditions were strict. Furthermore, the diffracting quality of the *Tb*PYK/F26BP/Mg crystals was poor and the data were relatively difficult to process (see section **4.3.3**).

A total of 18 dyes were used to crystallise *Tc*PYK and *Tb*PYK in the presence or absence of effector (F26BP) and/or substrates or analogues (oxalate, ATP). From the dye screening trials, the addition of ponceau S was capable of improving crystal growth and diffracting quality for both *Tc*PYK and *Tb*PYK. The ligand ponceau S could be identified in the *Tc*PYK structure determined from the crystal grown in the presence of ponceau S, where it acts as an intermolecular bridge between macromolecules in the crystal. The *Tc*PYK crystal grown in the presence of ponceau S is referred to as *Tc*PYK/ponceau S. For the *Tb*PYK/F26BP/Mg structure determined from a crystal grown with ponceau S and F26BP, the position of the ponceau S molecule could not be clearly identified, although the space group was different from that of the structure from a crystal grown only with F26BP but not ponceau S. The *Tb*PYK crystal grown in the presence of F26BP and ponceau S is referred to as *Tb*PYK/F26BP/Mg(/ponceau S).

In the next two sections, detailed comparisons between structures determined

from crystals grown with or without ponceau S will be discussed.

4.3.2. Ponceau S improves the crystallisation of TcPYK

4.3.2.1. Dye screening result for the crystallisation of *Tc*PYK

i) Apo *Tc*PYK crystals grown without dyes

Apo *Tc*PYK crystals were grown in 11-15% PEG8000, 10% glycerol, and standard buffer [50 mM TEA (pH 7.2), 100 mM KCl, and 50 mM $MgCl_2$] at 4^{o}C. Crystals appeared after three months and grew to maximum dimensions between 0.1x0.1x0.1 mm - 0.3x0.15x0.1 mm (**Figure 4.2a**). The coverslip holding the crystals were transferred to a new well filled with 1 ml of well solution containing 20% PEG8000, 20% glycerol and standard buffer, and equilibrated overnight before being flash frozen in liquid nitrogen. The bigger crystal (**Figure 4.2a**) was shown by preliminary scanning to have better diffracting quality and was chosen for x-ray data collection to determine the crystal structure of apo *Tc*PYK (**Figure 4.2c**). Crystals were not grown in other crystallisation conditions.

ii) *Tc*PYK/ponceau S crystal grown with ponceau S

Ponceau S is the only dye molecule that showed significant improvement in *Tc*PYK crystallisation in the absence of other ligands. *Tc*PYK/ponceau S crystals were red in colour, and were grown in a rather broad range of conditions of 6-20% PEG8000, 10-20% glycerol, and standard buffer [50 mM TEA (pH 7.2), 100 mM KCl, and 50 mM $MgCl_2$] at 4^{o}C. The bigger crystals with higher diffracting quality were grown in conditions containing 20% glycerol. Crystals appeared the next day after setting the crystal trays and took 3-5 days to grow to maximum dimensions (**Figure 4.2b**). Crystals grown in different conditions were screened first, and one of the crystals grown in standard buffer condition containing 18% PEG8000 and 20% glycerol was shown to have the best diffracting quality. This crystal was chosen for x-ray data

collection to determine the crystal structure of *Tc*PYK/ponceau S (**Figure 4.2d**).

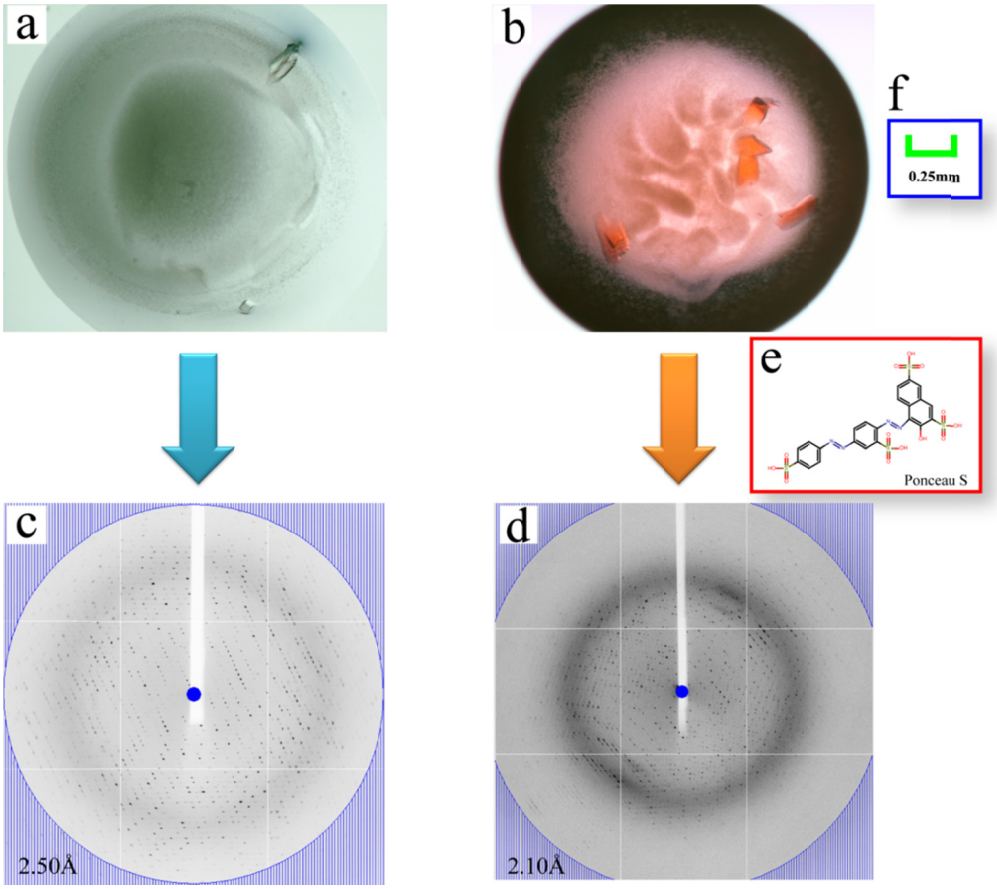


Figure 4.2 Improvement of His$_6$-*Tc*PYK crystal quality in the presence of ponceau S.

(a, b) Light-microscope photographs of His$_6$-*Tc*PYK crystals in the absence (a) and presence (b) of ponceau S. The photos were taken using the same scale shown in (f). Two crystals were observed in (a), where one crystal (0.3x0.15x0.1 mm) was at about one o'clock and the other (0.1x0.1x0.1 mm) at five o'clock. It took three months to grow these crystals without ponceau S. Red crystals (0.3x0.15x0.1 mm) were found in (b), which only took 3-5 days to grow to maximum dimensions. The red colour of the crystals is an indication of the binding of ponceau S.

(c, d) The oscillation images during data collection from His$_6$-*Tc*PYK crystals grown in the absence (c) and presence (d) of ponceau S. The images shown here correspond to the diffraction patterns used for structure determinations. The edges of the oscillation images are indicated by blue circles, and the maximum resolution of the images are shown in the lower left corner.

(e) Two-dimensional representation of ponceau S, which was used as an additive to improve the crystallisation of both *Tc*PYK and *Tb*PYK.

4.3.2.2. X-ray data analysis: data collection and processing

Data for apo *Tc*PYK and *Tc*PYK/ponceau S were collected on beamline I04 and I02, respectively. For the apo *Tc*PYK crystal, the data were collected to a maximum resolution of 2.50 Å as 1^{o} oscillation images over 180^{o}. Data were indexed in the monoclinic space group of *C*2, with unit cell dimensions of a=113.78 Å, b=121.42 Å, c=97.16 Å and angles of $\alpha=90.00^{o}$, $\beta=115.73^{o}$, $\gamma=90.00^{o}$. For the *Tc*PYK/ponceau S crystal, the data were collected to a maximum resolution of 2.10 Å as 0.8^{o} oscillation images over 240^{o}. Data were indexed in the monoclinic space group of *C*2, with unit cell dimensions of a=113.20 Å, b=121.83 Å, c=96.54 Å and angles of $\alpha=90.00^{o}$, $\beta=109.81^{o}$, $\gamma=90.00^{o}$. Other information on data collection and processing can be seen in **Table 4.2**. Apart from the same space group, the unit cells from these two structures are also quite similar as well, indicating that the presence of ponceau S in the crystallisation conditions possibly did not disrupt the original macromolecular arrangement in the crystal.

4.3.2.3. Content of the asymmetric unit and obtaining phases

For *Tc*PYK with a molecular weight of 56769.1 Da, the program Cell Content Analysis based on Matthew's equation (Matthews 1968) in CCP4 was used, indicating there were two monomers per asymmetric unit for space group *C*2 for both apo *Tc*PYK and *Tc*PYK/ponceau S. Phases for both structures were determined by PHASER (McCoy et al. 2007) using the method of molecular replacement, with the input search model of the monomer from the previously determined *Tc*PYK/F26BP/OX/Mg. Two monomers in the crystallographic asymmetric unit were identified for both structures as a result of molecular replacement. These two crystallographically unique monomers with crystallographic twofold symmetry generate a biologically relevant tetramer. The resulting electron density maps were clearly visible except that the densities for one of the two B-domains (residues 90-187) and the effector loops (residues 482-488) within the A-domains were poorly

identified at this stage.

4.3.2.4. Structure solutions for apo *Tc*PYK and *Tc*PYK/ponceau S

In the final model of apo *Tc*PYK after structure refinement, both B-domains were successfully built up except that the effector loops of both chains were still missing. The poor electron density for the B-domain is not uncommon for PYK structures, as it is a flexible domain in the absence of coordination with active-site ligands. The flexibility of the B-domain will be discussed in **Chapters 5**. A similar situation is observed with the effector loop, which is normally disordered without effector bound and could not be modeled.

Both chains in the structure were similar, with a root-mean-square (RMS) deviation of 0.20 Å for all C-α atoms of residues 2-89 and 188-499 (excluding the B-domains). The RMS deviation for all C-α atoms of residues 2-499 is much larger at 1.34 Å because of the flexible B-domains (**Table 4.3**). For *Tc*PYK/ponceau S, two monomers were identified in the final model where only one B-domain was possible to build. Interestingly, one effector loop without a ligand bound to the effector site could eventually be built with an average B-factor of 55.7 $Å^2$, compared to the average B-factor of 31.0 $Å^2$ for the overall structure (protein only). The electron density for ponceau S was clearly visible, and was modelled in the structure. Both chains in the structure were similar, with an RMS deviation of 0.27 Å for all C-α atoms of residues 2-89 and 188-499 (excluding the B-domains) (**Table 4.3**).

4.3.2.5. Ponceau S stabilises lattice formation in *Tc*PYK/ponceau S crystals

Ponceau S molecules have been clearly modelled (**Figure 4.3c, d**) in the structure of the *Tc*PYK/ponceau S complex and shown to be located at the junction between two asymmetric units in the crystal to stabilise the packing of the crystal lattice (**Figure 4.3a, b**). Two crystallographically related ponceau S molecules provide crosslinks between two tetramers, and each of the ponceau S molecules was bound to the active

site of *Tc*PYK. The active site of *Tc*PYK/ponceau S was superposed onto that of the previously determined *Lm*PYK/F26BP/OX/ATP/Mg structure (PDB code: 3HQP) to reveal the overlapping binding sites of ponceau S and adenine ring of ATP (**Figure 4.4a**). This negatively charged dye (containing sulfonic acid groups) was bound to the positively charged area within the active site (**Figure 4.4b**) and further stabilised by hydrophobic interactions provided by Tyr60 and Pro30.

4.3.2.6. *Tc*PYK/ponceau S structure is in the inactive T-state conformation

The structures of apo *Tc*PYK and *Tc*PYK/ponceau S have the same space group and almost identical unit cells, and indicate that the ponceau S molecules did not introduce a new lattice packing to the crystal, but only provide bridges to tighten the packing between lattices. This is the potential mechanism for ponceau S to decrease the crystal growing time and to increase the diffracting resolution for *Tc*PYK crystals. Furthermore, the tetramer of *Tc*PYK/ponceau S was superposed onto the tetramer of apo *Tc*PYK, resulting in an RMS fit of only 1.20 Å for all C-α atoms (excluding the B-domains) (**Table 4.4**). Additionally, the tetramers of apo *Tc*PYK and *Tc*PYK/ponceau S have identical RMS fits (2.44 Å excluding B-domains) onto the tetramer of the active R-state *Tc*PYK/F26BP/OX/ATP/Mg structure (**Table 4.4**). These results from the RMS fits indicate that the *Tc*PYK/ponceau S structure was in the inactive T-state conformation. Therefore, ponceau S is an alternative additive to crystallise inactive T-state *Tc*PYK by providing crosslinks between *Tc*PYK macromolecules in the crystal to promote the rapid formation of a stable lattice, resulting in high-quality diffracting crystals.

Table 4.3 Average RMS difference (Å) for C-α atoms of residues

Subunit	**apo-*Tc*PYK**	***Tc*PYK/ ponceau S**	***Tb*PYK/ F26BP/Mg**	***Tb*PYK/ F26BP/Mg[a]**
Monomer[*]	1.34 Å	n.a.	0.16 Å	0.40 Å
Monomer (excluding B-domains)	0.20 Å	0.27 Å	0.15 Å	0.20 Å
Tetramer[†]	n.a.[‡]	n.a.	0.55 Å	n.a.
Tetramer (excluding B-domains)	n.a.	n.a.	0.38 Å	n.a.

[*] For each structure, the RMS difference numbers between monomers were obtained by the superposition (C-α atoms fit) of one single chain (+/- B-domain) to the other chain (+/- B-domain) using CCP4 superpose (Potterton et al. 2002). The B-domain contains residues 90-187 for each subunit.

[a] This structure is determined from the crystal growing with ponceau S.

[†] For the structure of *Tb*PYK/F26BP/Mg which contains three non-crystallography symmetric tetramers, the RMS difference numbers between tetramers were obtained by the superposition (C-α atoms fit) of one tetramer (+/- B-domains) to the other tetramer (+/- B-domains) using CCP4 superpose.

[‡] n.a., not applicable

Table 4.4 Average RMS difference (Å) for C-α atoms of residues between structures from +/- ponceau S crystals [†]

	Tetramer	**Tetramer excluding B-domains**
apo-*Tc*PYK ***vs*** *Tc*PYK/ponceau S	1.52 Å	1.20 Å
apo-*Tc*PYK ***vs*** *Tc*PYK/F26BP/OX/Mg	3.01 Å	2.44 Å
*Tc*PYK/ponceau S ***vs*** *Tc*PYK/F26BP/OX/Mg	2.70 Å	2.44 Å
*Tb*PYK/F26BP/Mg ***vs*** *Tb*PYK/F26BP/Mg [‡]	0.53 Å	0.40 Å

[†] The RMS difference numbers were obtained by the superposition (C-α atoms fit) of tetramers or tetramers excluding B-domains from *Tc*PYK structures (apo-*Tc*PYK, *Tc*PYK/ponceau S and *Tc*PYK/F26BP/OX/Mg), or *Tb*PYK structures [*Tb*PYK/F26BP/Mg and *Tb*PYK/F26BP/Mg(/ponceau S)] using CCP4 superpose (Potterton et al. 2002). The B-domain contains residues 90-187 for each subunit.

[‡] This structure is determined from the crystal growing with ponceau S.

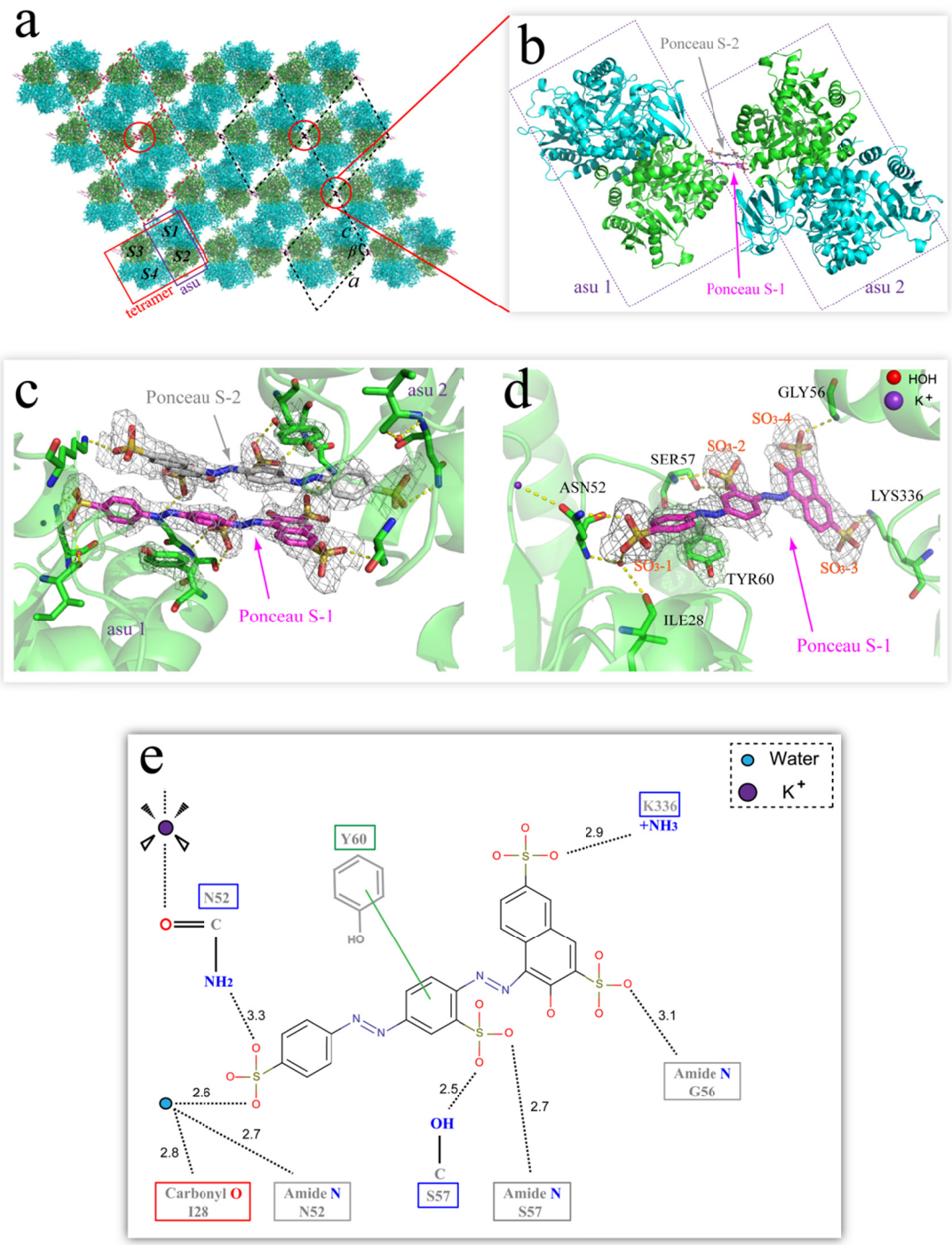


Figure 4.3 Stabilisation by ponceau S of the *Tc*PYK lattice packing as an intermolecular bridge between unit cells.

(a) The arrangement of *Tc*PYK molecules along vector (a, 0, c) in a crystal stabilised by ponceau S is shown to represent the lattice packing in two-dimensions. Two choices of unit cell along vector (a, 0, c) are indicated by black dashed lines and brown dashed lines. *Tc*PYK molecules are represented by lines, while ponceau S is shown as sticks. *Tc*PYK is a homotetramer enzyme as indicated by the red box. A biologically relevant tetramer is composed of four subunits of identical sequence

indicated as S1, S2, S3 and S4, where two of the four subunits are bound by ponceau S indicated in green; the other two subunits are shown in cyan. The asymmetric unit (asu) is contained within the purple box. The packing of unit cells stabilised by ponceau S is indicated by red circles. (b) Close-up of the intermolecular bridge between two unit cells mediated by ponceau S. The polypeptide chains of two asymmetric units (asu) are shown as cartoons while the ponceau S molecules are shown as sticks. (c, d) The interactions between the active sites of *Tc*PYK and bound ponceau S are shown to describe the stabilisation role of ponceau S. The ponceau S molecules are shown with electron density from a F_o-F_c map, which was calculated at 2.1 Å and is contoured at 2σ. These two ponceau S molecules are crystallographically symmetrical molecules and have ~50% occupancy for each molecule in the crystal. Each dye molecule has similar interactions to the active site of the enzyme. (e) A schematic drawing showing the interactions (dotted lines) and interatomic distances (given in Å) to represent **Figure 4.3d**.

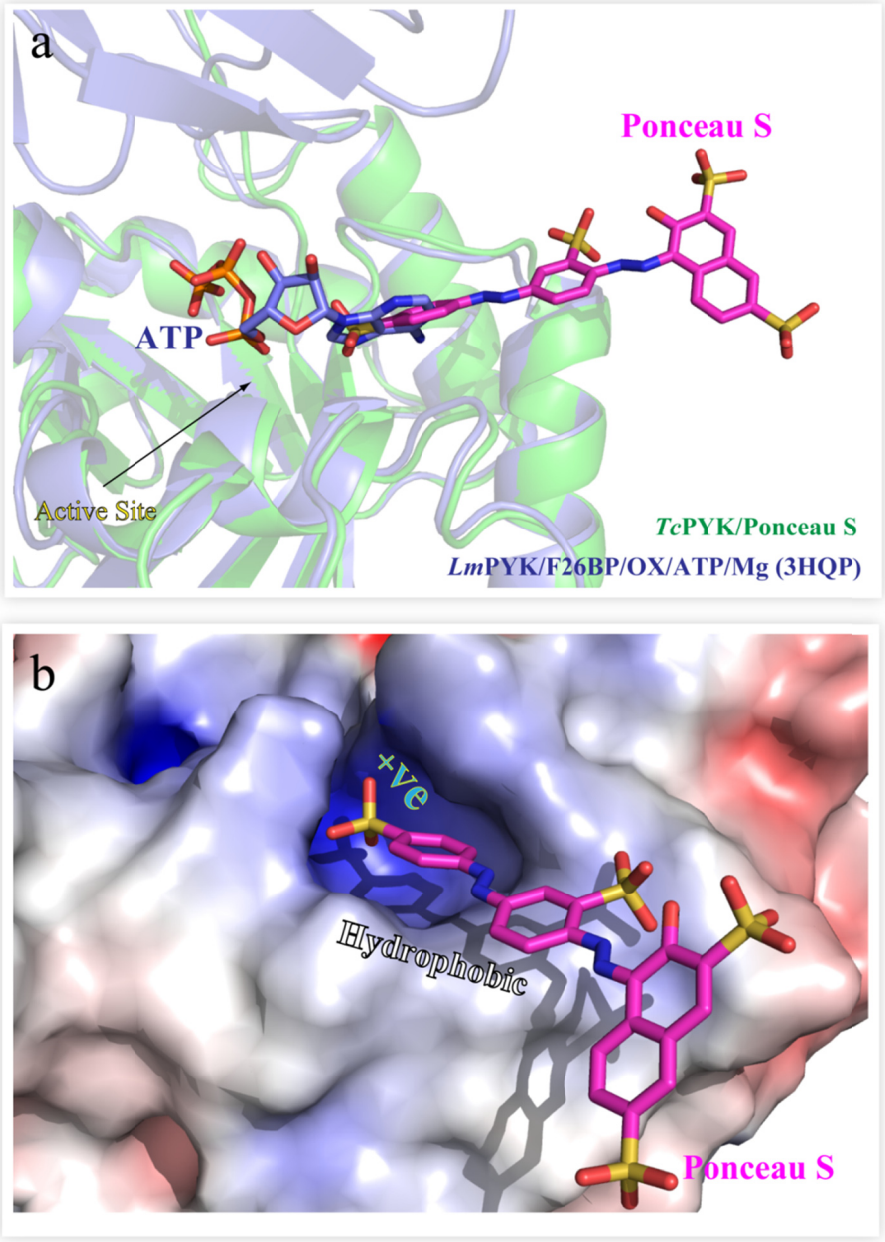


Figure 4.4 The overlap between the ponceau S and substrate binding sites.
(a) An overlap of the binding sites for ponceau S and ATP was observed when *Tc*PYK/ponceau S monomer (green) was superposed onto *Lm*PYK/F26BP/OX/ATP/Mg (3HQP) monomer (blue) based on their A- and C-domains (RMS fit of the c-α atoms is 0.54 Å). (b) The electrostatic surface of the ponceau S binding site of *Tc*PYK. The electrostatic potentials were calculated by PDB2PQR and ABPS (Dolinsky et al. 2007; Dolinsky et al. 2004; Baker et al. 2001) and visualised by PyMol (Delano 2002). The areas of positive charge (+Ve) are represented by blue, which interact with the negatively charged sulfono group of ponceau S. The hydrophobic areas further stabilise the binding of ponceau S, and are contributed by Tyr60 (shown in **Figure 4.3e**) and possibly Pro30. Areas of negative charge are shown in red.

4.3.3. Ponceau S improves the crystallisation of TbPYK/F26BP/Mg

4.3.3.1. Dye screening result for the crystallisation of *Tb*PYK

i) *Tb*PYK/F26BP/Mg crystals grown without dyes

Crystals of *Tb*PYK in complex with F26BP can be easily obtained in conditions of 15-20% PEG8000, 10-20% glycerol, and standard buffer [50 mM TEA (pH 7.2), 100 mM KCl, and 50 mM $MgCl_2$] at 4^{o}C. The crystals appeared the next day and took 5-7 days to grow to maximum dimensions. Unfortunately, the crystals observed were small, thin and fragile (**Figure 4.5a**). The coverslip holding the crystals were transferred to a new well filled with 1 ml of well solution containing 20% PEG8000, 20% glycerol and standard buffer, and equilibrated overnight before flash freezing in liquid nitrogen. Large numbers of crystals were scanned for x-ray diffraction to choose the best diffracting crystal for determination of the structure of *Tb*PYK/F26BP/Mg (**Figure 4.5c**). Crystals were not observed in other crystallisation conditions.

ii) *Tb*PYK/F26BP/Mg grown with ponceau S (*Tb*PYK/F26BP/Mg(/ponceau S))

Interestingly, ponceau S is also the only dye offering significant improvement in *Tb*PYK crystallisation in the presence of the effector F26BP. No crystals were observed in the conditions with other dyes. Red *Tb*PYK/F26BP/Mg(/ponceau S) crystals were grown in conditions of 10-20% PEG8000, 10-20% glycerol, and standard buffer [50 mM TEA (pH 7.2), 100 mM KCl, and 50 mM $MgCl_2$] at 4^{o}C. Crystals appeared the next day after setting the crystal trays and took 5-7 days to grow to maximum dimensions of 1.0x0.2x0.1 mm (**Figure 4.5b**). The conditions containing 20% glycerol usually gave better quality crystals. Crystals were screened for x-ray diffraction, and crystals grown in standard buffer condition containing 18% PEG8000 and 20% glycerol were eventually chosen for x-ray data collection to determine the structure of *Tb*PYK/F26BP/Mg(/ponceau S) (**Figure 4.5d**).

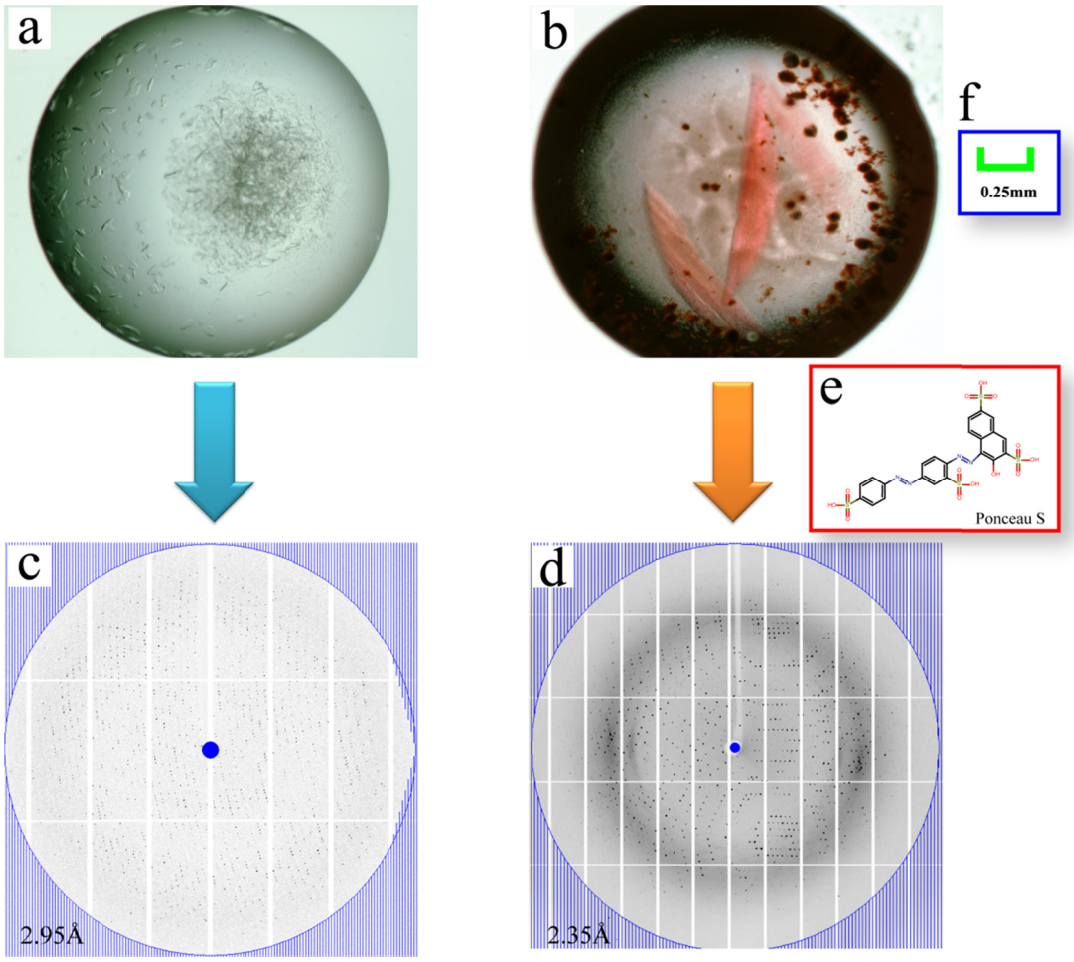


Figure 4.5 Improvement of *Tb*PYK/F26BP/Mg crystal quality in the presence of ponceau S.

(a, b) Light-microscope photographs of *Tb*PYK/F26BP/Mg crystals in the absence (a) and presence (b) of ponceau S. The photos were taken using the same scale shown in (e). Abundant micro crystals were observed in (a). Crystals appeared the next day after setting up crystal trays and grew to maximum dimensions after 5-7 days. Red crystals of large size were found in (b), which also took 5-7 days to grow to maximum dimensions. The average dimension of the crystals is 1.0x0.2x0.1 mm. The red colour of the crystals indicates the binding of ponceau S.

(c, d) The oscillation images during data collection from untagged *Tb*PYK crystals grown in the absence (c) and presence (d) of ponceau S. The images shown here belong to the diffraction patterns used for structure determinations. The edges of the oscillation images are indicated by blue circles and the maximum resolution of the images are shown in the lower left corner.

(e) Two-dimensional representation of ponceau S, which was used as an additive to improve the crystallisation of both *Tc*PYK and *Tb*PYK.

4.3.3.2. X-ray data analysis: data collection and processing

Data for the *Tb*PYK/F26BP/Mg and *Tb*PYK/F26BP/Mg(/ponceau S) structures were collected on beamline I04-2 and microfocus beamline I24, respectively. For the *Tb*PYK/F26BP/Mg crystal, the data were collected to a maximum resolution of 2.95 Å as 0.2^{o} oscillation images over 360^{o}. Data were indexed in the monoclinic space group of $P2_1$, with unit cell dimensions of a=177.86 Å, b=110.07 Å, c=185.47 Å and angles of α=90.00^{o}, β=105.57^{o}, γ=90.00^{o}. For the *Tb*PYK/F26BP/Mg(/ponceau S) crystal, the data were collected to a maximum resolution of 2.35 Å as 0.5^{o} oscillation images over 200^{o}. Data were indexed in the orthorhombic space group of *I*222, with unit cell dimensions of a=103.68 Å, b=109.00 Å, c=268.38 Å and angles of α=90.00^{o}, β=90.00^{o}, γ=90.00^{o}. Further information on data collection and processing can be seen in **Table 4.2**.

4.3.3.3. Content of the asymmetric unit and phase determination

For *Tb*PYK with a molecular weight of 54466.6 Da, the program Cell Content Analysis based on Matthew's equation (Matthews 1968) in CCP4 was used, indicating there were twelve and two monomers per asymmetric unit for *Tb*PYK/F26BP/Mg (space group $P2_1$) and *Tb*PYK/F26BP/Mg(/ponceau S) (space group *I*222) structures, respectively. Phases for both structures were determined by PHASER (McCoy et al. 2007) using the method of molecular replacement, with the input search model of the monomer from the apo *Tc*PYK structure which was previously determined in section 4.3.2. Three tetramers and two monomers in the crystallographic asymmetric unit were identified for *Tb*PYK/F26BP/Mg and *Tb*PYK/F26BP/Mg(/ponceau S), respectively, as a result of molecular replacement. For *Tb*PYK/F26BP/Mg(/ponceau S), these two crystallographically unique monomers with crystallographic two-fold symmetry generate a biologically relevant tetramer. The resulting electron density maps were clearly visible except that the electron density for one B-domain (residues 90-187) from each complex structure

was poor at this stage.

4.3.3.4. Structure solutions of *Tb*PYK/F26BP/Mg and *Tb*PYK/F26BP/Mg(/ponceau S)

For the *Tb*PYK/F26BP/Mg structure, three tetramers were built in the final model where the electron density for one B-domain was still poor. All the chains were similar with an RMS deviation of 0.16 Å for all C-α atoms of residues 2-499 and an RMS deviation of 0.15 Å for all C-α atoms of residues 2-89 and 188-499 (excluding the B-domains) (**Table 4.3**). The three tetramers in the asymmetric unit were similar as well, with an RMS deviation of 0.55 Å for all C-α atoms of residues 2-499 and an RMS deviation of 0.38 Å for all C-α atoms of residues 2-89 and 188-499 (excluding the B-domains) (**Table 4.3**). In the final model of *Tb*PYK/F26BP/Mg(/ponceau S) after structure refinement, two monomers with completed B-domains were successfully built up with an average B-factor of 51.91 $Å^2$ (protein only) or 42.34 $Å^2$ (protein only excluding B-domains). The average B-factors for these two B-domains are 56.4 $Å^2$ and 129.9 $Å^2$, respectively. Both chains in the structure were similar, with an RMS deviation of 0.40 Å for all C-α atoms of residues 2-499 and an RMS. deviation of 0.20 Å for all C-α atoms of residues 2-89 and 188-499 (excluding B-domains) (**Table 4.3**). Both structures have complete effector loops with F26BP bound to the effector sites. The additive ponceau S could not be modeled in the *Tb*PYK/F26BP/Mg(/ponceau S) structure as there was no clear electron density to identify its position.

4.3.3.5. Ponceau S improves the crystallisation of *Tb*PYK/F26BP/Mg complex by introducing a new arrangement of macromolecules in the crystal

The *Tb*PYK/F26BP/Mg structure determined from the crystal grown in the absence of ponceau S belongs to the space group of $P2_1$ and has three tetramers in the asymmetric unit. Interestingly, under similar crystallisation conditions but adding

ponceau S, the *Tb*PYK/F26BP/Mg structure belongs to the space group *I*222 and has two monomers in the asymmetric unit. Additionally, the crystals grown in the presence of ponceau S have larger size and much better diffracting quality (**Figure 4.2**). The tetramer of *Tb*PYK/F26BP/Mg(/ponceau S) was superposed onto the tetramer of *Tb*PYK/F26BP/Mg, resulting in an RMS fit of 0.53 Å for all C-α atoms or 0.40 Å for all C-α atoms but excluding B-domains (**Table 4.4**), indicating these two structures are nearly identical in conformation, and that the ligand ponceau S did not have a significant effect on the structure conformation. Unlike the role of ponceau S in the *Tc*PYK/ponceau S structure, no electron density was clear enough to identify the position of the ponceau S molecule in the *Tb*PYK/F26BP/Mg(/ponceau S) complex. Therefore, a discussion on the role of ponceau S in the *Tb*PYK/F26BP/Mg(/ponceau S) structure could not be determined. It is possible that the negatively charged groups (sulfono groups) and/or the hydrophobic groups of ponceau S could be involved in lattice formation, but with multiple poses resulting in unclear electron density.

In short, ponceau S is a proven additive to encourage new and stable lattice formation for the crystallisation of *Tb*PYK/F26BP/Mg without influencing its conformation. In **Chapter 6, 7 and 8**, *Tb*PYK/F26BP/Mg crystals grown in the presence of ponceau S were successfully used in a series of crystal soaking experiments.

CHAPTER 5: X-ray Crystal Structures of *Tc*PYK and Its Allosteric Mechanism

5.1. Introduction

From the previous sections, pure and active *Tc*PYK has been expressed in *E. coli* and purified in high quality (**Chapter 3**), and the crystal structures of *Tc*PYK have been determined at high resolution (**Chapter 4**) (**Table 5.1**). This chapter will describe the structural and biophysical studies on *Tc*PYK to explore the structure for the first time and understand the allosteric mechanism to regulate enzyme activity of *Tc*PYK. Work from this chapter will form the basis of a publication:

Morgan, H.P. [*], **Zhong, W.** [*], *et al*. Evolutionary divergence of pyruvate kinase allostery. (Manuscript in preparation) [*] Co-first authors

5.2. Materials and Methods

The materials and the general methods used in this chapter have been described in **Chapter 2**. In brief, the structure of apoenzyme *Tc*PYK was solved by the methods described in **Section 4**; *Tc*PYK/F26BP/OX/Mg for comparison in this chapter was determined by colleague Dr. Hugh Morgan; the structure analysis including rigid-body rotation, B-domain movement and interface interactions followed the methods in **Section 5**; the thermal stability of *Tc*PYK in the presence of various ligands was investigated by thermal shift assay described in **Section 3**.

Table 5.1 X-ray data collection parameters and refinement statistics

	apo *Tc*PYK	*Tc*PYK/F26BP/OX/Mg [†]
Data collection		
Space group	*C*2	*I*422
Cell dimensions		
a, *b*, *c* (Å)	113.78,121.42,97.16	173.77,173.77,211.86
α, *β*, *γ* (°)	90.00,115.73,90.00	90.00,90.00,90.00
Solvent content (%)	53	65
Wavelength (Å)	0.98	0.98
Resolution (Å)	43.76-2.50	55.90-2.80
No. reflections	252349(31324)	553645(31324)
No. unique reflections	38483(5416)	40095(5770)
Wilson B-factor	37.4	69.7
R_{merge} (%)	9.0(25.7)	12.1(51.5)
R_{meas} (%)	10.5(30.6)	12.6(54.3)
$R_{p.i.m.}$(%)	5.5(16.4)	3.3(17.0)
$\langle I/\sigma I \rangle$	13.7(5.8)	20.1(3.7)
% completeness	93.7(90.5)	100.0(100.0)
Multiplicity	6.6(5.8)	13.8(10.1)
Refinement statistics		
Monomers in ASU	2	2
No. reflections	36536	38066
R_{work} / R_{free}	16.15/22.12	16.73/21.65
No. of atoms		
Protein	7566	7666
Ligands	6	52
Metal ions	2	4
Water	533	243
Average B-factors (Å^2)		
Protein	31.03	57.01
Protein (excluding B-domains)	27.00	45.60
Ligand	38.47	42.00
Metal ion	40.10	48.33
Water	29.07	39.72
R.m.s. deviations		
Bond lengths (Å)	0.01	0.01
Bond angles (degree)	0.99	1.84
Ramachandran		
Favoured (%)	96.2	94.8
Allowed (%)	99.9	99.7
No. of outliers	1	3

Values in parentheses are for the highest resolution shell.

[†] Values determined by Dr. Hugh Morgan; manuscript in preparation

[‡]The outlier residue in each monomer is Thr296, a key active-site residue that is commonly found in this configuration in PYK structures. The outline residue Ile134 is modelled in the poor electron density.

5.3. Results and Discussion

5.3.1. Overall structure of TcPYK

*Tc*PYK is a tetrameric enzyme composed of four identical protein chains (**Figure 5.1a, 5.1b**). The amino acid sequence of a single chain is shown in **Figure 5.2** (the 20 amino acids of the N-terminal His_6-tag sequence are not presented in this figure. Each chain contains four domains, which are defined as N-terminal domain (N), A-domain (A), B-domain (B) and C-domain (C) (**Figure 5.1c, 5.1d**).

The secondary structures of the enzyme are represented in **Figure 5.2** (only α-helices and β-strands are shown). The N-terminal domain, corresponding to residues 2-18, contains a single well-defined α-helix. The A-domain is the largest domain of the four and has a parallel $(\alpha/\beta)_8$ barrel with additional helices, corresponding to two separate stretches of polypeptide chain residues 19-89 and 188-358. The A-domains make contacts between adjacent subunits within a tetramer to form a large interface in *Tc*PYK which is also called the A-A interface. The B-domain contains nine β-strands with only one short α-helix and is formed from residues 90-187. The active site of *Tc*PYK is located in the cleft between A- and B-domains. The C-domain corresponds to the remaining residues 359-499 and contains a core of β-strands flanked by α-helices, harbouring an activator F26BP binding site (effector site). The C-domains make contacts between adjacent subunits within a tetramer to form a small interface which is also named the C-C interface.

For apo *Tc*PYK, the enzyme stays in its inactive state; for *Tc*PYK/F26BP/OX/Mg, the enzyme is activated by F26BP and stays in its active state. The conformational changes between these two states and the movement of the B-domain will be discussed in sections **5.3.4** & **5.3.5**. However, the overall structural architecture of the individual chains of the T- and R-sate tetramers are the same. However, as discussed later, the shape of the tetramer are different.

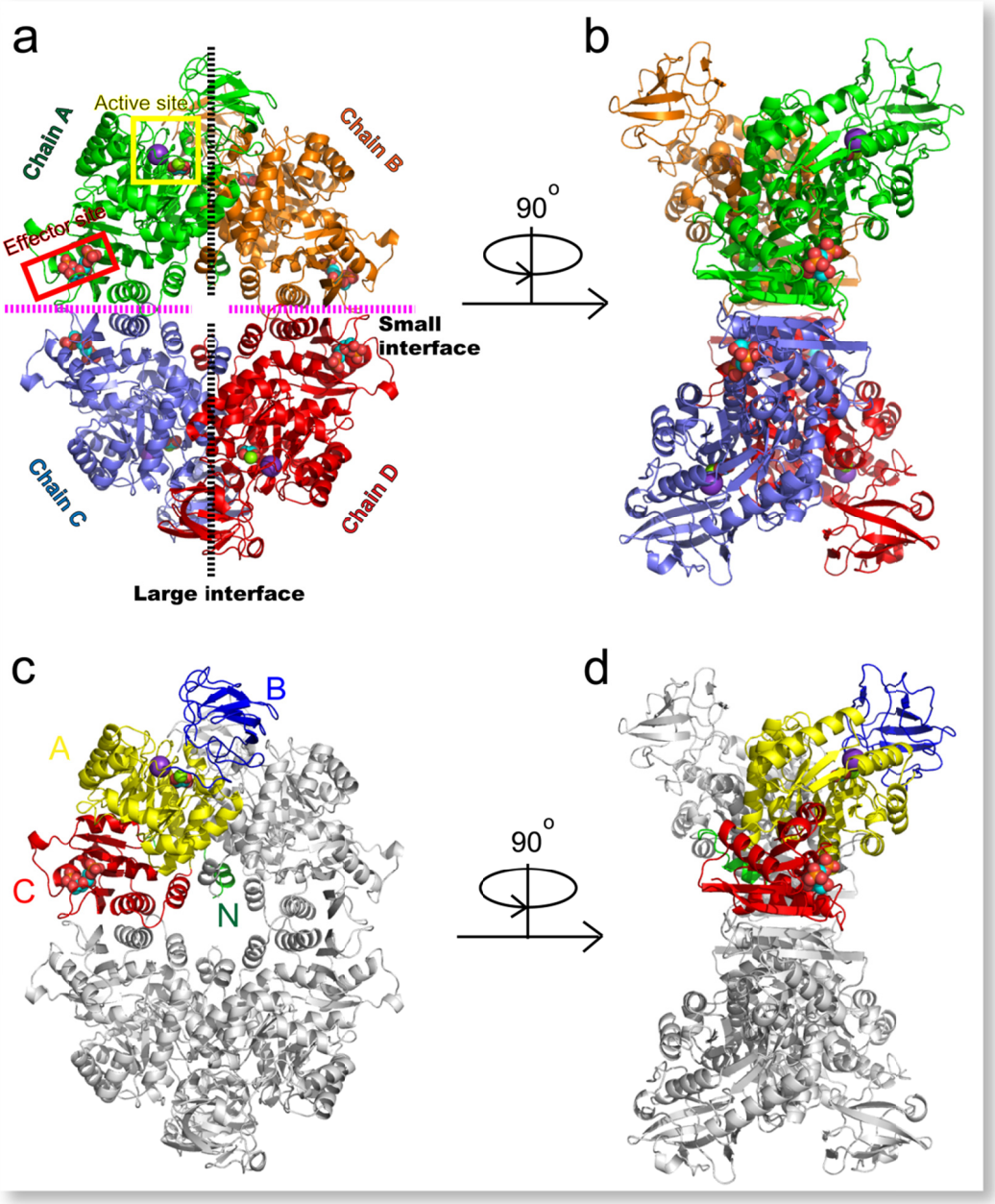


Figure 5.1 Structure of *Tc*PYK/F26BP/OX/Mg showing the tetramer architecture and domain boundaries. Metals and ligands are represented by spheres.

(a, b) Two orthogonal views of the *Tc*PYK/F26BP/OX/Mg showing the tetramer architecture, active site and effector site. Each monomer has been coloured to aid the identification of subunit interfaces. The large (A-A) and small (C-C) interfaces between subunits are shown as dashed lines. (c, d) Two orthogonal views of the *Tc*PYK/F26BP/OX/Mg tetramer where one subunit is coloured to show domains: N-terminal (green = residues 2-18), A-domain (yellow = residues 19-89, 188-358), B-domain (blue = residues 90-187), C-domain (red = residues 359-499). *Tc*PYK/F26BP/OX/Mg has the same tetramer architecture and domain boundaries as apo *Tc*PYK.

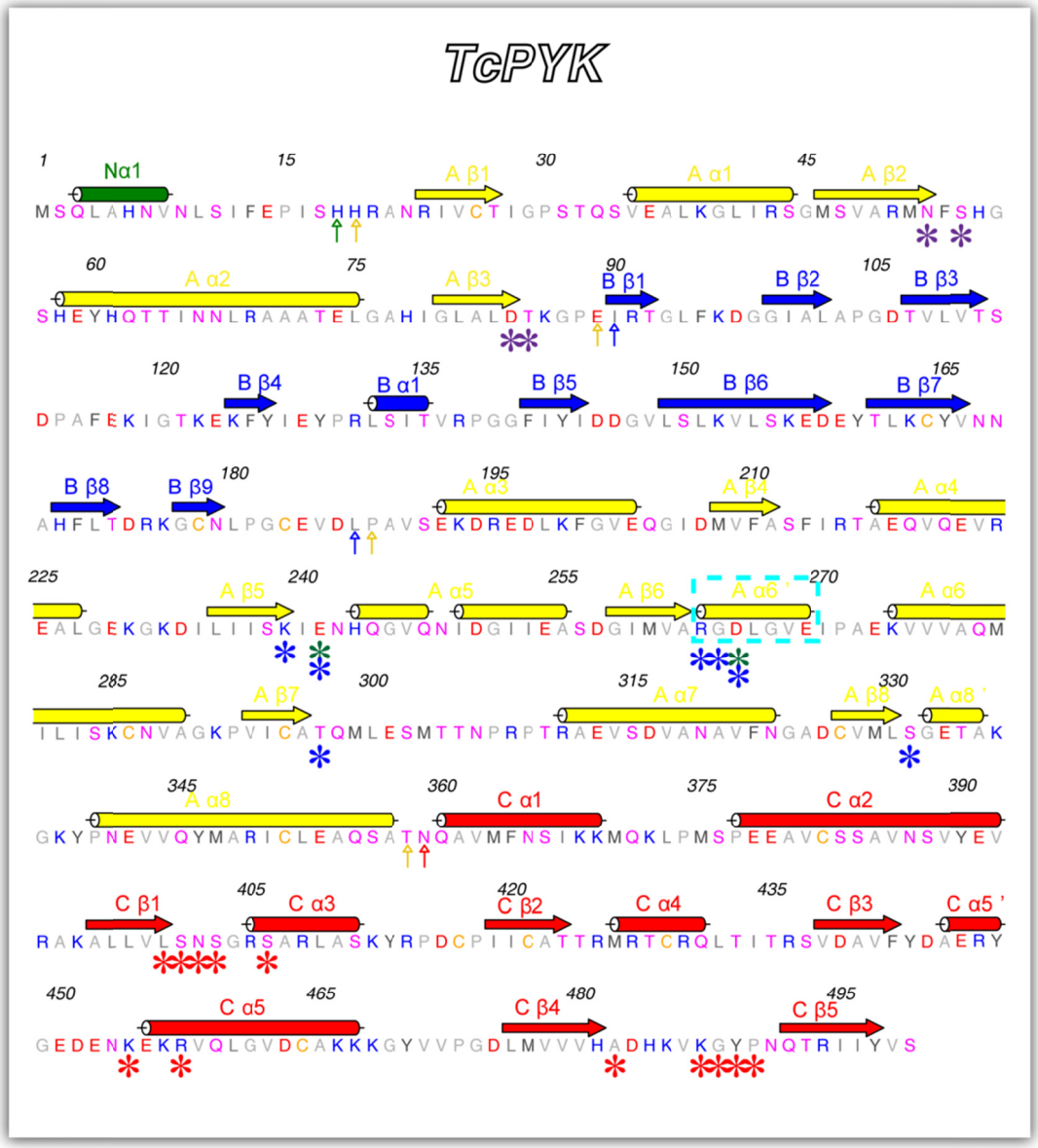


Figure 5.2 The amino acid sequence of *Tc*PYK.
Secondary structural elements defined in *Tc*PYK/F26BP/OX/Mg by DSSP (Kabsch and Sander 1983; Joosten et al. 2011) are indicated (only α-helix and β-strand are shown). Secondary structural elements are labeled by different colours corresponding to their domain regions: N-terminal domain (green), A-domain (yellow), B-domain (blue) and C-domain (red). Domain boundaries are indicated by vertical arrows in domain-specific colours. Residue numbers corresponding to the *Tc*PYK sequence are shown above the secondary structural elements. The active site amino acids involved in divalent metal binding (oxalate-coordinating metal) (*), potassium metal ion binding (*), substrate analogue oxalate binding (*) and effector F26BP binding (*) are indicated by asterisks. The small α-helix Aα6' (residues Arg263-Glu269) which is involved in allosteric activation, divalent metal and oxalate binding is indicated by a dashed box (cyan). The figure was generated using Aline (Bond and Schüttelkopf 2009) and all residues are coloured by residue types.

5.3.2. Effector site of TcPYK

5.3.2.1. The effector loop is stabilised by F26BP binding

The allosteric effector site is located in a pocket within the C-domain which is ~40 Å distance away from the active site (**Figure 5.1a**). In the absence of activator F26BP binding, the effector loop (residues 482-488) cannot be clearly identified in the apo *Tc*PYK structure, where the electron densities for residues 485-489 of chain A and residues 483-487 of chain B are unclear. In the F26BP-bound structure of *Tc*PYK/F26BP/OX/Mg, the effector loops in both crystallographically unique chains are clearly identified in the electron density map with an average B-factor of 51.17 $Å^2$ compared to the overall protein B-factor of 57.01 $Å^2$. The superposition of effector sites from structures of apo *Tc*PYK and *Tc*PYK/F26BP/OX/Mg is shown in **Figure 5.3**, which indicates nearly identical conformations of the effector site (with an average RMS fit of 0.495 Å for C-α atoms) but adds the missing effector loop in the apo structure.

5.3.2.2. The interactions between F26BP and the effector site

The effector F26BP molecule with two negatively charged phospho groups is buried in a positively charged effector binding pocket (**Figure 5.4**), and forms multiple interactions with enzyme residues and water molecules (**Figure 5.5**). The 2'-phospho group forms strong electrostatic interactions with the positively charged side chains of residues Lys454 and Arg457. Additionally, more hydrogen bond interactions contribute to the effector binding. The side chains of Ser401, Asn402, Ser403 and Ser406, and the main chain atoms of residues Leu400, Asn402, Ser403, Ser406, Ala482, Gly488 and Tyr489 form hydrogen bonds with F26BP (**Figure 5.5**). Some interactions are mediated by water molecule networks as well, such as Lys454, Lys487 and Gly488, although the water molecules are not always there. One potential interaction is between the furanose ring of F26BP and the ring of Pro490.

Residues Lys487, Gly488, Tyr489 and Pro490 are responsible for stabilising the effector loop (residues 482-488). The interactions along the small (C-C) interface which further stabilise the effector loops will be discussed in section **5.3.4.2**.

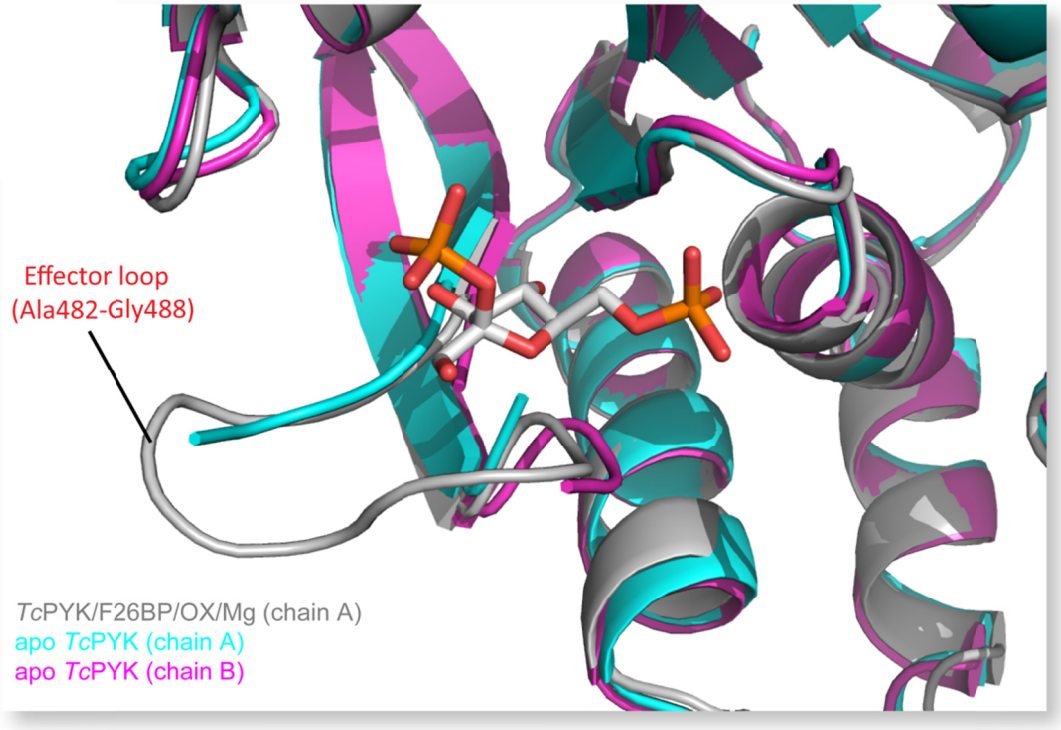


Figure 5.3 Superposition of effector sites from the structures of apo *Tc*PYK and *Tc*PYK/F26BP/OX/Mg.

The two asymmetric unique chains (chain A and chain B) of apo *Tc*PYK were superposed onto *Tc*PYK/F26BP/OX/Mg (chain A), resulting in an RMS fit of 0.49 Å and 0.50 Å for all C-α atoms of residues 359-499 (C-domain), respectively. Chain A and Chain B of apo *Tc*PYK are represented as cartoons in cyan and purple, respectively. The *Tc*PYK/F26BP/OX/Mg structure is shown as a cartoon in grey, except the effector site ligand F26BP which is represented as sticks. The other asymmetric unique chain of *Tc*PYK/F26BP/OX/Mg is not shown in this comparison as it has a nearly identical effector site to that of chain A of *Tc*PYK/F26BP/OX/Mg (RMS. fit of 0.20 Å for all C-α atoms of C-domain). The average B-factor of effector loops is 51.17 $Å^2$ compared to the overall protein B-factor of 57.01 $Å^2$ in *Tc*PYK/F26BP/OX/Mg.

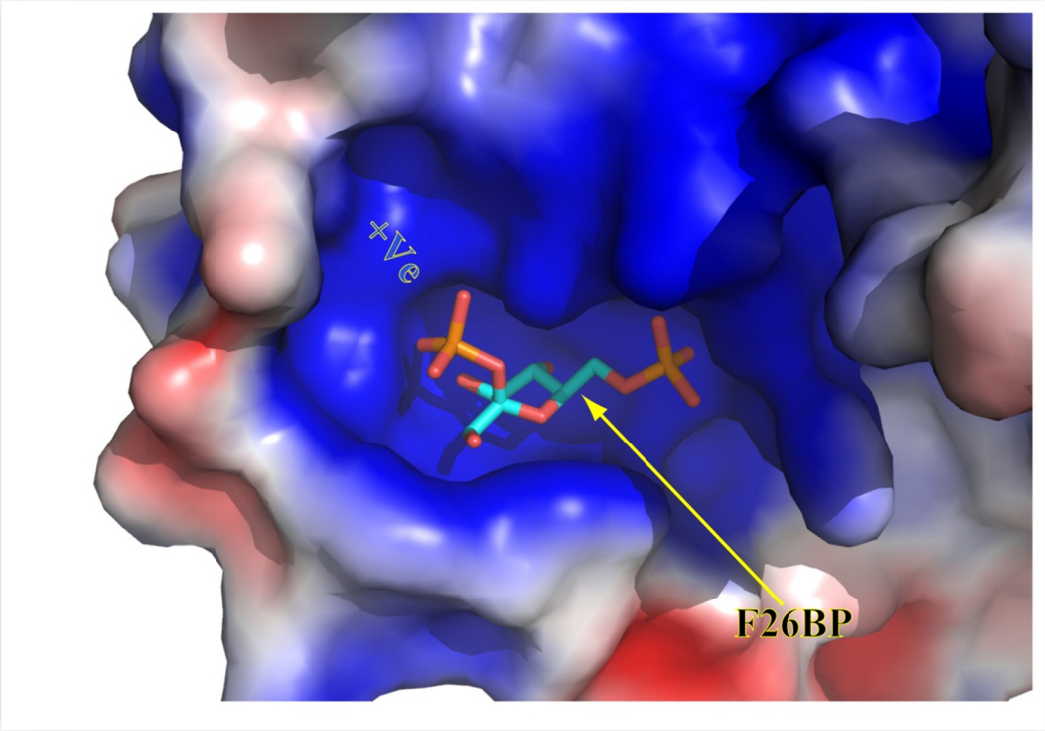


Figure 5.4 The electrostatic surface of *Tc*PYK showing the effctor F26BP binding site.

The electrostatic potentials were calculated by PDB2PQR and ABPS (Baker et al. 2001; Dolinsky et al. 2004; Dolinsky et al. 2007) and visualized by PyMol (Delano 2002). The areas of positive charge (+Ve) are represented by blue, which are essential for binding the negatively charged phospho groups of F26BP. The areas of negative charge and the hydrophobic areas are shown in red and white, respectively.

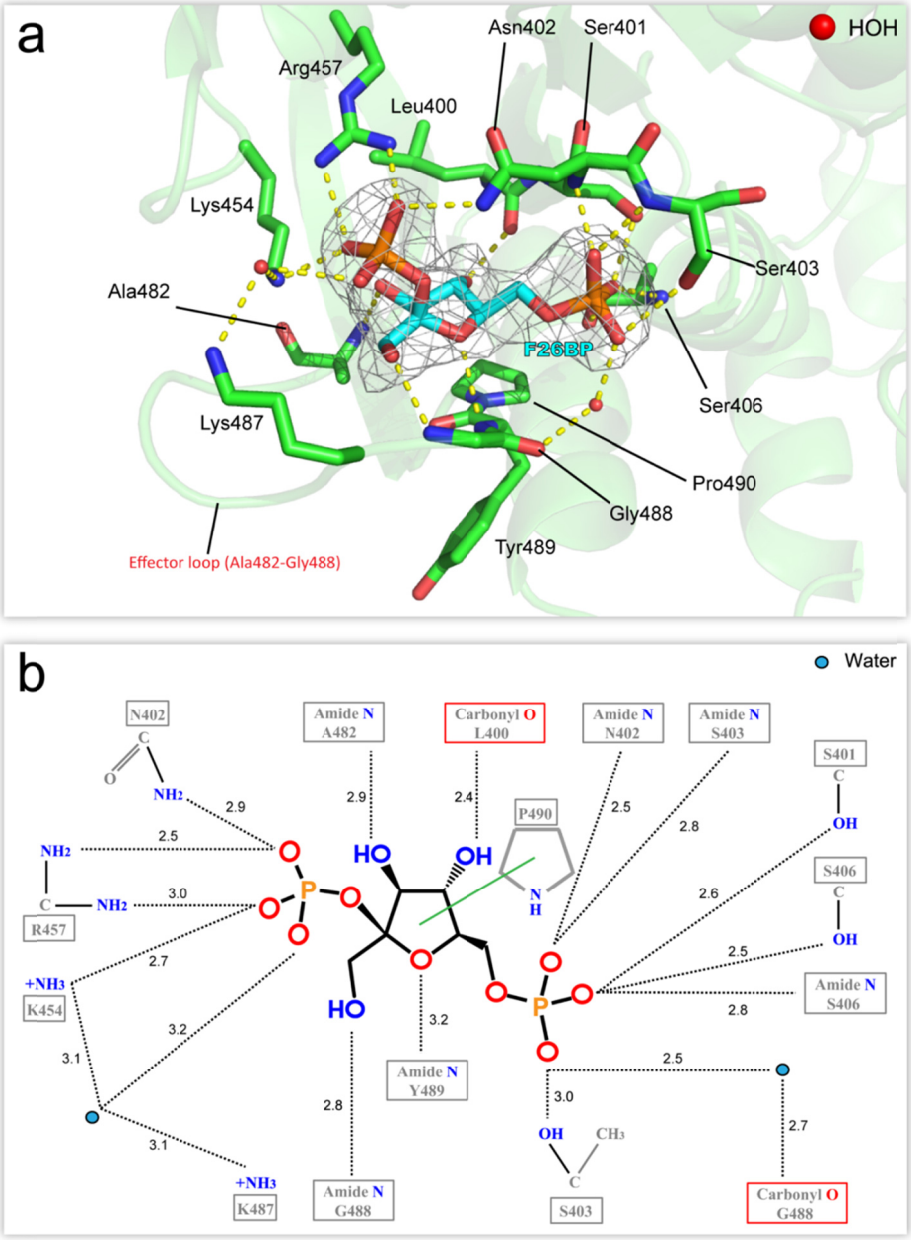


Figure 5.5 The effector site of *Tc*PYK with bound F26BP.
(a) Close-up of the effector site of *Tc*PYK (*Tc*PYK/F26BP/OX/Mg, chain A) showing the activator F26BP binding mode. Protein is shown as a cartoon, except the residues involved in F26BP binding which are shown as sticks. The F26BP molecule is shown as sticks with an unbiased *Fo-Fc* election density map contoured at 4.0 σ (grey). Water molecules are shown as red spheres. Possible interactions involved in F26BP binding are indicated by dashed lines in yellow. The ordered effector loop (Ala482-Gly488) is indicated as well. (b) A schematic drawing showing the estimated interactions at the *Tc*PYK effector site. The interatomic distances for the interactions are given in Ångstroms.

5.3.3. Active site of TcPYK

5.3.3.1. Monovalent K^+ is bound at the active site of apo *Tc*PYK

A close-up view of the active site of apo *Tc*PYK is shown in **Figure 5.6**. The RMS fit for the A-domains from the two asymmetrically unique chains is 0.19 Å for all C-α atoms or 0.29 Å for all atoms, which indicates nearly identical active sites. Both the divalent ion Mg^{2+} and the monovalent ion K^+ were present in the conditions to grow the crystal, however, only the electron density for K^+ was observed with an average B-factor of 40.10 $Å^2$ in the apo structure, compared to the overall B-factor of 31.03 $Å^2$ for the protein. The presence of K^+ was further confirmed by its coordinating environment in each chain. Four residues directly coordinate with K^+, namely Asn52, Ser54, Asp84 and Thr85 in both chains. The ligand glycerol, which favours the growth of crystals at high concentrations (~1.35 - 2.70 M) is modeled only in one chain (chain A, **Figure 5.6a**), showing interactions with protein residues and the K^+ atom, but was not observed in the other chain (**Figure 5.6b**).

In chain A, the coordination sphere of K^+ has distorted octahedral geometry and is formed by three side-chain oxygen atoms from residues Asn52, Ser54 and Asp84, one main-chain carbonyl oxygen atom from residue Thr85, one oxygen atom from glycerol and one water molecule. The radius of K^+ coordination ranges from 2.6-3.2 Å.

In chain B, only four atoms (from protein residues Asn52, Ser54, Asp84 and Thr85) can be identified to form the coordination sphere of K^+. The coordinating distances range from 2.7-3.1 Å. The K^+ ions in both chains have strong electron densities with *Fo-Fc* signals over 3.0 σ at the resolution of 2.50 Å. The detailed coordination sphere of K^+ in the active site can be seen in **Figure 5.7**. Although the octahedral geometry for K^+ in chain B is not completed, the presence of K^+ in many other PYK structures including apo *Lm*PYK (3HQN) and the reasonable B-factor of K^+ (38.04 $Å^2$) indicate the presence of K^+ in this chain. The 'missing' coordination is possibly due to the quality of the data.

Residues Phe213, Glu241, Asp265, Thr297 and Arg311 are indicated as sticks in **Figure 5.6**. These five residues will have strikingly different side-chain orientations in *Tc*PYK/F26BP/OX/Mg compared to the side-chain orientations in the apo structure. The role of the side-chain re-orientations for enzyme regulation will be discussed in section **5.3.3.2**.

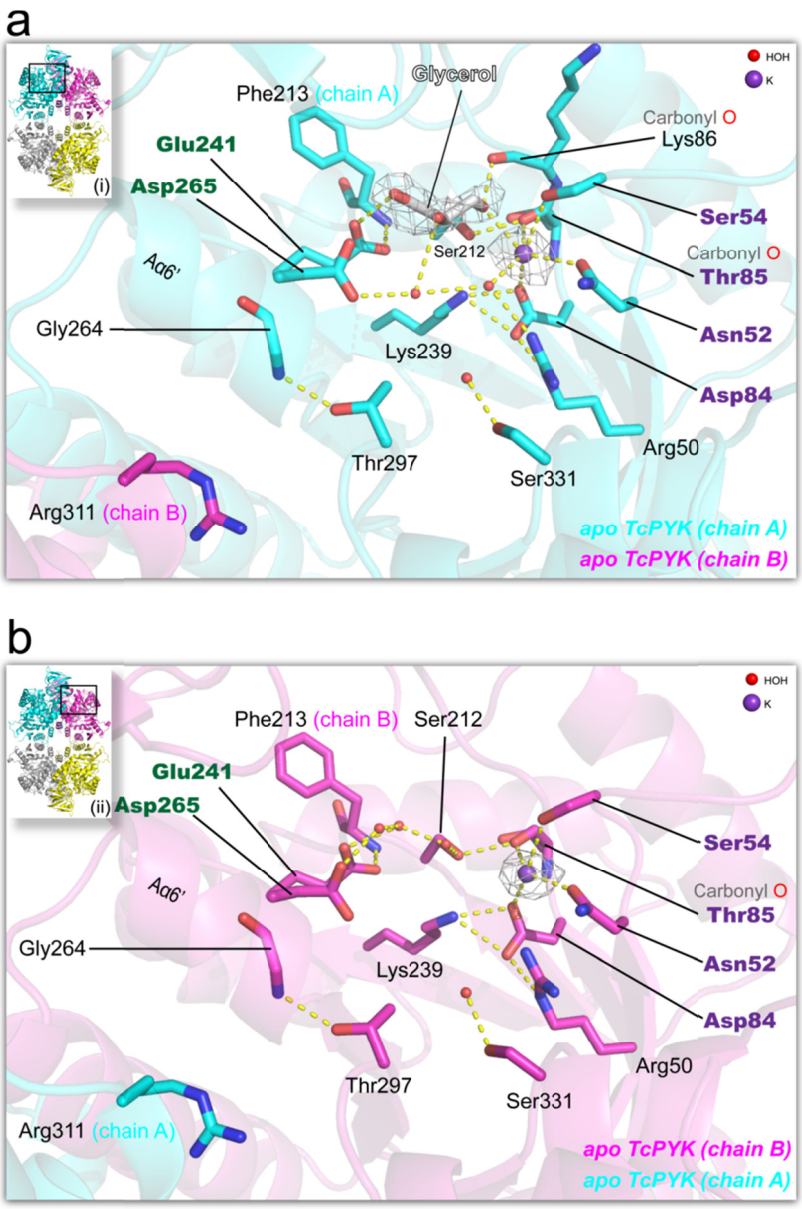


Figure 5.6 Close-up of the active site of apo *Tc*PYK.
The polypeptide chain is shown as a cartoon in cyan (chain A) or purple (chain B). Chain A and chain B are asymmetrically unique in the crystal structure. Interacting residues and the ligand glycerol are shown as sticks. Residue Phe213 which potentially favours the binding of substrate or substrate analogues in the correct

position is shown in stick format as well. The short α-helix Aα6' (residues Arg263-Glu269) within the A-domain is indicated. The monovalent ion K^+ is shown as a purple sphere. The *Fo-Fc* electron density for the K^+ ion is shown as a grey mesh contoured at 3.0 σ. The water molecules are shown as red spheres. Potential interatomic interactions are indicated by yellow dashed lines. (a) The active site of chain A is shown. The position of the active site (chain A) is indicated in the inset (i). The *Fo-Fc* electron density for the ligand glycerol is shown as a grey mesh contoured at 3.0 σ. (b) The active site of chain B is shown. The position of the active site (chain B) is indicated in the inset (ii).

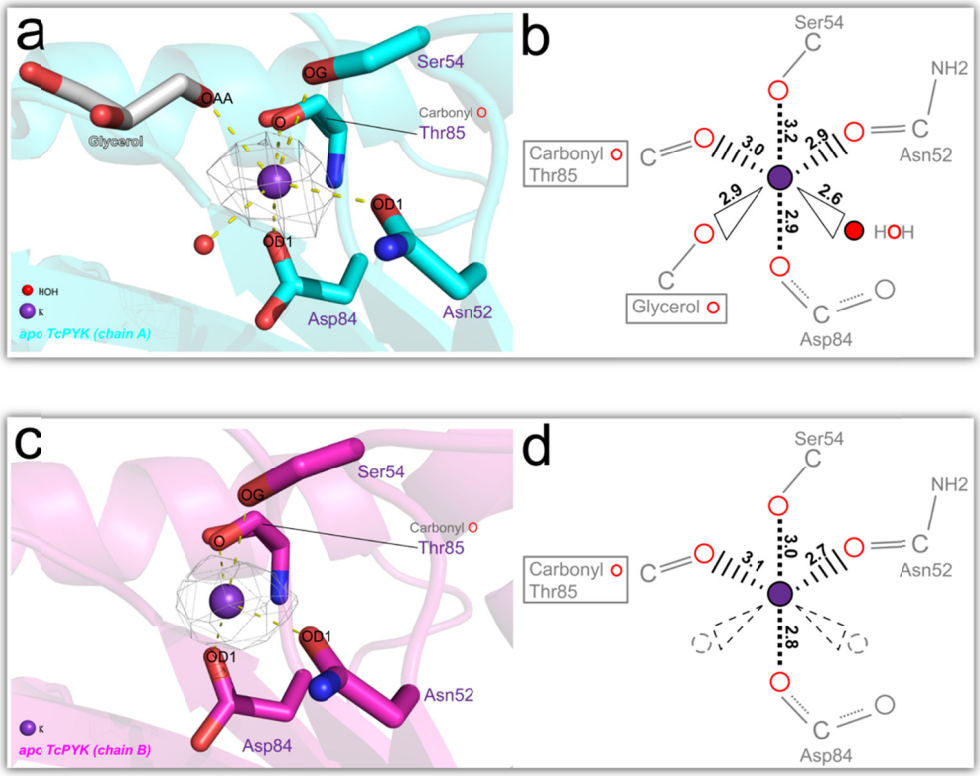


Figure 5.7 Close-up of K^+ coordination at the active site of apo *Tc*PYK.
(a, c) The polypeptide chain is shown as a cartoon in cyan (chain A in **Figure 5.6a**) or purple (chain B in **Figure 5.6b**). Chain A and chain B are asymmetrically unique in the crystal structure. Interacting residues and the ligand glycerol are shown as sticks. The monovalent ion K^+ is shown as a purple sphere. The *Fo-Fc* electron density for the K^+ ion is shown as a grey mesh contoured at 3.0 σ. The water molecules are shown as red spheres. Interatomic interactions are indicated by yellow dashed lines. (b, d) Schematic representations of the coordination sphere of the active site K^+. The octahedral geometry for K^+ in chain B is not completed and the missing coordination is represented by dashed lines. The interatomic distances for the interactions are given in Ångstroms.

5.3.3.2. Substrate analogue oxalate binds to the active site in *Tc*PYK/F26BP/OX/Mg

A close-up view of the active site of *Tc*PYK/F26BP/OX/Mg is shown in **Figure 5.8**. The RMS fit for the A-domains from the two asymmetrically unique chains is 0.20 Å for all C-α atoms or 0.32 Å for all atoms, which indicates nearly identical active sites for both chains.

(i) Both Mg^{2+} and K^{+} metal ions bind to the active site

Unlike the apo *Tc*PYK structure, both Mg^{2+} and K^{+} metal ions have been clearly identified in *Tc*PYK/F26BP/OX/Mg, with an average B-factor of 39.63 $Å^2$ and 55.13 $Å^2$, respectively. Both metal ions have strong electron densities with *Fo-Fc* signals over 3.0 σ at the resolution of 2.80 Å (**Figure 5.9** and **5.10**).

The divalent ion Mg^{2+} is located in a highly negatively charged sphere formed by the negatively charged carboxyl groups of oxalate and the negatively charged side-chain carboxyl groups of residues Asp265 and Glu241. The distorted octahedral coordination sphere of Mg^{2+} is formed by the two side-chain carboxylate oxygen atoms of residues Asp265 and Glu241, two of the carboxylate oxygen atoms of the oxalate molecule and two water molecules. The radius of Mg^{2+} coordination ranges from 1.9-2.3 Å in the two asymmetrically unique chains in this structure. The detailed metal coordination can be found in **Figure 5.9 (chain A)** and **5.10 (chain B)**.

Similar to apo *Tc*PYK, the expected octahedral geometry for K^{+} in both chains are not completed (**Figure 5.9** and **5.10**) but K^{+} with an average B-factor of 56.40 $Å^2$ has still been modeled. Four protein residues (Asn52, Ser54, Asp84 and Thr85) consistently coordinate with K^{+} with a coordinating radius of 2.7-2.9 Å in chain A, and 2.7-3.1 Å in chain B. Only one water molecule is identified in the structure which coordinates with K^{+} in chain B. No additional water molecules or other potential molecules are apparent in either chain in order to complete the octahedral coordination sphere of K^{+}. The detailed metal coordination can be seen in **Figure 5.9**

and **5.10.**

(ii) Oxalate coordinates with Mg^{2+} and amino acid residues

Oxalate which is an analogue of the substrate PEP is anticipated to bind to the PEP binding site with a similar binding mode as the true substrate. This has indeed been confirmed by the PEP-bound complex structure of *Tb*PYK (*Tb*PYK/F26BP/PEP) and will be described in **Chapter 6**. In the oxalate-bound structure of *Tc*PYK (*Tc*PYK/F26BP/OX/Mg), the electron density for oxalate is well defined in the active sites of both asymmetrically unique chains, with the unbiased *Fo-Fc* signal over 5.0 σ at 2.80 Å resolution (**Figure 5.8**). The average B-factor of oxalate is 55.28 $Å^2$ compared to the overall protein B-factor of 57.01 $Å^2$. The two carboxyl groups of oxalate coordinate with the divalent ion Mg^{2+} and thereby contribute to its metal coordination sphere. Apart from the divalent metal coordination, oxalate forms direct interactions with the enzyme as well. The carboxyl group of oxalate forms interactions with main chain atoms of residues Gly264 and Asp265, located in the small α-helix Aα6' which is stabilised by residue Arg311 from the adjacent subunit, and also forms a hydrogen bond with the side chain of residue Thr297. In chain A, the interaction between oxalate and residue Ser331 is mediated by a water molecule, which is not observed in chain B. The positively charged residue Lys239 which participates in phospho transfer during the PYK enzyme reaction forms an interaction with the carboxyl group of oxalate as well. The detailed binding mode of oxalate in the active site can be seen in **Figure 5.9** and **5.10**.

(iii) Side-chain re-orientations in *Tc*PYK/F26BP/OX/Mg

Active sites from apo *Tc*PYK and *Tc*PYK/F26BP/OX/Mg are superposed to indicate the differences in amino acid side-chain orientations between these two structures (**Figure 5.11a**). A total of five residues are found to have significant differences in side-chain orientation: (1) Phe213 is in an 'open' conformation in apo *Tc*PYK, but in

a ‘closed’ conformation in *Tc*PYK/F26BP/OX/Mg; (2, 3) the orientations of Glu241 and Asp265 correlate with their binding with the divalent ion Mg^{2+}; (4) Thr297 forms a hydrogen bond with the small α-helix Aα6’ in the absence of oxalate binding, but interacts with the carboxyl group of oxalate after side-chain re-orientations; (5) Arg311 from the adjacent chain forms two hydrogen bonds with the small α-helix Aα6’ in *Tc*PYK/F26BP/OX/Mg, but is buried inside the large (A-A) interface in the apo structure. The side-chain re-orientation of Arg311 directly correlates with enzyme allosteric regulation, and will be discussed in the next section.

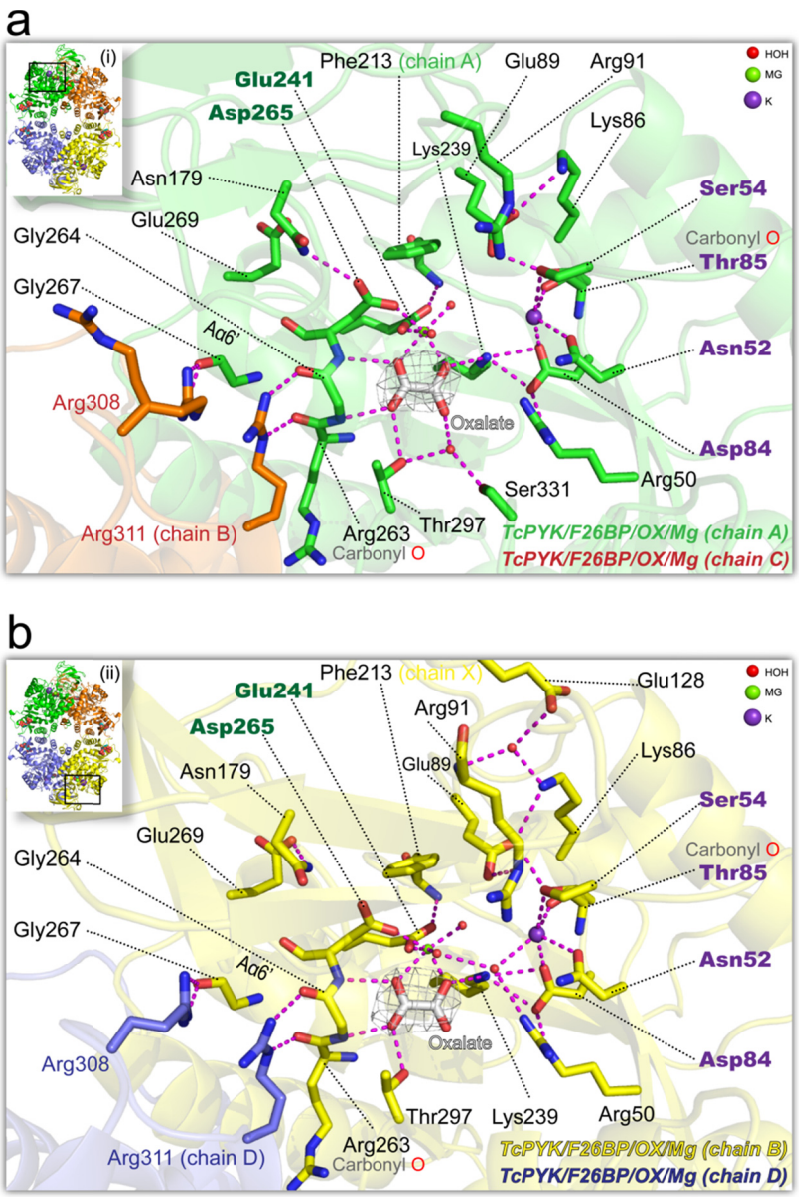


Figure 5.8 Close-up of the active site of *Tc*PYK/F26BP/OX/Mg.
The polypeptide chain is shown as a cartoon in green (chain A), brown (chain C), yellow (chain B) or blue (chain D). Chain A and chain B are asymmetrically unique in the crystal structure. Interacting residues and the ligand oxalate are shown as sticks. Residue Phe213 which potentially favours the binding of substrate or substrate analogue in the correct position is shown in stick format as well. The short α-helix Aα6' (residues Arg263-Glu269) within the A-domain is indicated. The divalent ion Mg^{2+} and the monovalent ion K^{+} are shown as a green sphere and a purple sphere, respectively. Potential interatomic interactions are indicated by purple dashed lines. (a) The active site of chain A is shown. The position of the active site (in chain A) is indicated in the inset (i). The *Fo-Fc* electron density for oxalate is shown as a grey mesh contoured at 5.0 σ. (b) The active site of chain B is shown. The position of the active site (in chain B) is indicated in the inset (ii). The *Fo-Fc* electron density for oxalate is shown as a grey mesh contoured at 5.0 σ.

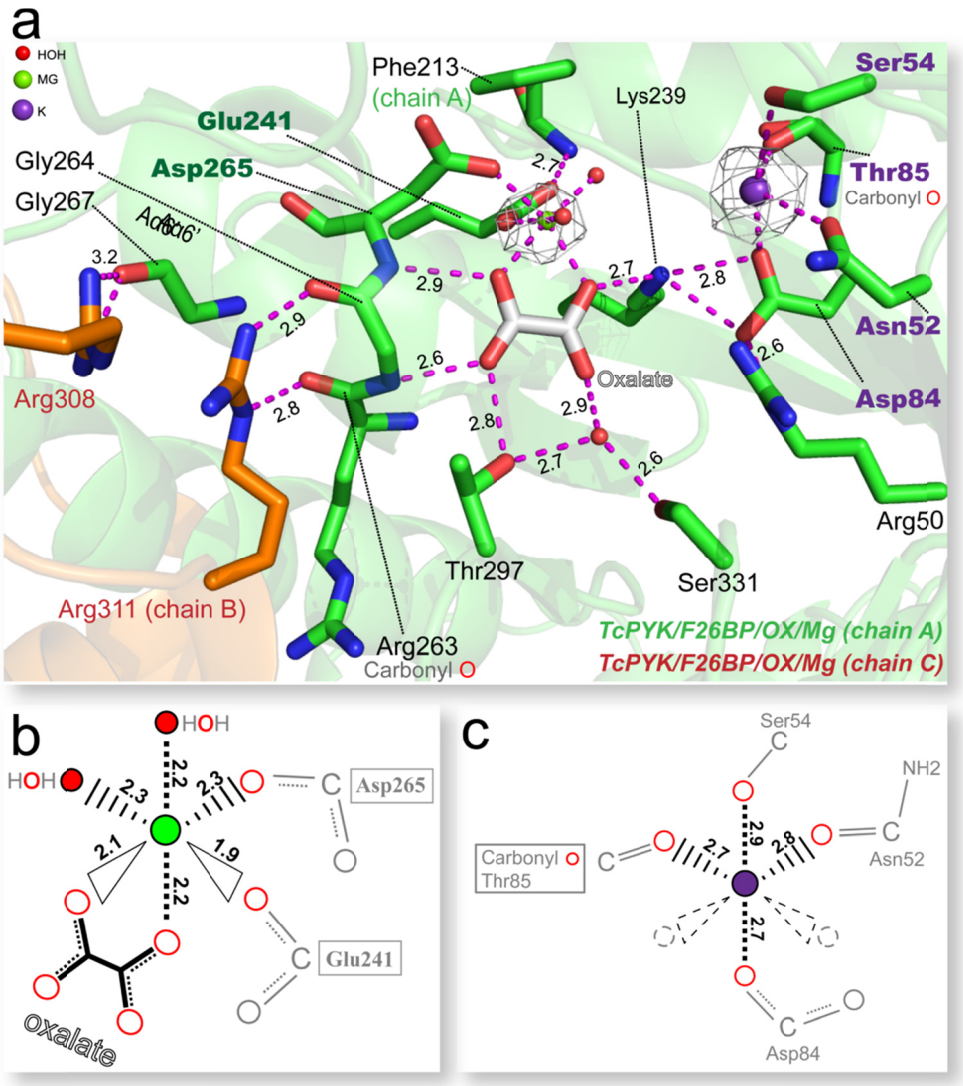


Figure 5.9 Metal ion coordination at the active site of *Tc*PYK/F26BP/OX/Mg (chain A).

(a) Close-up of the active site of *Tc*PYK/F26BP/OX/Mg (**Figure 5.8a**) showing the metal coordination and small α-helix Aα6' stabilisation by oxalate binding. The polypeptide chain is shown as a cartoon. Ligands and interacting residues are shown as sticks. Residue Phe213 which potentially favours the binding of oxalate in the correct position is shown in stick format as well. Chain A of protein is coloured green and chain C is brown. The divalent metal ion Mg^{2+} and monovalent ion K^{+} are shown as a green sphere and a purple sphere, respectively. The *Fo-Fc* electron densities for Mg^{2+} and K^{+} are shown as a grey mesh contoured at 3.0 σ and 4.0 σ, respectively. Water molecules are shown as red spheres. Potential interatomic interactions are indicated by purple dashed lines and the relevant distances are given in Ångstroms. (b, c) Schematic representations are shown for Mg^{2+} coordination (b) and K^{+} coordination (c) in the active site, respectively. The octahedral geometry for K^{+} is not completed and the missing coordination is represented by dashed lines. The interatomic distances for the interactions are given in Ångstroms.

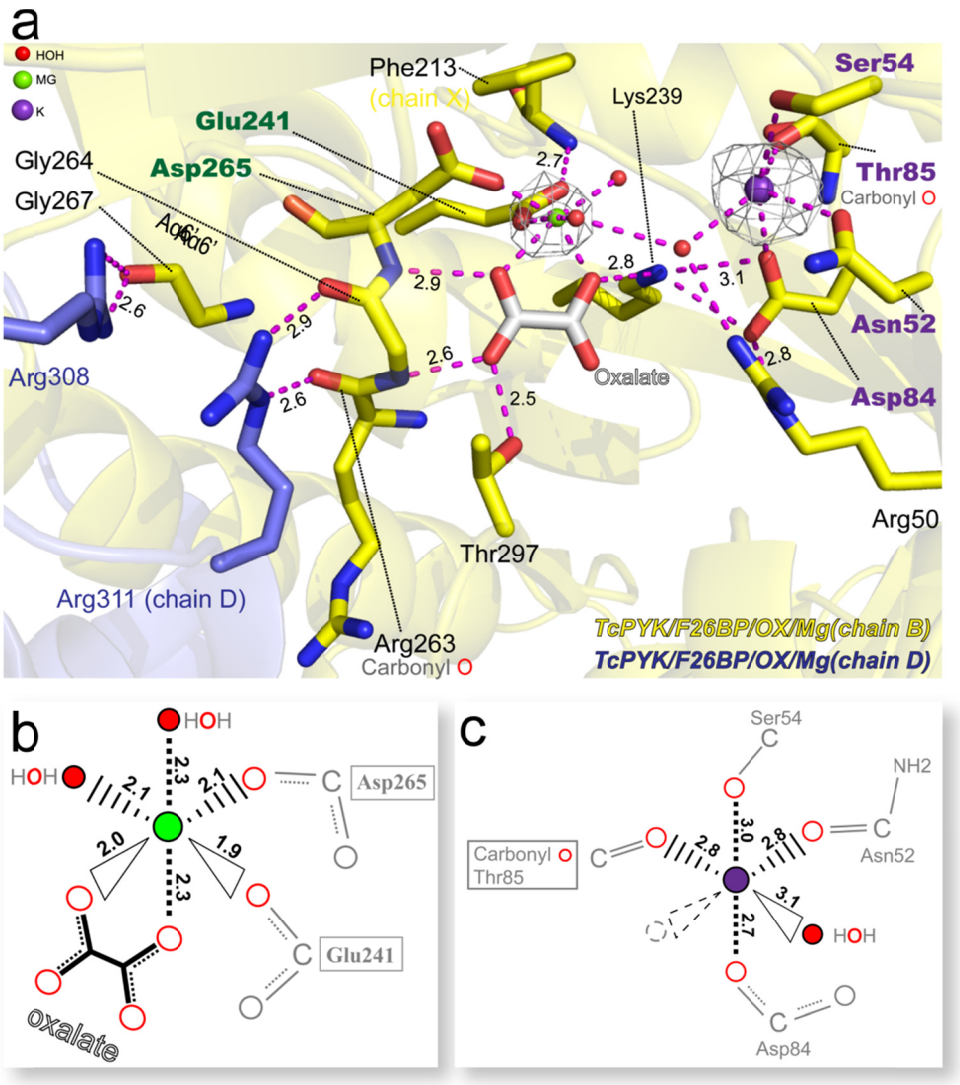


Figure 5.10 Metal ion coordination at the active site of *Tc*PYK/F26BP/OX/Mg (chain B).

(a) Close-up of the active site of *Tc*PYK/F26BP/OX/Mg (**Figure 5.8b**) showing the metal coordination and small α-helix Aα6' stabilisation by oxalate binding. The polypeptide chain is shown as a cartoon. Ligands and interacting residues are shown as sticks. Residue Phe213 which potentially favours the binding of oxalate in the correct position is shown in stick format as well. Chain B of protein is coloured yellow and chain D is blue. The divalent metal ion Mg^{2+} and monovalent ion K^{+} are shown as a green sphere and a purple sphere, respectively. The *Fo-Fc* electron densities for Mg^{2+} and K^{+} are shown as a grey mesh contoured at 3.0 σ and 4.0 σ, respectively. Water molecules are shown as red spheres. Potential interatomic interactions are indicated by purple dashed lines and the relevant distances are given in Ångstroms. (b, c) Schematic representations are shown for Mg^{2+} coordination (b) and K^{+} coordination (c) in the active site, respectively. The octahedral geometry for K^{+} is not completed and the missing coordination is represented by dashed lines. The interatomic distances for the interactions are given in Ångstroms.

5.3.3.3. A flexible protein motif from the active site

(i) Movement of a protein motif is observed at the active site in *Tc*PYK/F26BP/OX/Mg compared to the apo structure

Previously, five side-chain re-orientations have been shown by superposition of the active sites of apo *Tc*PYK and *Tc*PYK/F26BP/OX/Mg. Moreover, a shift of a small flexible motif containing the small α-helix Aα6' is also observed from this comparison (**Figure 5.11b**), and includes Asp265 which is important for active site Mg^{2+} binding. The C^{α} RMS difference of residue Asp265 is 0.8 Å, while no significant C^{α} RMS difference of Glu241 is observed, which is the other Mg^{2+} coordinated residue. This small shift of the motif cooperating with the side-chain re-orientations of residues Asp265 and Glu241 possibly favours the binding of Mg^{2+} in the active site.

(ii) Further identification of the small motif

To further identify and characterise the shifted motif, the C^{α} RMS difference for each residue of A-domain from the superposed structures was calculated and the results were plotted in **Figure 5.12a**. The residues lying at the domain boundary normally have relatively high RMS difference numbers as a result of the mobile B-domain (residues 90-187). Therefore, those residues were not considered further in the analysis. Apart from the domain boundary regions, residues 262-277 at the active site show significant shift (with an average RMS fit of 1.12 Å) between apo *Tc*PYK and *Tc*PYK/F26BP/OX/Mg structures compared to the overall RMS fit of 0.42 Å for the superposed A-domains (**Figure 5.12**). In this case, residues 262-277 are recognised as a flexible motif at the active site from a comparison between apo *Tc*PYK and *Tc*PYK/F26BP/OX/Mg structures.

(iii) The mechanism of this motif shift and its role in enzyme function

From the superposed structures in **Figure 5.11b**, the movement of this motif seems to

be directly triggered and stabilised by additional interactions across the large interface, namely the hydrogen bonding interactions between the small motif (residues Arg263 and Gly264) and the side chain of Arg311 from the adjacent subunit. In the apo structure, the side chain of Arg311 is buried inside the large interface and does not possess the interaction with the small motif. The interactions between the small motif and active site ligands (Mg^{2+} and oxalate) may in addition stabilise the conformation of the motif (**Figure 5.11a**). Therefore, the position of this small motif seems to be essential for the transition between inactive state (apo *Tc*PYK) and active state (*Tc*PYK/F26BP/OX/Mg) structures.

A question is raised by these observations: what initially triggers the movement of that small motif? It could be the effector F26BP which controls the allosteric state of *Tc*PYK, or the active site ligand oxalate could directly coordinate with the small motif. In **Chapter 6**, it will be proved that the small motif shift at the active site only correlates with the allosteric regulation of the enzyme.

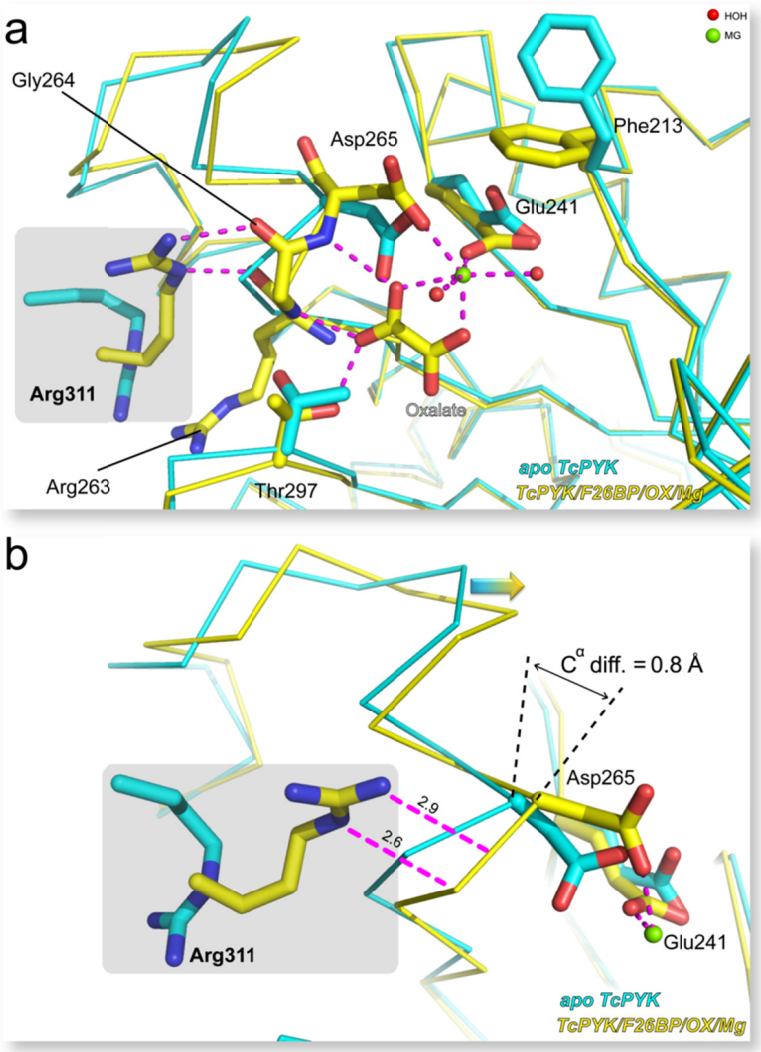


Figure 5.11 Superposition of the A-domains (residues 19-89, 188-358) of apo *Tc*PYK and *Tc*PYK/F26BP/OX/Mg to demonstrate the shift of a small enzyme motif and side-chain reorientations.

The superposed structures are apo *Tc*PYK (cyan) in inactive T-state and *Tc*PYK/F26BP/OX/Mg (yellow) in active R-state. The superposition was performed on the A-domains (residues 19-89, 188-358) of PYK structures in monomeric form (chain A from apo *Tc*PYK and chain B from *Tc*PYK/F26BP/OX/Mg). Interatomic interactions are indicated by purple dashed lines and the relevant distances are given in Ångstroms. The Arg311 residues (shaded by grey boxes) which stabilize the small α-helix Aα6'are from adjacent chains within a tetramer. (a) The polypeptide chains are shown as ribbons while relevant residues and oxalate are shown as sticks. Mg^{2+} is shown as a green sphere, while water molecules are represented as red spheres. (b) Rotated and enlarged version of the structures shown in (a). The shift of the small motif including Aα6' is indicated by an arrow. The C^{α} atom of residue Asp265 which coordinates Mg^{2+} in *Tc*PYK/F26BP/OX/Mg has a shift of 0.8 Å compared to the residue in apo *Tc*PYK. The interactions between Arg311 and the small motif are highlighted by purple dashed lines.

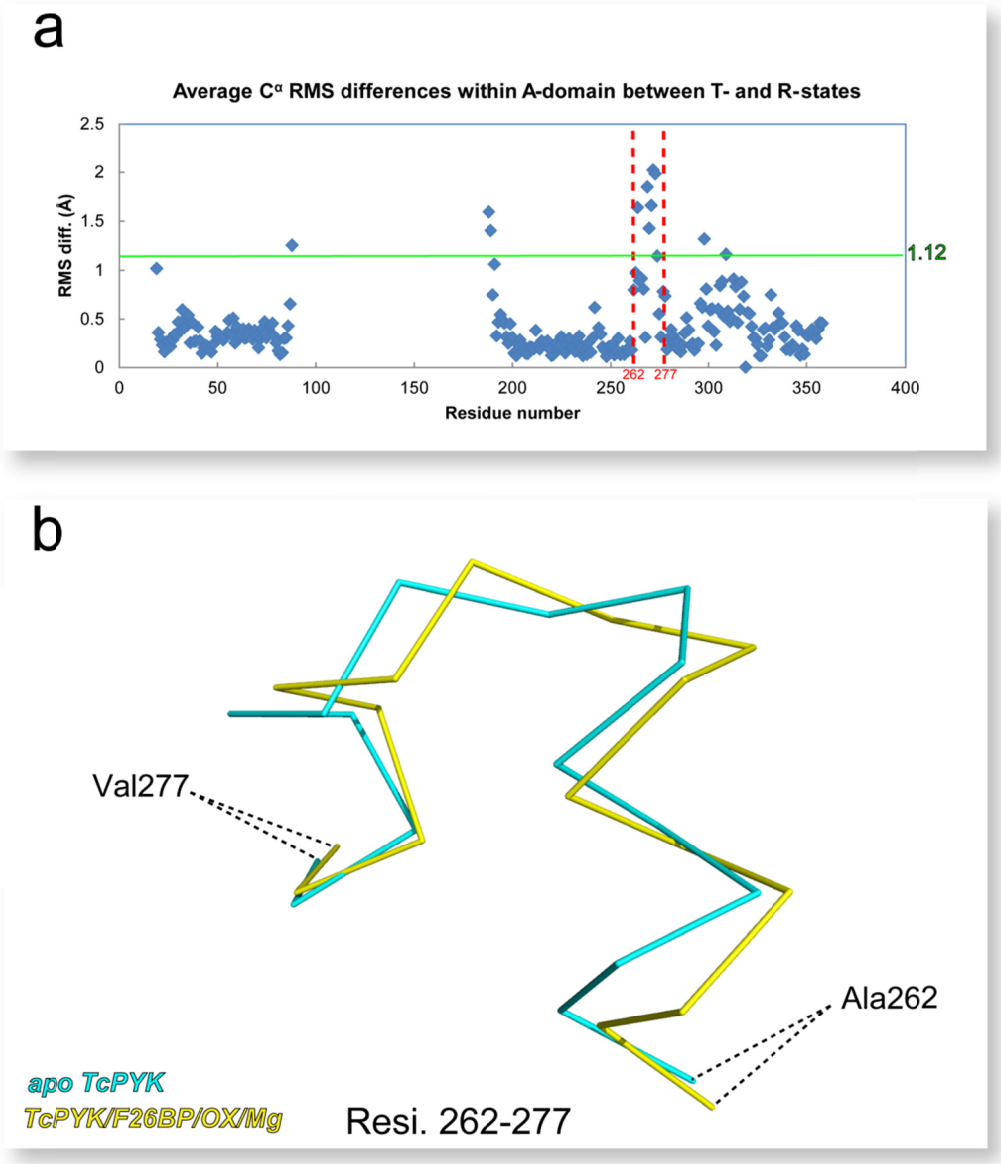


Figure 5.12 The C^{α} RMS differences identify a significant shift from a small motif within the A-domain as a result of allosteric activation or substrate analogue binding.

(a) The calculation for RMS differences was performed by the superposition of the A-domains (19-89, 188-358) from inactive state (apo *Tc*PYK) and active state (*Tc*PYK/F26BP/OX/Mg) structures. The RMS difference values were plotted as a function of residue number. A small motif (residues 262-277) was identified with large RMS differences, and is indicated by dashed lines in red. The average RMS difference for all residues of this small motif is 1.12 Å as indicated by the green line. In PYK structures, residues located atdomain boundaries normally have high RMS differences, as can be seen in the plot. (b) The polypeptide chain of the identified motif from residues 262 to 277 is shown as ribbons. The inactive state apo *Tc*PYK and active state *Tc*PYK/F26BP/OX/Mg structures are shown in cyan and yellow, respectively

5.3.4. Allosteric mechanism of TcPYK

5.3.4.1. The T- to R-states transition: rigid-body rotation function

(i) *Tc*PYK is allosterically regulated by the activator F26BP

In **Chapter 3**, it has been shown that *Tc*PYK displays sigmoidal kinetics with respect to the substrate PEP, and that F26BP is capable of coverting this response to hyperbolic and to significantly increase the affinity for PEP (**Table 3.2**). In the structures of *Tc*PYK, the effector site where F26BP is bound lies ~40 Å away from the active site where the reaction proceeds. Both the kinetic properties and the structures of *Tc*PYK suggest that *Tc*PYK is an allosteric enzyme, and that F26BP allosterically regulates its activity.

A question is raised by these observations: how does the activator F26BP manage to influence the enzyme kinetic behaviour for *Tc*PYK by binding to the effector site which is far away from its active site? It has been reported by Dr. Hugh Morgan (Morgan et al. 2010b) that in *Lm*PYK an AC-core (A- and C-domains) rotation (6^{o}) is responsible for the T- and R-state transition which is initiated by the binding of effector (F26BP) or active site ligand (PEP or PEP analogue); meanwhile, the binding of nucleotide ligand (ADP or ATP) is not responsible for this rigid-body motion but is required for the fully closed B-domains.

(ii) The F26BP-bound structure is held in a different tetrameric conformation compared to the apo structure of *Tc*PYK

The crystal structures of apo *Tc*PYK and F26BP-bound *Tc*PYK/F26BP/OX/Mg provide an opportunity to compare their differences in structure. To answer the question raised in section (i) above, the monomer of apo *Tc*PYK was superposed onto the monomer of *Tc*PYK/F26BP/OX/Mg, resulting in an average RMS fit of 0.64 Å for all C-α atoms (**Table 5.2**). The mobile B-domains (residues 90-187) were removed from the comparison. This relatively small difference (0.64 Å) between monomers indicates there is no significant conformational change for the monomeric

structure after the binding of F26BP and oxalate to the enzyme.

However, *Tc*PYK is a homotetramer enzyme: will the effector F26BP alter the tetrameric conformation? Complete tetramers were generated from apo *Tc*PYK and *Tc*PYK/F26BP/OX/Mg from their symmetry related monomers, and superposed onto each other resulting in an average RMS fit of 2.40 Å for all C-α atoms (**Table 5.3**). Mobile B-domains and the short N-terminal domains (residues 1-18) were removed from the comparison. Compared to the RMS fit of 0.64 Å for the monomer structure, the RMS fit of 2.40 Å for the tetramer structure is significant, thus indicating conformational differences between the tetrameric structures of apo *Tc*PYK and *Tc*PYK/F26BP/OX/Mg. These differences can also been recognised visually in **Figure 5.14a** and **5.14b**. The structure without bound effector and active-site ligand is considered to be in the inactive T-state conformation, while the structure with bound F26BP and active-site ligand is considered to be in the active R-state conformation (**Figure 5.14c**). The transition between T- and R-states is found to be regulated by a rigid-body motion.

(iii) An 8^{o} rigid-body rotation regulates the transition between the T- and R-states of *Tc*PYK

The AC-core rigid-body rotation was observed by two steps of superposition, as shown in **Figure 5.13**. The mobile B-domains and N-terminal domains which are not directly related to the rigid-body motion are not included in this rotation function study. Only the A- and C-domains (AC-core) which are relatively rigid within each chain are considered to be involved in the rotation analysis. The movement of B-domain regulated by the binding of substrate or substrate analogue will be discussed in section **5.3.5**. As has been described in the Materials and Methods section, the AC cores from two tetramers (apo *Tc*PYK and *Tc*PYK/F26BP/OX/Mg) were superposed (C-α atoms fit) simultaneously onto each other in the first step resulting in an average RMS fit of 2.40 Å for all Cα atoms. In the second step,

individual chains from the T-state structure were superposed to the corresponding chain from the R-state structure using the new coordinates generated from the first step, resulting in an average RMS fit of 0.54 Å for all Cα atoms. An 8^{o} rotation motion of each AC-core was observed from the rotation matrix in the second step of the superposition process. **Figure 5.14** summarises this AC-core rigid-body rotation to elucidate the conformational transition between the T- and R-states of *Tc*PYK.

To simply locate the rotation pivot of each chain, the RMS differences between the T- and R-state tetramer structures after the first superposition step were plotted with corresponding residue number **(Figure 5.15)**. A group of residues 429-433 with the lowest average RMS fit of 0.52 Å for Cα atoms was considered to be the rotation pivot on each chain as a result of its RMS minimum difference between T- and R-state structures.

Table 5.2 The RMS differences (Å) for C-α atoms between monomer structures of apo *Tc*PYK and *Tc*PYK/F26BP/OX/Mg †

		*Tc*PYK/F26BP/OX/Mg		Average
		Chain A	Chain B	
apo *Tc*PYK	Chain A	0.63 Å	0.66 Å	0.64 Å
	Chain B	0.65 Å	0.65 Å	

†B-domains were removed from the comparisons because of their multiple positions in different structures.

Table 5.3 The AC-core rigid-body rotation from T- to R-state of *Tc*PYK

T-state	R-state	Average RMS (Å) †	Rotation Angle ‡
apo *Tc*PYK	*Tc*PYK/F26BP/OX/Mg	2.40	$8.2^{o}\pm0.3^{o}$

†Average C^{α} RMS differences between T- and R-state tetramer structures (AC cores). All four AC cores (A- and C-domains) from the apo *Tc*PYK tetramer were simultaneously superposed onto all four AC cores from *Tc*PYK/F26BP/OX/Mg tetramer.

‡Calculated rotation angle with standard deviation.

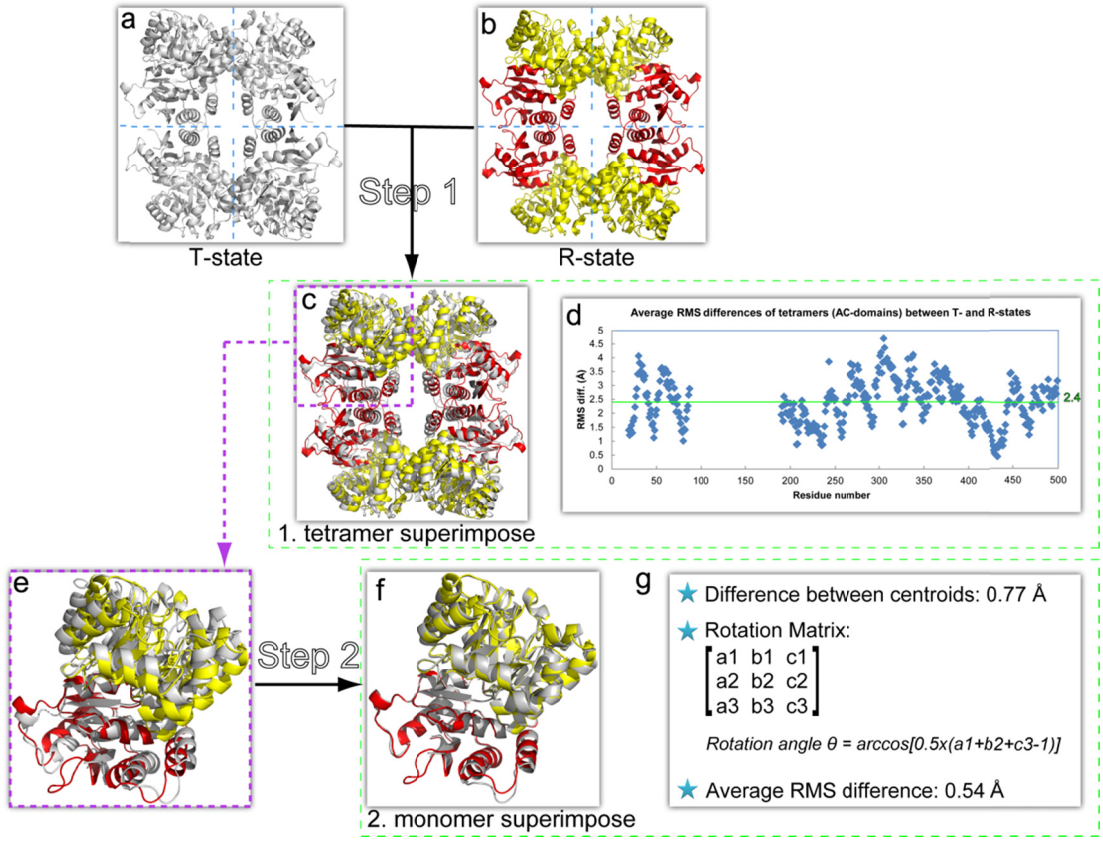


Figure 5.13 A flow-chart showing the procedure to discover rigid-body rotations between T- and R-state structures of *Tc*PYK.

The polypeptide chains are shown as cartoons, and domains are highlighted by different colours (A-domains: yellow; C-domains: red). (a, b) T- and R-state tetramers of *Tc*PYK are from the structures of apo *Tc*PYK and *Tc*PYK/F26BP/OX/Mg, respectively. Mobile B-domains and N-terminal domains are removed from the models. The tetramer is generated from symmetry related monomers. The four subunits (AC cores) of each tetramer are indicated by dashed lines. (c, d) The first step was to superpose the T-state tetramer onto the R-state tetramer resulting in new coordinates (c). The RMS differences were plotted with residue numbers in (d). The overall RMS difference (for Cα atoms) is 2.4 Å which is indicated by green horizontal line (f, g) The second step was to superpose single chains of T- and R-state structures individually using the coordinates generated from the first step. Figures (e) and (f) represent one example of the superposition process in the second step. The small difference (0.77 Å) between centroids of these two models indicates an almost pure rotation in this superposition process. The rotation angle applied in the superposition process can be calculated from the rotation matrix (g) obtained in the second step superposition process. The average RMS difference (for Cα atoms) after superposition is 0.54 Å. The method described here to analyse the rigid-body motion of trypanosomatid PYKs has been published by Dr. Hugh Morgan to describe the rigid-body rotation of *Lm*PYK (Morgan et al. 2010b). Refer to section **2.4.13** in **Chapter 2** for more details.

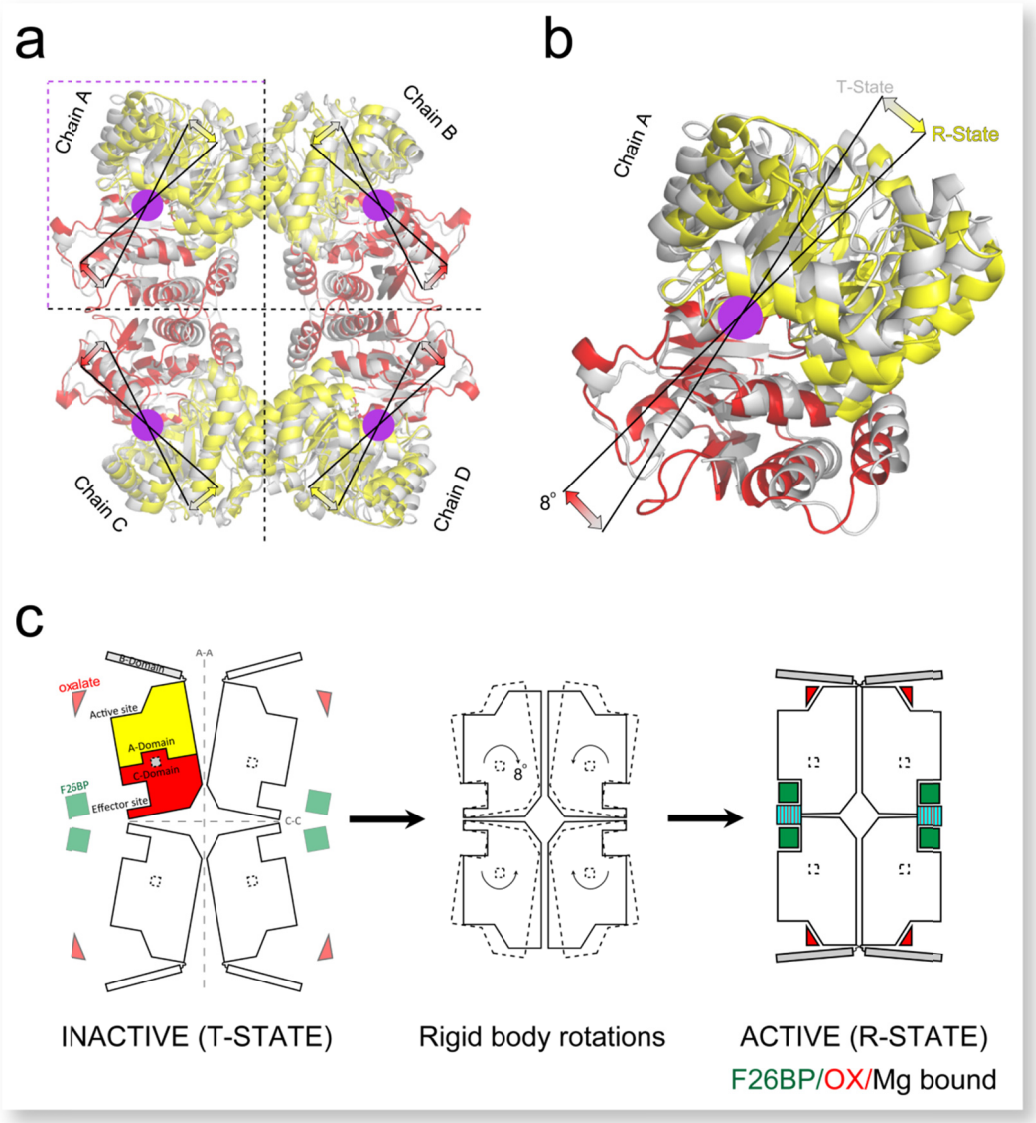


Figure 5.14 Rigid body rotation of *Tc*PYK between T- and R-states.
(a) All B-domains and N-terminal domains are removed, and interface boundaries are highlighted by dashed lines. The superposed polypeptide chains are shown as cartoons, and domains are highlighted by different colours (A-domains: yellow; C-domains: red). The C-α atoms of the AC cores (A- and C- domains) of the inactive T-state apo *Tc*PYK tetramer (grey) were superposed onto the active R-state *Tc*PYK/F26BP/OX/Mg tetramer with an RMS fit of 2.4 Å. The T- (grey) to R-states transition: the directions of the 8° rigid body rotation of the AC-core around the central pivots (highlighted in purple) are shown by arrows. (b) An enlargement of chain A showing the 8° rotation around the pivot. (c) Schematic representation of the T- and R-states of *Tc*PYK structures, and the motion of rigid body rotation between these two states initiated by the allosteric effector F26BP and substrate analogue oxalate. This modification of the scheme is originally presented by colleague Dr. Hugh Morgan (Morgan et al. 2010b).

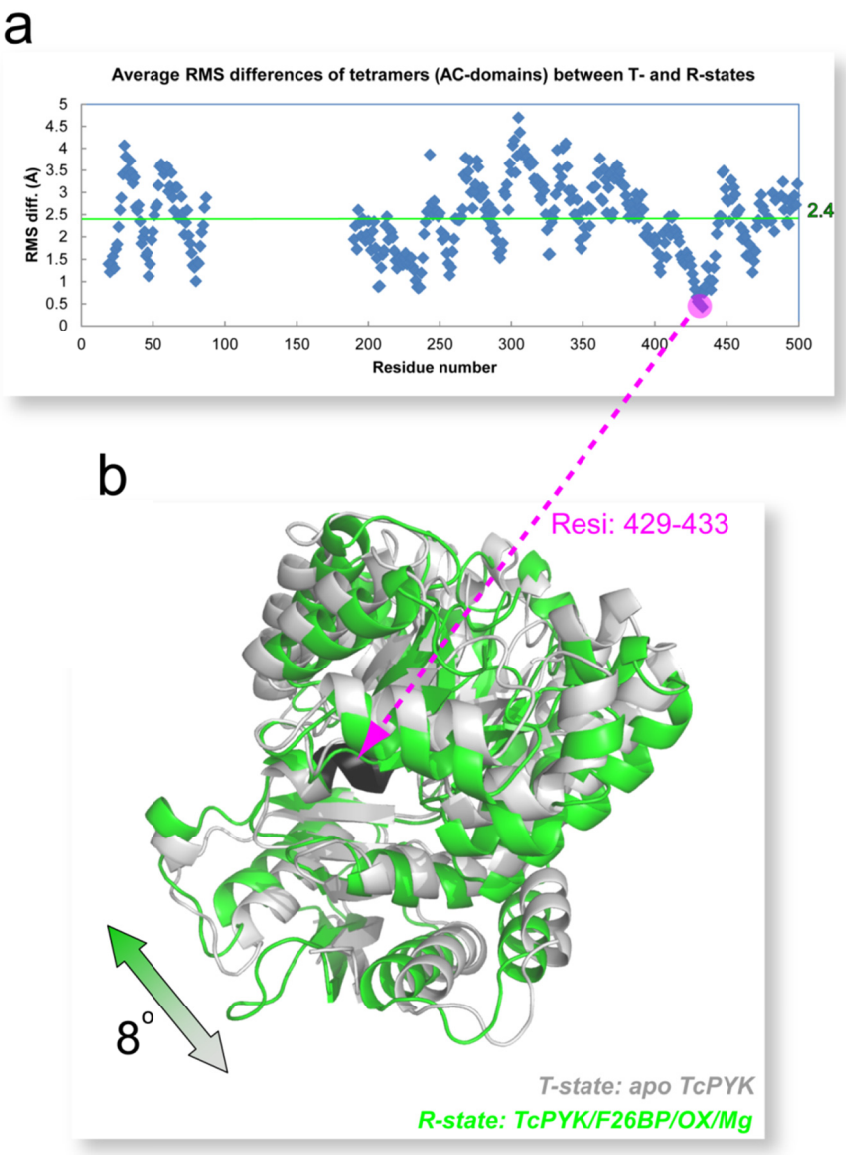


Figure 5.15 The rotation pivot of the rigid-body (AC-domains) rotation for *Tc*PYK is located at residues 429-433 as observed from the analysis of RMS differences (C^{α}) of each residue between the T- and R- state structures.

(a) Average RMS differences between the AC cores (A-domain and C-domain) of T-state apo *Tc*PYK and R-state *Tc*PYK/F26BP/OX/Mg structures. The calculation for RMS differences was performed by simultaneously superposing the AC-cores of the tetramer from apo *Tc*PYK onto the tetramer of *Tc*PYK/F26BP/OX/Mg. The numbers of RMS differences were plotted as a function of residue numbers. A rotation pivot (residues 429-433) with the lowest RMS difference (average 0.52 Å) was identified and shaded in purple. The average RMS difference for all residues of the AC cores is 2.40 Å as indicated by the green line. (b) The rigid body rotation for each subunit, where the rotation is indicated by an arrow from T-state (grey) to R-state conformation (green). The pivot point (residues 429-433) located on an α-helix of each C-domain is highlighted in black.

5.3.4.2. Additional interface interactions lock structure in R-state

(i) Summary of interface interactions between T- and R-state structures

Interface interactions play a main role in stabilising the tetrameric conformation of *Tc*PYK in both T- and R-states, and include large (A-A) interface interactions and small (C-C) interface interactions (**Figure 5.16** and **5.18**). The residues involved in the interface interactions (forming hydrogen bonds and salt bridges) and the interatomic distances are analysed by the web-based program PDBePISA (Krissinel and Henrick 2007) and are summarised in **Table 5.4** (large interface) and **Table 5.5** (small interface) for both the T- and R-states of *Tc*PYK.

(ii) Changes of interactions from T- to R-states of *Tc*PYK across the large interface

During the transition from T-state (apo *Tc*PYK) to R-state (*Tc*PYK/F26BP/OX/Mg), the average interface area along the large interface increases from 2163.8 $Å^2$ to 2726.3 $Å^2$, which indicates more and stronger interactions in an R-state conformation. Generally, most of the interface interactions along the large interface are contributed by residues of the A-domains, and only a few residues from the other domains are involved (**Table 5.4**). It has been observed that Arg311 changes its side-chain position dramatically during the T- to R-state transition (**Figure 5.11**). In the T-state structure, the side-chain of Arg311 is buried inside the large interface to interact with Gln298 and Asp316 from the adjacent chain. In the R-state structure, the side-chain of Arg311 performs a nearly 180^o re-orientation and flips out to make new hydrogen bonding interactions with the small α-helix Aα6' (residues Arg263 and Gly264) from the adjacent chain. Arg311 has been proved to be crucially important for enzyme activity in other PYKs from *E. coli* (Valentini et al. 2000) and humans (Pendergrass et al. 2006).

(iii) Changes of interactions from T- to R-state *Tc*PYK across the small interface

During the transition from T- to R-state, the average interface area along the small

interface also increases, from 1017.7 Å^2 to 1143.7 Å^2, and thereby contributes to increased interface interactions. The small interface interactions rely totally on residues of the C-domains (**Table 5.5**). It has been discussed in section **5.3.2** that the effector loop (residues 482-488) is normally disordered in the absence of effector binding. However, in the presence of F26BP, the effector forms additional hydrogen bonds with the effector loop to lock it in place (**Figure 5.5**). Each ordered effector loop of the enzyme forms four additional salt bridges (Asp483…Arg494) across the small (C-C) interface in a *Tc*PYK tetramer to further stabilise the effector loops and the R-state conformation (**Figure 5.17**). Another residue in the effector loop, Lys485, potentially forms a hydrogen bond with Ser499 across the C-C interface, although the interaction is not always present. Additional hydrogen bonding interactions across the C-C interface also contribute to the stability of R-state *Tc*PYK. For example, Tyr489 makes connections with Ile495 and Tyr497 across the C-C interface via water molecules shown in **Figure 5.17**.

(iv) Potential hydrophobic interactions across interfaces

Some hydrophobic residues along the interfaces of the tetramer may contribute to the conformational stability for both T- and R-state *Tc*PYKs (**Figure 5.18**). Along the large interface, residues potentially involved in hydrophobic interactions are Leu4, Ala5, Val8, Leu10, Ile12, Phe13, Val246, Ile270, Ala272, Val276, Val277, Met280, Ile281, Val288, Pro309, Val314, Ala318, Phe322, Tyr346, Ile350, Ala354, Ala357, Met 363 and Met370. Along the small interface, residues potentially involved in hydrophobic interactions are Pro374, Pro377, Ala380, Val381, Val392, Leu476 and Ile495, Tyr489 and Tyr497.

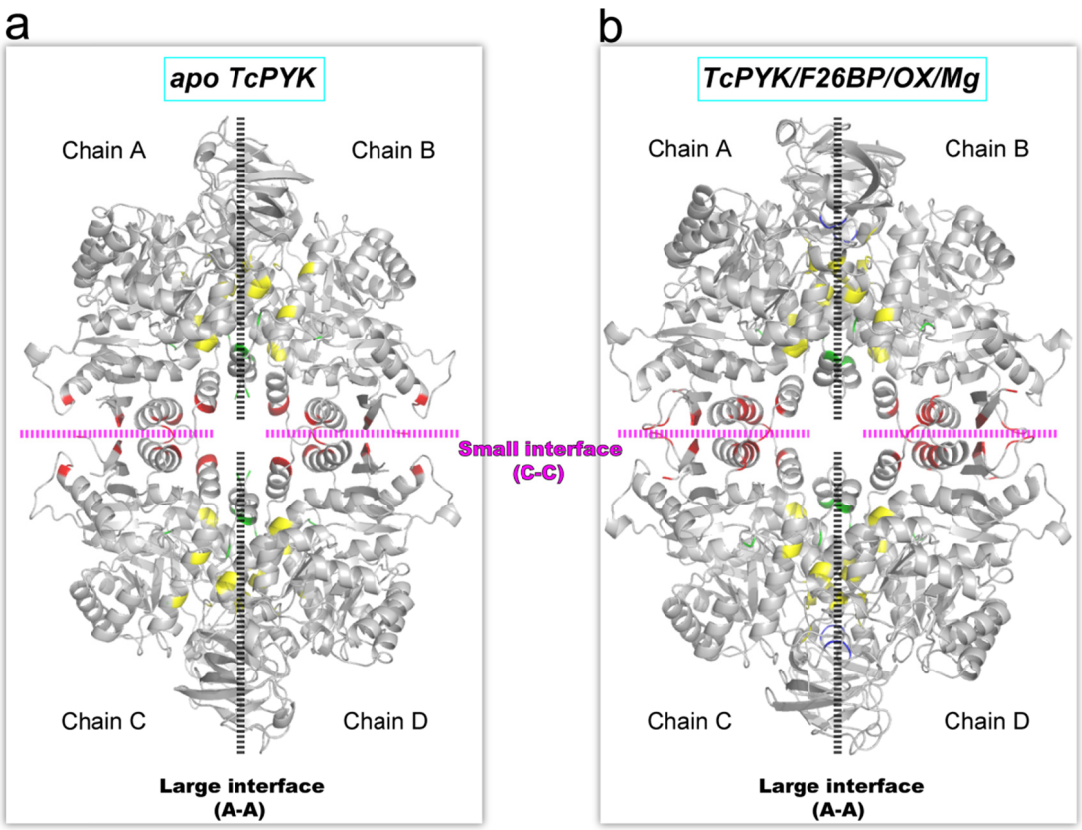


Figure 5.16 Residues involved in interface polar interactions between subunits of the *Tc*PYK tetramer.

The inactive T-state structure of the *Tc*PYK tetramer (apo *Tc*PYK) is shown in (a) while the active R-state structure of the *Tc*PYK tetramer (*Tc*PYK/F26BP/OX/Mg) is shown in (b). Tetramers were made by symmetric mates. Both structures are represented as cartoons in grey. The large (A-A) and small (C-C) interfaces between subunits are shown as dashed lines, and indicate the four subunits (monomers) in one tetramer. Residues involved in the interface interactions (excluding hydrophobic interactions which will be shown in **Figure 5.18**) are indicated by colours. Residues belonging to domains A, B, C and N-terminal are coloured in yellow, blue, red and green, respectively.

Table 5.4 Comparison of large (A-A) interface polar interactions of T- and R-state structures of *Tc*PYK

		apo *Tc*PYK T-state	***Tc*PYK/F26BP/OX/Mg** R-state	
Interface[†]		A/B or C/D	A/B	C/D
Chain 1[‡]	**Chain 2**	**Distance (Å)**		
Lys369[NZ]*	Met1[O]	3.40		
Lys285[NZ]	Val8[O]	3.00	3.57	3.87
Lys274[NZ]	Ile12[O]	2.76	2.95	3.43
Tyr346[OH]	Glu273[OE1]	2.79		
Asn7[ND2]	Ser284[OG]	2.96	3.31	3.10
Arg311[NH2]	Gln298[O]	3.04		
Arg311[NE]	Gln298[OE1]	2.98		
Arg311[NH1]	***Asp316[OD1]***	2.83		
Ser315[OG]	Asp316[OD1]	3.52	2.43	2.74
Arg311[NH2]	***Asp316[OD2]***	2.76		
Arg311[NH1]	***Asp316[OD2]***	3.53		
Arg311[NH2]	***Asp316[OD1]***	3.63		
Lys274[NZ]	***Glu353[OE2]***	3.04		
Lys274[NZ]	***Glu353[OE1]***	3.02	3.64	3.00
Met1[O]	Lys369[NZ]	3.86		
Val8[O]	Lys285[NZ]	2.94	3.57	3.87
Ile12[O]	Lys274[NZ]	2.92	2.95	3.43
Glu273[OE1]	Tyr346[OH]	3.02		
Ser284[OG]	Asn7[ND2]	2.96	3.31	3.10
Gln298[O]	Arg311[NH2]	3.15		
Gln298[OE1]	Arg311[NE]	2.84		
Asp316[OD1]	***Arg311[NH1]***	2.89		
Asp316[OD1]	Ser315[OG]	3.82	2.43	2.74
Asp316[OD2]	***Arg311[NH2]***	2.84		
Asp316[OD2]	***Arg311[NH1]***	3.39		
Glu353[OE1]	***Lys274[NZ]***	3.09	3.64	3.00
Glu353[OE2]	***Lys274[NZ]***	2.84		
Arg311[NH1]	Asp146[O]			3.23
Arg311[NH1]	Gly264[O]		2.60	2.91
Arg308[NH2]	Gly267[O]		2.18	2.58
Arg311[N]	Gln298[OE1]		2.60	2.76

Table 5.4 ***(continued)***

Arg311[NH2]	Gln298[OE1]		3.41	3.48
Asp146[O]	Arg311[NH1]			3.23
Gly264[O]	Arg311[NH1]		2.60	2.91
Gly267[O]	Arg308[NH2]		2.18	2.58
Gln298[OE1]	Arg311[NH2]		3.41	3.48
Gln298[OE1]	Arg311[N]		2.60	2.76
Arg311[NE]	Arg263[O]		2.79	2.63
Arg263[O]	Arg311[NE]		2.79	2.63
Average interface area (Å^2)‡		2163.8	2726.3	

† The interfaces between subunits shown here are those corresponding to **Figure 5.14**. Chain 1 and Chain 2 which form half of the large interface correspond to: chain A and chain B for the T-state structure; chain A and chain B, or chain C and chain D for the R-state structure.

‡ The interface areas, interacting residues and distances are obtained from the online service PDBePISA (Krissinel and Henrick 2007). Atom distances beyond 3.5 Å are not shown in the table unless they are corresponding to other interacting distances.

* Residues belonging to domains A, B, C and N-terminal are shaded in yellow, blue, orange and green, respectively. Residues which can potentially form salt bridges are shown in bold and italic.

Table 5.5 Comparison of small (C-C) interface polar interactions of T- and R-state structures of *Tc*PYK

		apo *Tc*PYK T-state	***Tc*PYK/F26BP/OX/Mg** R-state
Interface[†]		A/C or B/D	A/C or B/D
Chain 1	**Chain 2**	**Distance (Å)**	
Ser384[OG][*]	Ser384[OG]	2.86	
Met375[N]	Glu391[O]	2.93	2.79
Lys368[NZ]	***Glu391[OE2]***	2.55	3.63
Ile495[N]	Asn491[O]	3.13	2.89
Thr493[N]	Thr493[O]	3.00	2.77
Asn491[ND2]	Ile495[O]	3.15	2.91
Glu391[OE2]	***Lys368[NZ]***	2.59	3.51
Glu391[O]	Met375[N]	2.92	2.92
Ile495[O]	Asn491[ND2]	2.86	3.13
Gln492[OE1]	Gln492[NE2]	2.52	3.20
Thr493[O]	Thr493[N]	3.00	3.02
Asn491[O]	Ile495[N]	3.01	2.87
Ser388[OG]	Ser384[OG]		3.03
Asn387[ND2]	Ser384[OG]		3.13
Lys368[NZ]	Asn387[OD1]		2.60
Ser384[OG]	Ser388[OG]		2.94
Ser384[OG]	Glu391[OE1]		3.23
Ser383[OG]	Glu391[OE2]		2.88
Arg494[NE]	***Asp483[OD1]***		2.93
Asp483[OD1]	***Asp494[NE]***		3.10
Ser384[OG]	Asn387[ND2]		3.17
Asn387[OD1]	Lys368[NZ]		2.76
Glu391[OE1]	Ser384[OG]		3.14
Glu391[OE2]	Ser383[OG]		2.74
Asp483[OD2]	***Arg494[NH2]***		3.30
Arg494[NH2]	***Asp483[OD1]***		3.40
Interface area (Å^2)[‡]		1017.7	1143.7

[†] The interfaces between subunits shown here correspond to **Figure 5.14**. Chain 1 and Chain 2 which form half of small interface correspond to: chain A and chain C, or chain B and chain D for the T- or R-state structure.

[‡] The interface areas, interacting residues and distances are obtained from the online service PDBePISA (Krissinel and Henrick 2007). Atom distances beyond 3.5 Å are not shown in the table unless they are corresponding to other interacting distances.

[*] All residues here belong to domains C and are shaded in blue. Residues which can

potentially form salt bridges are shown in bold and italic. Residues in the effector loop (residues 482-488) are coloured purple.

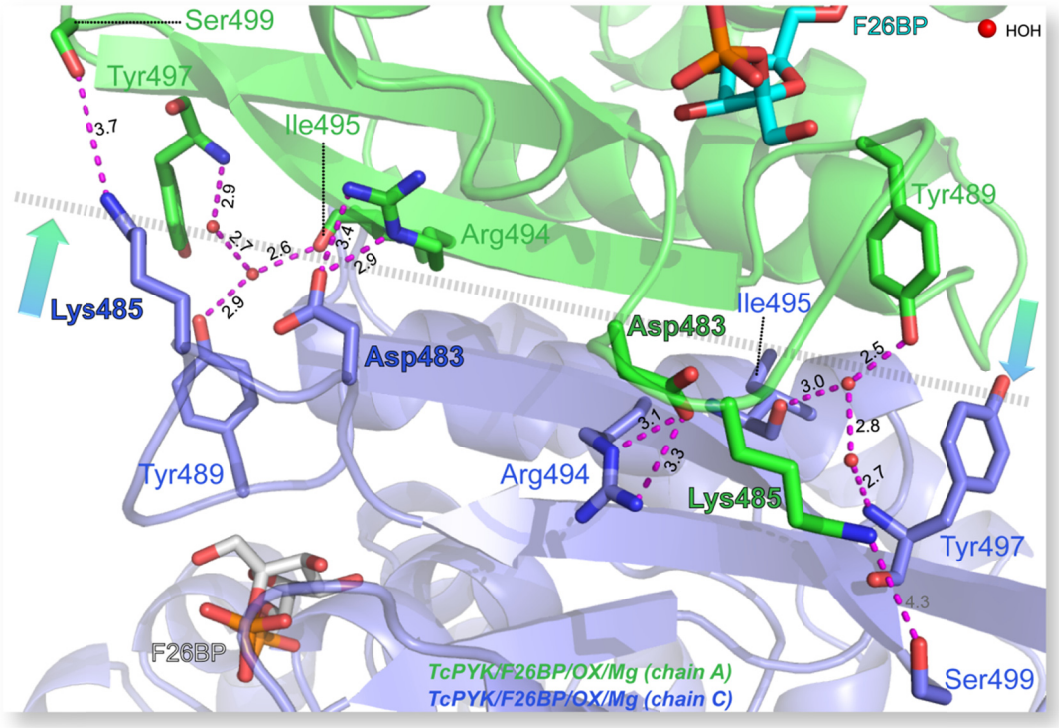


Figure 5.17 Close-up of the A/C interface of *Tc*PYK/F26BP/OX/Mg showing the additional salt bridges formed across the small (C-C) interface as a result of F26BP binding.

The polypeptide chain is represented as cartoons in green (chain A) or blue (chain C). These two chains correspond to the chains discussed in **Figure 5.14** and **Table 5.3**. The interacting residues which are located on or close to the effector loop are shown as sticks. F26BP molecules are also represented as sticks. Water molecules are shown as spheres in red. Interatomic interactions are indicated by purple dashed lines and the interacting distances are given in Ångstroms. F26BP binding caused additional salt bridges to be formed between Asp483…Arg494. Other additional hydrogen bond interactions were observed between Lys485 and Ser499, although not always present and relatively weak (3.7 Å). The hydrogen bond interactions via water molecules were Tyr489…water…Ile495, and Tyr489…water…water…Tyr497. The small (C-C) interface between chain A and chain C is shown as a dashed line in grey, while the movements of the effector loops are indicated as arrows.

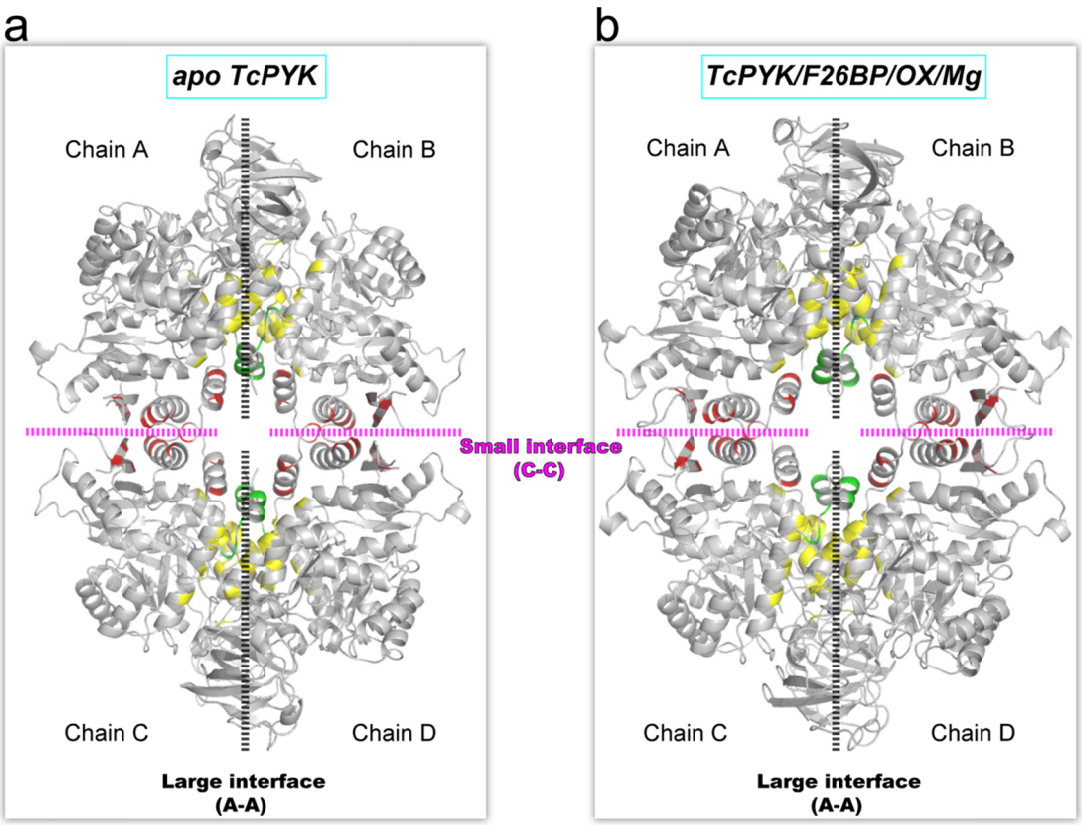


Figure 5.18 Residues potentially involved in hydrophobic interactions along the interfaces of the *Tc*PYK tetramer.

The inactive T-state structure of the *Tc*PYK tetramer (apo *Tc*PYK) is shown in (a) while the active R-state structure of the *Tc*PYK tetramer (*Tc*PYK/F26BP/OX/Mg) is shown in (b). Both structures are represented as cartoons in grey. The large (A-A) and small (C-C) interfaces between subunits are shown as dashed lines, and indicate the four subunits (monomers) in one tetramer. Residues potentially involved in hydrophobic interactions along the interfaces are indicated by different colours. These residues are as follows: for the large interface: Leu4, Ala5, Val8, Leu10, Ile12, Phe13, Val246, Ile270, Ala272, Val276, Val277, Met280, Ile281, Val288, Pro309, Val314, Ala318, Phe322, Ile350, Ala354, Ala357, Met363 and Met370; for the small interface: Pro374, Pro377, Ala380, Val381, Val392, Leu476 and Ile495. Residues belonging to domains A, B, C and N-terminal are coloured in yellow, blue, red and green, respectively.

5.3.4.3. The active R-state structure is in a thermally more stable conformation

As has been discussed in the preceding sections, the interface areas of *Tc*PYK increase after the conformational change from T- to R-state, and involve an 8^{o} rigid-body rotation of the AC-core. Additional interactions across the interfaces are formed in the R-state *Tc*PYK, such as the four additional salt bridges across the small interface formed between F26BP-bound effector loops. Therefore, the R-state structure of *Tc*PYK is predicted to be more rigid and stable. Biophysical measurements confirming this prediction are presented in this section. In the case of *Tb*PYK which is a proven anti-parasite drug target, more detailed biophysical studies have been carried out and the results will be presented in the next chapter (**Chapter 6**). Besides using the thermal-shift assay, high-resolution size-exclusion chromatography can also be considered to distinguish T-state *Tc*PYK and R-state *Tc*PYK.

(i) R-state *Lm*PYK is also in a thermally more stable conformation

This type of allosteric regulation which involves conformational change is almost identical to that of *Lm*PYK (Morgan et al. 2010b) which shares 75% amino acid sequence identity with *Tc*PYK. For *Lm*PYK, the active R-state conformation can be initiated and stabilised in three ways: 1. effector F26BP binding to the effector site; or 2. substrate PEP (or its analogue oxalate) cooperatively binding to the active site and the binding of nucleotide ligand which is required for fully closed B-domains and further stabilises the structure; or 3. the combination of effector-site ligand (F26BP) and active-site ligands including nucleotide ligands. All these three different ligated states of *Lm*PYK have been proven to be in more thermally stable conformations compared to the inactive T-state *Lm*PYK (Morgan et al. 2010b). Additionally, divalent metal Mg^{2+} normally presents in the R-state structures to stabilise the active conformation or/and participate in the binding of active site ligands to the protein. The roles of Mg^{2+} for allosteric *Tb*PYK will be discussed in

the next chapter (**Chapter 6**).

(ii) Three ligands were selected for thermal-shift assays

Thermal-shift assays were performed to analyse the stabilising effects of ligands (active site ligands and effector site ligand) on *Tc*PYK. F26BP was selected as an effector-site ligand, while the substrate PEP and its analogue oxalate were selected as two active-site ligands for these assays. In **Chapter 3**, *Tc*PYK was shown to behave with a sigmoidal response to the substrate PEP, but this response was converted to hyperbolic in the presence of the effector F26BP. Therefore, the substrate PEP may stabilise *Tc*PYK in a cooperative manner. The effector F26BP is currently unavailable commercially, so it was not possible to do the assay for the ligand combination of F26BP and PEP (or oxalate).

(iii) R-state *Tc*PYK is in a thermally more stable conformation compared to the T-state enzyme

The stabilities of *Tc*PYK in the presence of added ligands are given in **Figure 5.19**. Addition of the active-site ligand PEP to the apoenzyme increases the melting temperature (T_m) from 54.5°C to 57.5 °C. Addition of the other active-site ligand oxalate to the apoenzyme also increases T_m from 54.5°C to 61.0 °C. For the effector-site ligand F26BP (1.5 μM), the T_m of *Tc*PYK is increased from 54.5°C to 56.0 °C. All three ligands individually increase the thermal stability of *Tc*PYK indicating more rigidity and stability in the R-state conformation. Compared to substrate PEP, substrate analogue oxalate forms a thermally more stable complex with *Tc*PYK under the same ligand concentration [T_m (*Tc*PYK/oxalate) = 61.0 °C, compared to T_m (*Tc*PYK/PEP) = 57.5 °C]. One oxygen atom from the carboxyl group of oxalate can form indirect interaction with side chain oxygen of Ser331 mediated by a water molecule (**Figure 5.8a**). This indirect interaction between ligand and Ser311 cannot be formed in *Tc*PYK/PEP complex [The binding mode between

*Tc*PYK and PEP can be predicted by crystal structure of *Tb*PYK/F26BP/PEP/Mg which will be discussed in the next chapter (**Chapter 6**)]. In this case, this additional interaction in *Tc*PYK/oxalate complex might contribute to the higher T_m.

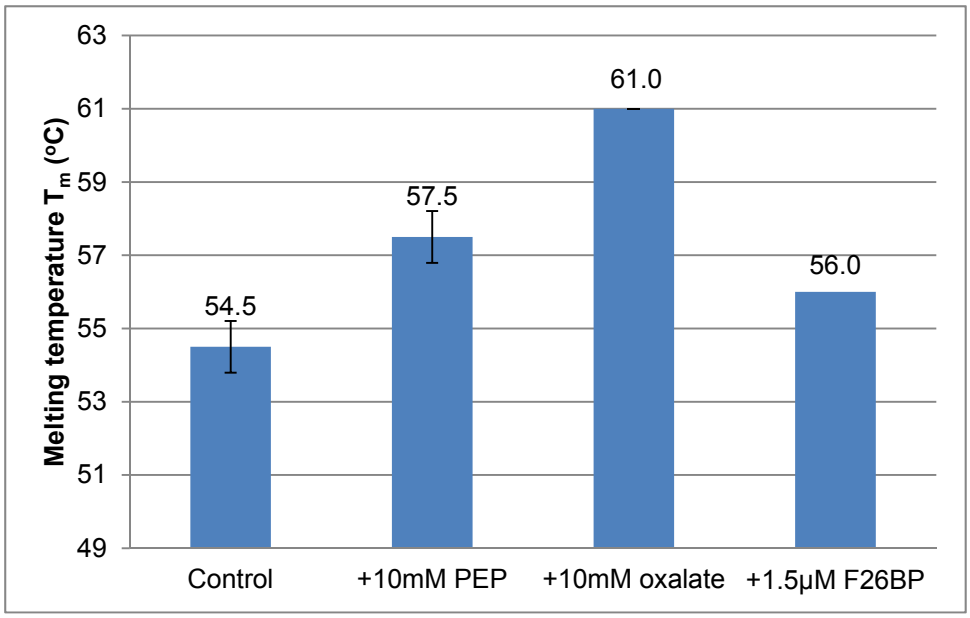


Figure 5.19 Thermal shift assay results for *Tc*PYK.
The error bars shown in the figure correspond to the standard deviations for duplicated results.
Control sample: the buffer condition in this sample contains 50 mM TEA (pH7.2), 100 mM KCl and 10 mM $MgCl_2$.
+10 mM PEP sample: the buffer condition here is the same to that of the control sample, but including 10 mM of the cooperative substrate PEP.
+10 mM oxalate sample: the buffer condition here is the same to that of the control sample, but including 10 mM of oxalate, a substrate analogue.
+1.5 μM F26BP sample: the buffer condition here is the same to that of the control sample, but including 1.5 μM of allosteric effector F26BP. The experiment for this sample was only done once because of the lack of a commercial source of F26BP.

5.3.5. Partially closed B-domain in TcPYK/F26BP/OX/Mg compared to the open B-domain in apo TcPYK

(i) The lid-like B-domain has multiple conformations in PYK and its closure movement is initiated by the binding of active-site ligands

The B-domains of PYK are relatively flexible, resulting in poor electron density for structure determination. B-domains without active-site ligand binding normally stay in an open conformation (Mattevi et al. 1995; Morgan et al. 2010b). Active-site ligands are capable of triggering the movement of the B-domain towards the A-domain to make a closed conformation. In the next chapter (**Chapter 6**), ligand soaking experiments are described which were performed to visualise and prove the movement of the B-domain *in crystallo* by active site ligand binding. The fully closed B-domains can only be observed in ATP-bound PYK structures, and the domain rotation angle is around 40^{o} compared to the fully open B-domain structures (Morgan et al. 2010b; Larsen et al. 1998).

(ii) The B-domain adopts a partially closed conformation in *Tc*PYK/F26BP/OX/Mg compared to apo *Tc*PYK where it stays in the fully open conformation

To analyse the B-domain movement between structures of apo *Tc*PYK and *Tc*PYK/F26BP/OX/Mg, a series of superpositions were performed according to the procedures in Materials and Methods. The B-domain rotation angles from the structure of apo *Tc*PYK to that of *Tc*PYK/F26BP/OX/Mg are summarised in **Table 5.6**. The B-domain rotation between these two structures is between 21.6 ^{o} and 35.2^{o}, depending on which two chains are used in the comparison. The difference of B-domain conformations in either apo *Tc*PYK or *Tc*PYK/F26BP/OX/Mg further demonstrates that the B-domain is a relatively flexible domain. The B-domain rotation can be visualised in **Figure 5.20a** and **5.21b**. However, whether the active-site ligand (oxalate in this structure) or effector-site ligand F26BP is responsible for this partial closure of the B-domain is still not clear just based on

these two structures. According to the ligand soaking results in **Chapter 6** and the work of Morgan et al 2010, B-domain movement is related only to active-site ligand binding and is independent from allosterical conformational changes.

(iii) Additional interactions lock the B-domain in its partially closed conformation in *Tc*PYK/F26BP/OX/Mg

To analyse the interactions which lock the B-domain in its partially closed conformation, monomers from apo *Tc*PYK and *Tc*PYK/F26BP/OX/Mg were superposed and the interactions involving the B-domain were examined (**Figure 5.20c**). Residues from the B-domain or on the boundary of A-B domains form interactions with residues from the A-domain to stabilise the partially closed conformation of the B-domain in *Tc*PYK/F26BP/OX/Mg. By contrast, these additional interactions are missing in the open B-domain structure (apo *Tc*PYK). For example, Arg91 from the B-domain forms a salt bridge with residue Glu89 which is on the boundary of the A-B domains. Glu128 from the B-domain makes interactions with the B-domain Lys86 via a water molecule network, although it is not always there. Both of these two major interactions between the A- and B-domains are missing in apo *Tc*PYK where the B-domain stays in its open conformation.

Table 5.6 B-domain rotation angles between the structures of apo *Tc*PYK and *Tc*PYK/F26BP/OX/Mg

		*Tc*PYK/F26BP/OX/Mg	
		Chain A	Chain B
apo *Tc*PYK	Chain A	21.6^{o}	21.6^{o}
	Chain B	35.2^{o}	32.5^{o}

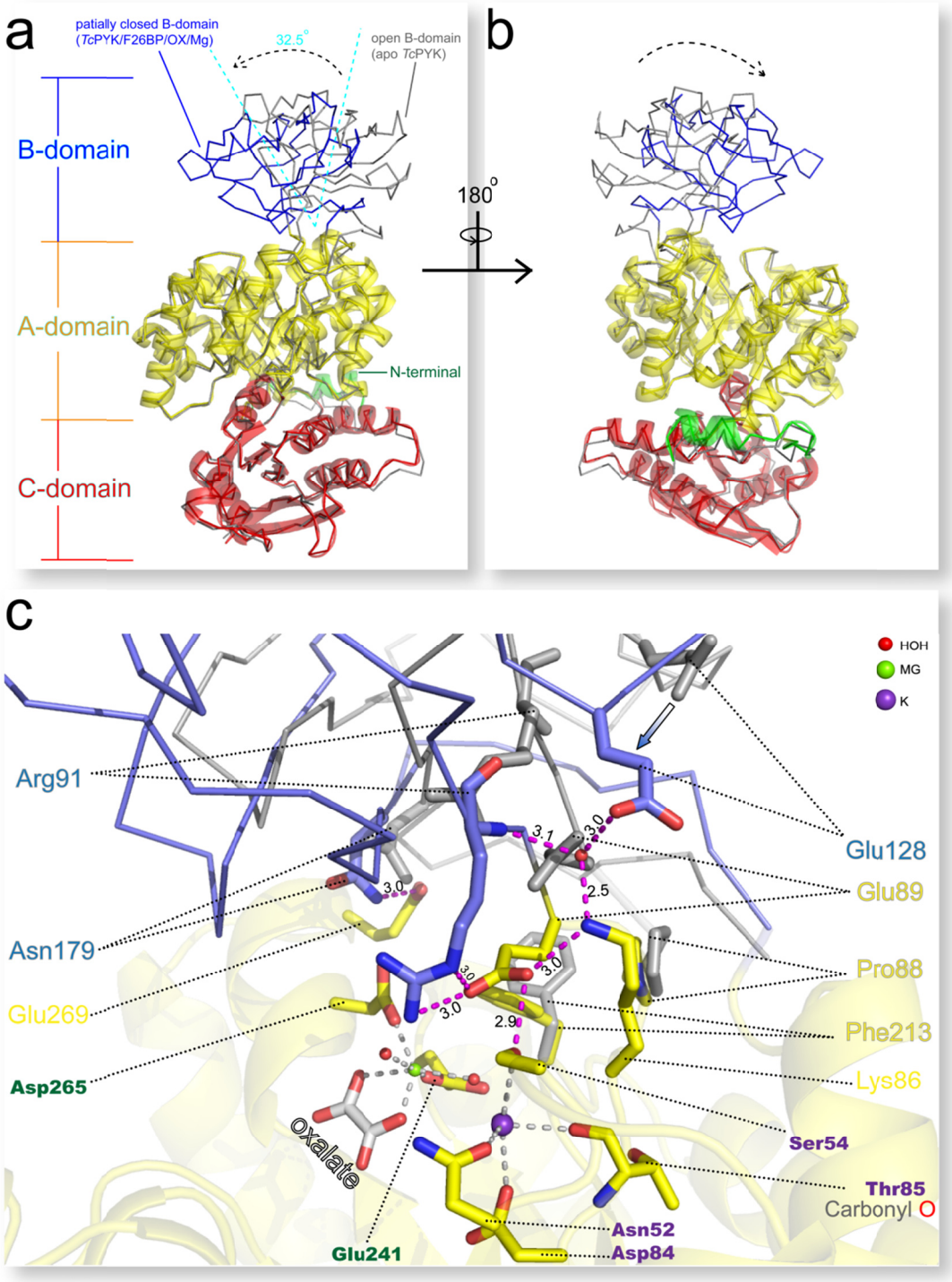

a
patially closed B-domain
(TcPYK/F26BP/OX/Mg)
32.5°
open B-domain
(apo TcPYK)
B-domain
A-domain
C-domain
N-terminal
b
180°
c
HOH
MG
K
Arg91
Asn179
Glu269
Asp265
Oxalate
Glu241
Glu128
Glu89
Pro88
Phe213
Lys86
Ser54
Thr85
Carbonyl O
Asn52
Asp84
3.1
3.0
3.0
2.5
3.0
3.0
3.0
2.9

Figure 5.20 B-domain movement is initiated by oxalate binding in *Tc*PYK/F26BP/OX/Mg compared to apo *Tc*PYK, and is locked in a partially closed conformation by interatomic interactions between the A- and B- domains. (a, b) Two views of the superposed chain B from apo *Tc*PYK and chain B from *Tc*PYK/F26BP/OX/Mg. The superposition was performed on the AC-cores (residues 19-89, 188-499) of the PYK structures in monomeric form. The B-domain from apo *Tc*PYK (chain B) experienced a domain rotation of 32.5^{o} compared to oxalate-bound *Tc*PYK/F26BP/OX/Mg (chain B). Structures of both chains are shown as ribbons, while the A-, C- and N-terminal domains of chain B from *Tc*PYK/F26BP/OX/Mg are shown in cartoon format. Apo *Tc*PYK is coloured in grey for all domains. For *Tc*PYK/F26BP/OX/Mg, the A-domain is coloured yellow, the B-domain is coloured blue, the C-domain is coloured red and the N-terminal domain is coloured green. The B-domain rotation function was analysed by the program CCP4 superpose (Potterton et al. 2002). (c) Close-up of the superposed active sites from apo *Tc*PYK (chain B) and *Tc*PYK/F26BP/OX/Mg (chain B) showing the interactions involved in B-domain movement. The interacting residues and the ligand (oxalate) are shown as sticks while metal ions and water molecules are shown as spheres. The potential interatomic interactions in *Tc*PYK/F26BP/OX/Mg are indicated as dashed lines, where the interactions directly involved in B-domain movement are highlighted by pink dashed lines. The interatomic distances for the interactions are given in Ångstroms. The movement of the B-domain is indicated by arrow.

5.4. Conclusions

To summarise, *Tc*PYK is a homotetramer and each individual chain is composed of four domains (A-domain, B-domain, C-domain and N-terminal domain). *Tc*PYK is an allosteric enzyme and is regulated by the activator F26BP binding to the effector site of the enzyme. Apo *Tc*PYK is in its inactive T-state conformation while the F26BP-bound structure of *Tc*PYK/F26BP/OX/Mg stays in a thermally more stable active R-state conformation. The allosterical conformational change (between T- and R-states) of the *Tc*PYK tetramer encompasses an 8^{o} AC-core rigid-body rotation which can be initiated independently by the effector site ligand F26BP or by the active site ligands PEP or oxalate. A flexible protein motif (residues 262-277) which is potentially important for enzyme allosteric regulation has been identified in the active site. The previously described mobile B-domains have also been observed in *Tc*PYK structures.

The methods applied in *Tc*PYK studies and the corresponding results open a door to study *Tb*PYK (81% protein sequence identity with *Tc*PYK) which has been identified as a drug target to kill *Trypanosoma brucei* for curing the disease of sleeping sickness (**Chapter 1**). In the next chapter, a series of more rounded results and discussions on *Tb*PYK will be presented which potentially favours the design of specific inhibitors against *Tb*PYK.

CHAPTER 6: X-ray Crystal Structures of *Tb*PYK and the Regulatory Role of Magnesium at the Active Site

6.1. Introduction

"PYKs are either constitutively active (as exemplified by the muscle isoenzyme M1 in mammals) or are allosterically regulated (essentially all other forms of PYK). Rabbit muscle M1 PYK has been extensively studied (Larsen et al. 1997; Larsen et al. 1994; Wooll et al. 2001; Williams et al. 2006), and the many mechanistic insights have been derived from structures in complex with various substrate analogues (1AQF, 1A49, 1A5U, 1F3W and 2G50). These structures, however, provide little insight into the role of allosteric regulation by the activator fructose bisphosphate (FBP) or of the cooperative binding of the substrate PEP, neither of which is observed in constitutively active PYKs. In particular, a structural understanding of the communication links between the active site and the 40Å distant effector site has remained incomplete. A high-resolution structure of *Tb*PYK in complex with its allosteric activator F26BP and Mg^{2+} has been determined in this work which is the first 'effector only' structure of an allosterically regulated PYK. This structure, together with the structure with added PEP obtained by crystal soaking, helps to provide a vivid picture of the structural changes during the catalytic cycle of an allosterically regulated PYK."

Work from this chapter forms the basis of a publication:

Zhong, W. *, Morgan, H. P. *, McNae, I. W., Michels, P. A., Fothergill-Gilmore, L. A., Walkinshaw, M. D. *'In crystallo'* substrate binding triggers major domain movements and reveals magnesium as a co-activator of *Trypanosoma brucei* pyruvate kinase. (Manuscript submitted to *Acta Crystallogr. D*) * Co-first authors

Text in this chapter which is directly taken from this manuscript is indicated by quotation marks.

Two X-ray crystal structures determined in this chapter are deposited in the PDB: *Tb*PYK/F26BP/Mg (PDB ID: 4HYW); *Tb*PYK/F26BP/PEP/Mg (PDB ID: 4HYV).

6.2. Materials and Methods

The materials and the general methods used in this chapter have been described in **Chapter 2**. In brief, the structures of *Tb*PYK/F26BP/Mg and *Tb*PYK/F26BP/PEP/Mg were solved by the methods described in **Section 4**; *Tc*PYK/F26BP/OX/Mg for comparison in this chapter was determined by my colleague Dr. Hugh Morgan; the structure analysis including rigid-body rotation, B-domain movement and interface interactions were followed the methods in **Section 5**; the thermal stability of *Tb*PYK in the presence of various ligands were carried out by method described in **Section 3**.

6.3. Results and Discussion

6.3.1. The structures of TbPYK/F26BP/Mg with and without PEP

Previously, the only available crystal structure of any PYK with its physiological activator, but without active-site ligands was the low resolution (5 Å) structure of *Lm*PYK/F26BP (Morgan et al. 2010b). The F26BP-bound *Tb*PYK crystal structure (*Tb*PYK/F26BP/Mg) at a resolution of 2.35 Å with two subunits per asymmetric unit has been determined in the work reported in this thesis (**Table 6.1**). The dye ponceau S was used in the crystallisation conditions as an additive to improve the crystal packing (see **Chapter 4**), as previously observed for other sulfonic acid-containing dyes (Morgan et al. 2011). Without this additive a different crystal form with three tetramers per asymmetric unit diffracted only to 2.95 Å (see **Chapter 4**). As well as the substrate-free structure, a complex of *Tb*PYK/F26BP/PEP/Mg was also obtained by soaking PEP into the *Tb*PYK/F26BP/Mg crystals (**Table 6.1**).

Table 6.1 Data collection, refinement and Ramachandran plot statistics for *Tb*PYK/F26BP/Mg, *Tb*PYK/F26BP/PEP/Mg and *Tb*PYK/F26BP/EDTA/Mg crystals †

	***Tb*PYK/ F26BP/Mg**	***Tb*PYK/ F26BP/PEP/Mg**	***Tb*PYK/ F26BP/EDTA**
Data collection and processing			
Space group	*I*222	*I*222	*I*222
Cell dimensions			
a, *b*, *c* (Å)	103.68,109.00,268.38	103.41,108.93,263.34	103.41,108.89,268.63
α, *β*, *γ* (°)	90.00,90.00,90.00	90.00,90.00,90.00	90.00,90.00,90.00
Solvent content (%)	65	64	64
Wavelength (Å)	0.98	0.98	0.98
Resolution	54.50-2.35	65.83-2.30	74.98-2.40
No. reflections	411424 (60956)	536643 (78473)	407292 (59876)
No. unique reflections	63644 (9232)	66318 (9572)	59635 (8628)
Wilson B-factor (Å^2)	42.2	37.1	38.3
Rmerge (%)	8.8 (54.6)	12.6 (67.0)	13.4 (76.7)
$\langle I/\sigma I \rangle$	12.4 (3.1)	11.3 (3.2)	11.4 (3.6)
Rmeas (%) (within I+/I-)	10.5 (64.8)	14.4 (76.7)	14.5 (82.8)
Rmeas (%) (all I+ & I-)	10.3 (64.2)	14.1 (75.9)	14.5 (82.8)
Rpim (%) (within I+/I-)	5.6 (34.4)	7.0 (37.1)	5.5 (31.0)
Rpim (%) (all I+ & I-)	4.0 (24.6)	5.0 (26.3)	5.5 (31.0)
% completeness	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
Multiplicity	6.5 (6.6)	8.1 (8.2)	6.8 (6.9)
Refinement statistics			
Monomers in ASU	2	2	2
No. reflections	411424 (60417)	536643 (62951)	407292 (56622)
R_{work} / R_{free}	16.38/19.87	14.62/18.84	17.77/21.58
No. of nonhydrogen atoms	8289	8594	7526
Protein	7653	7666	7143
Water	585	857	301
Ligands	51	71	82
Average B-factor (Å^2)			
Overall	50.85	38.06	35.00
Overall (exclude B-domain of chain A)	42.92	34.34	33.46
Protein (chain A and chain B)	51.60	37.56	34.95
Protein (chain A and chain B exclude B-domain of chain A)	43.02	33.27	33.33
B-domain of chain A	130.59	76.11	77.38
B-domain of chain B	56.87	44.02	39.74
Water	42.03	43.07	35.06
Ligands	39.71	32.39	45.44
No. Residues (protein)	994	996	932
No. Ligands	8	10	6
No. Water	585	857	301
R.m.s deviations			
Bond lengths (Å)	0.0104	0.0117	0.0104
Bond angles (degree)	0.9859	1.0976	1.0947
Ramachandran plots			
Favoured (%)	97.2	97.2	97.4
Allowed (%)	99.8	99.8	99.8
No. of outliers[a]	2	2	2

Values in parentheses are for the highest resolution shell.

[a] The outlier residue in each monomer is Thr296, a key active-site residue that is commonly found in this configuration in PYK structures.

† Part of this table (not including the data for EDTA complex) will be submitted for publication (see section **6.1**).

6.3.2 Overall structures of TbPYK

The *Tb*PYK/F26BP/Mg and *Tb*PYK/F26BP/Mg/PEP X-ray structures show the expected homotetrameric architecture (**Figure 6.1a**) which is similar to the published *Lm*PYK R-state structure (Morgan et al. 2010b) and *Tc*PYK R-state structure *Tc*PYK/F26BP/OX/Mg determined by my colleague Dr. Hugh Morgan (see **Chapter 5**). The average C^{α} RMS differences between R-state tetrameric structures of *Tb*PYK, *Lm*PYK and *Tc*PYK are less than 1 Å for the AC-cores, while the average C^{α} RMS differences between *Tb*PYK (R-state) and T-state structures of *Lm*PYK and *Tc*PYK are more than 2.5 Å for the AC-cores (**Table 6.2**).

Each chain of *Tb*PYK is composed of four domains which are similar to *Lm*PYK and *Tc*PYK (**Figure 6.1a**). Refer to section **5.3.1** in **Chapter 5** for more detailed descriptions on the trypanosomatid PYK structures. The secondary structure of *Tb*PYK is shown in **Figure 6.2** (only α-helices and β-strands are shown), which is not affected by binding of the substrate PEP.

The B-domains of the two crystallographically independent chains of the *Tb*PYK/F26BP/Mg structure show closely similar orientations, however there is a significant difference in temperature factors with average values of 130 and 56 $Å^2$, respectively for the B-domains of chains A and B compared to an average B-factor of 51 $Å^2$ for the tetramer (**Table 6.1**). Interestingly, PEP was soaked into crystals of *Tb*PYK/F26BP/Mg resulting in different orientations of the B-domains in each asymmetric unit which will be discussed in section **6.3.5**. The PYK structures for comparison in figures throughout this chapter have been summarised in **Table 6.3**.

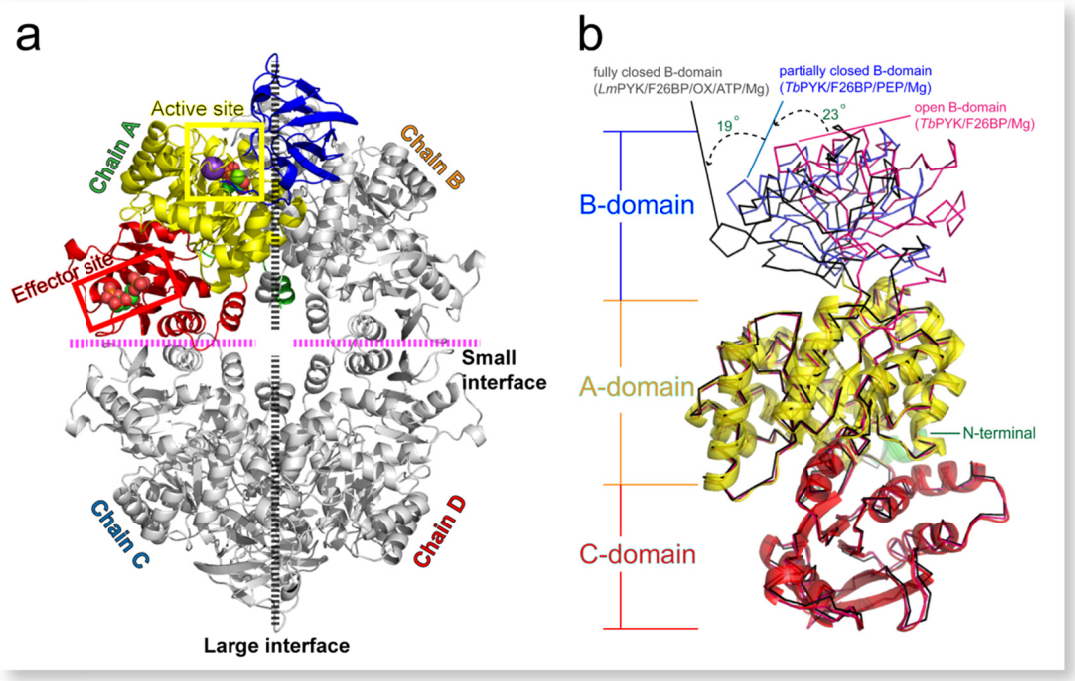


Figure 6.1 Tetrameric structure of *Tb*PYK and the disposition of domains. †
(a) The four polypeptide chains of *Tb*PYK/F26BP/PEP/Mg are represented by cartoons, with ligands and metal ions shown as spheres. The large (A-A) and small (C-C) interfaces are indicated by dashed lines, and the active and effector sites are boxed. One chain (chain A) has been coloured to aid the identification of domains: N-terminal domain (green = residues 2-18), A-domain (yellow = residues 19-89, 188-358), B-domain (blue = residues 90-187), C-domain (red = residues 359-499). (b) The relatively static AC-domains and the multiple positions of the B-domains of subunits from three different PYK crystal structures are shown. The superposed structures are: *Tb*PYK/F26BP/Mg (pink backbone, chain B), *Tb*PYK/F26BP/PEP/Mg (blue backbone, chain A) and *Lm*PYK/F26BP/OX/ATP/Mg (black backbone, chain K, PDB code: 3HQP). All structures are shown as ribbons with the regular secondary structures of the A-, C- and N-terminal domains of *Tb*PYK/F26BP/PEP/Mg indicated by cartoons coloured yellow, red and green, respectively. The B-domain rotations of these PYK complexes are indicated by arrows and rotation angles. The B-domain in *Tb*PYK/F26BP/Mg (chain B) is in the open conformation, while the B-domain in *Tb*PYK/F26BP/PEP/Mg is in the partially or half-closed conformation. The fully ligand-bound structure of *Lm*PYK/F26BP/OX/ATP/Mg has the B-domain in its fully closed position. The domain rotation function was analysed by the program DynDom (Hayward et al. 1997; Hayward and Berendsen 1998). Refer to **Table 6.3** for the corresponding B-domain positions. This figure will be submitted for publication (see section **6.1**).

Table 6.2 Average RMS difference (Å) of PYK tetrameric structures (AC-cores)†

	*Tb*PYK/ F26BP/Mg	*Tb*PYK/ F26BP/PEP/Mg	*Tb*PYK/ F26BP/EDTA
*Tb*PYK/F26BP/Mg	0	0.20	0.13
*Tb*PYK/F26BP/PEP/Mg	0.20	0	0.18
*Lm*PYK/F26BP/OX/ATP/Mg (3HQP)	0.82	0.87	0.60
apo *Lm*PYK (3HQN)	2.54	2.52	2.36
*Tc*PYK/F26BP/OX/Mg	0.53	0.60	0.35
apo *Tc*PYK	2.51	2.52	2.21

† Part of this table (not including data for *Tc*PYK and EDTA complex) will be submitted for publication (see section **6.1**).

Table 6.3 Summary of B-domain positions in trypanosomatid PYK structures†

PYK structure	PDB ID	State	Chains in AS	B domains in tetramer
*Tb*PYK/F26BP/Mg	4HYW	R	2	chains A&D - open chains B&C - open
*Tb*PYK/F26BP/PEP/Mg	4HYV	R	2	chains A&D - partly closed chains B&C - open
*Tb*PYK/F26BP/EDTA	n.a.	R	2	chains A&D - partly closed chains B&C - open
*Lm*PYK/F26BP/OX/ATP	3HQP	R	16	all chains - closed
apo *Lm*PYK	3HQN	T	2	chains A&D - open chains B&C - open

Abbreviation: AS, asymmetric unit

† Part of this table (not including data for apo *Lm*PYK and EDTA complex) will be submitted for publication (see section **6.1**).

n.a., not applicable

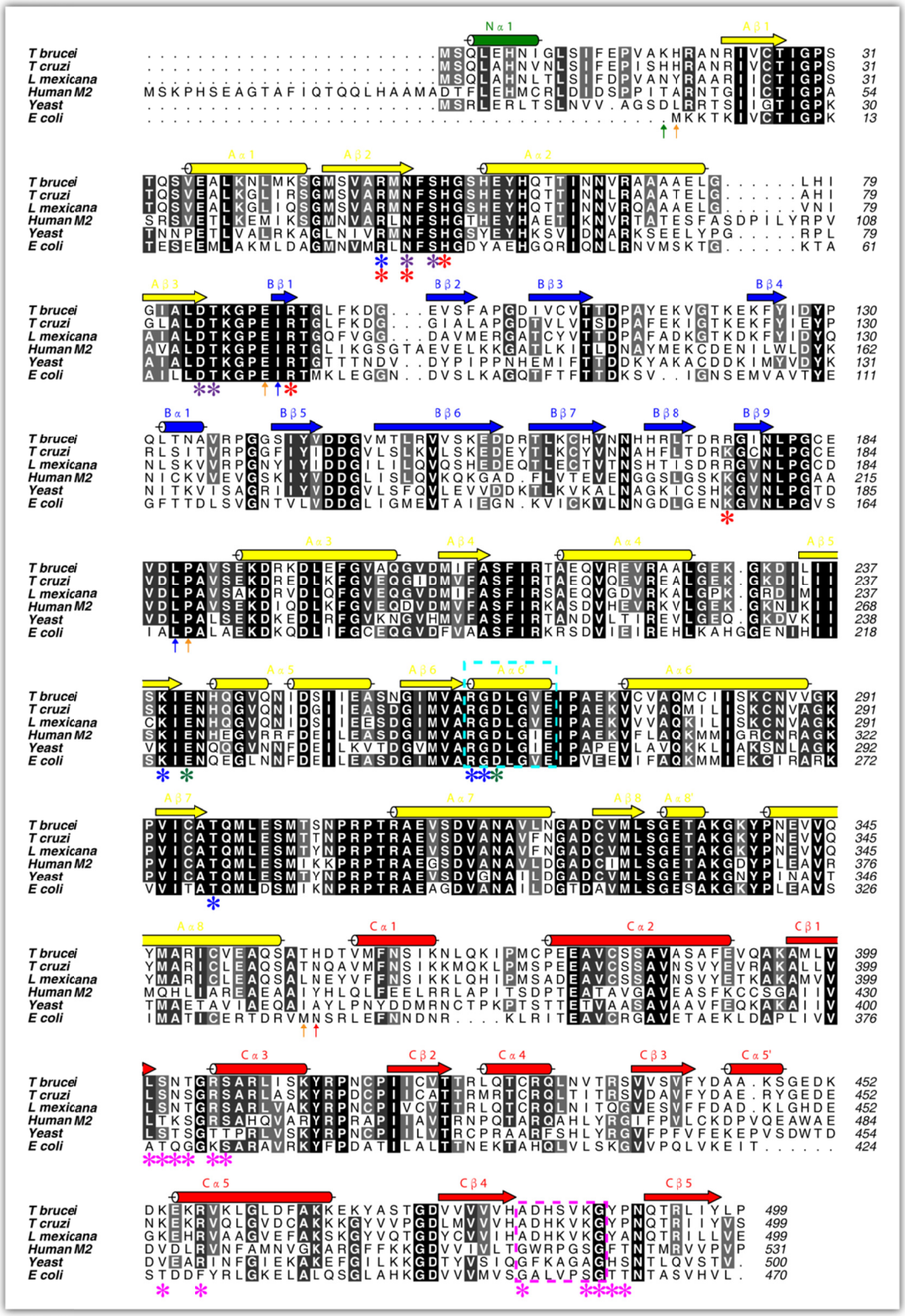

T brucei
T cruzi
L mexicana
Human M2
Yeast
E coli

Figure 6.2 Sequence alignment of pyruvate kinases from *T. brucei, T. cruzi, L. mexicana, Homo sapiens* M2, *S. cerevisiae* and *E. coli.* †

The sequence alignment was performed using the program Clustal Omega at the European Bioinformatics Institute (Goujon et al. 2010; Sievers et al. 2011). Secondary structural elements defined in *Tb*PYK/F26BP/Mg by DSSP (Kabsch and Sander 1983; Joosten et al. 2011) are shown above the sequences (only α-helices and β-strands are shown). Secondary structural elements are labelled in different colours corresponding to their domain regions: N-terminal domain (green), A-domain (yellow), B-domain (blue) and C-domain (red). Domain boundaries are indicated by vertical arrows in domain-specific colours. The conservation of the residues is indicated by shading from black (identical in six sequences) to dark grey (conserved in five) to light grey (conserved in four) to white (low or no conservation). Residue numbers corresponding to each PYK are listed after the sequences. In *Tb*PYK, the amino acids involved in divalent metal binding (PEP-coordinating metal, Mg-1 site) (*), potassium metal ion binding (*), substrate PEP binding (*) and effector F26BP binding (*) are indicated by asterisks. The red asterisks (*) indicate product ATP binding residues in *Lm*PYK. Residues 263-269 of the small α-helix Aα6′ which are involved in allosteric regulation and in binding divalent metal and the substrate PEP are indicated by a dashed box (cyan). The effector loop residues are indicated by a pink dashed box. The figure was generated using the program Aline (Bond and Schüttelkopf 2009).

† This figure will be submitted for publication (see section **6.1**).

6.3.3. The effector binding site of TbPYK

F26BP has been clearly identified in both crystallographically independent effector sites of *Tb*PYK/F26BP/Mg and *Tb*PYK/F26BP/PEP/Mg, and possesses almost identical binding modes. Similar to *Tc*PYK which has been described in **Chapter 5**, F26BP binds to a highly positively charged effector site in *Tb*PYK as well (**Figure 6.3**).

The F26BP effector molecule is held in place by a network of salt bridges and hydrogen bonds (**Figure 6.4b** and **6.4c**). For both crystallographically independent F26BP binding sites, the 2′-phospho group interacts with residues Lys454 and Arg457 with an additional direct hydrogen bond to the side chain of Asn402. The 6′-phospho group makes direct hydrogen bonds to the side chains and backbones of Ser401, Asn402, Thr403, Arg405 and Ser406. The conformation of Arg405 is relatively flexible and not the same even between asymmetrically otherwise unique chains. Additionally, some interactions are mediated by water molecule networks involving Asn402, Glu450, Lys454, Lys487 and Gly488.

Those residues involved in F26BP binding are highly conserved between trypanosomatid PYKs, but show much lower conservation compared with the mammalian, yeast and bacterial PYKs that are activated by F16BP (**Figure 6.4a**). Similar to the R-state structure of *Tc*PYK (*Tc*PYK/F26BP/OX/Mg) described in **Chapter 5**, there is a conserved salt bridge (Asp483…Arg494) formed across the C-C interface as a result of the loop stabilisation by the binding of F26BP, which helps to lock the enzyme in its active state. The formation of additional salt bridges will be discussed in section **6.3.7** (**Figure 6.18**). Despite considerable efforts from various laboratories, no well-diffracting crystals of a complex of FBP with pyruvate kinase have previously been obtained without additional substrate analogue or nucleotide binding at the active site. The *Tb*PYK/F26BP/Mg structure presented here shows for the first time that the presence of FBP and Mg^{2+} alone are enough to stabilise the tetramer in an active R-state conformation.

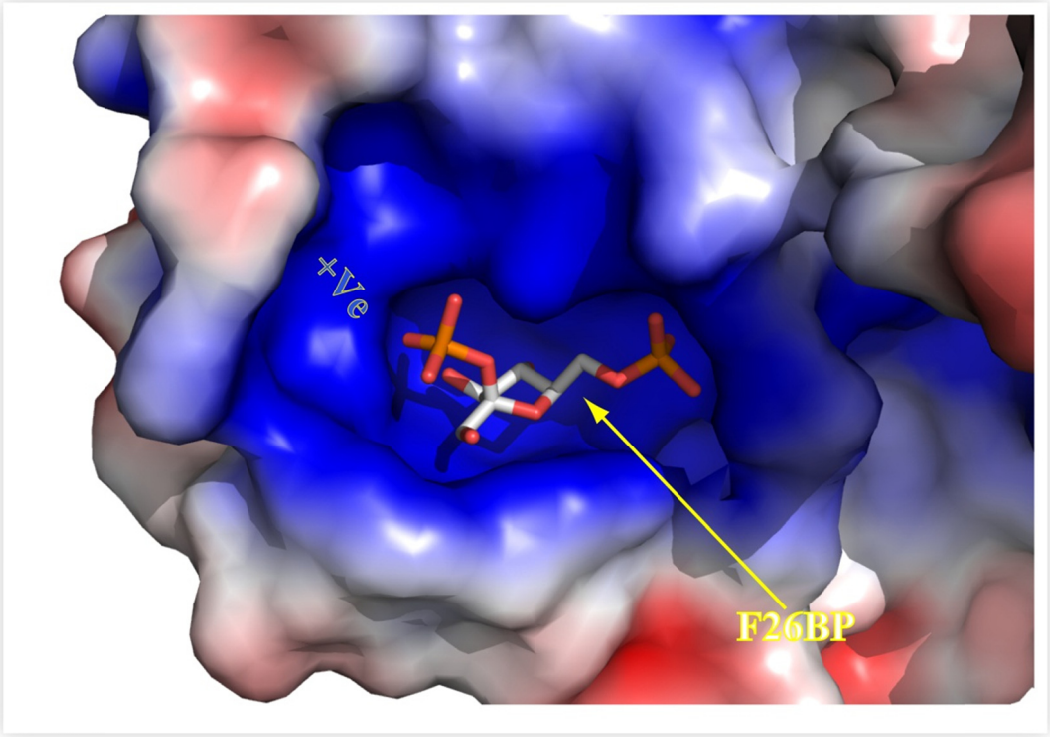
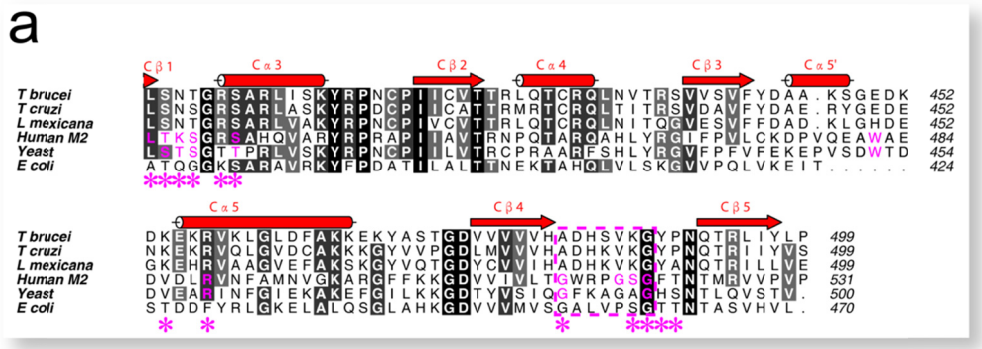
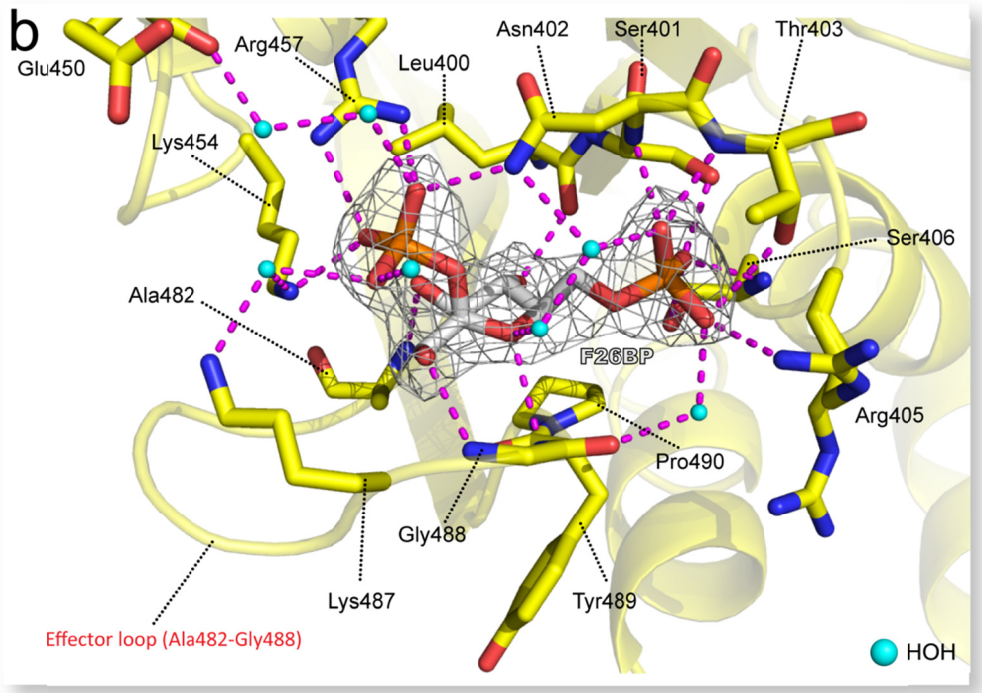


Figure 6.3 The electrostatic surface of *Tb*PYK/F26BP/Mg showing the effctor F26BP binding site of *Tb*PYK.

The electrostatic potentials were calculated by programs of PDB2PQR and ABPS (Baker et al. 2001; Dolinsky et al. 2004; Dolinsky et al. 2007) and visualized by PyMol (Delano 2002). The areas of positive charge (+Ve) are represented by blue, are essential for binding the negatively charged phospho groups of F26BP. Areas of negative charge and the hydrophobic areas are shown in red and white, respectively.


a
Cβ1
Cα3
Cβ2
Cα4
Cβ3
Cα5'
T brucei
T cruzi
L mexicana
Human M2
Yeast
E coli
452
452
452
484
454
424
Cα5
Cβ4
Cβ5
499
499
499
531
500
470


b
Glu450
Arg457
Leu400
Asn402
Ser401
Thr403
Lys454
Ser406
Ala482
F26BP
Arg405
Pro490
Gly488
Lys487
Tyr489
Effector loop (Ala482-Gly488)
HOH

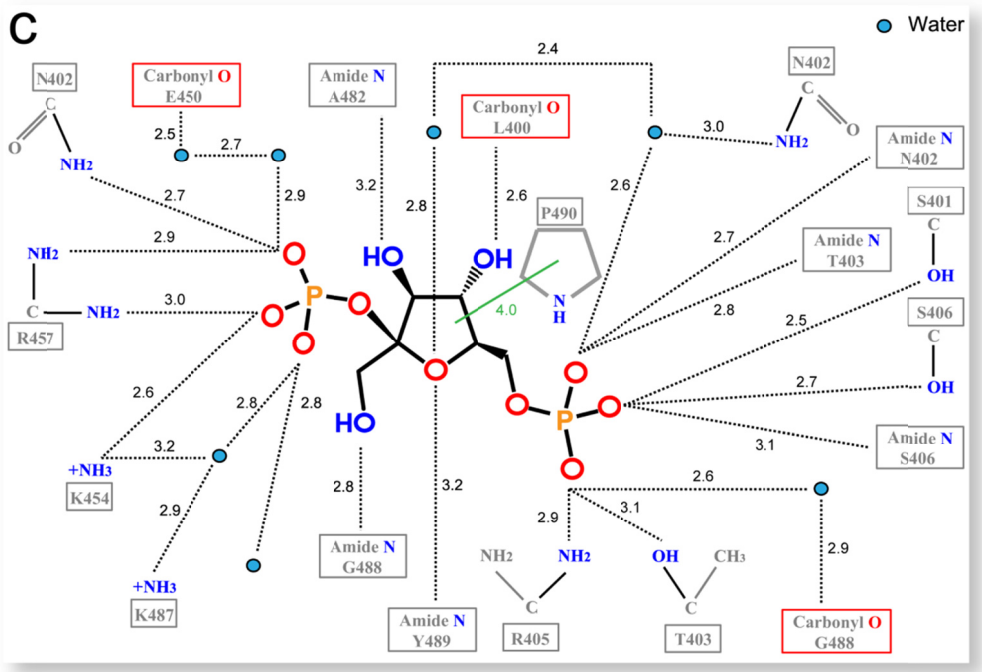

c
Water
N402
Carbonyl O E450
Amide N A482
Carbonyl O L400
N402
Amide N N402
S401
Amide N T403
S406
Amide N S406
P490
R457
K454
K487
Amide N G488
Amide N Y489
R405
T403
Carbonyl O G488

Figure 6.4 The effector site of *Tb*PYK with bound F26BP. †

(a) Sequence alignment of pyruvate kinases from *T. brucei, T. cruzi, L. mexicana,* human M2, *S. cerevisiae* and *E. coli* (starting from residue 400 of *Tb*PYK). The sequence alignment was performed using the program Clustal Omega (Sievers et al 2011; Goujon et al. 2010). Secondary structural elements defined in *Tb*PYK/F26BP/Mg by DSSP (Kabsch and Sander 1983; Joosten et al. 2011) are shown above the sequences (only α-helices and β-strands are shown). The conservation of the residues is indicated by shading from black (identical in six sequences) to dark grey (conserved in five) to light grey (conserved in four) to white (low or no conservation). Residue numbers corresponding to each PYK are listed after the sequences. In *Tb*PYK, the amino acids involved in effector F26BP binding (*) are indicated by asterisks. The amino acids involved in effector F16BP binding for human M2PYK and yeast PYK are coloured pink. The effector loop residues are indicated by a pink dashed box. (b) Close-up of the effector site of *Tb*PYK/F26BP/Mg, chain B showing the binding mode of the activator F26BP. The polypeptide chain is shown in cartoon format, with residues involved in F26BP binding shown as sticks. The F26BP molecule is shown with an unbiased *Fo-Fc* election density map contoured at 5.0 σ (grey). Water molecules are shown as cyan spheres. Interactions (within 3.2 Å) are indicated by dashed lines in pink. The position of the stabilised effector loop (Ala482-Gly488) is indicated as well. (c) Schematic drawing showing the interactions at the *Tb*PYK effector site. Water molecules are shown as cyan spheres. The interatomic distances for the interactions are given in Ångstroms.

† Part of this figure (not including the data for *Tc*PYK) will be submitted for publication (see section **6.1**).

6.3.4. The active site of TbPYK

The unique active site of the first 'effector only' structure *Tb*PYK/F26BP/Mg is described in this section, and comparison with the first PEP-bound PYK structure *Tb*PYK/F26BP/PEP/Mg reveals a series of '*in crystallo*' concerted movements at/around the active site which are involved in the catalytic mechanism.

6.3.4.1. A third Mg^{2+} binding site is revealed adjacent to the active site of R-state *Tb*PYK in the absence of substrates

Divalent metal ions are essential for phospho transfer during the catalytic cycle of PYKs, where Mg^{2+} is generally considered to be the physiologically relevant metal. **Figure 6.5a** shows the positions of the two canonical Mg^{2+} ions (Mg-1 binds PEP, and Mg-2 binds to ATP by tridentate interactions) found in PYK structures crystallised with a variety of non-productive ligand mixtures (Larsen et al. 1997; Morgan et al. 2010b; Larsen et al. 1994; Wooll et al. 2001; Larsen et al. 1998; Dombrauckas et al. 2005; Christofk et al. 2008b; Fenton et al. 2010; Jurica et al. 1998; Valentini et al. 2002). The Mg-2 position is only observed in structures that contain ATP (Morgan et al. 2010b; Larsen et al. 1998). Divalent metal ions have never been found in T-state PYK structures, although divalent metal ions at relatively high concentrations were present in the buffers from which the PYKs were crystallised (Morgan et al. 2010b; Rigden et al. 1999; Morgan et al. 2011; Mattevi et al. 1995).

The crystal structure of *Tb*PYK/F26BP/Mg reported here shows for the first time the presence of Mg^{2+} (Mg-3) adjacent to the active site of R-state PYK in the absence of substrates or substrate analogues (**Figure 6.5a** and **6.5b**). Interestingly, this Mg-3 binding site is distinct from both of the previously described canonical positions and lies ~3 Å away from the canonical Mg-1 position (**Figure 6.5a**). It is nevertheless coordinated by the same side-chain carboxyl groups of residues Glu241

and Asp265. Distorted octahedral coordination geometry around this Mg^{2+} is formed by interactions with the two negatively charged side chains and four water molecules (**Figure 6.5b** and **Figure 6.6**). The coordinating distances of the Mg^{2+} at the Mg-3 binding site of *Tb*PYK/F26BP/Mg range from 2.0-2.3 Å (**Figure 6.6b**). The small α-helix Aα6′ which is usually coordinated with substrate PEP or its analogue, forms interactions with hydroxyl oxygen of Thr197 and a water molecule in *Tb*PYK/F26BP/Mg (**Figure 6.5b**).

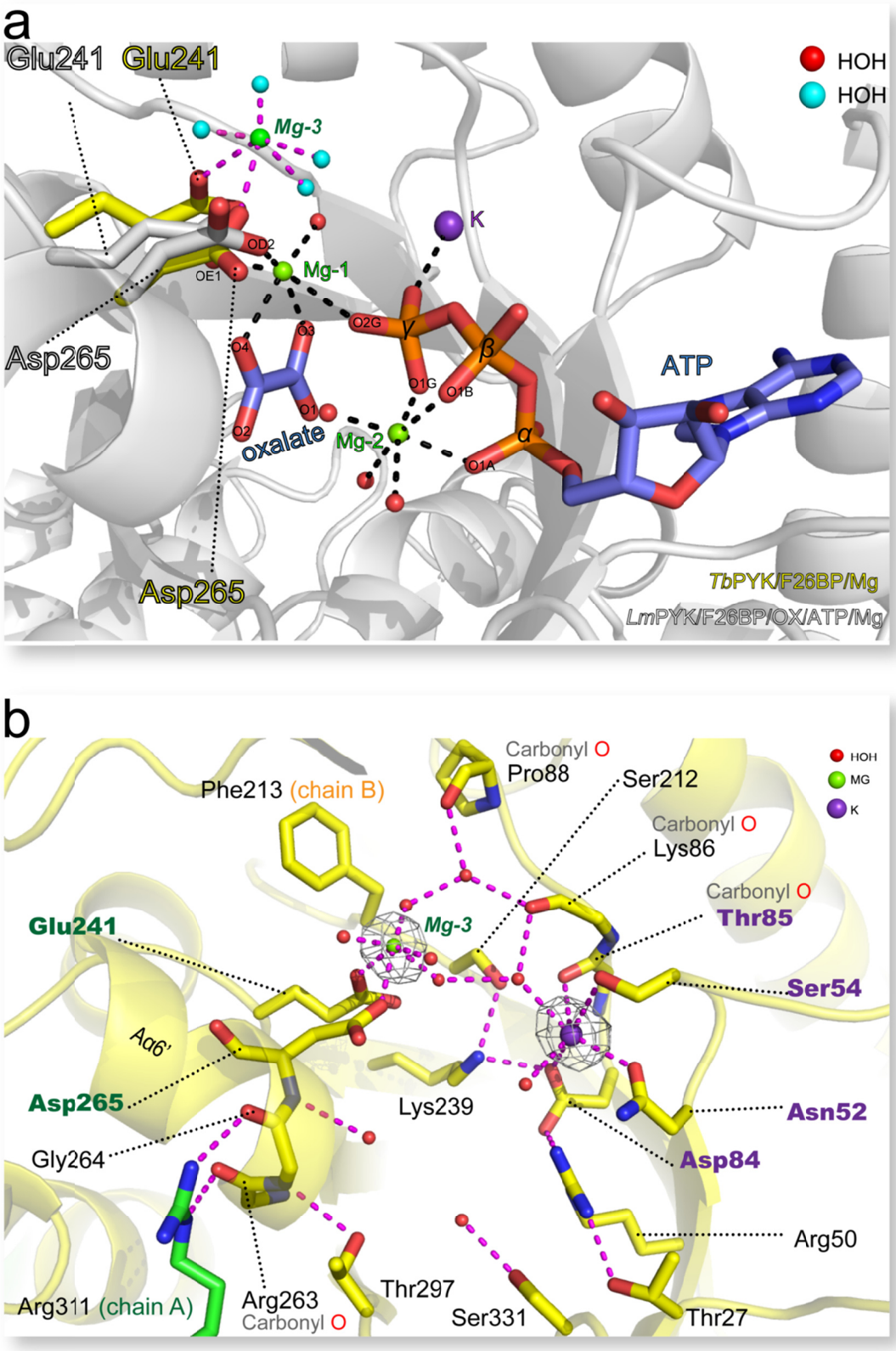


Figure 6.5 Comparison of the active sites of *Lm*PYK/F26BP/OX/ATP/Mg and *Tb*PYK/F26BP/Mg. †

(a) Close-up of the *Lm*PYK/F26BP/OX/ATP/Mg (chain C; PDB code: 3HQP)

active site showing two Mg^{2+} binding sites (Mg-1 and Mg-2) and the binding modes of the substrate analogue oxalate and product ATP. The polypeptide chain of *Lm*PYK is shown in cartoon format, and coloured white. The two Mg-coordinating residues (Glu241, Asp265), and oxalate and ATP are shown as sticks. The interatomic interactions are shown as black dashed lines. In this model the canonical divalent metal binding site, Mg-1 coordinates with two enzyme residues Glu241 (OE1) and Asp265 (OD2), the oxalate molecule (O3 and O4), ATP (O2G) and one water molecule (red sphere). The second canonical divalent metal binding site, Mg-2 coordinates with the α-, β-, and γ- phospho groups of ATP and three water molecules (red spheres). Both divalent metals have distorted octahedral coordination spheres. After superposition of *Tb*PYK/F26BP/Mg (chain B) active site on the *Lm*PYK structure (chain C) active site, only the unique non-canonical divalent metal binding site (Mg-3) was observed, together with altered positions of the coordinating residues Glu241 and Asp265 (shown as sticks in yellow) and coordinating water molecules (cyan spheres). The interactions between the non-canonical divalent metal (shown as a green sphere) and two residues (Glu241 and Asp265) as well as four water molecules (cyan spheres) are shown as pink dashed lines. (b) Close-up of the active site of *Tb*PYK/F26BP/Mg with the polypeptide chain shown as a cartoon, with chain B coloured yellow and chain A green. Interacting residues are shown as sticks, as is Phe213 which potentially favours the binding of the substrate PEP in its correct position. The short α-helix Aα6′ (residues Arg263-Glu269) within the A-domain forms two hydrogen bonds with the side chain of residue Arg311 from the neighbouring subunit within the same tetramer. The divalent metal ion Mg^{2+}, monovalent ion K^{+} and water molecules are shown as green, purple and red spheres, respectively. The *Fo-Fc* electron densities for Mg^{2+} and K^{+} metal ions are shown as grey meshes contoured at 4.0 σ. Interatomic interactions (within 3.3 Å) are indicated by pink dashed lines. Refer to **Table 6.3** for the corresponding structures.

† This figure and its legend will be submitted for publication (see section **6.1**).

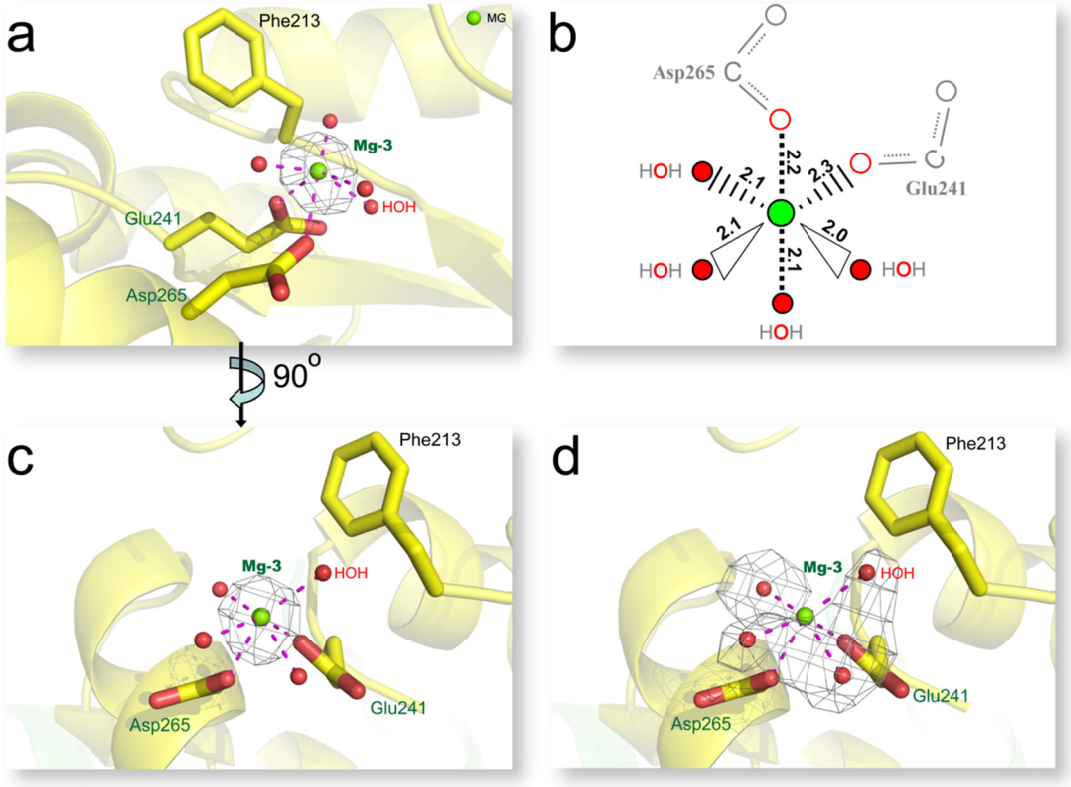


Figure 6.6 Close-up of the Mg-coordinating site Mg-3 adjacent to the active site of *Tb*PYK/F26BP/Mg (chain B). †

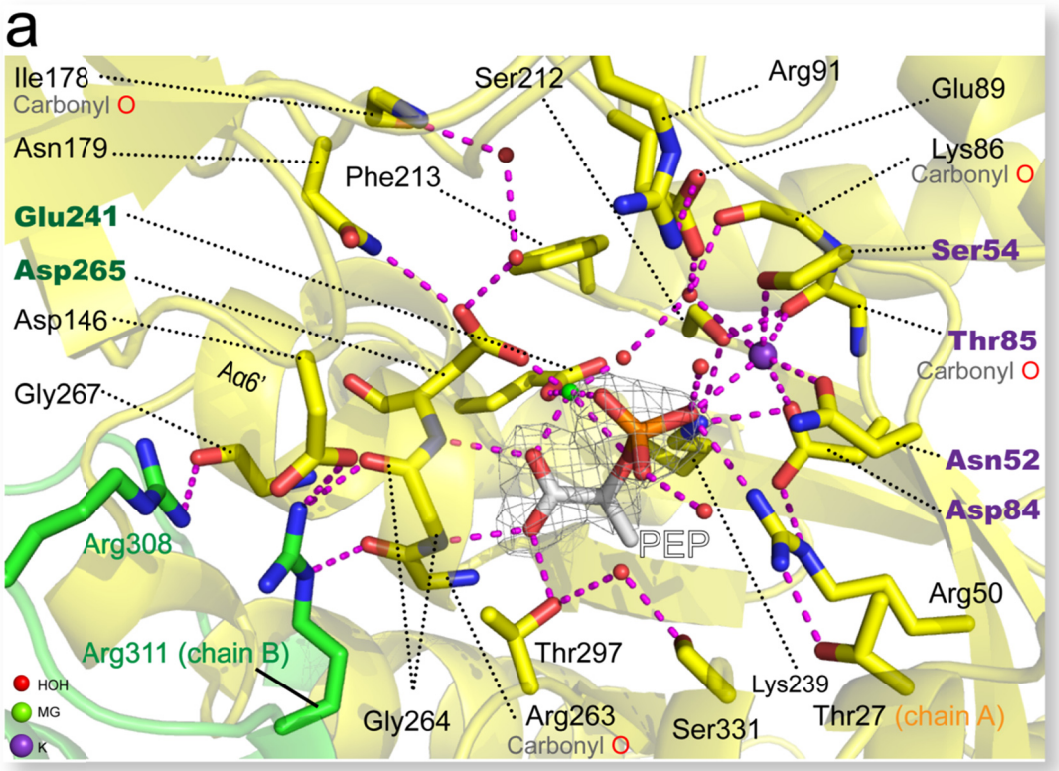
(a, c) Orthogonal views of the Mg-coordinating site, Mg-3. The polypeptide chain is shown as a cartoon while Mg-coordinating residues and Phe213 are shown as sticks. The divalent metal ion Mg^{2+} and water molecules are shown as green spheres and red spheres, respectively. The *Fo-Fc* electron density for Mg^{2+} is shown as a grey mesh contoured at 4.0 σ. Interatomic interactions are indicated by pink dashed lines. (b) Schematic representation of the coordination sphere of the Mg^{2+}. The interatomic distances for the interactions are given in Ångstroms. (d) The *Fo-Fc* electron density for water molecules involved in Mg-coordination is shown as a grey mesh contoured at 3.5 σ.

† This figure and its legend will be submitted for publication (see section **6.1**).

6.3.4.2. Protein-ligand interactions at the PEP binding site

PEP binds in a similar position at the active site (**Figure 6.7**) as the previously reported substrate analogues including oxalate (Morgan et al. 2010b; Larsen et al. 1998; Dombrauckas et al. 2005; Christofk et al. 2008b), pyruvate (Larsen et al. 1994; Wooll et al. 2001; Williams et al. 2006; Fenton et al. 2010), phospholactate (Larsen et al. 1997) and phosphoglycollate (Jurica et al. 1998; Valentini et al. 2002). The divalent metal is coordinated by the phospho and carboxyl groups of PEP and by two carboxyl groups from residues Asp265 and Glu241 (**Figure 6.7; Figure 6.8a** and **6.8b)**.

In the structure of *Tb*PYK/F26BP/PEP/Mg, the phospho group of PEP coordinates directly with both the Mg^{2+} and K^{+} metal ions (**Figure 6.7a**). The octahedral coordination geometries of Mg^{2+} in the two crystallographically independent chains of *Tb*PYK/F26BP/PEP/Mg are shown in **Figure 6.8a** and **6.8b**. The two Mg^{2+} coordination spheres have slight differences which may be related to the conformation of the B-domain and the side-chain orientation of Phe213. The monovalent K^{+} forms a distorted octahedral coordination sphere which is similar to that of the structure of the PEP-free complex (*Tb*PYK/F26BP/Mg), but one coordinating water molecule is replaced by the oxygen of the phospho group of PEP (**Figure 6.7**). The carboxyl group of PEP forms interactions with the small α-helix Aα6′ and with the side chain of residue Thr297. An overlay of the structure of rabbit muscle M1 PYK complexed with L-phospholactate (Larsen et al. 1997) with *Tb*PYK/F26BP/PEP/Mg shows reasonable (~2Å) correspondence of the carboxyl groups of PEP and L-phospholactate, but a closer (~1Å) fit of the phospho groups. In both structures the phospho group interacts with Arg50 and Lys239 (Arg72 and Lys269 in rabbit muscle M1 PYK), which are conserved even among diverse species (**Figure 6.2**), and may play an essential role in phospho transfer.


a
Ile178
Carbonyl O
Asn179
Glu241
Asp265
Asp146
Gly267
Arg308
Arg311 (chain B)
HOH
MG
K
Ser212
Arg91
Glu89
Phe213
Lys86
Carbonyl O
Ser54
Thr85
Carbonyl O
Asn52
Asp84
Aα6'
PEP
Arg50
Thr297
Lys239
Gly264
Arg263
Carbonyl O
Ser331
Thr27 (chain A)

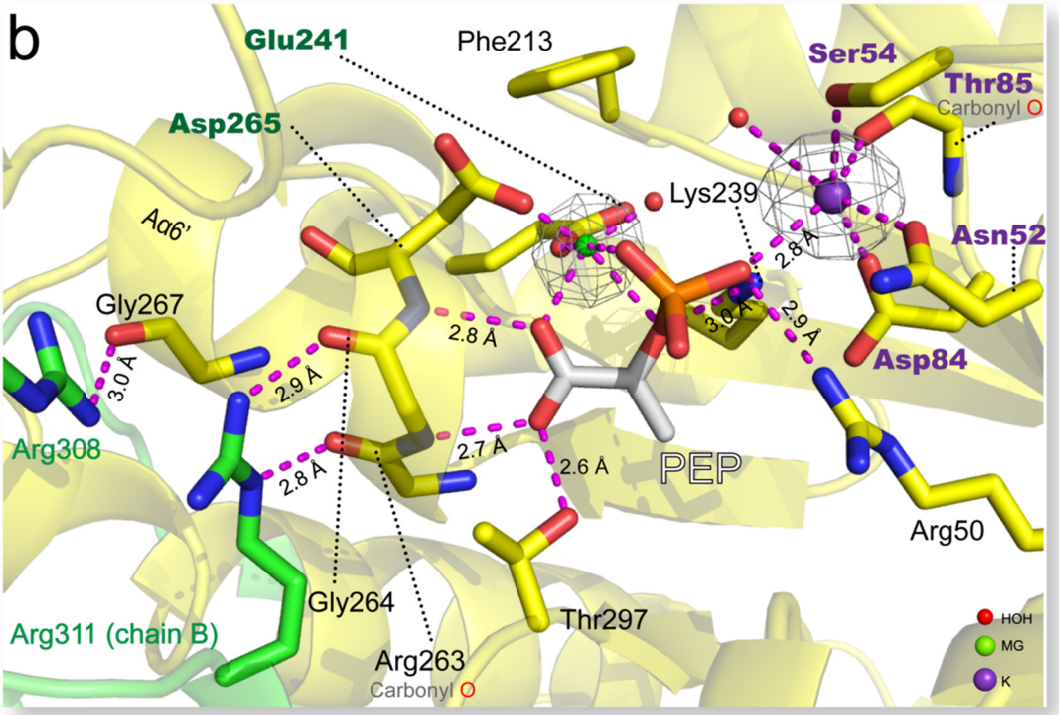

b
Glu241
Phe213
Ser54
Thr85
Carbonyl O
Asp265
Aα6'
Lys239
Asn52
Gly267
3.0 Å
2.9 Å
2.8 Å
2.8 Å
3.0 Å
2.9 Å
Asp84
Arg308
2.8 Å
2.7 Å
2.6 Å
PEP
Arg50
Gly264
Thr297
Arg311 (chain B)
Arg263
Carbonyl O
HOH
MG
K

Figure 6.7 The active site of *Tb*PYK with bound PEP. †

(a) The polypeptide chain at the active site of *Tb*PYK/F26BP/PEP/Mg is shown in cartoon format, with ligands and interacting residues shown as sticks. Residue Phe213 which potentially favours the binding of the substrate PEP in its correct position is also shown. Chain A of the PYK tetramer is coloured yellow and chain B green. The divalent metal ion Mg^{2+} is shown as a green sphere, and is coordinated by the two residues labelled in dark green. The monovalent ion K^{+}, is a purple sphere, and is coordinated by the residues labelled in purple. Water molecules are shown as red spheres. The *Fo-Fc* electron density for the substrate PEP molecule is shown as a grey mesh contoured at 5.0 σ. Interatomic interactions are indicated by pink dashed lines. (b) Close-up of the active site to highlight the metal coordination in the complex with PEP. The colour scheme is the same as in panel (a). The *Fo-Fc* electron densities for the Mg^{2+} and K^{+} metal ions are shown as a grey mesh contoured at 6.0 σ. Interatomic interactions are indicated by pink dashed lines and the corresponding distances are given in Ångstroms.

† This figure and its legend will be submitted for publication (see section **6.1**).

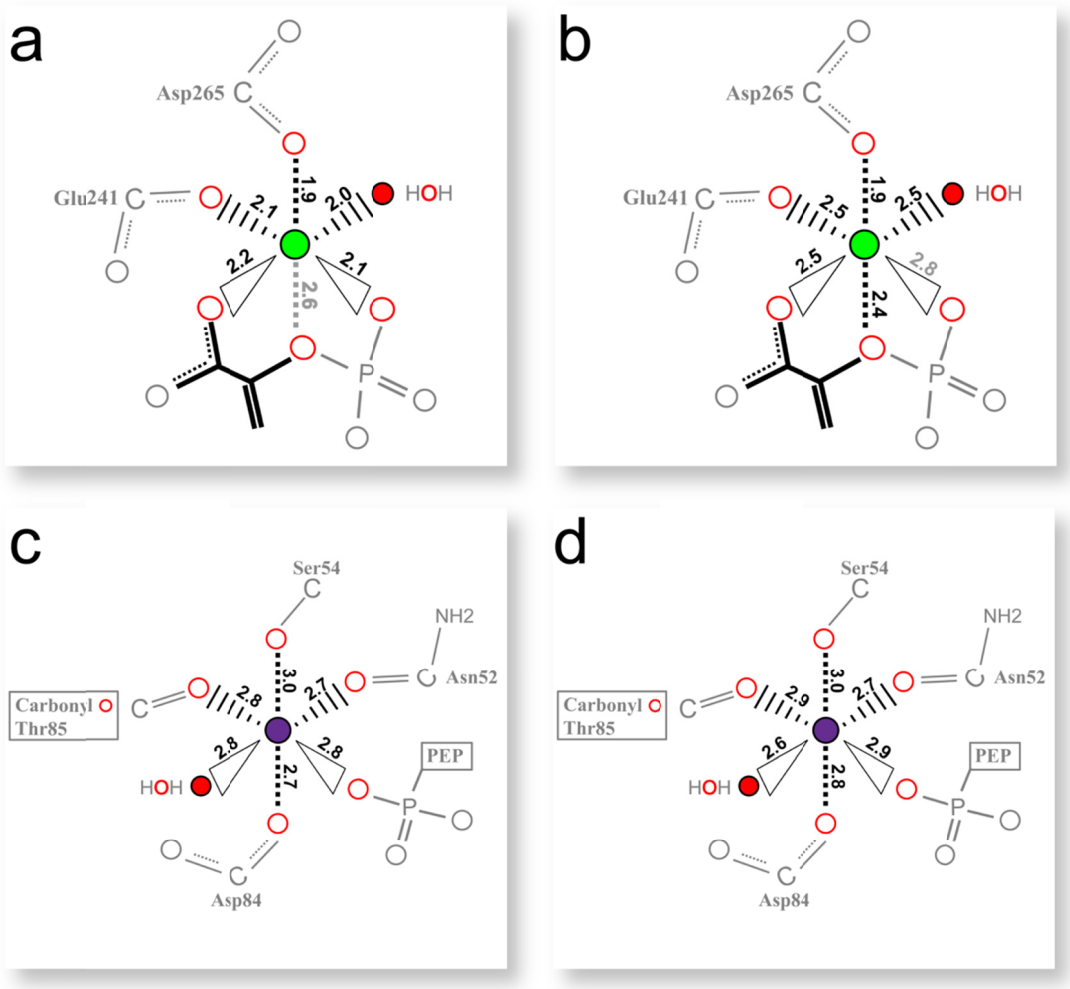


Figure 6.8 Schematic representations of metal ion coordination at the active site of *Tb*PYK/F26BP/PEP/Mg. †

(a, b) Mg^{2+} (green spheres) coordination in chain A and chain B, respectively. The interatomic distances for the interactions are given in Ångstroms. (c, d) K^{+} (purple spheres) coordination in chain A and chain B, respectively. As in panels a and b, the interatomic distances for the interactions are given in Ångstroms. Refer to **Table 6.3** for the corresponding B-domain positions.

† This figure will be submitted for publication (see section **6.1**).

6.3.5. 'In crystallo' substrate binding triggers major domain movements and reveals magnesium as a co-activator of TbPYK

6.3.5.1. Crystal soaking with substrate induces a large B-domain movement

The two structures *Tb*PYK/F26BP/Mg and *Tb*PYK/F26BP/PEP/Mg described here provide a unique example of large scale *'in-crystallo'* domain movements induced by soaking the substrate into the crystal. The program DynDom (Hayward et al. 1997; Hayward and Berendsen 1998) was used to analyse these movements, and **Figure 6.1b** shows the 23° rotation of the B-domain with respect to the AC-core induced by PEP binding. The domain rotation angle has also been calculated by a few steps of CCP4 superpose which has been described in **Chapter 2**, resulting in the same 23° rotation.

It is apparent that in the presence of F26BP, 5 mM PEP is sufficient to cause partial domain closure. Only one of the two crystallographically independent B-domains shows this movement, as crystal packing interactions lock the other B-domain in an open conformation. The conformation of the B-domain observed in chain A can be classified as half-closed. Fully closed B-domain conformations (**Figure 6.1b**) differ in rotation angle by about 40^{o} compared to open-form structures (Morgan et al. 2010b; Larsen et al. 1998) and have only been observed in PYK structures with bound ATP.

6.3.5.2. Changes in Mg^{2+} coordination help trigger closure of the B-domain

Superposition of the active sites from the two structures *Tb*PYK/F26BP/Mg and *Tb*PYK/F26BP/PEP/Mg (**Figure 6.9**) reveals the concerted movements that occur when the PEP substrate binds. The movement of Mg^{2+} from its Mg-3 location to the PEP-coordinated Mg-1 location induces a number of side-chain conformational changes. As well as the rotation of the side chains of Asp265 and Glu241, there is a large side-chain movement of Phe213 in which the dihedral angle is rotated by 95^{o} to

fill the vacated Mg-3 space (**Figure 6.9**). Interestingly, the structure of the constitutively active rabbit muscle M1 PYK co-crystallised with L-phospholactate also showed a range of open and partially closed B-domain conformations, with Mg^{2+}-coordination to the phospho group lacking when the B-domain was in the open conformation (Larsen et al. 1997). In this conformation with bound L-phospholactate the Mg^{2+} was located instead at a site adjacent to the active site in a similar position to Mg-3 observed in *Tb*PYK in the presence of F26BP, but in the absence of PEP. In this way the structure of the constitutively active rabbit muscle M1 PYK in the presence of a substrate analogue can resemble the structure of the effector-activated *Tb*PYK in the absence of substrate.

The different side chain conformations of Phe213 correlate with the state of closure of the B-domain. **Figure 6.9** shows the close steric interaction between Phe213 and Pro88 when the B-domain is in its 'open' conformation (**Figure 6.1b**). The movement of the side chain of Phe213 may provide a mechanism which allows the linker residue Pro88 to move and close down the B-domain lid. The relatively small movement of Pro88 (of less than 2 Å) allows much larger side chain movements of Asp146, Arg91, Asn179 and Glu89 which form a network of hydrogen bonds to trap the B-domain in a partially closed conformation (**Figure 6.10**).

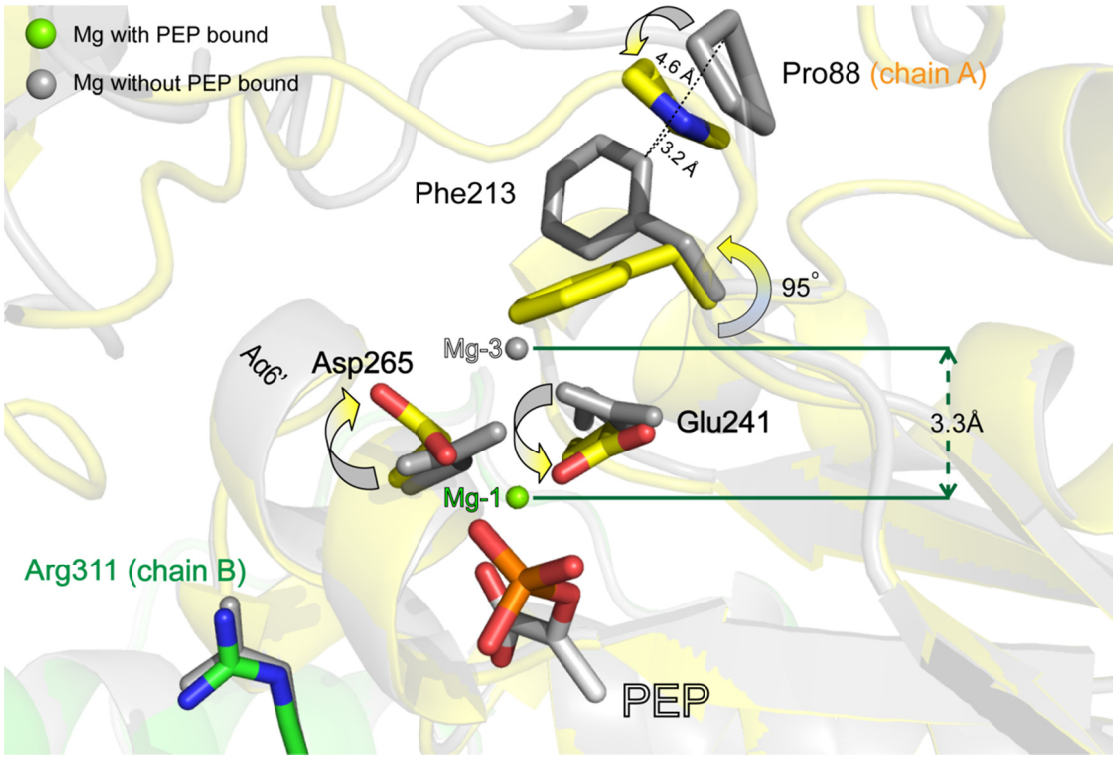


Figure 6.9 Side-chain reorientations at the active site of *Tb*PYK. †
The active site of *Tb*PYK/F26BP/PEP/Mg (coloured side chains: yellow, chain A; green, chain B; white, substrate PEP) is shown superposed on *Tb*PYK/F26BP/Mg (grey side chains). The average RMS difference is only 0.20 Å (**Table 6.2**). Polypeptide chains are shown in cartoon format, while relevant residues and the substrate PEP are shown as sticks. The two Mg^{2+} positions are shown as spheres, with the grey sphere in the 'priming' position (Mg-3 position), and the green sphere in the canonical position (Mg-**1** position) for coordination with PEP. The distance between these two Mg^{2+} ions is indicated as 3.3 Å. The reorientations of the side chain of Phe213 and the carboxyl side chains of residues Asp265 and Glu241 are indicated by arrows. The small movement of Pro88 as a result of B-domain movement is indicated as well. The distances between Pro88 and Phe213 are shown as black dotted lines. Refer to **Table 6.3** for the corresponding B-domain positions.
† This figure will be submitted for publication (see section **6.1**).

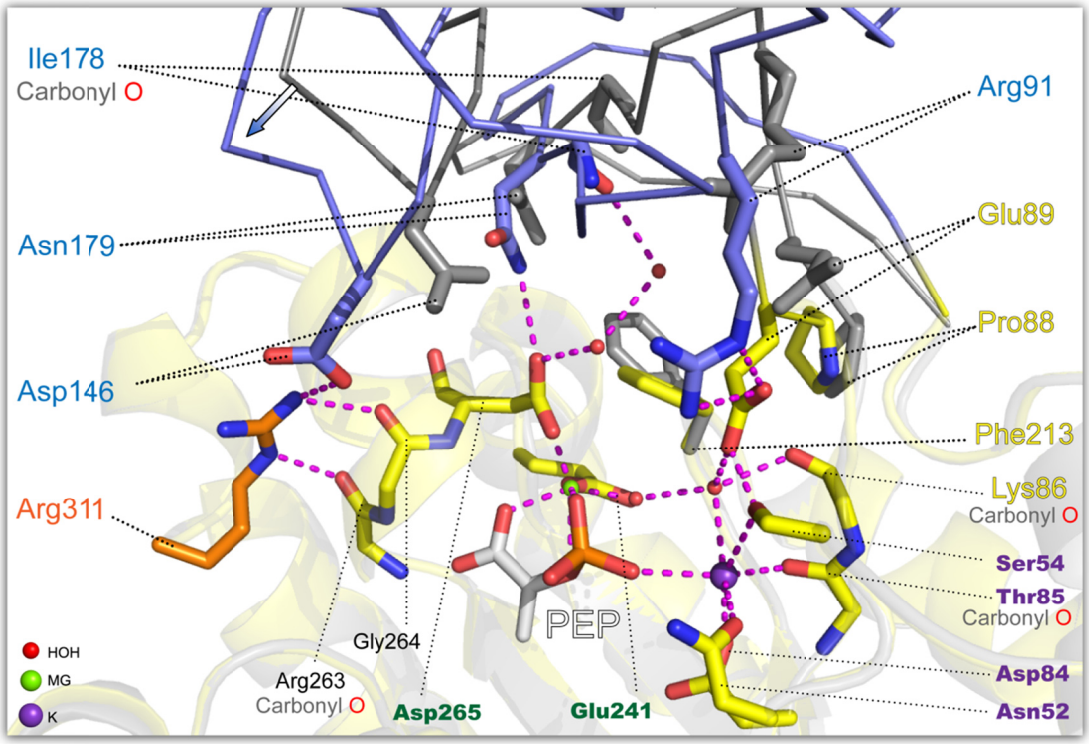


Figure 6.10 Close-up of the superposed active sites from the B-domain of chain A of *Tb*PYK/F26BP/PEP/Mg (blue) and chain B (grey) of *Tb*PYK/F26BP/Mg showing the interactions involved in B-domain movement. †

The interacting residues and PEP are shown as sticks, while metal ions and water molecules are shown as spheres. The interatomic interactions (within 3.2 Å) are indicated as pink dashed lines. The open form B-domain structure (grey) does not possess these interactions. Refer to **Table 6.3** for the corresponding B-domain positions.

† This figure will be submitted for publication (see section **6.1**).

6.3.5.3. Mg^{2+} collaborates with F26BP as a natural activator of PYK

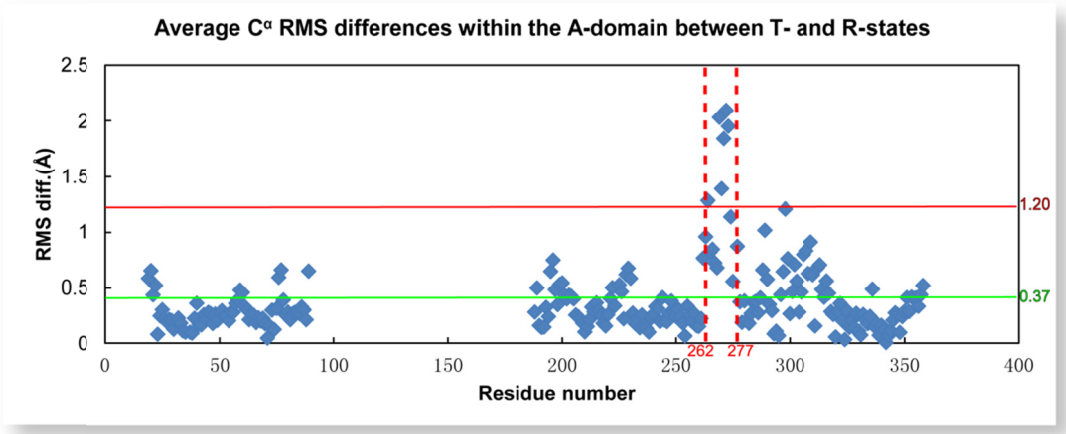
Binding of F26BP stabilises the R-state (active) conformation of trypanosomatid PYKs which allows more efficient binding of PEP ($S_{0.5}$ = 0.119 mM) compared to T-state PYK ($S_{0.5}$ = 1.03 mM) (see **Table 3.3** in **Chapter 3**). This difference in binding can be rationalised by subtle structural differences at their active sites: the flexible A 6′ helix is held rigid in an ideal binding geometry for the substrate in the R-state structure while the Aα6′ helix in the T-state structure adopts a slightly different conformation with a less optimal geometry for binding substrate. In **Figure 6.11**, a small motif (residues 262-277) with high RMS differences was identified, and is indicated by dashed lines in red. The average C^{α} RMS differences for all residues of this small motif is 1.20 Å as indicated by the continuous red line, compared to 0.37 Å for the average C^{α} RMS differences for all residues of the A-domains (indicated by the green line). No significant shift of this motif was found between the structures of *Tb*PYK/F26BP/Mg and *Tb*PYK/F26BP/PEP/Mg (with an average C^{α} RMS fit of 0.26 Å for the residues of this motif). A similar motif shift between the T- and R-states of *Lm*PYK was also observed (data not shown).

Interestingly, although the T-state structure was crystallised in the presence of 50 mM Mg^{2+}, no ions were located in the active site. In contrast, for both R-state structures presented here we see that Mg^{2+} (both in the canonical PEP-bound Mg-1 position and in the Mg-3 or 'priming' position) coordinates in such a way as to stabilise the R-state active site conformation (**Figure 6.5; Figure 6.9**). This suggests that both the effector molecule F26BP and the magnesium ions are stabilising the R-state conformation and that Mg^{2+} is playing an additional role in the allosteric mechanism over and above its catalytic function. Such a hypothesis is consistent with published results showing that *Tb*PYK has a sigmoidal response in activity with respect to Mg^{2+}, Mn^{2+} and Co^{2+} concentration (Callen et al. 1991).

It is noteworthy that in our crystallisation trials, crystals never grew in the presence of F26BP without Mg^{2+}, which indicates that *Tb*PYK might not be able to

adopt a stable R-state structure with F26BP alone. By contrast, crystals of the complex *Tb*PYK/F26BP/Mg were readily grown in the presence of the metal. Similar observations have been made with yeast PYK which shows enhanced binding of its natural effector F16BP in the presence of Mn^{2+} (Mesecar et al. 1997a; Mesecar et al. 1997b).

a



b

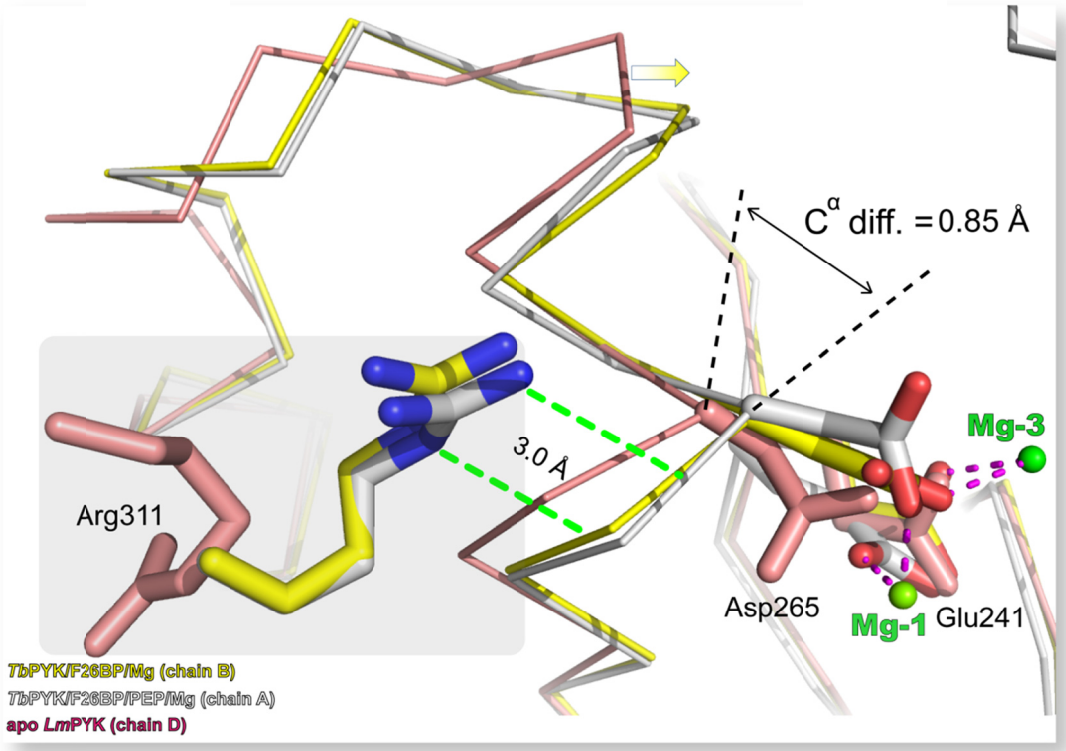


Figure 6.11 The C^{α} RMS differences identify a significant shift of a small motif within the A-domain between the T-state of *Lm*PYK and the R-state of *Tb*PYK. †
(a) The calculation for RMS differences was performed by the superposition of the A-domains (19-89, 188-358) from inactive T-state (apo *Lm*PYK) and active R-state (*Tb*PYK/F26BP/Mg) structures. The RMS differences were plotted as a function of residue numbers. The average C^{α} RMS differences for the small motif (dashed lines in red) and for all the residues are indicated by red line and green line, respectively. (b) The superposed structures are apo *Lm*PYK (pink) in the inactive T-state, and *Tb*PYK/F26BP/Mg (yellow) and *Tb*PYK/F26BP/PEP (white), both in the active R-state. The structures are shown as a ribbon while relevant residues are shown as sticks. The Mg^{2+} ions are shown as spheres in green. The interactions between the protein and Mg^{2+} ions are indicated by pink dashed lines. The shift of the small motif including Aα6′ is indicated by the arrow. The C^{α} atom of residue Asp265 which coordinates the Mg^{2+} ion in *Tb*PYK/F26BP/Mg or *Tb*PYK/F26BP/PEP/Mg has a similar shift of 0.85 Å compared to the T- state structure of apo *Lm*PYK. The interactions between Arg311 (in the neighbouring chain) and the small motif are indicated by green dashed lines and the interaction distance is about 3.0 Å. Refer to **Table 6.3** for the corresponding structures.
† This figure will be submitted for publication (see section **6.1**).

6.3.5.4. Crystal of *Tb*PYK/F26BP/Mg soaked with EDTA has no Mg^{2+} at or near the active site and forms a *Tb*PYK/F26BP/EDTA complex

EDTA (ethylenediamine tetraacetic acid) is capable of sequestering metal ions such as Ca^{2+}, Mg^{2+} and Fe^{3+}. Only seven protein structures with bound EDTA have been deposited in the PDB so far (PDB ID: 3TBN, 3RNJ, 3L7K, 3L7M, 2AXN, 1ZLQ, 1NNF), only two of which show metal-EDTA interactions (1ZLQ and 1NNF). This section will describe the first EDTA-bound PYK structure (*Tb*PYK/F26BP/EDTA) obtained by crystal soaking. A crystal of *Tb*PYK/F26BP/Mg was soaked with 25 mM EDTA and the structure was determined to 2.4 Å, and contained two chains in the asymmetric unit. The AC-cores of the *Tb*PYK/F26BP/EDTA and *Tb*PYK/F26BP/Mg tetramers were superposed resulting in a 0.13 Å C^{α} RMS difference (**Table 6.2**), indicating both structures possess almost identical R-state conformations.

(i) EDTA mimics the phospho groups of ATP bound to the active site

The Mg^{2+} at the Mg-3 position (**Figure 6.5**) in *Tb*PYK/F26BP/Mg is not observed in the EDTA soaked structure of *Tb*PYK/F26BP/EDTA (**Figure 6.12**), which suggests that PYK-bound Mg^{2+} has been successfully chelated and removed by the high concentration of EDTA during the crystal soaking procedure.

Furthermore, EDTA has been clearly identified at the active site of *Tb*PYK, partially overlapping with the ATP binding site. The overlaid structures of *Tb*PYK/F26BP/EDTA and *Lm*PYK/F26BP/OX/ATP are presented in **Figure 6.13** to reveal the similar ligand binding modes. The carboxyl groups C(1)OO, C(12)OO and C(10)OO of EDTA resemble the phospho groups of ATP at the α, β and γ positions, respectively. The carboxyl group C(1)OO forms ionic interactions with Arg50, while C(10)OO coordinates with Lys239 and K^{+}, and C(12)OO interacts with Asn52 and His55 (**Figure 6.12**). The carboxyl group C(5)OO is more flexible, as indicated by relatively poor density in **Figure 6.12**, and only forms a weak interaction with a water molecule. A series of water molecules form a network to stabilise EDTA

binding as well. No obvious Mg^{2+} density was observed possibly due to the divalent metal chelation by free EDTA in solution but not PYK-bound EDTA.

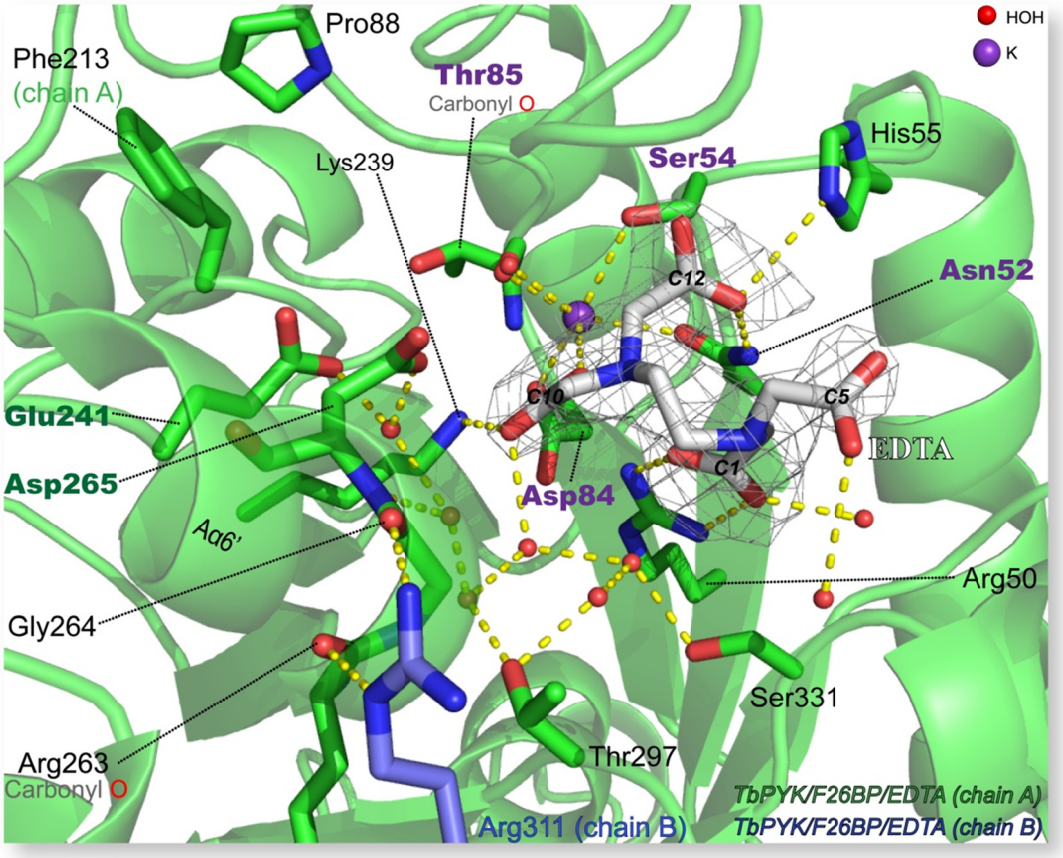


Figure 6.12 The active site of *Tb*PYK with bound EDTA.
The polypeptide chain at the active site of *Tb*PYK/F26BP/EDTA is shown in cartoon format, with the ligand EDTA (white) and relevant residues shown as sticks. Chain A is coloured green and chain B blue. The monovalent ion K^+ is a purple sphere, and is coordinated by the residues labelled in purple. Water molecules are shown as red spheres. The *Fo-Fc* electron density for the EDTA molecule is shown as a grey mesh contoured at 2.5 σ. Interatomic interactions are indicated by yellow dashed lines.

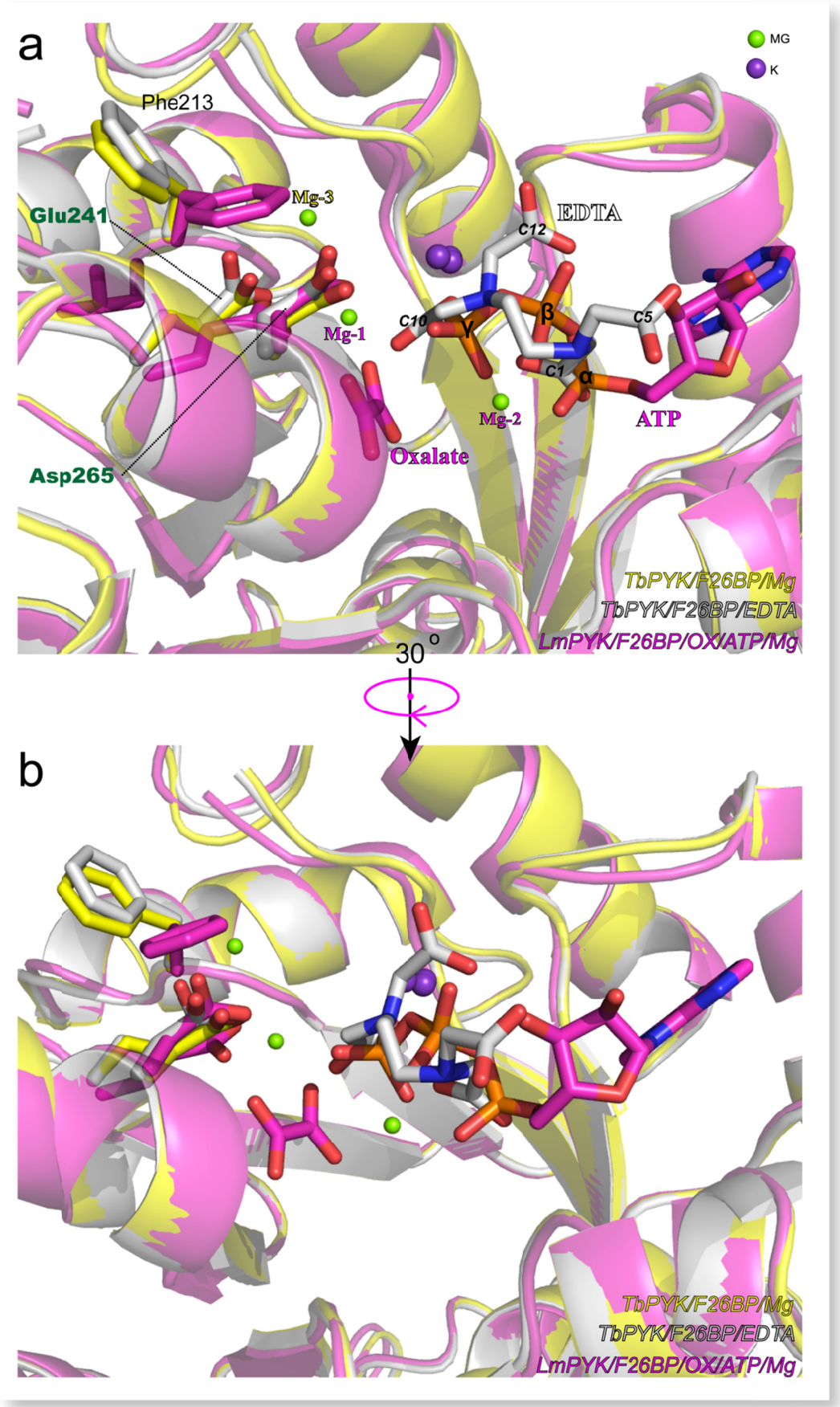

a
MG
K
Phe213
Glu241
Mg-3
EDTA
C12
C10
γ
β
C5
Mg-1
C1
α
Mg-2
ATP
Oxalate
Asp265
TbPYK/F26BP/Mg
TbPYK/F26BP/EDTA
LmPYK/F26BP/OX/ATP/Mg
30 °
b
TbPYK/F26BP/Mg
TbPYK/F26BP/EDTA
LmPYK/F26BP/OX/ATP/Mg

Figure 6.13 Comparison of the active sites of *Tb*PYK/F26BP/Mg, *Tb*PYK/F26BP/EDTA and *Lm*PYK/F26BP/OX/ATP/Mg.

(a) The polypeptide chains of *Tb*PYK/F26BP/Mg (chain A), *Tb*PYK/F26BP/EDTA (chain A) and *Lm*PYK/F26BP/OX/ATP/Mg (chain C; PDB code: 3HQP) are shown in cartoon format, and coloured yellow, white and pink, respectively. The metals Mg^{2+} and K^{+} are shown as spheres coloured in green and purple, respectively. Mg^{2+} observed in *Lm*PYK/F26BP/OX/ATP/Mg are in positions Mg-1 and Mg-2. Mg^{2+} in the 'priming' Mg-3 position is from *Tb*PYK/F26BP/Mg. The flexible Phe213, two Mg-coordinating residues (Glu241, Asp265), oxalate and ATP are shown as sticks. The superposed structures are rotated by 30^{o} and shown in (b). Refer to **Table 6.3** for the corresponding structures.

(ii) The chelation of Mg^{2+} by EDTA induces movement of the B-domain

Both crystallographically independent B-domains have been modelled in *Tb*PYK/F26BP/EDTA but have different orientations. Compared to B-domains in the open conformation from *Tb*PYK/F26BP/Mg, the B-domain (chain A) of the EDTA-bound structure undergoes a 12^{o} rotation (**Figure 6.14**) which is not as significant as the B-domain movement from the PEP-bound structure (23^{o}). Crystal packing interactions lock the other B-domain (chain B) of the EDTA-bound structure in an open conformation.

(iii) The correlation between PEP, Mg^{2+}, Phe213 and the B-domain

The correlation between substrate PEP, Mg^{2+}, Phe213 and the B-domain has been discussed in section **6.3.5**. It has been shown that the different side-chain conformations of Phe213 correlate with the state of closure of the B-domain, where the rotation of the B-domain requires the movement of the side chain of Phe213 (**Figure 6.9**). Interestingly, the B-domain in chain A of *Tb*PYK/F26BP/EDTA possesses a more modest movement (12^{o}) compared to the open-form B-domain (**Figure 6.14**), but the side chain of Phe213 in chain A is still in the open conformation (**Figure 6.13**).

Although the side chain of Phe213 in the EDTA-bound structure (chain A) is modelled in an open conformation, it shows a relatively high average B-factor of 66 Å^2 compared to the average B-factor of 45 Å^2 for the side chain of Phe213 from chain B, suggesting that the side chain of Phe213 is relatively flexible in chain A. This flexible Phe213 with a high B-factor may explain the small movement (12^{o}) of the B-domain. The side chain of Phe213 is proposed to possess a closed conformation only in the presence of the Mg^{2+}-PEP complex or Mg^{2+}-PEP analogue complex (**Figure 6.9**). Without this complex, the side chain of Phe213 will not possess a closed conformation to induce a 23^{o} movement of B-domain in *Tb*PYK/F26BP/EDTA, where only the 12^{o} rotation can be achieved. Therefore, the

correlation between PEP (or its analogue), Mg^{2+}, Phe213 and B-domain is further confirmed by this EDTA-bound structure.

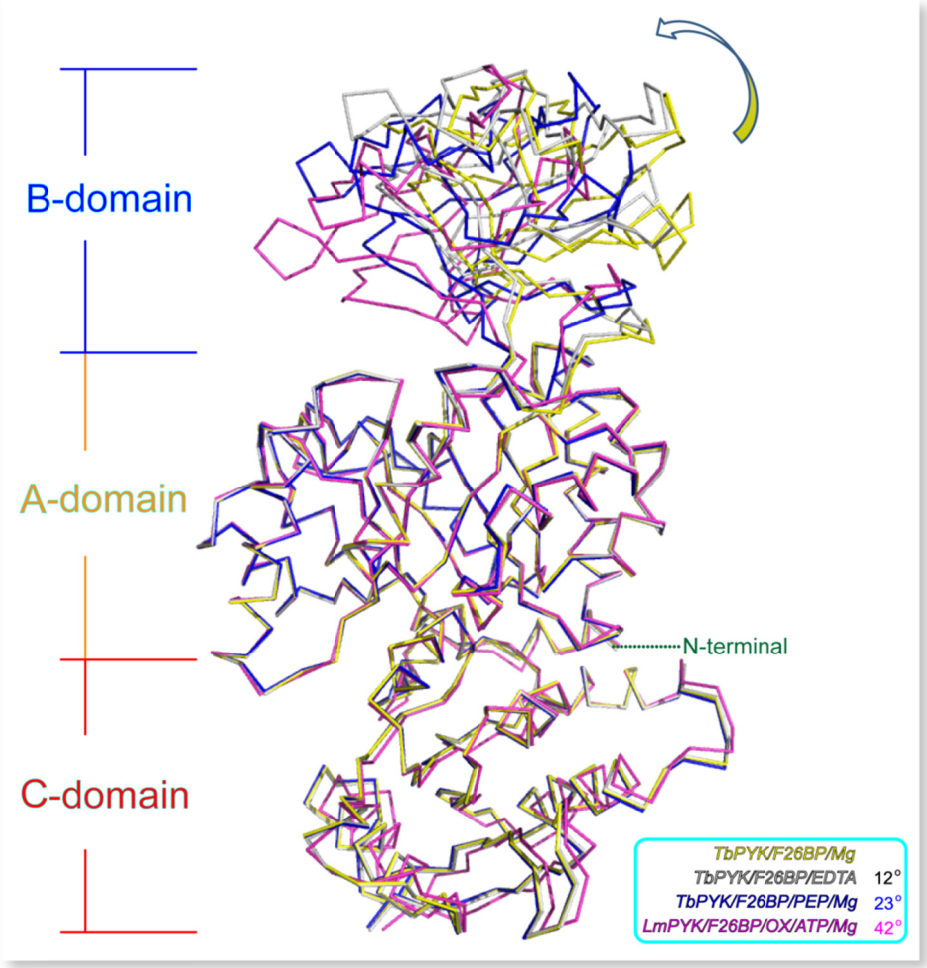


Figure 6.14 The B-domain from the EDTA-bound structure possesses a 12° rotation compared to the B-domain in the open position from *Tb*PYK/F26BP/Mg.

The superposed structures are: *Tb*PYK/F26BP/Mg (yellow backbone, chain A), *Tb*PYK/F26BP/EDTA (white backbone, chain A), *Tb*PYK/F26BP/PEP/Mg (blue backbone, chain A) and *Lm*PYK/F26BP/OX/ATP/Mg (pink backbone, chain C, PDB code: 3HQP). The relatively static AC-domains and the multiple positions of the B-domains of subunits are shown as ribbons. The B-domain rotations of these PYK complexes are indicated by arrows and the rotation angles are shown in the cyan box. The B-domain in *Tb*PYK/F26BP/Mg (chain A) is in the open conformation, while the B-domains in *Tb*PYK/F26BP/EDTA and *Tb*PYK/F26BP/PEP/Mg are in the partially or half-closed conformations. The fully ligand-bound structure of *Lm*PYK/F26BP/OX/ATP/Mg has the B-domain in its fully closed position. Refer to **Table 6.3** for the corresponding structures.

6.3.6. Mg^{2+} and K^{+} are retained at or near the active site of PYK in the presence of F26BP to facilitate PEP binding

The comparison of the *Tb*PYK/F26BP/Mg and *Tb*PYK/F26BP/PEP/Mg structures shows that when PEP binds there is a reorientation of the side chains of Phe213, Asp265 and Glu241 such that the Mg^{2+} ion moves 3.3 Å from the Mg-3 pocket to coordinate with the phospho group of PEP and occupy the canonical Mg-1 pocket (**Figure 6.5**; **Figure 6.9**). The movement of Mg^{2+} between these two sites is therefore acting like a two-way toggle switch. When the **Mg-3** site is occupied, the B-domain is held open but the active site is still maintained in its R-state conformation. When the **Mg-1** site is occupied the B-domain is pulled shut and the magnesium coordinates directly to PEP. Importantly the structures show that occupation of either the Mg-1 site or the Mg-3 site would stabilise the enzyme in its active R-state conformation. In this way the enzyme could retain the R-state conformation after product release, and the active site would be primed to accept the next PEP substrate molecule.

6.3.7. Allosteric mechanism of TbPYK

6.3.7.1. Rigid-body rotation of AC-domains regulates allosteric activity of *Tb*PYK

An 8^{o} rigid-body rotation which regulates the conformational transition between the T- and R-states of *Tc*PYK has been described in **Chapter 5** by comparing the T-state structure (apo *Tc*PYK) and R-state structure (*Tc*PYK/F26BP/OX/Mg) of *Tc*PYK. This rigid-body motion was fully characterised for the first time for *Lm*PYK another trypanosomatid PYK (Morgan et al. 2010b). The kinetics for *Tb*PYK, *Tc*PYK and *Lm*PYK which have 74-81% sequence identity (**Table 6.4**) suggest that all three PYKs are allosterically activated by F26BP (see **Table 3.3** in **Chapter 3**). Furthermore, the R-states of these three trypanosomatid PYKs are structurally similar

to each other (**Table 6.2**; **Figure 6.1a**). In this way it is plausible to use the T-state structures of *Tc*PYK and *Lm*PYK as T-state models for *Tb*PYK, although the T-state structure of *Tb*PYK is not available currently.

Fits of the *Tb*PYK/F26BP/Mg or the *Tb*PYK/F26BP/Mg/PEP structure onto the AC-cores of the T-state apoenzyme *Tc*PYK structure or T-state apoenzyme *Lm*PYK (3HQN) structures give rigid body rotation angles of ~8^{o} (**Table 6.5**) which is similar to the rotation angle for *Tc*PYK (see **Chapter 5**). The superposition of tetramers (AC-cores) of apo *Tc*PYK (T-state) and *Tb*PYK/F26BP/Mg (R-state) are presented in **Figure 6.15** showing the rotation movement between these two conformations. The rotation pivot in each rigid body (AC-core) of a tetramer has been predicted to be residues 429-433 (**Figure 6.16**) which is similar to the predicted pivots of *Tc*PYK and *Lm*PYK.

Table 6.4 Pairwise protein sequence comparisons of trypanosomatid, yeast, human M2 and *E. coli* PYKs †

	*Tb*PYK	*Tc*PYK	*Lm*PYK	*Hs*M2PYK	*Sc*PYK	*E.coli* PYK
*Tb*PYK	100	81	74	48	48	42
*Tc*PYK		100	76	47	48	42
*Lm*PYK			100	47	48	42
*Hs*M2PYK				100	49	44
*Sc*PYK					100	43
E. coli PYK						100

The pairwise sequence analysis was obtained from the EMBL-EBI web server (EMBOSS Stretcher): http://www.ebi.ac.uk/Tools/psa/emboss_stretcher/
Values are overall percent sequence identities.
Hs, Homo sapiens; Sc, Saccharomyces cerevisiae
† This table will be submitted for publication (see section **6.1**).

Table 6.5 Angles of AC-core rigid body rotation from T- to R-state of *Tb*PYK †

T-state	R-state	Rotation Angle [a]	Average RMS [b]
apo *Tc*PYK	*Tb*PYK/F26BP/Mg	$8.3^{o}\pm0.2^{o}$	0.51 Å
apo *Lm*PYK (3HQN)	*Tb*PYK/F26BP/Mg	$8.3^{o}\pm0.2^{o}$	0.65 Å
apo *Tc*PYK	*Tb*PYK/F26BP/PEP/Mg	$8.1^{o}\pm0.5^{o}$	0.52 Å
apo *Lm*PYK (3HQN)	*Tb*PYK/F26BP/PEP/Mg	$8.3^{o}\pm0.2^{o}$	0.66 Å

[a] Calculated rotation angles with standard deviation
[b] Average C^{α} RMS difference after rotation
† Part of this table (not including *Tc*PYK data) will be submitted for publication (see section **6.1**).

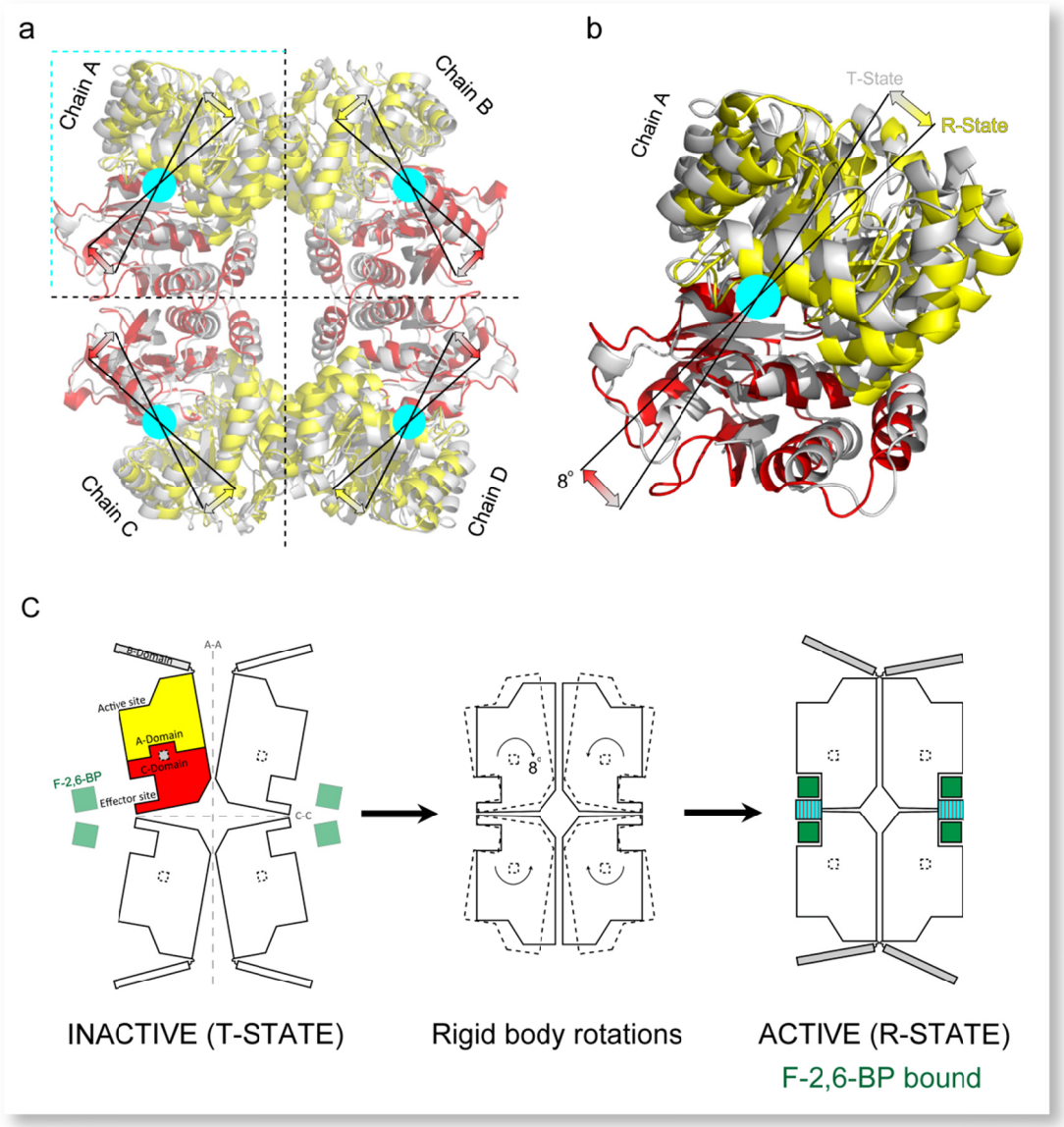


Figure 6.15 Rigid body rotation of *Tb*PYK between T- and R-states.
(a) All B-domains and N-terminal domains are removed, and interface boundaries are indicated by dashed lines. The superposed polypeptide chains are shown as cartoons, and domains are highlighted by different colours (A-domains: yellow; C-domains: red). The C-α atoms of the AC cores (A- and C- domains) of the R-state *Tb*PYK/F26BP/Mg tetramer were superposed onto inactive T-state apo *Tc*PYK (grey) with an RMS fit of 2.5 Å. The directions of the 8° rigid body rotation of the AC-cores around the central pivots (highlighted in cyan) for the T- (grey) to R-state transition are shown by arrows. (b) An enlargement of chain A showing the 8° rotation around the pivot. (c) Schematic representation of the T- and R-states of the *Tb*PYK structures, and the motion of rigid-body rotation between these two states initiated by the allosteric effector F26BP. This figure is a modification of the scheme originally presented by my colleague Dr. Hugh Morgan (Morgan et al. 2010b).

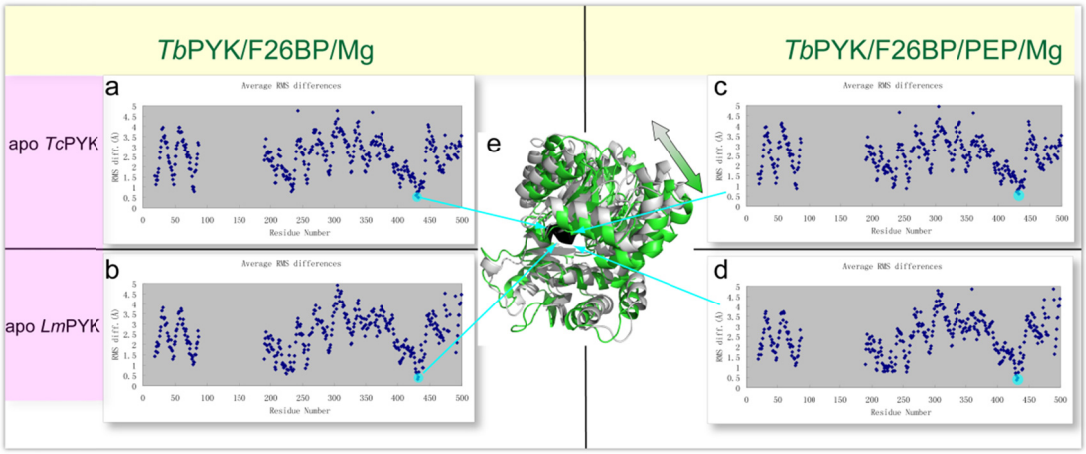


Figure 6.16 The rotation pivot of the rigid body (AC-domains) rotation for *Tb*PYK is located at residues 429-433 as observed from the analysis of RMS differences (C^{α}) of each residue between the T- and R-state structures.

(a) Average RMS differences between the AC cores of T-state apo *Tc*PYK and R-state *Tb*PYK/F26BP/Mg structures. (b) Average RMS differences between the AC cores of T-state apo *Lm*PYK and R-state *Tb*PYK/F26BP/Mg structures. (**c**) Average RMS differences between the AC cores of T-state apo *Tc*PYK and R-state *Tb*PYK/F26BP/PEP/Mg structures. (d) Average RMS differences between the AC cores of T-state apo *Lm*PYK and R-state *Tb*PYK/F26BP/PEP/Mg structures. A rotation pivot (residues 429-433) with the lowest RMS difference was identified and shaded in cyan. (e) The rigid body rotation for each subunit, where the rotation is indicated by an arrow from T-state (grey; apo *Tc*PYK) to R-state conformation (green; *Tb*PYK/F26BP/Mg). The pivot point (residues 429-433) located on an α-helix of each C-domain is highlighted in black.

6.3.7.2. Interface interactions of R-state *Tb*PYK

Interface interactions including hydrogen bonds, salt bridges and hydrophobic interactions between subunits within a tetramer play essential roles in the stabilisation of the R-state conformation of the *Tb*PYK tetramer (**Figure 6.14**). The detailed interactions across the large (A-A) and small (C-C) interfaces are shown in **Table 6.6** and **Table 6.7**, respectively. Some hydrophobic residues located at the interface are proposed to enhance the structural stability as well (**Figure 6.17b**). The loops (residues 482-488) at the effector sites are stabilised by F26BP binding, and form additional salt bridges along the small (C-C) interface which are possibly lost in the T-state form of *Tb*PYK (**Figure 6.16**). These interactions across interfaces are not affected by the binding of the substrate PEP (*Tb*PYK/F26BP/PEP/Mg) which is soaked into crystals of *Tb*PYK/F26BP/Mg as this R-state can be formed in the presence of F26BP and Mg^{2+} without substrate binding (section **6.3.5**).

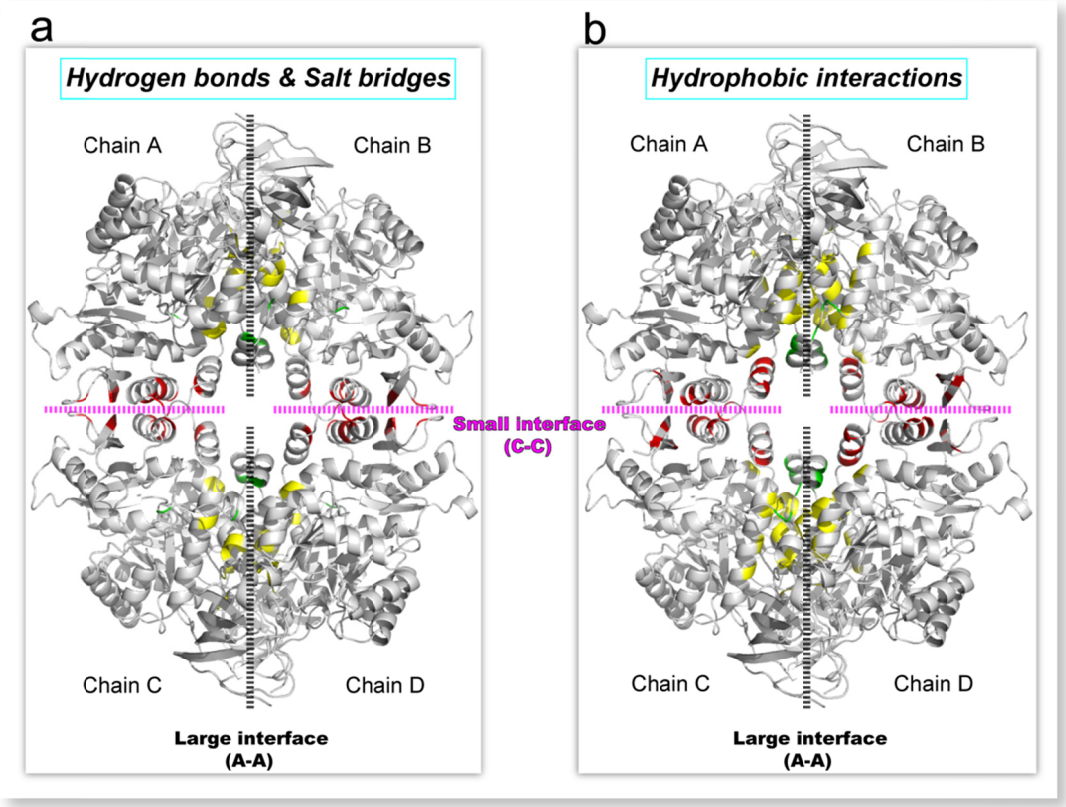


Figure 6.17 Residues involved in interface interactions between subunits of the R-state *Tb*PYK tetramer.

The structure of *Tb*PYK/F26BP/Mg is represented as cartoons in grey. The large (A-A) and small (C-C) interfaces between subunits are shown as dashed lines, and indicate the four subunits (monomers) in one tetramer. Chain A and chain B are asymmetrically unique chains in the crystal unit cell. The other half portion of the tetramer is generated from symmetry-related monomers. Chain C and chain D are the symmetry-related monomers of chain B and chain A, respectively. (a) Residues involved in the interface interactions forming hydrogen bonds and salt bridges are indicated by colours. Residues belonging to domains A, B, C and N-terminal are coloured in yellow, blue, red and green, respectively. These residues are listed in **Table 6.6 and 6.7**. (b) Residues potentially involved in hydrophobic interactions along the interfaces are indicated as belonging to different domains in the same way. These residues are as follows: for the large interface: Leu4, Val8, Leu10, Ile12, Phe13, Val246, Ile270, Ala272, Val277, Met280, Ile281, Val288, Pro309, Val314, Ala318, Leu322, Ile350, Ala354, Ala357, Met363, Ile367 and Leu370; for the small interface: Pro374, Pro377, Ala380, Val381, Val392, Leu476 and Leu495. Residues belonging to domains A, B, C and N-terminal are coloured in yellow, blue, red and green, respectively.

Table 6.6 Large (A-A) interface polar interactions of *Tb*/F26BP/Mg

Interface[†]		A/B or D/C
Chain 1[‡]	**Chain 2**	**Distance (Å)**
Lys274[NZ]	Ile12[O]	3.32
Asn7[ND2]	Ser284[OG]	3.12
Ser315[OG]	Asp316[OD1]	2.88
Lys274[NZ]	***Glu353[OE2]***	3.00
Lys274[NZ]	***Glu353[OE1]***	3.12
Val8[O]	Lys285[NZ]	3.47
Ser284[OG]	Asn7[ND2]	3.11
Asp316[OD1]	Ser315[OG]	2.59
Glu353[OE1]	***Lys274[NZ]***	3.24
Glu353[OE2]	***Lys274[NZ]***	3.02
Arg311[NH2]	Gly264[O]	2.95
Arg311[N]	Gln298[OE1]	2.79
Gly264[O]	Arg311[NH2]	2.94
Gln298[OE1]	Arg311[N]	2.83
Arg311[NE]	Arg263[O]	2.98
Arg263[O]	Arg311[NE]	2.88
Total interface area (Å^2)[‡]		2548.5

[†] The interfaces between subunits shown here are those corresponding to **Figure 6.17a**. Chain 1 and Chain 2 which form half of the large interface correspond to chain A and chain B, or to chain C and chain D of *Tb*PYK/F26BP/Mg.

[‡] The interface areas, interacting residues and distances are obtained from the online service PDBePISA (Krissinel and Henrick 2007). Potential hydrophobic interactions are not available from this analysis. Atom distances beyond 3.5 Å are not shown in the table.

[*] Residues belonging to domains A and N-terminal are shaded in yellow and green, respectively. Residues which can potentially form salt bridges are shown in bold and italic.

Table 6.7 Small (C-C) interface polar interactions of *Tb*PYK/F26BP/Mg

		***Tb*PYK/F26BP/Mg**
Interface[†]		A/C or D/B
Chain 1	**Chain 2**	**Distance (Å)**
Ala380[O]*	Ser388[OG]	3.32
Met375[N]	Glu391[O]	2.73
Lys368[NZ]	***Glu391[OE1]***	2.41
Ile495[N]	Asn491[O]	2.93
Thr493[N]	Thr493[O]	2.87
Asn491[ND2]	Ile495[O]	2.93
Glu391[O]	Met375[N]	2.79
Ile495[O]	Asn491[ND2]	3.01
Gln492[OE1]	Gln492[NE2]	3.50
Thr493[O]	Thr493[N]	2.87
Asn491[O]	Ile495[N]	2.85
Ser384[OG]	Ser384[O]	3.37
Ser384[OG]	Ser388[OG]	2.57
Ser383[OG]	Glu391[OE2]	3.05
Arg494[NH1]	***Asp483[OD2]***	3.00
Asp483[OD2]	***Asp494[NH1]***	3.01
Ser384[O]	Ser384[OG]	3.39
Glu391[OE1]	Ser384[OG]	2.87
Glu391[OE2]	Ser383[OG]	2.86
Total interface area (Å^2)[‡]		1057.1

[†] The interfaces between subunits shown here correspond to **Figure 6.17a**. Chain 1 and Chain 2 which form half of the small interface of a tetramer correspond to chain A and chain C, or to chain D and chain B for *Tb*PYK/F26BP/Mg.

[‡] The interface areas, interacting residues and distances are obtained from the online service PDBePISA (Krissinel and Henrick 2007). Potential hydrophobic interactions are not available from this analysis. Atom distances beyond 3.5 Å are not shown in the table.

[*] All residues here belong to domains C and are shaded in blue. Residues which can potentially form salt bridges are shown in bold and italic. Residues in the effector loop (residues 482-488) are coloured purple.

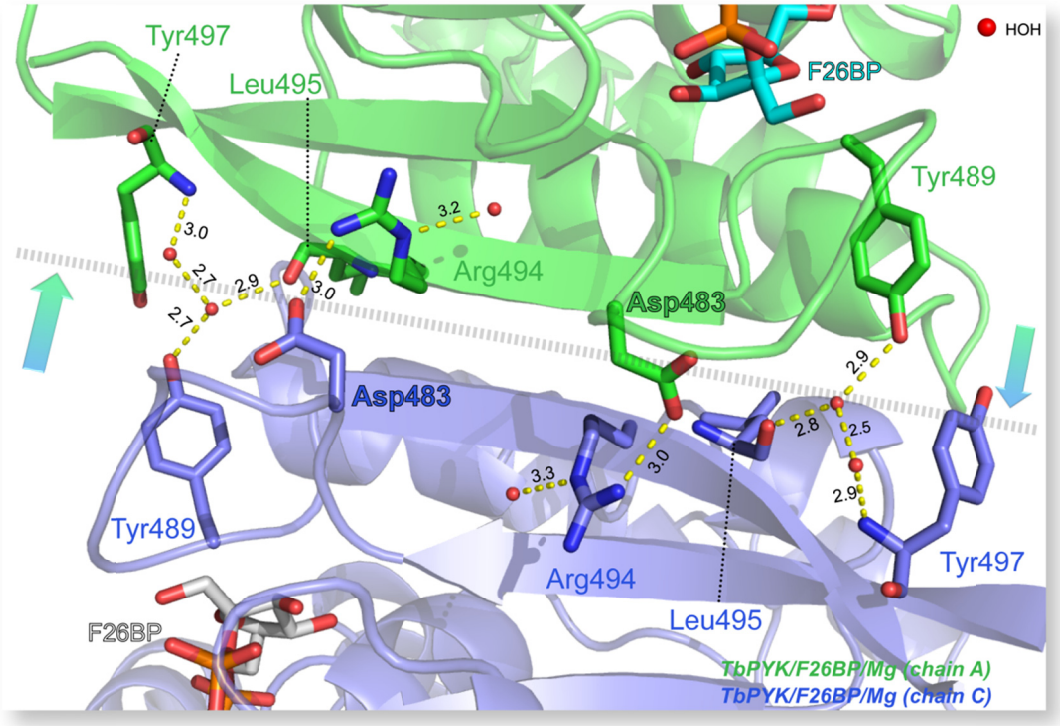


Figure 6.18 Close-up of the A/C interface of *Tb*PYK/F26BP/Mg showing the additional salt bridges formed across the small (C-C) interface as a result of F26BP binding.

The polypeptide chains are represented as cartoons in green (chain A) or blue (chain C). These two chains correspond to the chains discussed in **Figure 6.17** and **Table 6.6**. The interacting residues which are located on or close to the effector loop are shown as sticks. F26BP molecules are also represented as sticks. Water molecules are shown as spheres in red. Interatomic interactions are indicated by yellow dashed lines and the interacting distances are given in Ångstroms. F26BP binding caused additional salt bridges to be formed between Asp483…Arg494. The hydrogen bond interactions via water molecules were Tyr489…water…Leu495, and Tyr489…water…water…Tyr497. The small (C-C) interface between chain A and chain C is shown as a dashed line in grey, while the movements of the effector loops are indicated as arrows.

6.3.7.3. The active R-state structure is in a thermally more stable conformation

Previously, the R-states of *Tc*PYK (**Chapter 5**) and *Lm*PYK (Morgan et al. 2010b) have been shown to possess thermally more stable conformations as a result of additional interface interactions induced by effector or active-site ligand binding. The natural substrate PEP, oxalate (analogue of PEP), and the natural activator F26BP are each capable of inducing the T- to R-state transition of trypanosomatid PYKs in the presence of Mg^{2+}. They therefore stabilise the R-state conformation, and the increases in melting temperature have been studied in *Tc*PYK (**Figure 5.20** in **Chapter 5**). This analysis has been extended to *Tb*PYK in this section.

(i) Selection of ligands for thermal-shift assays

The natural substrates/products (PEP, ADP; ATP), the natural activator (F26BP), the analogue of PEP (OX), the analogue of F26BP (F16BP), and inorganic phosphate (Pi) were used in thermal-shift assays individually or with different combinations of ligands (**Figure 6.18**). The assay for F26BP was only done once without combining with the other ligands because of the lack of a commercial source of F26BP.

(ii) R-state *Tb*PYK is in a thermally more stable conformation compared to the T-state enzyme

The stabilities of *Tb*PYK in the presence of added ligands are given in **Figure 6.19**. Without the addition of any ligand, *Tb*PYK shows a melting temperature (T_m) of 45.5°C which is ~10°C lower than that of *Tc*PYK, although they have 81% sequence identity and almost identical active sites and effector sites (**Figure 6.2**). The difference in interface interactions may contribute to these significantly different T_m values for both T-state enzymes (**Figure 6.17, Table 6.6** and **6.7; Figure 5.18, Table 5.4** and **5.5** in **Chapter 5**). Addition of the active-site ligand PEP (10 mM) or oxalate (10 mM) to the apoenzyme of *Tb*PYK increases the T_m significantly from 45.5°C to 63.5°C and 62.5°C, respectively. It has been proved that oxalate alone can induce the

T- to R-state transition of *Lm*PYK and stabilise this active conformation (Morgan et al. 2010b). The sigmoidal response of *Tb*PYK with respect to the substrate PEP is completely converted to a hyperbolic response by the addition of the allosteric activator F26BP, indicates that PEP potentially can stabilise *Tb*PYK in the R-state conformation as well. The addition of the substrate ADP or product ATP to the apoenzyme fails to increase T_m which suggests that the nucleotide ligand itself is not capable of triggering the allosteric conformational transition.

The addition of the natural activator F26BP or its analogue F16BP increases the T_m as expected. The T_m values for both R-state *Tb*PYK and R-state *Tc*PYK are quite similar when comparing the complexes of *Tb*PYK/OX/Mg (T_m =62.5°C) and *Tc*PYK/OX/Mg (T_m =61°C), although the T_m values for the apoenzymes of *Tb*PYK and *Tc*PYK are ~10°C different.

Moreover, in the presence of inorganic phosphate (Pi), which is supposed to mimic the phospho groups of PEP or nucleotide ligands, *Tb*PYK has an increased T_m of 59.5°C compared to the T_m of the apoenzyme. The combination of different ligands can give even higher T_m values for *Tb*PYK. The most stable *Tb*PYK complex (T_m =67.5°C; Pi did not affect the T_m) is observed in the presence of a combination of F16BP, oxalate and ATP, which is *Tb*PYK/F16BP/OX/ATP/Mg. This most stable complex is similar to that of the crystal structure of *Lm*PYK/F26BP/OX/ATP/Mg (3HQP in Morgan et al. 2010b; T_m=62°C) that shows fully closed B-domains and R-state conformation, although F26BP is present in the structure instead of its analogue F16BP.

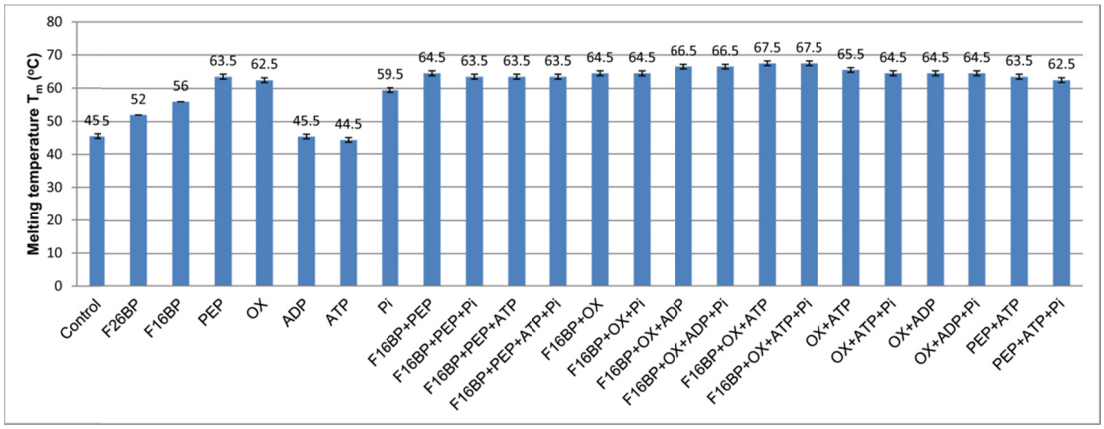


Figure 6.19 Thermal shift assay results for *Tb*PYK.

The error bars shown in the figure correspond to the standard deviations for duplicated results.

Control sample: the buffer in this sample of *Tb*PYK contains 50 mM TEA (pH7.2), 100 mM KCl and 10 mM $MgCl_2$.

The concentration for each ligand:

F26BP: 1.5 μM; F16BP: 4 mM; PEP: 10 mM; OX: 10 mM; ADP: 2 mM; ATP: 2 mM; Pi: 4mM. These concentrations are all saturated.

6.4. Conclusions

The crystal structures of the allosterically regulated *Tb*PYK in the presence and absence of the substrate PEP have revealed details of the conformational changes that occur during the enzyme's catalytic cycle. The active site of PYK is located between the AC-core of the enzyme and a mobile lid corresponding to domain-B. Many PYK structures have already been determined, but this thesis describes here the first 'effector only' structure, and the first with PEP, the true natural substrate. PEP soaked into crystals of the enzyme with bound allosteric activator F26BP and Mg^{2+} triggers a substantial 23° B-domain rotation '*in crystallo*' resulting in a partially closed active site. The interplay of side chains with Mg^{2+} and PEP explain the mechanism of the domain movement. The EDTA-bound structure of *Tb*PYK proves the capability of removing Mg^{2+} from the active site of F26BP-bound *Tb*PYK by the crystal soaking with EDTA.

Furthermore, it is apparent that when F26BP is present but PEP is absent, Mg^{2+} occupies a position distinct from the two canonical Mg^{2+} binding sites. This third site is adjacent to the active site, and acts to sequester Mg^{2+} in a 'priming' position such that the enzyme is maintained in its R-state conformation. In this way Mg^{2+} cooperates with F26BP to ensure that the enzyme is in a conformation that has high affinity for the substrate. On PEP binding, Mg^{2+} moves to its canonical position at the active site, and it is plausible that Mg^{2+} is retained in its 'priming' position after product release, as long as FBP is present.

Similar to *Tc*PYK and *Lm*PYK, the transition between T- and R-states of *Tb*PYK requires a concerted 8^{o} rigid body rotation of each of the four AC-cores in the tetramer and the R-state is stabilised by additional interface interactions. The thermal stability of *Tb*PYk has been extensively studied in this chapter.

CHAPTER 7: Pyruvate Kinases have an Inherent and Conserved Decarboxylase Activity

7.1. Introduction

Oxaloacetate decarboxylase (OAD; EC 4.1.1.3) catalyses the decarboxylation of oxaloacetate into pyruvate and CO_2 (**Figure 7.1a**). Typically these enzymes occur in bacteria, may be cytosolic or membrane bound, and comprise at least three structurally distinct families. Soluble cytosolic OADs have been characterised from *Azotobacter vinelandii* (Plaut and Lardy 1949), *Pseudomonas ovalis Chester* (Horton and Kornberg 1964), *Pseudomonas citronellolis* (O'Brien et al. 1977), *Acetobacter xylinum* (Benziman et al. 1978), *Veillonella parvula* (Ng et al. 1982), *Corynebacterium glutamicum* (Jetten and Sinskey 1995) and *Pseudomonas stutzeri* (Labrou and Clonis 1999). More recently the enzymes from *Pseudomonas aeruginosa* (PA4872; PDB ID 3B8I; Narayanan et al. 2008) and *Corynebacterium glutamicum* (cg1458; PDB ID 4DBF; Klaffl and Eikmanns 2010; Ran et al. 2011; Ran et al. 2012) have been expressed and purified from *E. coli*, and their X-ray crystal structures solved. The latter OAD does not share any sequence similarity with other known cytosolic OADs (which are members of the isocitrate lyase family), but instead appears to belong to the fumarylacetoacetate hydrolase family (Ran et al. 2012). The crystal structure of an OAD from *Vibrio cholera* is also available (Q6A1F6; PDB 2NX9; Studer et al. 2007), and corresponds to the subunit of a membrane-bound multiprotein complex that couples the oxaloacetate decarboxylase reaction with the transport of sodium ions across the membrane. It belongs to the sodium ion transport decarboxylase enzyme family found in anaerobic bacteria, and shares no similarity to either of the soluble bacterial OAD families. Interestingly, those identified OADs have a broad range of specific activity between species. For

example, OAD from *Corynebacterium glutamicum* has a k_{cat} value of 78 s^{-1} (Jetten and Sinskey 1995) in comparison with the k_{cat} value of 7500 s^{-1} from OAD in *Pseudomonas aeruginosa* (Narayanan et al. 2008).

It is relevant to mention that oxaloacetate is also decarboxylated at the beginning of the gluconeogenesis pathway, but in this case with concomitant phosphorylation to yield PEP. The reactions are catalysed by PEP carboxykinase (PEPCK), and by coincidence provide another example of an enzyme reaction that involves two distinct cytosolic and mitochondrial enzymes. There are thus at least five distinct enzyme families that are able to catalyse the decarboxylation of oxaloacetate, including pyruvate kinase (PYK) as described in this chapter.

Over 30 years ago, PYKs from codfish muscle (Creighton and Rose 1976a; Creighton and Rose 1976b) and rabbit muscle (Creighton and Rose 1976a; Jursinic and Robinson 1978) were reported to show an inherent oxaloacetate decarboxylase activity (**Figure 7.1b)**, with k_{cat} values of 0.95 s^{-1} and 1.68 s^{-1}, respectively. The observed decarboxylase activity was suggested to share a common or overlapping reaction site with the kinase activity (Creighton and Rose 1976a; Creighton and Rose 1976b; Jursinic and Robinson 1978).

PYKs are generally well conserved, with enzymes from even rather evolutionarily distant organisms showing 42-49% sequence identity (**Figure 7.2**; **Table 7.1**). In particular, the active sites of PYKs are highly conserved between species. To date no structural information from PYK in complex with the substrate oxaloacetate has been reported to provide further support for the idea of a common or overlapping active site for the decarboxylase and kinase activities of PYK.

However a number of observations support the hypothesis that these two active sites overlap. 1. Both decarboxylase and kinase activities show sigmoidal kinetics for their substrates in codfish muscle PYK (Creighton and Rose 1976b); 2. PEP which is the kinase substrate competitively inhibits decarboxylase activity of PYKs from codfish muscle and rabbit muscle (Creighton and Rose 1976a); 3. The substrate (PEP

or oxaloacetate) analogues oxalate, phosphoenol-α-ketobutyrate and 4-ethyloxalacetate competitively inhibit both activities (Creighton and Rose 1976a; Jursinic and Robinson 1978); 4. Phenylalanine is an allosteric inhibitor for both activities in rabbit muscle PYK, and exhibits similar kinetics (Jursinic and Robinson 1978); F16BP which is an allosteric effector for many PYKs also allosterically regulates decarboxylase activity of codfish muscle PYK and increases the velocity of the decarboxylase reaction catalysed by rabbit muscle PYK (Creighton and Rose 1976a; Jursinic and Robinson 1978); 5. Both activities from rabbit muscle PYK are inhibited by the same covalent modifiers which target the active site (e.g. dithionitrobenzoate) (Jursinic and Robinson 1978); 6. Divalent metal ions are shown to favour both activities (Creighton and Rose 1976a; Jursinic and Robinson 1978); 7. Pyruvate kinase stabilises enolpyruvate which is considered to be an intermediate in both kinase and decarboxylase activities (Robinson and Rose 1972; Creighton and Rose 1976a).

This chapter describes the application of ^{1}H NMR spectroscopy to observe the oxaloacetate decarboxylase activity in PYKs including allosteric (trypanosomatid) and non-allosteric (human M1) PYKs. Furthermore, crystal soaking enabled the determination of the first product (pyruvate)-bound structure of *Tb*PYK as well as a series of oxaloacetate analogue-bound structures of *Tb*PYK. These structures reveal the similar substrate recognition of the active site architecture for both activities, and for the first time propose a structure-based catalytic mechanism for the decarboxylase activity of PYKs. The work from this chapter forms the basis of a publication:

Zhong, W., *et al.* Pyruvate kinases have an inherent and conserved decarboxylase activity. (Manuscript in preparation)

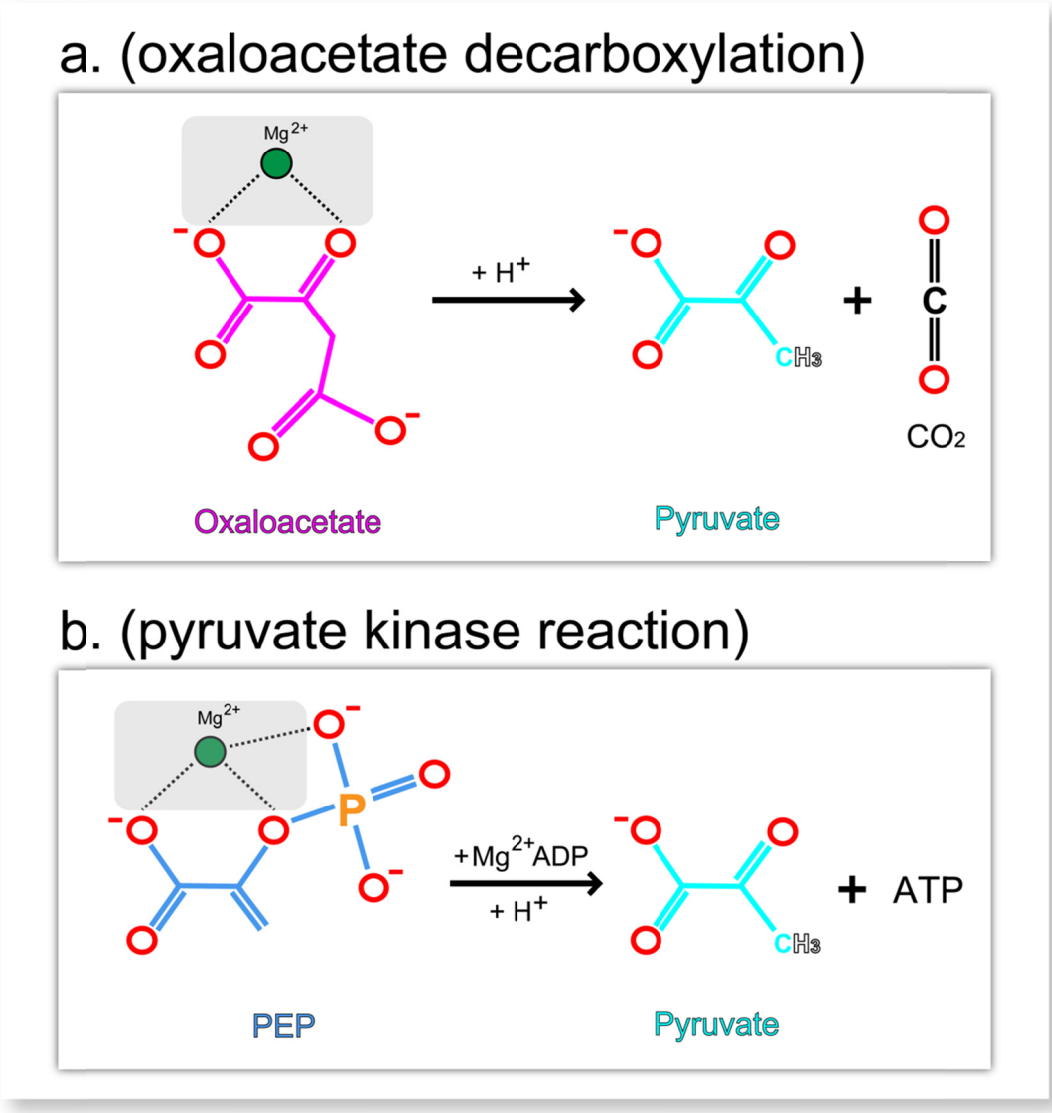


Figure 7.1 Schematic representations of the oxaloacetate decarboxylation and kinase reactions catalysed by pyruvate kinase.
The divalent metal ion Mg^{2+}, which is a common metal ion facilitating the decarboxylation of oxaloacetate and is also required for kinase activity in PYKs, is shown as a green sphere. The coordination between Mg^{2+} and oxaloacetate, and Mg^{2+} and PEP are indicated as dashed lines.

Table 7.1 Pairwise protein sequence comparisons of trypanosomatid, human M2, rabbit M1, yeast and *E. coli* PYKs

	*Tb*PYK	*Tc*PYK	*Lm*PYK	*Hs*M2PYK	*Hs*M1PYK	*Oc*M1 PYK	*Sc*PYK	*Ec*PYK
*Tb*PYK	100	81	74	48	47	48	48	42
*Tc*PYK		100	76	47	47	47	48	42
*Lm*PYK			100	47	48	47	48	42
*Hs*M2PYK				100	96	93	49	44
*Hs*M1PYK					100	97	48	44
*Oc*M1PYK						100	47	43
*Sc*PYK							100	43
*Ec*PYK								100

The pairwise sequence analysis was obtained from the EMBL-EBI web server (EMBOSS Stretcher): http://www.ebi.ac.uk/Tools/psa/emboss_stretcher/
Values are overall percent sequence identities.
Tb, Trypanosoma brucei; Tc, T. cruzi; Lm, Leishmania mexicana; Hs, Homo sapiens; Oc, Oryctolagus cuniculus, (rabbit), *Sc, Saccharomyces cerevisiae; Ec, Escherichia coli*.

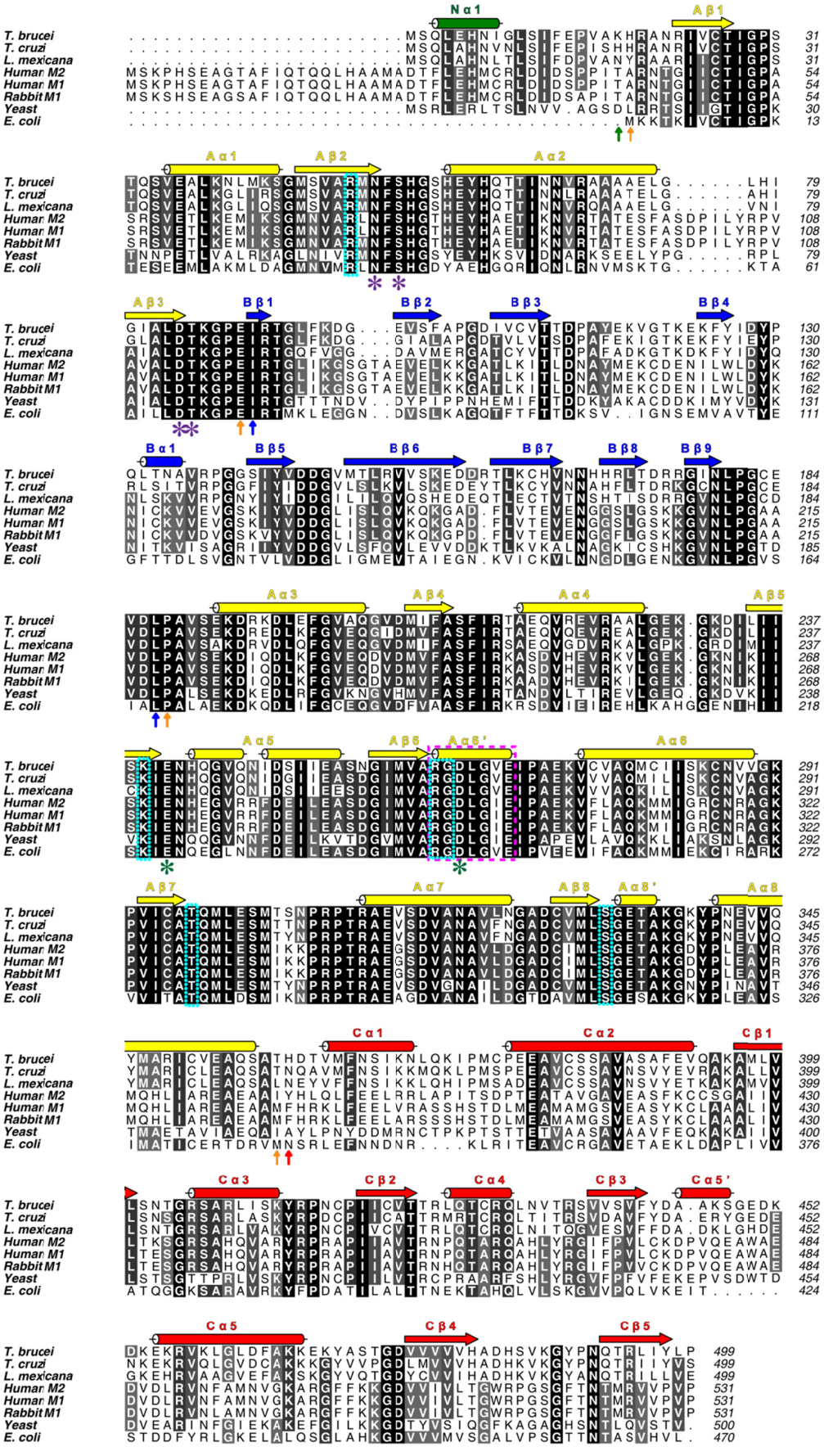

Nα1
Aβ1
T. brucei
T. cruzi
L. mexicana
Human M2
Human M1
Rabbit M1
Yeast
E. coli
Aα1
Aβ2
Aα2
Aβ3
Bβ1
Bβ2
Bβ3
Bβ4
Bα1
Bβ5
Bβ6
Bβ7
Bβ8
Bβ9
Aα3
Aβ4
Aα4
Aβ5
Aα5
Aβ6
Aα6'
Aα6
Aβ7
Aα7
Aβ8
Aα8'
Aα8
Cα1
Cα2
Cβ1
Cα3
Cβ2
Cα4
Cβ3
Cα5'
Cα5
Cβ4
Cβ5

Figure 7.2 Sequence alignment of pyruvate kinases from *T. brucei, T. cruzi, L. mexicana,* human M2, human M1, rabbit M1, *S. cerevisiae* and *E. coli*.
The sequence alignment was performed using the program Clustal Omega at the European Bioinformatics Institute (Sievers et al 2011; Goujon et al. 2010). Secondary structural elements defined in *Tb*PYK/F26BP/pyruvate/Mg by DSSP (Kabsch and Sander 1983; Joosten et al. 2011) are shown above the sequences (only α-helices and β-strands are shown). Secondary structural elements are labelled in different colours corresponding to their domain regions: N-terminal domain (green), A-domain (yellow), B-domain (blue) and C-domain (red). Domain boundaries are indicated by vertical arrows in domain-specific colours. The conservation of the residues is indicated by shading from black (identical in seven or eight sequences) to grey (conserved in five or six) to white (low or no conservation). Residue numbers corresponding to each PYK are listed after the sequences. In *Tb*PYK, the amino acids involved in divalent metal binding (PEP or pyruvate-coordinating metal, Mg-1 site) (*) and potassium metal ion binding (*) are indicated by asterisks. Residues involved in PEP, pyruvate or oxaloacetate analog binding are indicated by dotted boxes in cyan. Residues 263-269 of the small α-helix Aα6′ which are involved in allosteric regulation and in binding divalent metal and the substrate PEP or product pyruvate are indicated by a dashed box (pink). The figure was generated using the program Aline (Bond and Schüttelkopf 2009).

7.2. Materials and Methods

The materials and the general methods used in this chapter have been described in **Chapter 2**. In brief, the structures of *Tb*PYK in complex with pyruvate, D-malate, α-ketoglutarate or oxalate, were solved by the methods described in **Section 4** of **Chapter 2**, and the structure statistics are shown in **Table 7.2**. The ^{1}H NMR assay was carried out by the method in section **2.4.8**.

Table 7.2 Data collection, refinement and Ramachandran plot statistics

	*Tb*PYK/ F26BP/pyruvate/Mg	*Tb*PYK/ F26BP/D-malate/Mg	*Tb*PYK/ F26BP/α-ketoglutarate/Mg	*Tb*PYK/ F26BP/oxalate/Mg
Data collection				
Space group	*I*222	*I*222	*I*222	*I*222
Cell dimensions				
a, b, c (Å)	102.99,108.17,264.49	103.37,108.27,265.19	103.25,108.01,265.82	103.88,108.56,264.53
α, β, γ (°)	90.00,90.00,90.00	90.00,90.00,90.00	90.00,90.00,90.00	90.00,90.00,90.0
Solvent content (%)	64	64	64	64
Wavelength (Å)	0.98	0.98	0.98	0.98
Resolution	74.59-1.95	74.77-2.35	74.63-2.18	74.95-2.35
No. reflections	712620(78336)	418584	466551	427562
No. unique reflections	107112(15148)	62179	77484	62638
Wilson B-factor (Å^2)	20.0	36.70	24.41	38.8
Rmerge (%)	9.0 (47.0)	11.9 (70.3)	11.4 (40.9)	14.8 (94.0)
$\langle I/\sigma I \rangle$	12.7 (3.6)	10.9 (3.0)	11.1 (3.9)	10.2 (2.5)
Rmeas (%) (within I+/I-)	10.6 (57.5)	14.0 (83.3)	13.8 (50.8)	17.5 (110.4)
Rmeas (%) (all I+ & I-)	10.4 (56.0)	13.6 (80.3)	13.6 (50.3)	17.2 (108.7)
Rpim (%) (within I+/I-)	5.5 (32.4)	7.3 (43.8)	7.6 (29.6)	9.1 (57.2)
Rpim (%) (all I+ & I-)	4.0 (23.7)	5.2 (31.1)	5.5 (21.9)	6.5 (40.6)
% completeness	99.6 (97.5)	99.9 (99.7)	100.0 (99.9)	100.0 (100.0)
Multiplicity	6.7 (5.2)	6.7 (6.4)	6.0 (5.1)	6.8 (7.0)
Refinement				
Monomers in ASU	2	2	2	2
No. reflections	101757(78336)	418584 (59028)	466551 (73586)	427562 (59406)
R_{work} / R_{free}	15.26/18.25	16.13/20.34	15.72/19.38	16.91/21.51
No. of nonhydrogen atoms				
Protein	7663	7670	7670	7670
Water	785	431	751	442
Ligands	75	63	82	68
Average B-factor (Å^2)				
Overall	27.25	37.70	26.38	38.74
Overall (exclude B-domain of chain A)	25.39	33.94	24.51	35.08
Protein	26.38	37.64	25.70	38.96
Protein (exclude B-domain of chain A)	24.18	33.61	23.53	35.05
B-domain/Chain A	45.93	73.67	45.05	73.78
B-domain/Chain B	32.15	45.38	32.35	59.81
Water	35.83	36.64	32.88	35.78
Ligands	27.10	37.58	29.62	33.55
MG	22.89	38.07	36.56	36.93
Analog	23.88	38.15	39.62	40.15
No. Residues (protein)	996	996	996	996
No. Ligands	12	9	11	10
No. Water	785	431	751	442
R.m.s deviations				
Bond lengths (Å)	0.0123	0.0114	0.0110	0.0118
Bond angles (degree)	1.1669	1.1021	1.0749	1.0866
Ramachandran plots				
Favored (%)	97.5	96.9	97.2	97.2
Allowed (%)	99.7	99.8	99.8	99.8
No. of outliers †	3	2	2	2

Values in parentheses are for the highest resolution shell.

† The outlier residue in each monomer is Thr296, a key active-site residue that is commonly found in this configuration in PYK structures. Asp97 in *Tb*PYK/F26BP/pyruvate/Mg is an outlier residue which is modelled in the poor electron density.

7.3. Results and Discussion

7.3.1. Oxaloacetate decarboxylase activity of PYKs from trypanosomatids and humans

^{1}H NMR spectroscopy provides a convenient method to directly measure the conversion of oxaloacetate to pyruvate. The increase of the unique single peak on the NMR spectrum from the methyl group on pyruvate was monitored over time to give a conversion rate (**Figure 2.3** in **Chapter 2**).

The spontaneous decarboxylation rate of oxaloacetate under assay conditions was also measured by NMR to give a value of 0.162 μmol/min (**Table 7.3**). Under the experimental conditions described in section **2.4.8** (**Chapter 2**), *Tb*PYK and human M1PYK are shown to have oxaloacetate decarboxylase activities of 0.857 s^{-1} and 0.512 s^{-1}, respectively (**Table 7.3**). Interestingly, the addition of F26BP which is the natural activator of kinase activity also caused an increase in the decarboxylase activity of *Tb*PYK to 1.494 s^{-1}. These numbers are comparable to the reported values of 0.95 s^{-1} for codfish muscle PYK (Creighton and Rose 1976b) and 1.68 s^{-1} for rabbit muscle PYK (Creighton and Rose 1976a; Jursinic and Robinson 1978). Rabbit muscle LDH was used in the assay as a negative control showing no decarboxylase activity under the same condition as PYKs (**Table 7.3**).

Overall, PYKs from different species enhance the rate of the decarboxylation of oxaloacetate. As shown in **Figure 7.3**, for example, in the presence of 0.36 mg/ml *Tb*PYK the rate is enhanced by 2-fold over the spontaneous decarboxylation rate.

Furthermore, oxalate which is an analogue of enolpyruvate and a general competitive inhibitor for the kinase activity of PYKs was shown to inhibit decarboxylase activity of PYKs as well, where *Tb*PYK and human M1PYK were almost fully inhibited (**Table 7.3**; **Figure 7.3**). Additionally, oxalate at both 2 mM and 20 mM also partially inhibits the spontaneous decarboxylation of oxaloacetate (**Table 7.3**; **Figure 7.3**) possibly due to the chelation of Mg^{2+} by oxalate. This

obvious inhibition of decarboxylase activity of both *Tb*PYK and human M1PYK by oxalate suggests that oxalate either competitively binds to the oxaloacetate binding site, and/or is acting as a divalent metal Mg^{2+} chelator thus inhibiting the reaction.

Table 7.3 Conversion rates of oxaloacetate decarboxylation

	Conversion rate (μmol/min)		**References**
Spontaneous decarboxylation	0.162±0.010		This work [b]
Spontaneous decarboxylation +2 mM oxalate	0.141±0.004		
Spontaneous decarboxylation + 20 mM oxalate	0.089±0.014		
	Conversion rate [a] **(μmol/min/mg)**	**k_{cat} (s^{-1})**	
***Tb*PYK**	0.944±0.035	0.857±0.032	
***Tb*PYK**+18 μM F26BP	1.645±0.107	1.494±0.097	
***Tb*PYK**+20 mM oxalate	0.046±0.009	0.042±0.008	
Human M1PYK	0.510±0.045	0.512±0.045	
Human M1PYK + 2 mM oxalate	0.0625±0.027	0.063±0.027	
Rabbit muscle **LDH**	-0.005±0.008	-0.003±0.005	
Codfish muscle PYK	0.90	0.95[e]	Creighton and Rose 1976b [c]
Rabbit muscle PYK	1.7	1.68[e]	Jursinic and Robinson 1978 [d]

The values given are means ± SD.

[a] Conversion rate equates to the rate of production of pyruvate from oxaloacetate by decarboxylation. The spontaneous rate of oxaloacetate decarboxylation in solution has been subtracted.

[b] ^{1}H NMR assay conditions in this work (at 25^{o}C): 50 mM TEA (pH 7.2), 10 mM $MgCl_2$, 50 mM KCl, 50 mM oxaloacetate, 10% D_2O, 5 mM DMSO.

[c] LDH-coupled assay conditions: 50 mM TEA (pH7.6), 1.5 mM $MnSO_4$, 0.1 mM NADH, 3 U/ml LDH.

[d] LDH-coupled assay conditions: 100 mM TES (pH7.5), 1 mM $MgCl_2$, 1 mM FBP, 150 mM KCl, 0.5 mM NADH, 15 U/ml LDH.

[e] The number was calculated by using molecular masses of 63000 (codfish muscle PYK) and 59250 (rabbit muscle PYK), respectively.

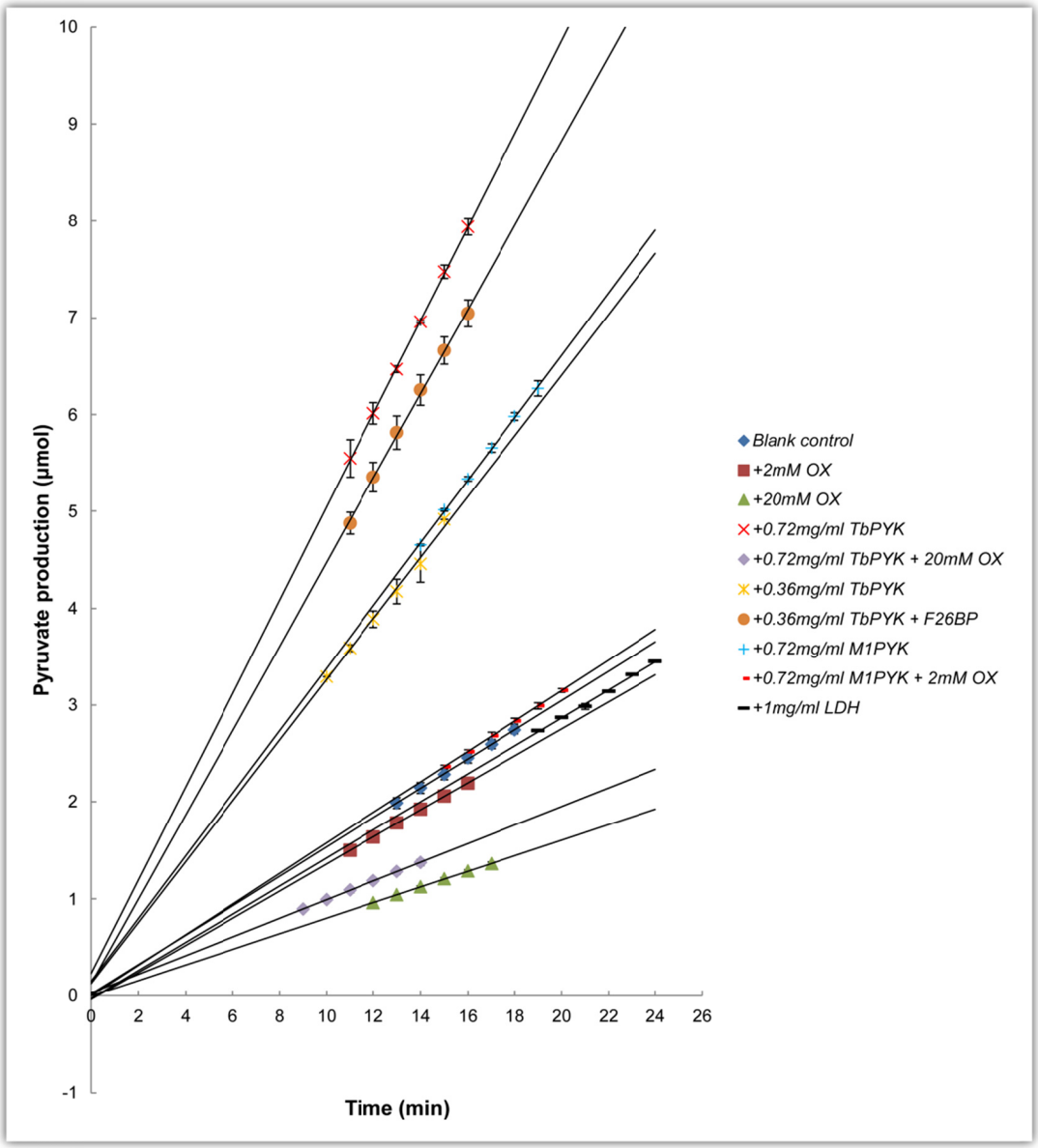


Figure 7.3 Pyruvate production from oxaloacetate under different experimental conditions.

Experiments were performed to monitor pyruvate production (-CH_3 group) in a 0.5 ml-system over time using a ^{1}H-NMR technique at 25^{o}C. The buffer conditions contained 50 mM TEA (pH7.2), 10 mM $MgCl_2$, 50 mM KCl, and 50 mM oxaloacetate. Three enzymes (human M1PYK, *Tb*PYK and rabbit muscle LDH) were added to the assay conditions individually to study their decarboxylase activity. 'Blank control' is the sample without adding enzyme. The activator F26BP was added to study its potential activation on the decarboxylase activity of *Tb*PYK. The substrate analogue oxalate (2 mM or 20 mM) was used to study its inhibition of the spontaneous decarboxylation of oxaloacetate or the enzyme catalysed decarboxylation. Error bars were calculated to represent the standard deviation but some of them are too small to show up.

7.3.2. First structure of the product pyruvate-bound TbPYK

7.3.2.1. Crystal soaking with oxaloacetate resulted in *Tb*PYK/F26BP/pyruvate/Mg

A series of crystal soaking and co-crystallisation experiments were set up to try to obtain a structure of *Tb*PYK with oxaloacetate as 'substrate'. In all soaking experiments only pyruvate, the product of oxaloacetate decarboxylation, could be identified at the active site (**Figure 7.4b**). The *Tb*PYK/F26BP/pyruvate/Mg X-ray structure has been refined at 1.95 Å resolution, and contains two protein chains in the asymmetric unit. The average RMS difference between the C^{α} atoms of the AC-cores (A- and C-domains) in the asymmetric unit is 0.22 Å (**Table 7.4**). The structure has a similar tetrameric architecture to the published *Tb*PYK/F26BP/Mg and PEP-bound *Tb*PYK/F26BP/PEP/Mg structures (**Chapter 6**; **Figure 7.4a**). The average C^{α} RMS difference between these three R-state *Tb*PYK tetramers (excluding the flexible B-domains) is less than 0.2 Å (**Table 7.5**). Similar to the substrate PEP-bound *Tb*PYK/F26BP/PEP/Mg, this newly determined *Tb*PYK/F26BP/pyruvate/Mg adopts two crystallographically independent B-domains showing different orientations (22^{o}) as a result of pyruvate binding to the active site (**Figure 7.5; Table 7.6**). The rotation of the B-domain regulates the binding of pyruvate in a manner similar to that for the binding of the substrate PEP (**Chapter 6**).

7.3.2.2. Active site of *Tb*PYK/F26BP/pyruvate/Mg

The active site of *Tb*PYK is located in the cleft between A- and B-domains **(Figure 7.4a)**. The product pyruvate is clearly identified in the structure (**Figure 7.4b**) and adopts an almost identical binding mode compared to the binding mode of the substrate PEP in *Tb*PYK/F26BP/PEP/Mg (**Chapter 6**), except that the phospho group of PEP has been replaced by water molecules in the pyruvate-bound structure for the coordination with the divalent metal Mg^{2+}. Additionally, the binding mode of

pyruvate in *Tb*PYK obtained from crystal soaking is almost identical to the pyruvate binding mode in co-crystallised PYK structures (Larsen et al. 1994; Wooll et al. 2001; Williams et al. 2006; Fenton et al. 2010).

Pyruvate binds in a highly polar cavity which is comprised by protein residues and metal ions (Mg^{2+} and K^{+}). Mg^{2+} is ligated to pyruvate at the active site in the Mg-1 position, which differs from the Mg-3 position in *Tb*PYK/F26BP/Mg determined from the crystal without pyruvate soaking (**Chapter 6**). The carboxyl group of pyruvate forms interactions with the short α-helix Aα6', the hydroxyl group of Thr293 and the divalent metal Mg^{2+} (**Figure 7.4b** and **7.4c).** The carbonyl oxygen atom C(2)O of pyruvate coordinates with the positively charged side chain of Lys239, Mg^{2+} and one water molecule (which also coordinates with K^{+}). The distorted octahedral coordination sphere of Mg^{2+} is formed by the two side-chain carboxylate oxygen atoms of Asp265 and Glu241, two oxygen atoms from the carboxyl group and carbonyl group of pyruvate, respectively, and two water molecules. The coordination sphere of K^{+} has distorted octahedral geometry as well, and is formed by three side-chain oxygen atoms from residues Asn52, Ser54 and Asp84, one main-chain carbonyl oxygen atom from Thr85, and two water molecules. The communication between Mg^{2+} and K^{+} is mediated by water molecules.

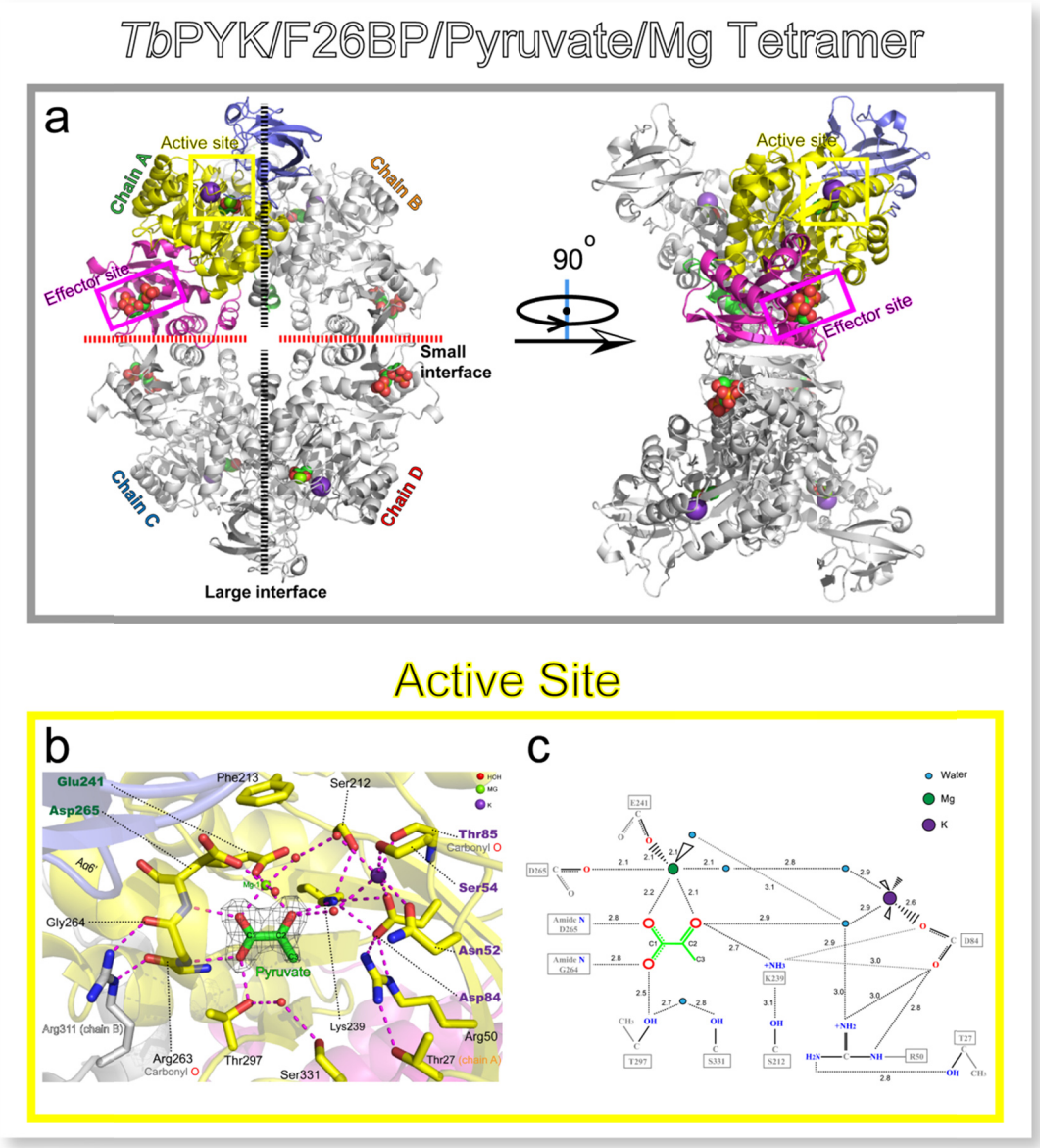


Figure 7.4 Structure of *Tb*PYK/F26BP/pyruvate/Mg showing the tetramer architecture, domain boundaries and the active site binding sphere.

(a) Two orthogonal views of *Tb*PYK/F26BP/pyruvate/Mg showing the tetramer architecture, active site and effector site. The polypeptide chain is shown as a cartoon. The large (A-A) and small (C-C) interfaces between subunits are shown as dashed lines. Each subunit contains four domains and one subunit is coloured to show domains: N-terminal (green = residues 2-18), A-domain (yellow = residues 19-89, 188-358), B-domain (blue = residues 90-187), C-domain (red = residues 359-499). Metals and ligands are represented by spheres. (b) Close-up of the active site of *Tb*PYK/F26BP/pyruvate/Mg, chain A showing the pyruvate binding mode. The polypeptide chain is shown as a cartoon, while the residues involved in pyruvate binding are shown as sticks. Pyruvate is shown as sticks with an unbiased *Fo-Fc* election density (grey) map contoured at 5.0 σ. Water molecules (red), Mg^{2+} (green)

and K^+ (purple) are shown as spheres. Potential interactions involved in pyruvate binding are indicated by dashed lines. (c) A schematic drawing showing the estimated interactions at the *Tb*PYK active site. The interatomic distances for the interactions are given in Ångstroms.

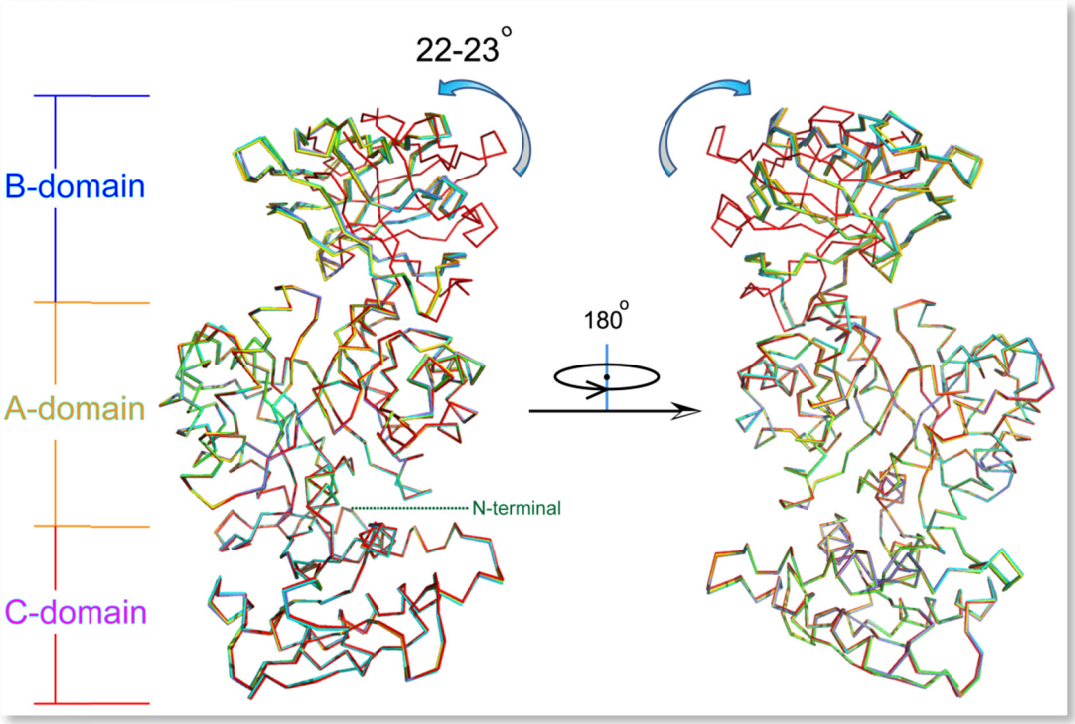


Figure 7.5 B-domain movement is initiated by active site ligand binding and is locked in a partially closed conformation.
Two views of the superposed monomers (chain A) from *Tb*PYK/F26BP/Mg (red; PDB ID: 4HYW), *Tb*PYK/F26BP/PEP/Mg (orange; PDB ID: 4HYV), *Tb*PYK/F26BP/pyruvate/Mg (green); *Tb*PYK/F26BP/D-malate/Mg (yellow), *Tb*PYK/F26BP/α-ketoglutarate/Mg (blue) and *Tb*PYK/F26BP/oxalate/Mg (cyan). The superposition was performed on the A/C/N-terminal domains (residues 2-89, 188-499) of the complex structures in monomeric form. The B-domain from active-site ligand-bound structures (*Tb*PYK/F26BP/PEP/Mg, *Tb*PYK/F26BP/pyruvate/Mg; *Tb*PYK/F26BP/D-malate/Mg, *Tb*PYK/F26BP/α-ketoglutarate/Mg and *Tb*PYK/F26BP/oxalate/Mg) experienced a domain rotation of 22-23° compared to *Tb*PYK/F26BP/Mg. Polypeptide chains are shown as ribbons. The B-domain rotation function was analysed by the program CCP4 superpose (Potterton et al. 2002).

Table 7.4 Average RMS differences between asymmetric units (chain A/chain B)

	C^{α} atoms		All atoms	
	Tetramer	AC-core	Tetramer	AC-core
*Tb*PYK/F26BP/Mg †	0.40 Å	0.18 Å	0.51Å	0.27 Å
*Tb*PYK/F26BP/PEP/Mg ‡	1.91 Å	0.21 Å	2.02 Å	0.32 Å
*Tb*PYK/F26BP/pyruvate/Mg	1.90 Å	0.22 Å	2.01 Å	0.31 Å
*Tb*PYK/F26BP/D-malate/Mg	1.94 Å	0.22 Å	2.05 Å	0.34 Å
*Tb*PYK/F26BP/α-ketoglutarate /Mg	1.90 Å	0.22 Å	2.03 Å	0.36 Å
*Tb*PYK/F26BP/oxalate/Mg	1.82 Å	0.22 Å	1.94 Å	0.31 Å

† The PDB ID for *Tb*PYK/F26BP/Mg is 4HYW.

‡ The PDB ID for *Tb*PYK/F26BP/PEP/Mg is 4HYV.

Table 7.5 Average RMS differences between *Tb*PYK tetramers

	*Tb*PYK/F26BP/PEP/Mg ‡			
	C^{α} atoms		All atoms	
	Tetramer	AC-core	Tetramer	AC-core
*Tb*PYK/F26BP/Mg †	1.26 Å	0.16 Å	1.37Å	0.26 Å
*Tb*PYK/F26BP/pyruvate/Mg	0.20 Å	0.11 Å	0.26 Å	0.17 Å
*Tb*PYK/F26BP/D-malate/Mg	0.18 Å	0.11 Å	0.25 Å	0.17 Å
*Tb*PYK/F26BP/α-ketoglutarate /Mg	0.22 Å	0.13 Å	0.33 Å	0.24 Å
*Tb*PYK/F26BP/oxalate/Mg	0.15 Å	0.12 Å	0.23 Å	0.18 Å

† The PDB ID for *Tb*PYK/F26BP/Mg is 4HYW.

‡ The PDB ID for *Tb*PYK/F26BP/PEP/Mg is 4HYV.

Table 7.6 '*In crystallo*' B-domain movements triggered by the binding of active-site ligands [a]

	*Tb*PYK/F26BP/Mg [†]	
	Rotation angles	RMSD (C^{α}) [b]
*Tb*PYK/F26BP/PEP/Mg [‡]	23°	0.712 Å
*Tb*PYK/F26BP/pyruvate/Mg	22°	0.712 Å
*Tb*PYK/F26BP/D-malate/Mg	22°	0.708 Å
*Tb*PYK/F26BP/α-ketoglutarate /Mg	22°	0.683 Å
*Tb*PYK/F26BP/oxalate/Mg	22°	0.701 Å

[a] The rotation angles of the B-domain were calculated by comparing the positions of the B-domains of *Tb*PYK/F26BP/Mg (chain A) with the B-domains of each of the active-site ligand bound structures (chain A). The detailed procedures can be found in **Materials and Methods** (section **2.4.14** in **Chapter 2**).

[b] RMSD (for all C^{α} atoms) between B-domains of two structures after rotation movements

[†] The PDB ID for *Tb*PYK/F26BP/Mg is 4HYW.

[‡] The PDB ID for *Tb*PYK/F26BP/PEP/Mg is 4HYV.

7.3.2.3. Effector site of *Tb*PYK/F26BP/pyruvate/Mg

The C-domain of *Tb*PYK habours the effector site for F26BP binding (**Figure 7.4a**). Comparison of the effector sites of *Tb*PYK/F26BP/Mg and *Tb*PYK/F26BP/pyruvate/Mg (**Figure 7.6**) shows that the effector site residues and the binding mode of F26BP remain unchanged before and after crystal soaking with pyruvate.

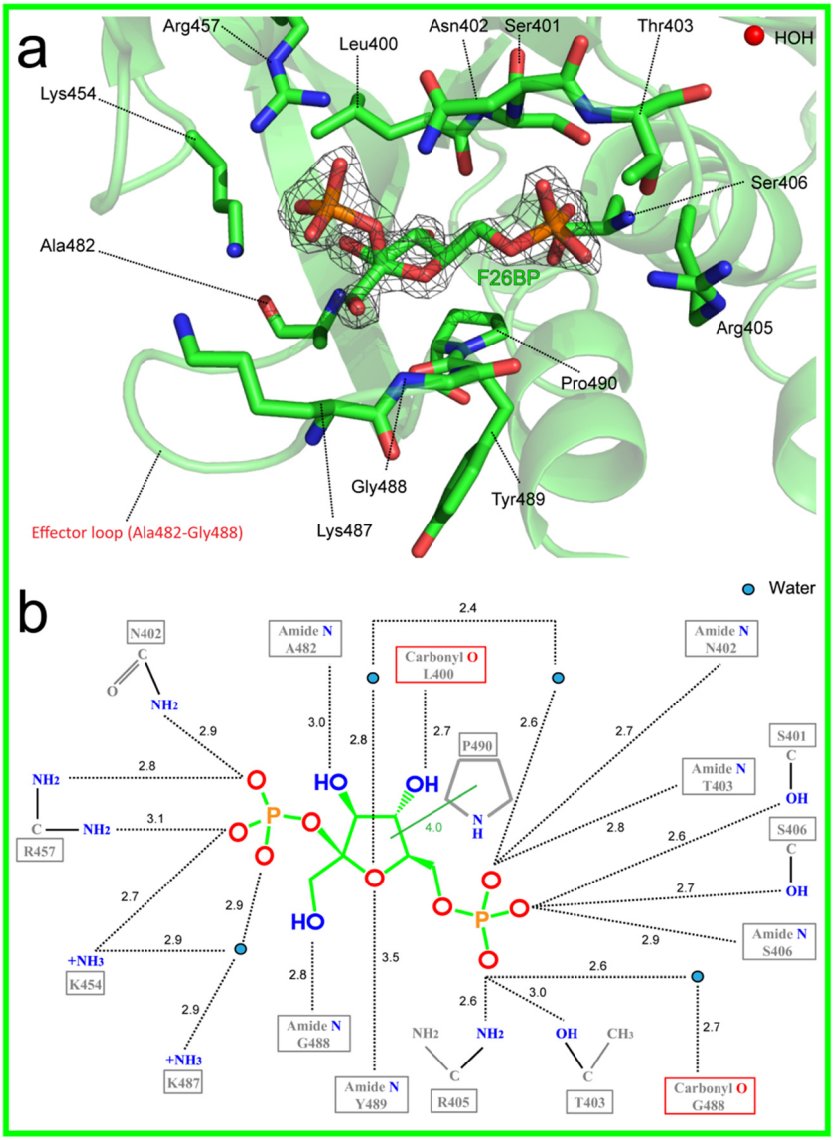

a
Arg457
Leu400
Asn402
Ser401
Thr403
HOH
Lys454
Ser406
Ala482
F26BP
Arg405
Pro490
Gly488
Tyr489
Effector loop (Ala482-Gly488)
Lys487
b
Water
N402
Amide N A482
Carbonyl O L400
Amide N N402
P490
S401
Amide N T403
S406
Amide N S406
R457
K454
K487
Amide N G488
Amide N Y489
R405
T403
Carbonyl O G488
c
Arg457
Leu400
Asn402
Ser401
Thr403
Lys454
Ser406
Ala482
F26BP
Arg405
Pro490
Gly488
Tyr489
Effector loop (Ala482-Gly488)
Lys487

Figure 7.6 Identical F26BP binding sites from *Tb*PYK complex structures before and after soaking with active-site ligands.
(a) Close-up of the effector site of *Tb*PYK (*Tb*PYK/F26BP/pyruvate/Mg, chain A) showing the F26BP binding mode. The polypeptide chain is shown as a cartoon, while the residues involved in F26BP binding are shown as sticks. F26BP is shown as sticks with an unbiased *Fo-Fc* election density (grey) map contoured at 8.0 σ. Water molecules are shown as red spheres. Potential interactions involved in F26BP binding are indicated by dashed lines. (b) A schematic drawing showing the estimated interactions at the *Tb*PYK effector site. The interatomic distances for the interactions are given in Ångstroms. (c) The effector sites from six *Tb*PYK complex structures were superposed onto each other: *Tb*PYK/F26BP/Mg (white; PDB ID: 4HYW), *Tb*PYK/F26BP/PEP/Mg (orange; PDB ID: 4HYV), *Tb*PYK/F26BP/pyruvate/Mg (green); *Tb*PYK/F26BP/D-malate/Mg (yellow), *Tb*PYK/F26BP/α-ketoglutarate/Mg (blue) and *Tb*PYK/F26BP/oxalate/Mg (cyan). F26BP and the residues involved in F26BP binding are shown as sticks. The effector loop (residues 482-488) is labelled. The side chain of Arg405 is relatively flexible even between asymmetrically unique chains in the structure.

7.3.3. Two oxaloacetate analogues bind to TbPYK at the kinase active site

Two oxaloacetate analogues, D-malate and α-ketoglutarate, were soaked into crystals of *Tb*PYK/F26BP/Mg and two structure complexes were determined: *Tb*PYK/F26BP/D-malate/Mg and *Tb*PYK/F26BP/α-ketoglutarate/Mg.

7.3.3.1. Overall structures of analogue-bound *Tb*PYK

*Tb*PYK/F26BP/D-malate/Mg and *Tb*PYK/F26BP/α-ketoglutarate/Mg have been refined at 2.35 Å and 2.18 Å, respectively, and contain two monomers (forming a biologically relevant dimer) in the asymmetric unit. These two asymmetrically unique protein chains in each structure were superposed onto each other showing almost identical overall AC-core structures (RMS difference for all C^{α} atoms of AC-cores = 0.22 Å), but differ in B-domain orientations (**Table 7.4; Figure 7.5**). Tetramer of each structure (generated from the crystallographic symmetry related dimer) is similar to both *Tb*PYK/F26BP/Mg and *Tb*PYK/F26BP/PEP/Mg in

quaternary structure excluding the mobile B-domains (**Table 7.5**). All four of the active-site ligand-bound structures obtained by crystal soaking (*Tb*PYK/F26BP/D-malate/Mg, *Tb*PYK/F26BP/α-ketoglutarate/Mg, *Tb*PYK/F26BP/pyruvate/Mg and *Tb*PYK/F26BP/PEP/Mg) are almost identical in tetrameric structure including the B-domains, as the B-domains from these structures all manifest the same domain movement induced by active-site ligand binding (**Table 7.5** and **7.6; Figure 7.5**).

7.3.3.2. Analogues bind to the kinase active site of *Tb*PYK

D-Malate and α-ketoglutarate are clearly identified in both of the crystallographically independent subunits of *Tb*PYK/F26BP/D-malate/Mg and *Tb*PYK/F26BP/α-ketoglutarate/Mg, respectively (**Figure 7.7a** and **7.7c**). Additionally, the binding mode for each analogue is similar to the binding mode of PEP (**Figure 7.7g** and **7.7h**) and pyruvate, which are the natural substrate and product for PYK, respectively (**Figure 7.8**). The similar binding poses of the analogues of the decarboxylase substrate (D-malate and α-ketoglutarate) and the kinase ligands (pyruvate and PEP) confirm that the kinase and decarboxylase activities share the same binding site,

In particular, the carboxyl group C(1)OO of D-malate and α-ketoglutarate coordinate with the short α-helix Aα6', the hydroxyl group of Thr293 and the divalent metal Mg^{2+} in an identical way to that observed in structures of PYK complexed with PEP or oxalate (**Figure 7.7**). D-malate and α-ketoglutarate have different geometries at *C*2 caused by sp3 and sp2 hybridisation states. Despite the difference in geometry the hydroxyl (2)OH oxygen atom of D-malate and the carbonyl C(2)O oxygen atom of α-ketoglutarate form similar interactions with Mg^{2+}, the side chain of Lys239 and one water molecule which coordinates with K^+ (**Figure 7.7a, 7.7b, 7.7c** and **7.7d**). Both carboxyl C(4)OO oxygen atoms of D-malate and the carboxyl C(5)OO oxygen atom of α-ketoglutarate interact with the hydroxyl oxygen

atom of Ser331. Additionally, the D-malate carboxyl group C(4)OO engages in an ionic interaction with the guanidinium group of Arg50. However, the α-ketoglutarate C(5)OO carboxylate flips away from Arg50 compared to the position of the D-malate carboxyl group C(4)OO. The coordination sphere of Mg^{2+} and K^{+} in *Tb*PYK/F26BP/D-malate/Mg and *Tb*PYK/F26BP/α-ketoglutarate/Mg are almost identical and similar to those observed in *Tb*PYK/F26BP/pyruvate/Mg and *Tb*PYK/F26BP/PEP/Mg.

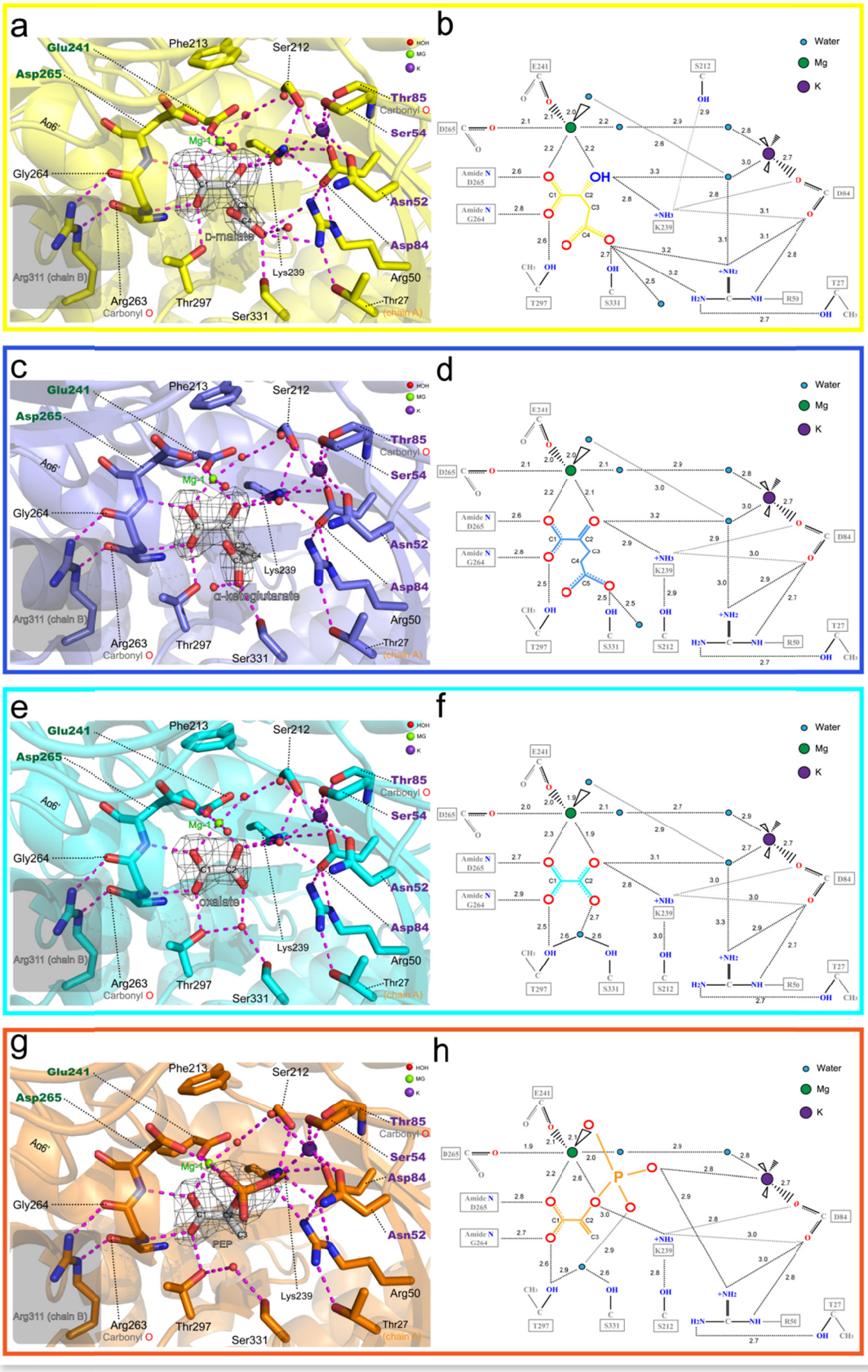

a
Glu241
Phe213
Ser212
Asp265
Thr85
Carbonyl O
Ser54
Aa6'
Mg-1
Gly264
Asn52
D-malate
Asp84
Arg311 (chain B)
Lys239
Arg50
Arg263
Carbonyl O
Thr297
Ser331
Thr27
(chain A)
HOH
MG
K
b
Water
Mg
K
c
α-ketoglutarate
d
e
oxalate
f
g
PEP
h

Figure 7.7 Similar binding modes of oxaloacetate analogues (D-malate, α-ketoglutarate or oxalate) or the substrate PEP at the active site of *Tb*PYK.
(a, c, e, g) Close-up the active site of *Tb*PYK in complex with active site ligands D-malate (a), α-ketoglutarate (c), oxalate (e) or PEP (g), respectively. The polypeptide chain is shown as a cartoon, while the residues involved in active-site ligand binding are shown as sticks. Active-site ligands are shown as sticks with an unbiased *Fo-Fc* election density (grey) map contoured at 3.5 σ (D-malate), 4.5 σ (α-ketoglutarate), 5.0 σ (oxalate) and 5.0 σ (PEP) respectively. Water molecules (red), Mg^{2+} (green) and K^{+} (purple) are shown as spheres. Potential interactions involved in active-site ligand binding are indicated by dashed lines. (b, d, f, h) Schematic drawings showing the estimated interactions between the active-site ligand and the active site of *Tb*PYK. The interatomic distances for the interactions are given in Ångstroms. The active-site interactions (g, h) are from *Tb*PYK/F26BP/PEP/Mg (PDB ID: 4HYV).

7.3.4. Crystal soaking with oxaloacetate analogues induces a large B-domain movement and a change in Mg coordination

7.3.4.1. B-domain movement

The B-domain of PYK has been shown to adopt multiple positions in different ligated states. Moreover, the position of the B-domain is involved in the regulation of kinase activity (Larsen et al. 1997; Morgan et al. 2010b; **Chapter 6**). A large scale *'in-crystallo'* B-domain movement with respect to the AC-core (A- and C-domains) has been observed by soaking the natural substrate PEP into the crystal of *Tb*PYK/F26BP/Mg resulting in *Tb*PYK/F26BP/PEP/Mg (**Chapter 6**). Additionally, this *'in-crystallo'* domain movement can also be triggered by the natural product pyruvate as shown in *Tb*PYK/F26BP/pyruvate/Mg (**Figure 7.5**). Structures of *Tb*PYK/F26BP/D-malate/Mg and *Tb*PYK/F26BP/α-ketoglutarate/Mg further show that the binding of oxaloacetate analogues can also induce the B-domain movement (**Figure 7.5**). The rotation angles of the B-domain triggered by kinase ligands (pyruvate and PEP) and analogues of the decarboxylase ligand are almost identical (22-23^{o}) (**Table 7.6**), and similarly yield a partially closed conformation. Additionally, this large domain rotation can only be observed in chain A from all *Tb*PYK soaking structures as crystal packing interactions lock the other B-domain in an open conformation.

The B-domain can only adopt a fully closed conformation when the enzyme binds nucleotide in the kinase reaction (Larsen et al. 1998; Morgan et al. 2010b). The fully closed B-domain locked in position by PEP and ADP can shield the substrates from the solvent to achieve the optimum environment for phospho transfer from PEP to ADP. However, nucleotide binding is not required in the decarboxylase activity of PYK (Creighton and Rose 1976b; Jursinic and Robinson 1978) so the B-domain can only adopt a partially closed conformation locked by the binding of oxaloacetate. The resulting active site has a partially solvent accessible environment which may substantially decrease the efficiency of the decarboxylase activity.

7.3.4.2. Change in Mg^{2+} coordination and a common substrate recognition

The movement of Mg^{2+} from its Mg-3 location to the PEP-coordinated Mg-1 location followed by a number of side-chain conformational changes has been revealed in **Chapter 6**. Superposition of the active sites of the five structures *Tb*PYK/F26BP/Mg (PDB ID: 4HYW), *Tb*PYK/F26BP/PEP/Mg (PDB ID: 4HYV), *Tb*PYK/F26BP/pyruvate/Mg, *Tb*PYK/F26BP/D-malate/Mg and *Tb*PYK/F26BP/α-ketoglutarate/Mg, reveals the same movements as those in the PEP-bound structure (**Figure 7.8**).

The binding modes of D-malate and α-ketoglutarate at the active site of *Tb*PYK (**Figure 7.7a** and **7.7c**) mimic the binding of the substrate oxaloacetate in *Tb*PYK. Overall, the active site residues of *Tb*PYK which form the substrate-binding sphere for both PEP and oxaloacetate can be divided into four categories (excluding the nucleotide binding residues): (1) divalent metal-binding residues (Glu241 and Asp265); (2) monovalent metal (K^+)-binding residues (Asn52, Ser54, Asp84 and Thr85); (3) substrate-binding residues (Gly264, Asp265 and Thr297 for both substrates PEP and oxaloacetate; Ser331 for oxaloacetate only). Two positively charged residues Lys239 and Arg50, which form ionic interactions with both substrates PEP and oxaloacetate, are potentially involved in the catalysis for both the kinase and decarboxylase activities; (4) Phe213 which potentially regulates the binding of substrate (**Chapter 6**).

This similar ligand recognition at the active site of PYK provides the opportunity to bind different substrates (PEP and oxaloacetate) which are structurally similar. Some metabolites, which have a similar scaffold with that of PEP or oxaloacetate, potentially are capable of interacting with the short α-helix Aα6', Mg^{2+}, Thr297, Ser331, Lys239 and Arg50 of PYK eventually bind at the active site. Malonate, L-tartrate and citrate which are structurally similar to PEP or oxaloacetate have been observed in the PEP/pyruvate/D-malate/α-ketoglutarate binding site of *Tb*PYK structures determined from crystals obtained by the same crystal soaking

method (see **Chapter 8**).

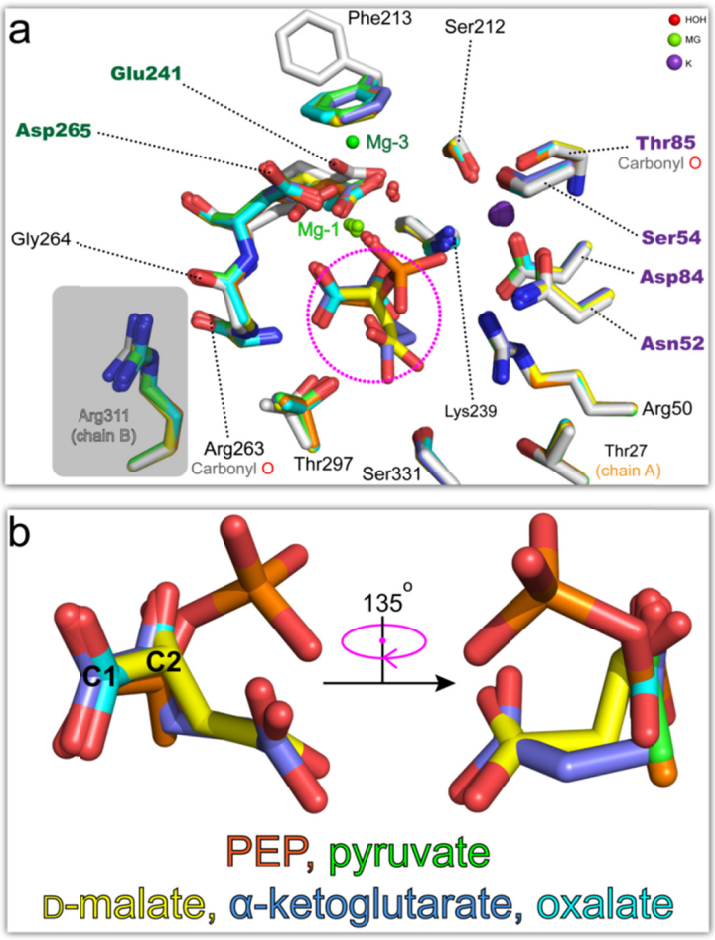


Figure 7.8 Similar binding modes of PEP, pyruvate and oxaloacetate analogues (D-malate, α-ketoglutarate or oxalate) at the active site of *Tb*PYK.
(a) Active sites of *Tb*PYK/F26BP/Mg (white; PDB ID: 4HYW), *Tb*PYK/F26BP/PEP/Mg (orange; PDB ID: 4HYV), *Tb*PYK/F26BP/D-malate/Mg (yellow), *Tb*PYK/F26BP/α-ketoglutarate/Mg (blue) and *Tb*PYK/F26BP/oxalate/Mg (cyan) were superposed onto the active site of *Tb*PYK/F26BP/pyruvate/Mg (green) to compare the ligand binding modes at the active site. Active site ligands and the residues involved in ligand binding are shown as sticks, while Mg^{2+} (green), K^{+} (purple) and water molecules (red) are shown as spheres. Two different Mg^{2+} binding sites (Mg-1 and Mg-3) which are located ~3.3-3.5 Å distance apart are labelled. Mg^{2+} locates in the Mg-3 binding site at the active site of *Tb*PYK/F26BP/Mg, but is in the position of the Mg-1 binding site when active-site ligands are bound (indicated by the dashed circle in pink). (b) Two views of an enlargement of the superposed active sites to visualise the overlapping ligand binding positions.

7.3.5. Oxalate acts as a competitive inhibitor for the decarboxylase and kinase activities of TbPYK

Oxalate has been shown to inhibit the decarboxylase activity of *Tb*PYK in the ^{1}H NMR results. Crystal soaking with oxalate reveals that the binding mode of oxalate at the active site of *Tb*PYK is almost identical to the binding mode of PEP, pyruvate and oxaloacetate analogues, which further confirm its competitive inhibition for both decarboxylase and kinase activities (**Figure 7.7e and 7.7f**; **Figure 7.8**). Moreover, the binding mode of oxalate from crystal soaking is also similar to other oxalate-bound PYK structures from co-crystallised crystals (Morgan et al. 2010b; Larsen et al. 1998; Dombrauckas et al. 2005; Christofk et al. 2008b). Nevertheless, the potential function of oxalate to chelate free Mg^{2+} may also contribute to the inhibition of decarboxylase activity.

7.3.6. Proposed catalytic mechanism of the oxaloacetate decarboxylase activity of TbPYK

We have reported a successful example of using crystal soaking to observe the binding details of the substrate PEP at the active site of *Tb*PYK (**Chapter 6**). A similar strategy to soak the decarboxylase substrate oxaloacetate into the crystal of *Tb*PYK/F26BP/Mg was used to try to elucidate the catalytic mechanism of decarboxylation for *Tb*PYK. Interestingly, only electron density for the product pyruvate can be clearly identified in the determined structure (**Figure 7.4b**), as discussed above. The reason for the lack of oxaloacetate in the refined structure is probably due to (1) the instability of oxaloacetate in the crystallisation conditions (with breakdown to pyruvate), (2) the instability of the *Tb*PYK/oxaloacetate complex, and/or (3) the relatively high binding affinity between pyruvate and *Tb*PYK. Therefore, two oxaloacetate analogues D-malate and α-ketoglutarate were soaked into crystals of *Tb*PYK/F26BP/Mg, and these resulted in two structure complexes

*Tb*PYK/F26BP/D-malate/Mg and *Tb*PYK/F26BP/α-ketoglutarate/Mg. The binding modes of these two analogues at the active site of *Tb*PYK (**Figure 7.7a** and **7.7c**) provide detailed templates to model the binding of the substrate oxaloacetate in *Tb*PYK (**Figure 7.9a**). Additionally, the binding mode of oxalate which potentially resembles the binding state of enolpyruvate can offer another template to elucidate the mechanism of decarboxylation by *Tb*PYK (**Figure 7.9b**).

Divalent metals play roles in the decarboxylation of oxaloacetate. Oxaloacetate is capable of spontaneous decarboxylation in aqueous solution, and the reaction is accelerated in the presence of multivalent metal ions, such as divalent Mg^{2+} and Mn^{2+}, and trivalent Fe^{3+} and Al^{3+} (Krebs 1942; Kornberg et al. 1948; Speck 1949; Tsai 1967). Furthermore, divalent metal ions are widely required for the activity of oxaloacetate decarboxylases. Cytosolic oxaloacetate decarboxylases from *Pseudomonas aeruginosa* (Narayanan et al. 2008) and *Corynebacterium glutamicum* (Ran et al. 2012), the membrane-bound oxaloacetate decarboxylase Na^{+} pump (Studer et al. 2007), and enzymes containing oxaloacetate decarboxylase activity such as malic enzymes (Change and Tong 2003) have been shown to coordinate divalent metal ions (e.g. Mg^{2+}, Mn^{2+}, Zn^{2+}) for the binding of the substrate oxaloacetate and for catalysis.

The proposed catalytic mechanism of oxaloacetate decarboxylation by *Tb*PYK is summarised in **Figure 7.9**. The carbon positions (C1, C2, C3 and C4) described here and in the figure are based on oxaloacetate. The substrate oxaloacetate modelled at the kinase active site of *Tb*PYK is stabilised by the short α-helix Aα6' (Gly264, Asp265), Thr297, Ser331, Mg^{2+}, and possibly also by Lys50. The side chain of Lys239 points toward the carbonyl oxygen C(3)O, and functions as a general acid by donating its proton to the carbonyl C(3)O to produce the enol form of pyruvate (enolpyruvate). This step is followed by the breakdown of the bond between C1 and C2 to produce CO_2. The position and function of the positively charged Lys239 in *Tb*PYK is similar to those of Arg159 in oxaloacetate decarboxylase from

Pseudomonas aeruginosa (PA4872; Narayanan et al. 2008) and Lys183 in human malic enzyme (Change and Tong 2003) (**Figure 7.10**). Following the formation of enolpyruvate and the release of CO_2, one water molecule is locked by the hydroxyl groups of Thr297 and Ser331 [proposed from the oxalate-bound structure of *Tb*PYK/F26BP/oxalate/Mg (**Figure 7.7e** and **7.7f**)]. This water molecule is capable of donating a proton to the C2 position, while Lys239 extracts the proton from the C3 hydroxyl, resulting in the stable product pyruvate. This hypothesis needs to be further tested and confirmed by more experiments such as mutagenesis.

7.4. Conclusions

PYKs from codfish muscle and rabbit muscle were reported to possess oxaloacetate decarboxylase activity over 30 years ago. Although the phospho transfer mechanism of kinase activity of PYKs has been studied in detail, the decarboxylase mechanism of PYKs is still unknown. Crystal structures of *Tb*PYK complexed with the product of oxaloacetate decarboxylation (pyruvate), and a series of substrate analogues (D-malate, α-ketoglutarate and oxalate) were solved. ^{1}H NMR was used in this work to follow oxaloacetate decarboxylation catalysed by PYKs from trypanosomatid and humans to confirm that the decarboxylase activity is conserved across distantly related species. Analogues of oxaloacetate and the PYK competitive inhibitor oxalate bind to the kinase active site of *Tb*PYK with similar binding modes, indicating that both decarboxylase and kinase activities share a common active site for substrate binding and catalysis. The structural information explains why the decarboxylation activity is restricted to substrates of α-keto acid analogues. A mechanism for decarboxylation is proposed to involve the conserved Lys239.

The k_{cat} for decarboxylation by human M1PYK is 0.512 s^{-1} (**Table 7.3**), compared to the k_{cat} for the kinase reaction of 347 s^{-1} (**Table 1.3** in **Chapter 1**). The k_{cat} for decarboxylation by *Tb*PYK in the absence and presence of F26BP is 0.857 s^{-1}

and 1.494 s^{-1} (**Table 7.3**), respectively, compared to the k_{cat} for the kinase reaction of 145 s^{-1} and 231 s^{-1}, respectively (**Table 1.3** in **Chapter 1**). It has been shown in section **7.1** that characterised OADs have a broad range of specific activity between species. However, a limited number of reported OADs are all obtained from bacteria and little information of ODAs is available from other species such as mammals. It is unclear whether this inherent decarboxylase activity of PYK is biologically relevant.

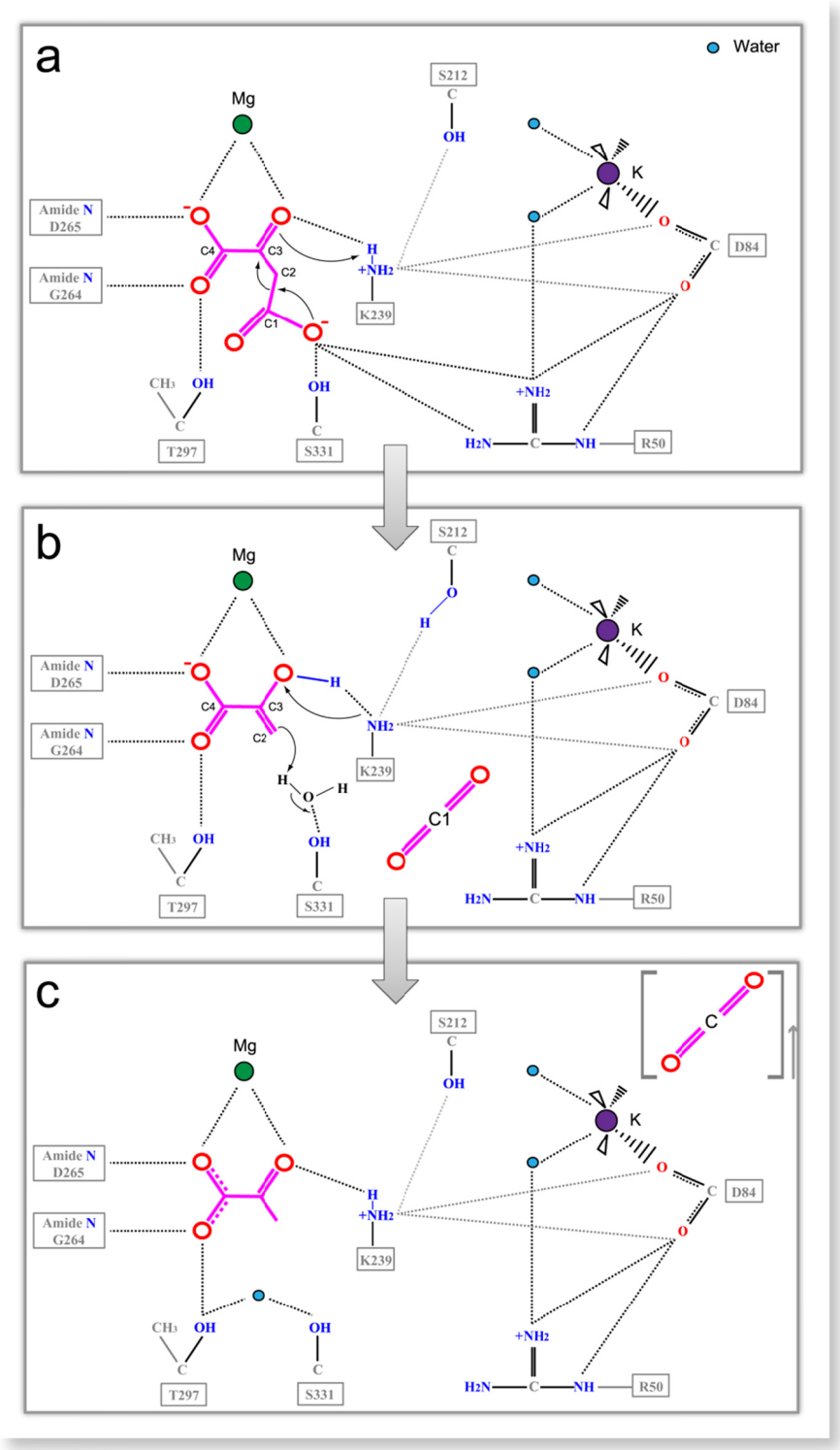


Figure 7.9 Schematic diagrams showing a proposed oxaloacetate decarboxylation catalytic mechanism for *Tb*PYK.

The binding mode of oxaloacetate at the active site of *Tb*PYK is modelled according to the binding modes of pyruvate, D-malate and α-ketoglutarate in the determined structures of *Tb*PYK from soaked crystals in this work. (a) The substrate oxaloacetate binds to the active site of *Tb*PYK and is stabilised by Gly264, Asp265,

Thr297, Ser331 and Mg^{2+}. The side chain of Lys239 points to the carbonyl oxygen C(3)O and is proposed to behave as a functional residue to catalyse the decarboxylation of oxaloacetate. Lys50 may have an ionic interaction with the carboxyl group C(1)OO. (b) Lys239 serves as a general acid and donates its proton to the carbonyl C(3)O to produce enolpyruvate, followed by the breakdown of the bond between C1 and C2 to produce CO_2. (c) A water molecule, which is attracted by the hydroxyl groups of Thr297 and Ser331, is proposed to donate a proton to the C2 position, while Lys239 extracts the proton from the C3 hydroxyl to produce pyruvate.

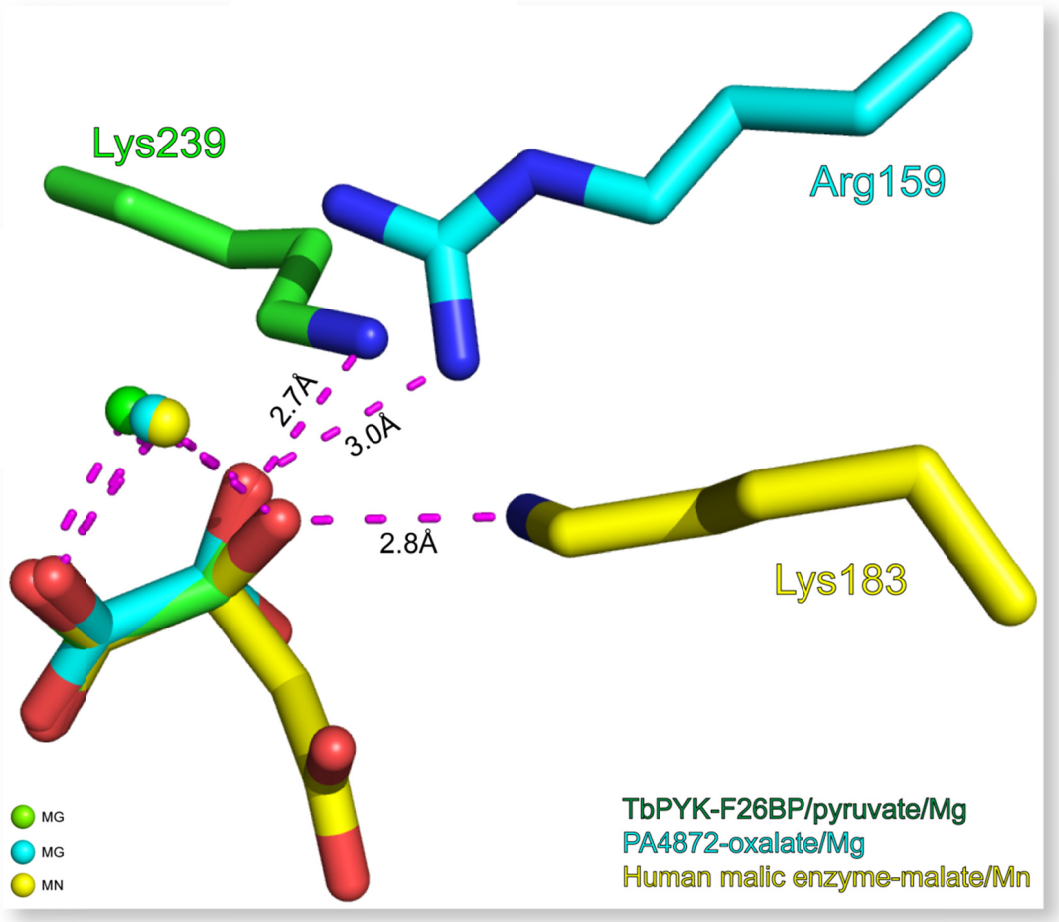


Figure 7.10 Similar functional residues and divalent metal ions for decarboxylase catalysis at the active sites of *Tb*PYK, oxaloacetate decarboxylase from *Pseudomonas aeruginosa* (*Pa*4872) and human malic enzyme.
The structures of *Tb*PYK/F26BP/Mg (chain A), *Pa*4872-oxalate/Mg (chain A) and human malic enzyme-malate/Mn (chain A) were superposed onto each other based on active site ligands (pyruvate, oxalate and D-malate; shown as sticks) and divalent metal ions (shown as spheres) to compare the functional residues around the active-site ligands. The functional residues Lys239 of *Tb*PYK/F26BP/Mg (green), Arg159 of *Pa*4872-oxalate/Mg (cyan) and Lys183 of human malic enzyme-malate/Mn (yellow) are represented as sticks. The interatomic interactions are shown as dashed lines in pink.

CHAPTER 8: Proposed Strategies for Designing Novel Inhibitors against Trypanosomatid PYKs

8.1. Introduction

Both *Tc*PYK and *Tb*PYK have been carefully characterised by kinetic and structural studies as described in previous chapters. In particular the role of the divalent metal ion in the Mg-3 position (**Figure 6.9** in **Chapter 6**) has been revealed to participate in a novel Mg-regulated mechanism for *Tb*PYK, as well as potentially other PYKs. These studies combined with the published work on *Lm*PYK (Ernest et al. 1994; Rigden et al. 1999; Hannaer et al. 2002; Tulloch et al. 2008; Morgan et al. 2010b)show that (1) PYKs from *Trypanosoma* and *Leishmania* have 74-81% sequence identity **(Table 8.1)**, (2) share similar kinetics and structures, and (3) experience rigid-body rotations of ~8^{o} which lock the enzyme in a thermally stable conformation (active R-state) by the binding of the effector F26BP at the effector site. These obvious similarities provide an opportunity to develop a unique drug to target trypanosomatid PYKs for the treatment of diseases caused by these parasites.

Although the active sites of PYKs are highly conserved between species, the effector sites, A-A interface and C-C interface which collaborate to stabilise PYK conformations are relatively varied between species (**Figure 8.1; Figure 2** in **Chapter 7**). These differences, especially the differences between trypanosomatid parasites and humans, offer interesting 'hot-spots' for the design of novel parasite-specific inhibitors.

This chapter will present three proposed strategies for the design of novel inhibitors against trypanosomatid PYKs (**Figure 8.2**):

(1) The binding of a modulator at the effector site, A-A interface or C-C interface in order to increase the inhibitor binding affinity (**Figure 8.2b, 8.2c and 8.2d**). A series

of substrate PEP analogues, which competitively bind to the active site of *Tb*PYK, are observed to have higher affinities for *Tb*PYK to compete out the substrate PEP when F16BP (analogue of F26BP) is bound to the effector site of the enzyme.

(2) Allosteric inhibitors targeting the effector site, A-A interface or C-C interface lock the enzyme in an inactive state where PYK has rather low affinity for substrate binding (**Figure 8.2e, 8.2f** and **8.2g**). A series of nanomolar bis-indole inhibitors (and their derivatives) for methicillin-resistant *Staphylococcus aureus* (MRSA) PYK were reported to selectively bind to the C-C interface of the enzyme (**Figure 8.2i**) (Zoraghi et al. 2011; Axerio-Cilies et al. 2012). Additionally, nanomolar activators were reported to bind to the A-A interface of human M2PYK (*Hs*M2PYK) which was potentially locked in an active state (**Figure 8.2j**) (Anastasiou et al. 2012).

(3) A third strategy is to generate a 'cocktail'-like inhibition system combining different inhibitors and modulators described above aiming to increase the drug efficiency and selectivity (**Figure 8.2h**). The concept is similar to that of 'cocktail therapy' for some serious diseases like HIV/AIDS.

*Tb*PYK is a validated drug target against *T. brucei* (Drew et al. 2003; Albert et al. 2005), and therefore will be used as a main example of trypanosomatid PYKs to be discussed in this chapter.

8.2. Materials and Methods

The materials and the general methods used in this chapter have been described in **Chapter 2**. In short, the structures of *Tb*PYK in complex with malonate, L-tartrate or citrate, were solved by the methods described in **Section 4** of **Chapter 2**, and the structure statistics are shown in **Table 8.3**. *Tb*PYK activity assay in the presence of substrate analogues was carried out by the method in section **2.4.7.2** of **Chapter 2**.

Table 8.1 Pairwise protein sequence comparisons of trypanosomatid, human M2, rabbit M1, yeast and *E. coli* PYKs

	*Tb*PYK	*Tc*PYK	*Lm*PYK	*Hs*M2PYK	*Hs*M1 PYK	*Gs*PYK	*Sa*PYK
*Tb*PYK	100	81	74	48	47	35	34
*Tc*PYK		100	76	47	47	35	35
*Lm*PYK			100	47	48	34	34
*Hs*M2PYK				100	96	36	34
*Hs*M1 PYK					100	36	34
*Gs*PYK						100	60
*Sa*PYK							100

The pairwise sequence analysis was obtained from EMBL-EBI web server (EMBOSS Stretcher): http://www.ebi.ac.uk/Tools/psa/emboss_stretcher/
Values are overall percent sequence identities.
Hs, Homo sapiens; Gs, Geobacillus stearothermophilus; Sa, Staphylococcus aureus subsp. aureus

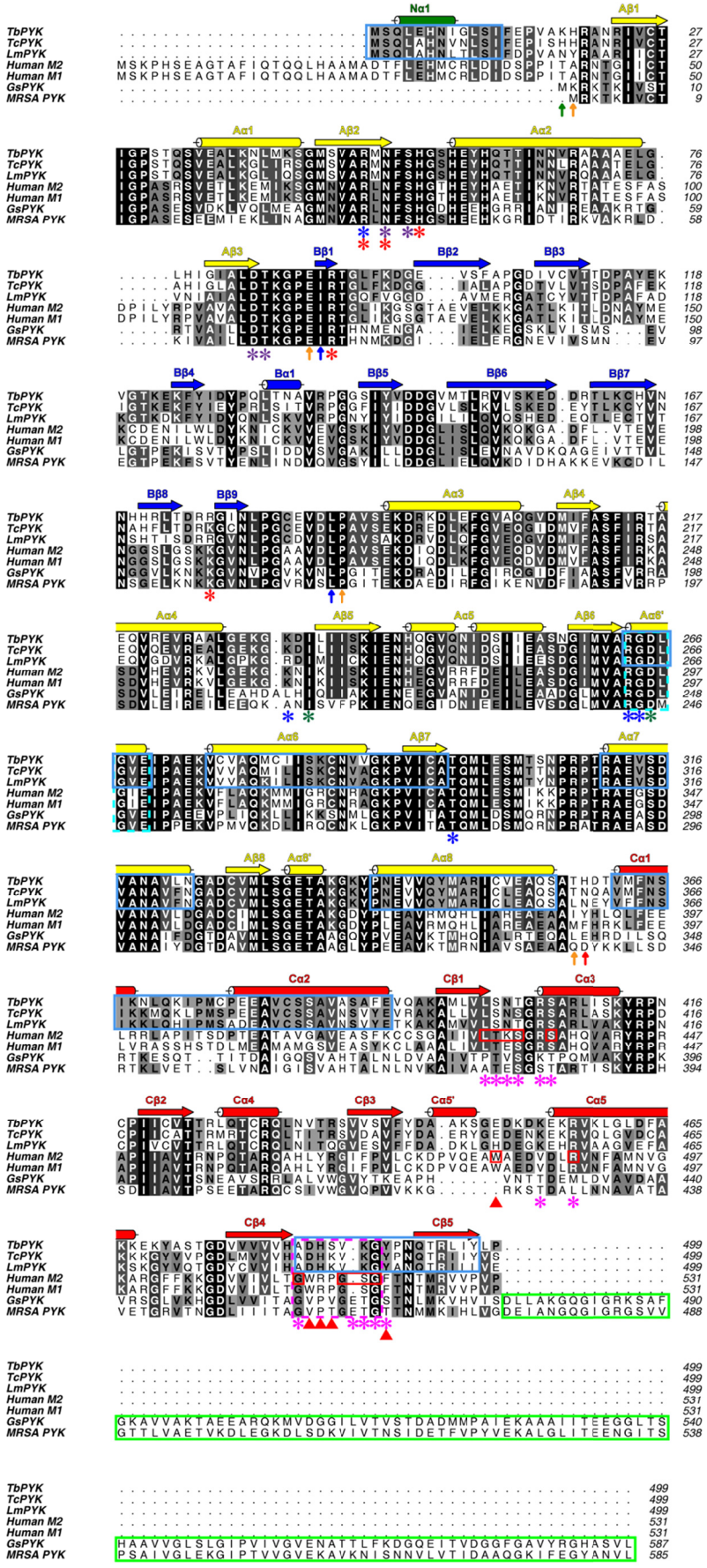

TbPYK
TcPYK
LmPYK
Human M2
Human M1
GsPYK
MRSA PYK

Figure 8.1 Sequence alignment of pyruvate kinases from *T. brucei*, *T. cruzi*, *L. mexicana*, *Homo sapiens* M2, *Homo sapiens* M1, *Geobacillus stearothermophilus* and methicillin-resistant *Staphylococcus aureus*.

The sequence alignment was performed using the program Clustal Omega at the European Bioinformatics Institute (Sievers et al 2011; Goujon et al. 2010). Secondary structural elements defined in *Tb*PYK/F26BP/PEP/Mg by DSSP (Kabsch and Sander 1983; Joosten et al. 2011) are shown above the sequences (only α-helices and β-strands are shown). Secondary structural elements are labelled in different colours corresponding to their domain regions: N-terminal domain (green), A-domain (yellow), B-domain (blue) and C-domain (red). Domain boundaries are indicated by vertical arrows in domain-specific colours. The conservation of the residues is indicated by shading from black (identical in seven or six sequences) to grey (conserved in five or four) to white (low or no conservation). Residue numbers corresponding to each PYK are listed after the sequences. For trypanosomatid PYKs, the amino acids involved in divalent metal ion binding (Mg-1 site) (*), potassium metal ion binding (*), substrate PEP binding (*), nucleotide binding (*) and effector F26BP binding (*) are indicated by asterisks. Residues 263-269 of the small α-helix Aα6′ which are involved in allosteric regulation and in binding divalent metal and the substrate PEP are indicated by a dashed box (cyan). The effector loop residues are indicated by a pink dashed box. The F16BP binding residues in human M2 are shown by red boxes. The residues involved in stacking interactions to stabilise the effector loop of human M2 are indicated by triangles in red. The blue boxes indicate the residue regions which are important for interface interactions. The extra C-terminal sequence (ECTS) domain (C'-domain) which potentially plays a role in tetramer stabilisation is shown by a green box. The figure was generated using the program Aline (Bond and Schüttelkopf 2009).

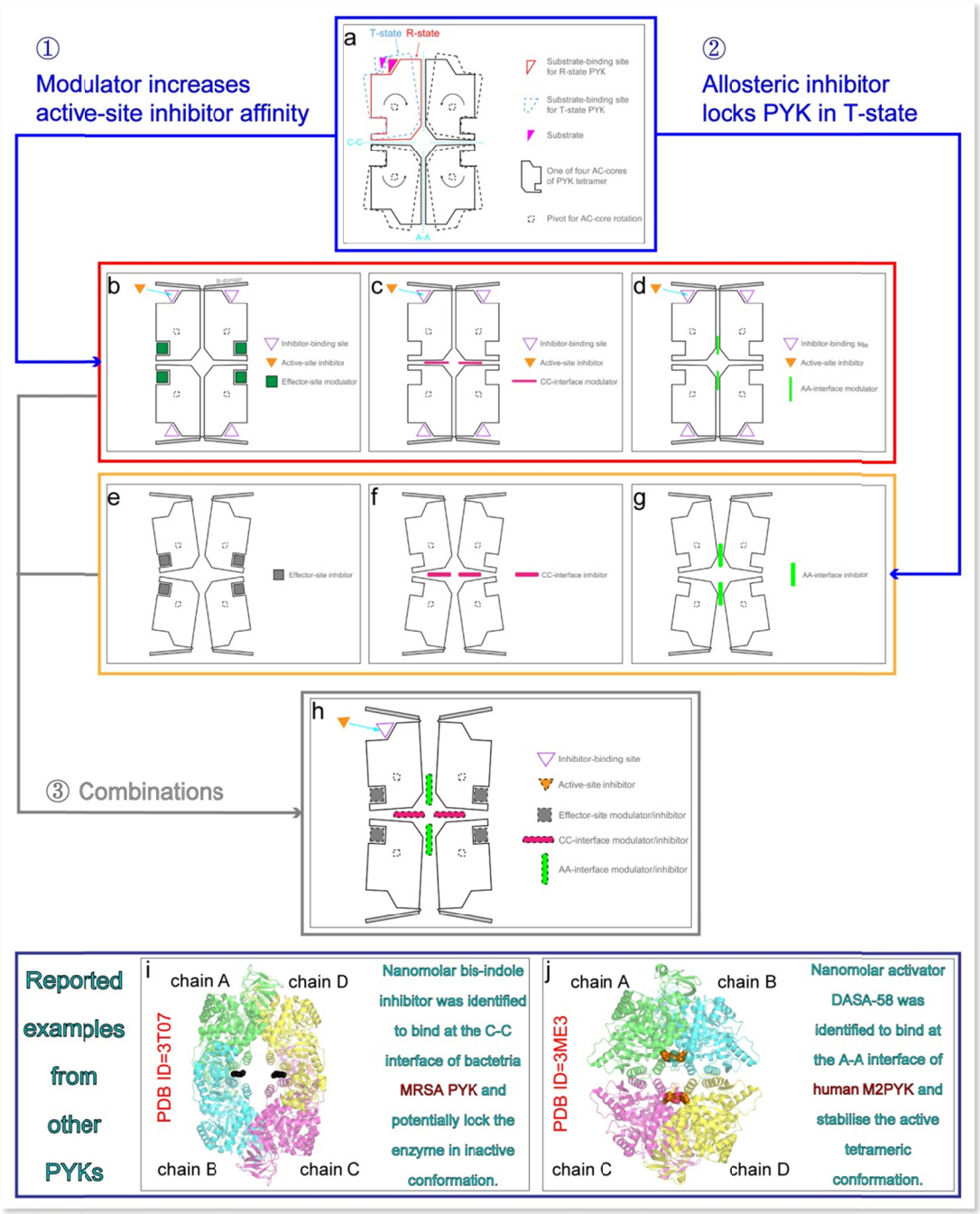


Figure 8.2 Proposed strategies for the design of trypanosomatid PYK-specific inhibitors.

Schematic representations describe the potential options to design novel inhibitors against trypanosomatid PYKs. (a) Both inactive T-state PYK (with less favoured substrate-binding site) and active R-state PYK (with favoured substrate-binding site) from trypanosomatid PYKs are shown. The rigid-body (AC-core) rotation between T- and R-states is indicated by arrows and the flexible B-domains are not shown. Two interfaces (A-A interface and C-C interface) are indicated by dashed lines in cyan. (b, c, d) An option for inhibitor development is to design a modulator to lock

allosteric PYK in a certain conformation where the active-site inhibitor has higher binding affinity to the enzyme. The target of this modulator can be (b) the F26BP-binding site, (c) the C-C interface or (d) the A-A interface. The studies on a series of substrate PEP analogues in this thesis (**Figure 8.6**) indicate the feasibility of this hypothesis. (e, f, g) A second option is to lock allosteric PYKs in the T-state (inactive) conformation to substantially reduce the affinity for the substrate. The target of this allosteric inhibitor on PYK can be (e) the F26BP-binding site (competitive binding to the effector site without stabilising the effector loop), (f) the C-C interface or (g) the A-A interface. Similar ligands have been reported and identified (i) for MRSA PYK (Zoraghi et al. 2011) and (j) for *Hs*M2PYK (Anastasiou et al. 2012). (h) A third option is to combine different inhibitors and modulators to further increase the inhibition efficiency and selectivity.

8.3. Results and Discussion

8.3.1. Structural studies and TbPYK inhibition assays to identify PEP analogues

8.3.1.1. Selected PEP analogues contain carboxyl groups

Substrate analogues are generally considered to be interesting starting points to develop specific inhibitors for enzymes. The carboxyl group of the PYK substrate PEP or of the product pyruvate has been proved to be required for binding by coordination with the small α-helix Aα6′ and hydroxyl group of Thr297 at the active site (**Figure 6.7** in **Chapter 6; Figure 7.4** in **Chapter 7**). Many small molecules including some intermediates in metabolism contain carboxyl groups therefore a total of six analogues were selected to study their effects on *Tb*PYK (**Table 8.2**; **Figure 8.3**). All the selected analogues were dissolved in TEA buffer (50 mM TEA, pH7.2, 100 mM KCl, 10 mM $MgCl_2$) and stored at -20^{o}C. NaOH (5 M) was used to adjust the pH to 7.2.

Table 8.2 Selected PEP analogues

	Analogue	Supplier& Product number	Formula	Molecular Weight
1	Oxalate	Sigma, O0136	NaOOCCOONa	134.00
2	α-ketoglutarate	Sigma, 75890	$HOOCCH_2CH_2COCOOH$	146.10
3	D-malate	Sigma, 46940U	$HO_2CCH_2CH(OH)CO_2H$	134.09
4	Malonate	Sigma, M1875	$CH_2(CO_2Na)_2 \cdot H_2O$	166.04
5	L-tartrate	Sigma, 09985	$(NH_4)_2C_4H_4O_6$	184.15
6	Citrate	Sigma, C8385	$C_6H_5K_3O_7 \cdot H_2O$	324.41

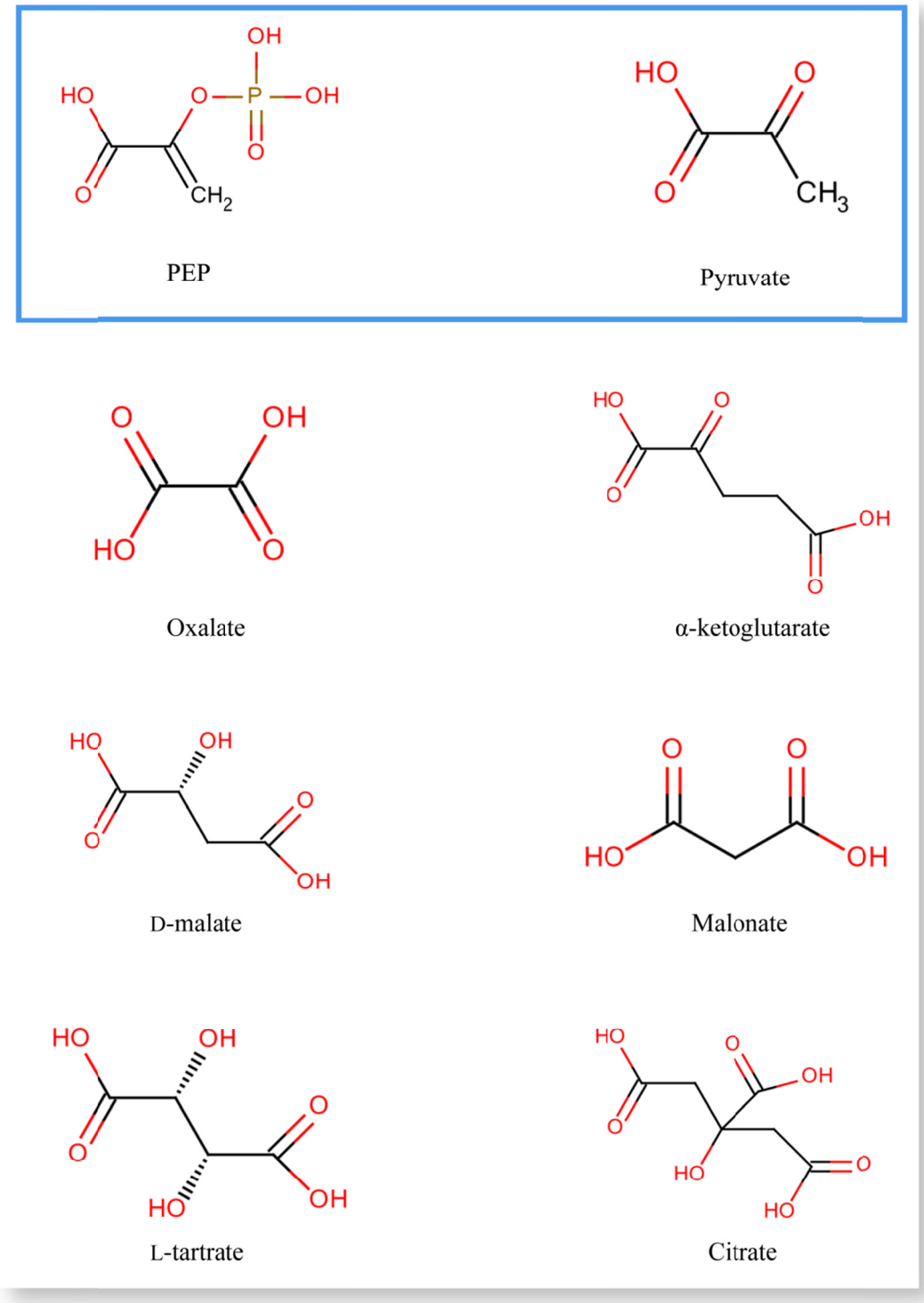


Figure 8.3 Two-dimensional representations of selected PEP analogues, corresponding to Table 8.2.
The substrate PEP and product pyruvate are shown for comparison in the blue box.

8.3.1.2. PEP analogues were soaked into *Tb*PYK crystals

It has been shown that both the substrate PEP (**Chapter 6**) and the product pyruvate (**Chapter 7**) could be soaked into crystals of *Tb*PYK/F26BP/Mg individually, and the corresponding structure complexes were successfully determined at high resolution. Both soaked ligands were identified clearly in the structures and their binding modes at the active site of *Tb*PYK were revealed. Crystals of *Tb*PYK/F26BP/Mg were incubated with 10 mM of the analogue overnight and flash frozen in liquid nitrogen. The procedures for crystal data collection, data processing and structure determination can be found in **Section 4** of **Chapter 2**. The data collection and processing statistics and the model refinement statistics are summarised in **Table 8.3 and Table 7.2 in Chapter 7**.

8.3.1.3. Overall structures of *Tb*PYK complexed with PEP analogues

Six *Tb*PYK structures complexed with PEP analogues are reported in this thesis: *Tb*PYK/F26BP/oxalate/Mg, *Tb*PYK/F26BP/α-ketoglutarate/Mg and *Tb*PYK/F26BP/D-malate/Mg have been discussed in **Chapter 7** to elucidate the inherent decarboxylase activity of PYK; *Tb*PYK/F26BP/malonate/Mg, *Tb*PYK/F26BP/L-tartrate and *Tb*PYK/F26BP/citrate are presented for the first time in this chapter.

For each structure complex, one asymmetric unit contains a dimer (forming the A-A interface between two identical subunits; **Table 8.4**) and comprises a biologically relevant tetramer with a crystallographic symmetry related dimer. These complexes share the same overall fold as *Tb*PYK/F26BP/Mg and the PEP complex (*Tb*PYK/F26BP/PEP/Mg), and the superposition of these structures (AC-cores) gave less than 0.2 Å of average RMS differences (C^{α} atoms) (**Table 8.5**; **Table 7.5** in **Chapter 7**). The soaking procedures also trigger a 22-23^{o} rotation for the B-domains from chain A (**Table 8.6; Table 7.6** in **Chapter 7**) which is similar to the B-domain movement in the PEP complex (*Tb*PYK/F26BP/PEP/Mg).

Table 8.3 Data collection, refinement and Ramachandran plot statistics

	*Tb*PYK/ F26BP/malonate/Mg	*Tb*PYK/ F26BP/L-tartrate	*Tb*PYK/ F26BP/citrate
Data collection			
Space group	*I*222	*I*222	*I*222
Cell dimensions			
a, b, c (Å)	103.53,108.79,265.51	103.78,108.43,266.30	103.57,108.87,267.46
α, β, γ (°)	90.00,90.00,90.00	90.00,90.00,90.00	90.00,90.00,90.00
Solvent content (%)	64	64	64
Wavelength (Å)	0.98	0.98	0.98
Resolution	66.38-2.30	54.21-2.45	75.04-2.30
No. reflections	489949 (71559)	376253 (55519)	464067 (66063)
No. unique reflections	66861 (9662)	55430 (8025)	67397 (9739)
Wilson B-factor (Å^2)	31.6	38.66	45.85
Rmerge (%)	15.4 (69.3)	14.4 (64.1)	7. (66.0)
<I/σI>	9.9 (3.4)	7.9 (3.0)	12.6 (2.8)
Rmeas (%) (within I+/I-)	18.0 (81.4)	17.0 (75.5)	8.4 (78.0)
Rmeas (%) (all I+ & I-)	17.8 (80.6)	16.9 (74.7)	8.3 (77.1)
Rpim (%) (within I+/I-)	9.3 (42.1)	8.9 (39.1)	4.4 (41.1)
Rpim (%) (all I+ & I-)	6.6 (29.6)	6.5 (28.0)	3.1 (29.2)
% completeness	100.0 (100.0)	99.8 (99.8)	99.9 (100.0)
Multiplicity	7.3 (7.4)	6.8 (6.9)	6.9 (6.8)
Refinement statistics			
Monomers in ASU	2	2	2
No. reflections	489949 (63472)	376253 (52600)	464067 (63985)
R_{work} / R_{free}	16.57/20.55	17.02/21.61	18.28/22.08
No. of nonhydrogen atoms			
Protein	7676	7633	6913
Water	593	372	365
Ligands	58	62	68
Average B-factor (Å^2)			
Overall	31.09	37.20	56.88
Overall (exclude B-domain of chain A)	27.74	31.52	n.a.‡
Protein	31.00	37.50	57.07
Protein (exclude B-domain of chain A)	27.33	31.50	n.a.
B-domain/Chain A	64.04	91.28	n.a.
B-domain/Chain B	38.99	37.58	63.75
Water	32.69	30.56	52.00
Ligands	25.56	40.12	63.23
MG	32.12	n.a.	n.a.
Analogue	36.72	57.34	86.50
No. Residues (protein)	996	996	905
No. Ligands	8	6	6
No. Water	593	372	365
R.m.s deviations			
Bond lengths (Å)	0.0105	0.0109	0.0110
Bond angles (degree)	1.0433	1.0829	1.0792
Ramachandran plots			
Favored (%)	97.2	96.3	97.2
Allowed (%)	99.6	99.7	99.8
No. of outliers †	4	3	2

Values in parentheses are for the highest resolution shell.

† The outlier residue in each monomer is Thr296, a key active-site residue that is commonly found in this configuration in PYK structures. The outlier residues Lys96 and Asp97 in *Tb*PYK/F26BP/tartrate, Asp146 in *Tb*PYK/F26BP/citrate are modelled by poor electron density.

‡ n.a., not applicable. The B-domain of chain A is missing in *Tb*PYK/F26BP/citrate as a result of poor density.

Table 8.4 Average RMS differences between chains in an asymmetric unit

Structures	AC-core (C^{α} atoms)
*Tb*PYK/F26BP/malonate/Mg	0.21 Å
*Tb*PYK/F26BP/L-tartrate	0.23 Å
*Tb*PYK/F26BP/citrate	0.18 Å

Table 8.5 Average RMS differences between *Tb*PYK tetramers (C^{α} atoms of AC-cores)

	*Tb*PYK/F26BP/PEP/Mg ‡
*Tb*PYK/F26BP/Mg †	0.16 Å
*Tb*PYK/F26BP/malonate/Mg	0.11 Å
*Tb*PYK/F26BP/L-tartrate	0.16 Å
*Tb*PYK/F26BP/citrate	0.13 Å

† The PDB ID for *Tb*PYK/F26BP/Mg is 4HYW.

‡ The PDB ID for *Tb*PYK/F26BP/PEP/Mg is 4HYV.

Table 8.6 '*In crystallo*' B-domain movements triggered by the binding of active site ligands [a]

	*Tb*PYK/F26BP/Mg †	
	Rotation angles	RMSD (C^{α}) [b]
*Tb*PYK/F26BP/malonate/Mg	23°	0.72 Å
*Tb*PYK/F26BP/L-tartrate/Mg	22°	0.71 Å

[a] The rotation angles of the B-domain were calculated by comparing the positions of B-domains between *Tb*PYK/F26BP/Mg (chain A) and each of the active-site ligand-binding structures (chain A). The detailed procedures can be found in **Materials and Methods**.

[b] RMSD (for all C^{α} atoms) between B-domains of two structures after rotation movements

† The PDB ID for *Tb*PYK/F26BP/Mg is 4HYW.

8.3.1.4. PEP analogues bind to the PEP-binding site with similar binding modes

It has been shown in **Chapter 7** that the PEP analogues oxalate, α-ketoglutarate and D-malate bind at the active site of *Tb*PYK and have similar binding modes to that of PEP (**Figure 7.7** in **Chapter 7**). In addition, another three PEP analogues malonate, L-tartrate and citrate are reported in this chapter showing similar PEP-binding modes (**Figure 8.4**; **Figure 8.5**). Crystal soaking for these analogues also induced a large side-chain movement of Phe213 (**Figure 8.5**).

The carboxyl group C(1)OO from each analogue forms interactions with the small α-helix Aα6′ and the hydroxyl group of Thr297. All PEP analogues coordinate with Mg^{2+} in the Mg-1 position, except L-tartrate and citrate. The side chain of Asp265 without Mg^{2+}-binding is likely to be more flexible in the L-tartrate complex compared to that from other complexes (**Figure 8.5**). However, the side chain of Asp265 in the citrate complex forms a hydrogen bond with the hydroxyl group C(3)OH of citrate (**Figure 8.4**). Similar to oxalate (**Figure 7.7** in **Chapter 7**), the carboxyl group C(3)OO from malonate forms ionic interactions with Mg^{2+}, Lys239 and Arg50, and forms hydrogen bonding interactions with two water molecules. Carboxyl groups C(4)OO from L-tartrate and C(5)OO from citrate directly coordinate with the monovalent K^{+} at the active site.

In conclusion, six PEP analogues are clearly identified at the substrate PEP-binding site of *Tb*PYK. The interactions between analogues, active site residues, metal ions and water molecules are expected to favour the stabilisation of the active site and of the overall conformation of *Tb*PYK. Additionally, the overlapping binding sites of analogues and the substrate PEP suggest the potential competitive inhibition of *Tb*PYK activity by these substrate analogues. In the next two sections, the results of PYK activity assays to test the effects of substrate analogues on *Tb*PYK activity are reported.

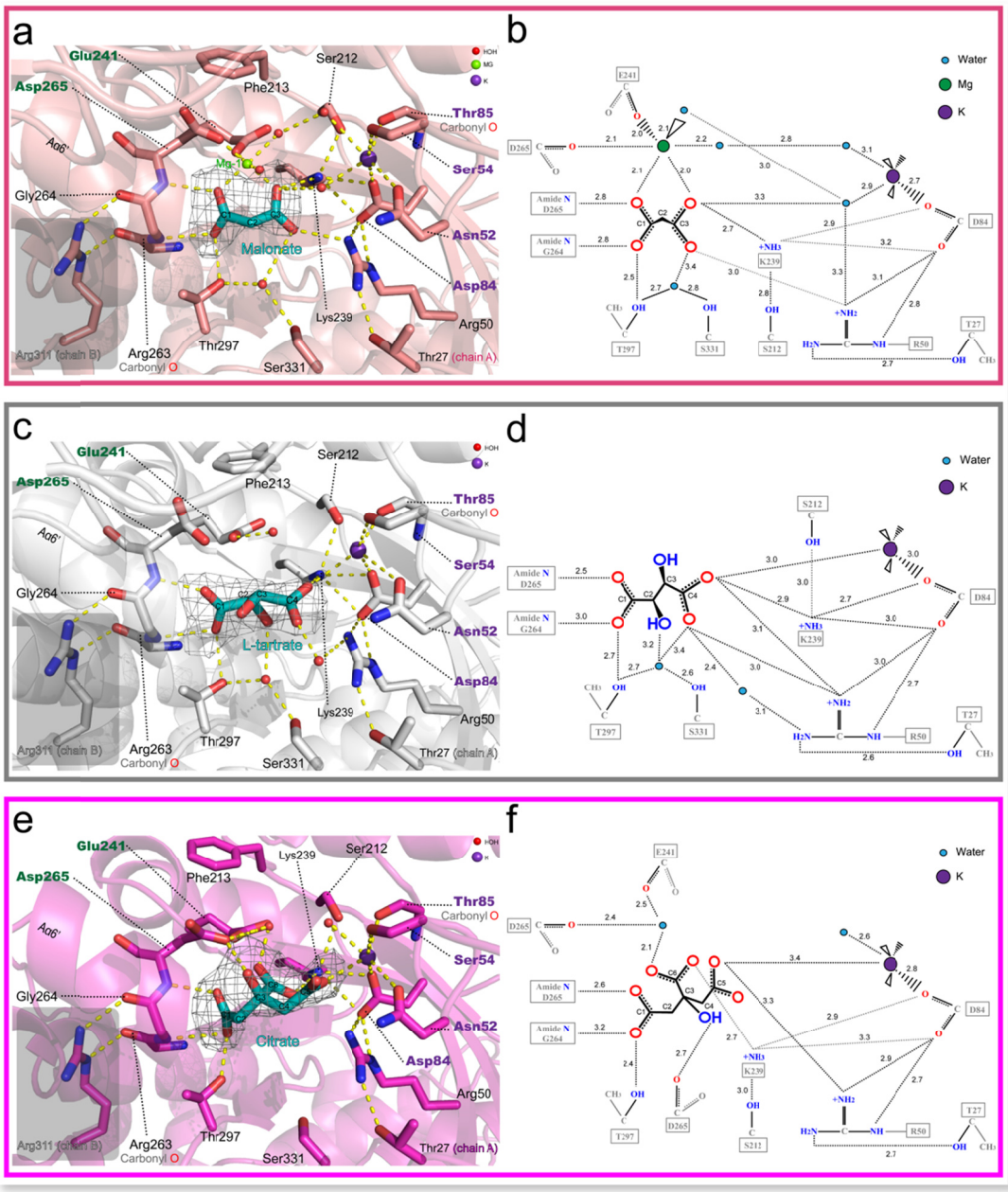


Figure 8.4 Similar binding modes of PEP analogues (malonate, L-tartrate or citrate) at the active site of *Tb*PYK.

(a, c, e) Close-ups of the active sites of *Tb*PYK in complex with the active site ligands (a) malonate, (c) L-tartrate or (e) citrate. The polypeptide chain is shown as a cartoon, while the residues involved in active-site ligand binding are shown as sticks. Active-site ligands are shown as sticks with an unbiased *Fo-Fc* election density (grey) map contoured at 3.0 σ (malonate), 3.0 σ (L-tartrate) and 2.0 σ (citrate), respectively. Water molecules (red), Mg^{2+} (green) and K^{+} (purple) are shown as spheres. Potential interactions involved in active-site ligand binding are indicated by dashed lines. (b, d, f) Schematic drawings showing the estimated interactions between the active-site ligand and the active site of *Tb*PYK. The interatomic distances for the interactions are given in Ångstroms.

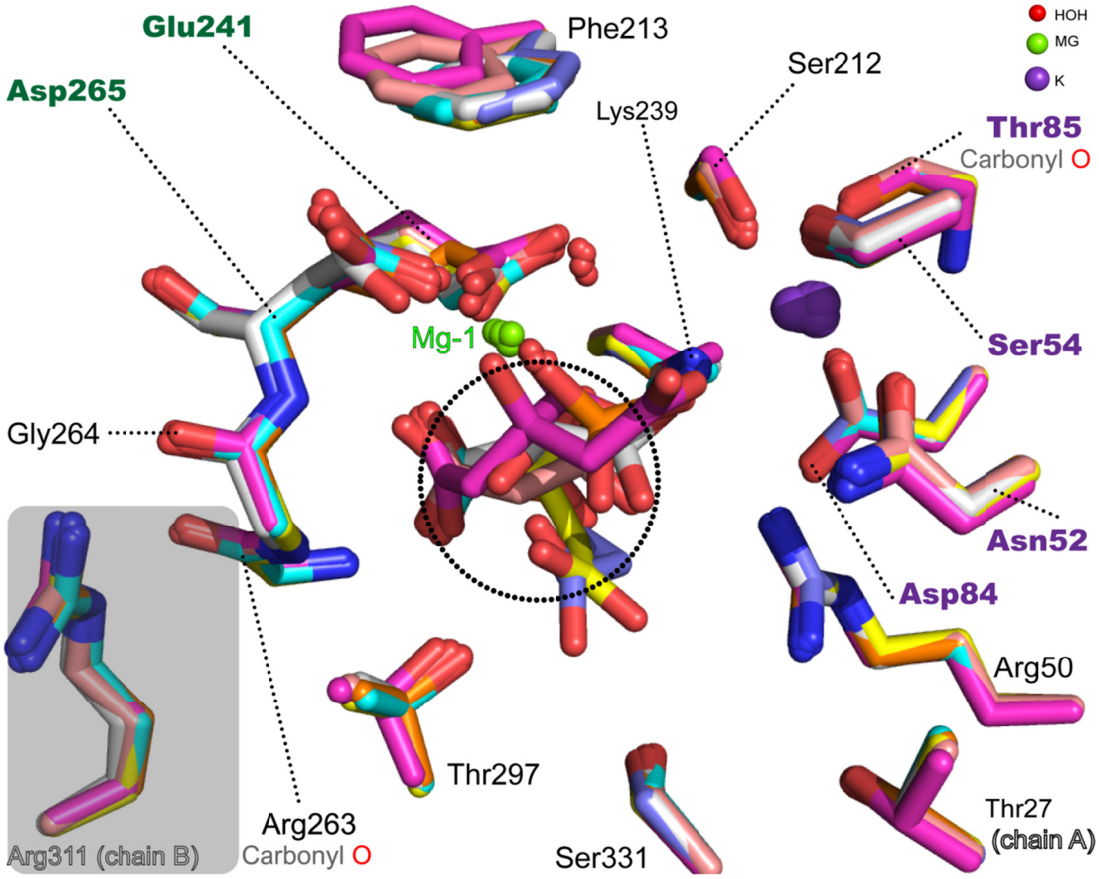


Figure 8.5 Similar binding modes of PEP and its analogues (malonate, L-tartrate, citrate, D-malate, α-ketoglutarate and oxalate) at the active site of *Tb*PYK.
The active sites of *Tb*PYK/F26BP/PEP/Mg (orange; PDB ID: 4HYV), *Tb*PYK/F26BP/malonate/Mg (salmon), *Tb*PYK/F26BP/L-tartrate/Mg (white), *Tb*PYK/F26BP/citrate/Mg (pink), *Tb*PYK/F26BP/D-malate/Mg (yellow), *Tb*PYK/F26BP/α-ketoglutarate/Mg (blue) and *Tb*PYK/F26BP/oxalate/Mg (cyan) were superposed onto each other to compare the ligand binding modes at the active site. Active site ligands and the residues involved in ligand binding are shown as sticks, while Mg^{2+} (green), K^{+} (purple) and water molecules (red) are shown as spheres. Mg^{2+} is in the position of the Mg-1 binding site when active site ligands are bound (indicated by dotted circle in black).

8.3.1.5. Low micromolar activator F16BP was used as a F26BP analogue in *Tb*PYK activity assay

Trypanosomatid PYKs normally switch conformations between two states (T-state and R-state) which are allosterically regulated by the effector-site ligand (e.g. F26BP) or active-site ligand (e.g. oxalate). The conformational differences will probably affect the binding of potential ligands. For example, trypanosomatid PYKs in the R-state have a 10-fold higher affinity for PEP compared to the T-state enzymes (**Table 3.3** in **Chapter 3**). Therefore, to fully characterise the role of a ligand on trypanosomatid PYKs, both T-state and R-state enzymes should be tested in parallel. Unfortunately, the natural allosteric effector F26BP for trypanosomatid PYKs is not commercially available. Only limited experiments were carried out on the available supply.

Although F26BP is the natural effector of trypanosomatid PYKs, F16BP has been shown to have low micromolar affinity ($K_{a0.5} \approx 47\mu M$) to *Tb*PYK (**Table 3.3** in **Chapter 3**). Additionally, the sigmoidal response to the substrate PEP is partially abolished by F16BP (Hill coefficient *h* is reduced from 1.88 to 1.31) indicating that probably *Tb*PYK is allosterically regulated by F16BP as well, although it is not as efficient as that of F26BP. Furthermore, F16BP is capable of increasing the melting temperature of *Tb*PYK by ~10 ^{o}C (**Figure 6.19** in **Chapter 6**) which suggests that the enzyme is at least partially transformed to the thermally stable R-state by F16BP binding.

In the assays reported here, F16BP was added for comparison with the sample without added F16BP when interesting ligands (e.g. PEP analogues) were tested on *Tb*PYK activity.

8.3.1.6. The behaviour of PEP analogues on *Tb*PYK activity are regulated by F16BP

The PEP analogues (**Figure 8.3**) were expected to be competitive inhibitors of

*Tb*PYK as they were shown to bind at the PEP-binding site (**Figure 8.5, Figure 8.6b**). Therefore, the LDH-coupled enzyme assay was performed to examine the potential inhibition of these analogues on *Tb*PYK in the presence and absence of F16BP under conditions of subsaturating substrates.

(i) Activation of *Tb*PYK by PEP analogues was observed in the absence of F16BP

Unexpectedly, four out of six analogues (oxalate, α-ketoglutarate, D-malate and L-tartrate) clearly behaved as activators of *Tb*PYK in the assay conditions without F16BP (**Figure 8.6a**). Oxalate activates *Tb*PYK at concentrations below 1.25 mM, but shows inhibition at 2.5 mM or above. Malonate does not have obvious effects on *Tb*PYK activity. Citrate fails to affect *Tb*PYK activity when its concentration is lower than 5 mM, however, shows substantial inhibition at 10 mM and 20 mM. This 'abrupt' inhibition by citrate is possibly due to its potential role of metal (e.g. Mg^{2+}) chelation. Divalent metal ions (10 mM of $MgCl_2$ in this assay) are necessary for kinase activity. The other three analogues (α-ketoglutarate, D-malate and L-tartrate) increase the enzyme activity by at least two fold. No inhibition is found from these three analogues. In particular, α-ketoglutarate and L-tartrate increase *Tb*PYK activity by up to five fold and four fold, respectively.

The results presented here disagree what the initial hypothesis because these PEP analogues can behave as activators under the conditions used.

(ii) PEP analogues only show inhibition of *Tb*PYK in the presence of F16BP

The expected inhibition on *Tb*PYK from PEP analogues can be obtained when F16BP was added to the assay (**Figure 8.6a**). All the selected analogues behaved as inhibitors of *Tb*PYK while oxalate had the most dramatic inhibition compared to the other analogues. Four of the analogues, α-ketoglutarate, D-malate, L-tartrate and malonate showed similar inhibition curves, and no more than 50% inhibition was observed although the analogue concentrations were as high as 20 mM. Citrate gave

a marked inhibition of *Tb*PYK from 5 mM, similar to the conditions without F16BP. *Tb*PYK was almost completely inactivated by 20 mM citrate both in the presence and absence of F16BP. The inhibition from citrate seems to be independent from the effect of F16BP indicating that the metal chelation by citrate plays an essential role in *Tb*PYK inhibition.

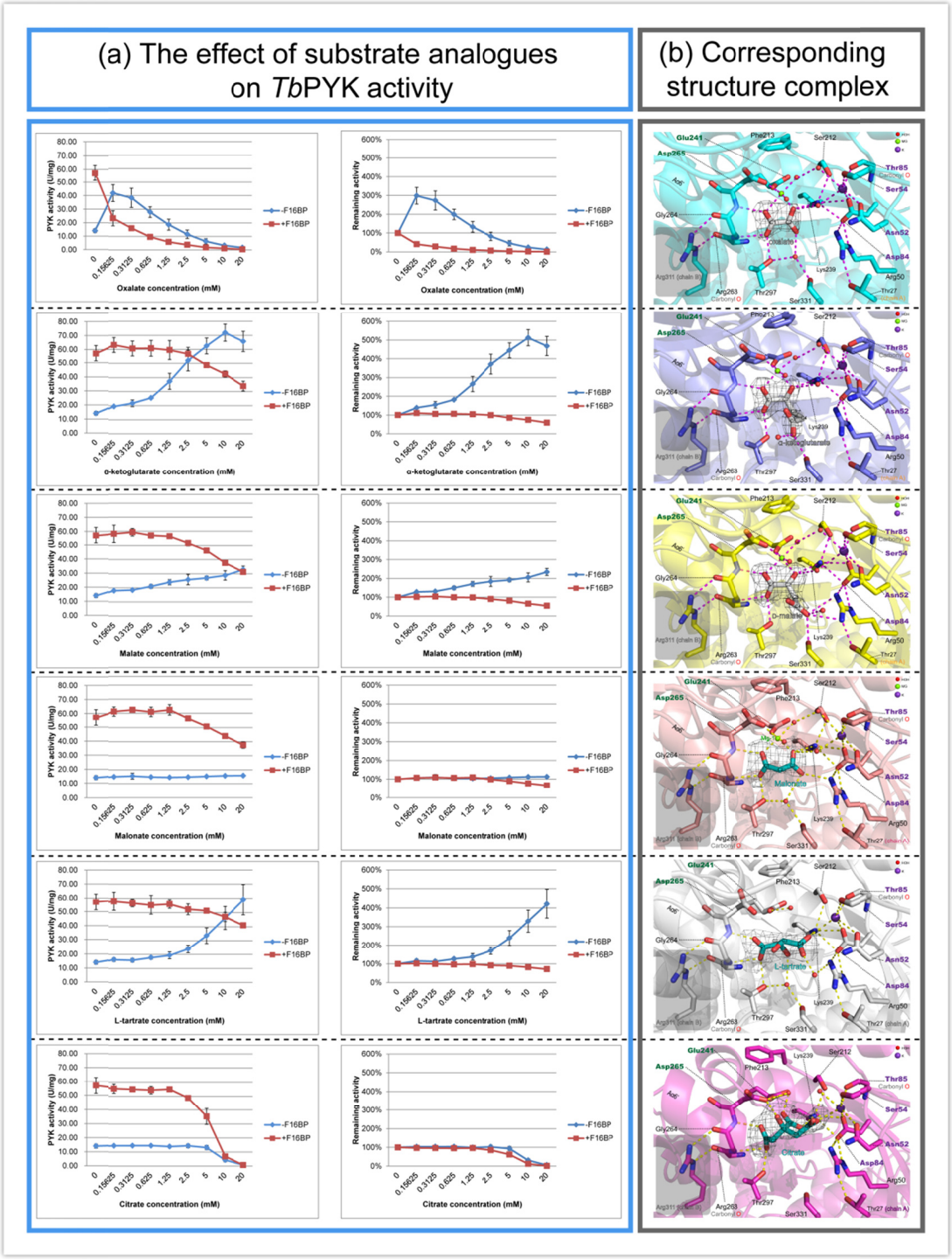


Figure 8.6 The double role (activation/inhibition) of PEP analogues on *Tb*PYK activity.

(a) Concentration-response curves observed for the titration of each analogue against *Tb*PYK activity in the absence (in brown colour) and presence (in blue colour) of the activator analogue F16BP (1 mM). The left column indicates the changes of PYK activity against the increasing analogue concentration. The right column shows the changes of remaining PYK activity (in percentage) against the increasing analogue concentration. All assays were performed in triplicate and the error bars indicate the

standard deviations. (b) Close-ups of the active sites of *Tb*PYK in complex with different analogues.

8.3.1.7. Double role of PEP analogues

(i) The active conformation of *Tb*PYK is stabilised by PEP analogues and results in enzyme activation

Oxalate has been shown in the *Lm*PYK structure (*Lm*PYK/OX/Mg; PDB ID: 3HQO) where it is capable of locking the enzyme in its active R-state without F26BP binding (Morgan et al. 2010b). Additionally, the thermal stabilisation of trypanosomatid PYKs enhanced by oxalate has been confirmed by melting shift assays (**Figure 6.19** in **Chapter 6**; **Figure 5.20** in **Chapter 5**; Morgan et al. 2010b). Furthermore, oxalate and the other five analogues show similar binding modes at the active site of *Tb*PYK which link the small α-helix Aα6′and K^+ binding residues (**Figure 8.6b**). Therefore, it is reasonable to predict that all these analogues play roles in the stabilisation of *Tb*PYK. This enzyme stabilisation by analogues may contribute to the increase of enzyme activity as long as the stabilisation role is more important than the binding competition with the substrate PEP. However, the analogues would show inhibition when their concentrations go high enough. For example, at around 2.5 mM oxalate started to show inhibition of *Tb*PYK in the absence of F16BP. The effect of citrate on *Tb*PYK activity is more complicated as it possibly chelates Mg^{2+}, as discussed in section **8.3.1.6**.

(ii) Enzyme activation by F16BP overwhelms the activations by PEP analogues and enhances analogue binding affinity to *Tb*PYK

F16BP is a low micromolar activator of *Tb*PYK (**Table 3.3** in **Chapter 3**) that is thought to mimic the natural activator F26BP and to stabilise the active form of *Tb*PYK. In the presence of 1 mM F16BP, PEP analogues no longer showed

activation in the assay compared to the conditions without added F16BP (**Figure 8.6b**). It is possible that F16BP is a much better activator of *Tb*PYK and can stabilise the enzyme in its active state better than are the PEP analogues. Furthermore, the presence of F16BP increases PEP affinity (**Table 3.3** in **Chapter 3**) and possibly also enhances the binding of PEP analogues as they have rather similar binding modes. The clear inhibition by PEP analogues (especially oxalate) in the presence of F16BP may further prove this hypothesis about the increase of analogue affinity. Once again, the inhibition from citrate seems to rely more on its role in metal chelation (**section 8.3.1.6**).

8.3.1.8. The communication between effector site and active site becomes an interesting drug target

Although the active sites of PYKs are highly conserved between species (**Figure 8.1**), active-site inhibitors of trypanosomatid PYKs are still potential templates for designing novel drugs (Morgan et al. 2012a). The results in section **8.3.1.6** and section **8.3.1.7** indicate that the affinity of active-site inhibitors may be regulated by effector-site ligand binding. The enzyme with a bound effector-site ligand may be locked in a conformation which has higher affinity for an active-site inhibitor (**Figure 8.2b**).

Compared to the conserved active site, the effector site is more varied between species (**Figure 8.1**). **Figure 8.7** shows a comparison between the effector sites of *Tc*PYK and Human M2PYK (*Hs*PYK), which require F26BP and F16BP as their natural activators, respectively. The differences at the effector site include the effector loop (residues Ala482-Gly488 for trypanosomatid PYKs and residues Gly514-Gly520 for *Hs*PYK) correlate with the specificity of effector recognition at the effector site. Therefore, an active-site inhibitor can be regulated by an effector-site ligand which is selectively recognized by varied effector sites. In other

words, the specificity of effector site is 'transferred' to the conserved active site by allosteric regulation (**Figure 8.2a**).

Otherwise, any other ligand (e.g. ligand located at the interface between subunits) which can allosterically control the inhibitor affinity to the enzyme also potentially can play the same role as the effector-site ligand (**Figure 8.2c and 8.2d**). In the next section, the possibility to design allosteric inhibitors which lock the enzyme in an inactive state will be discussed (**Figure 8.2e, 8.2f** and **8.2g**).

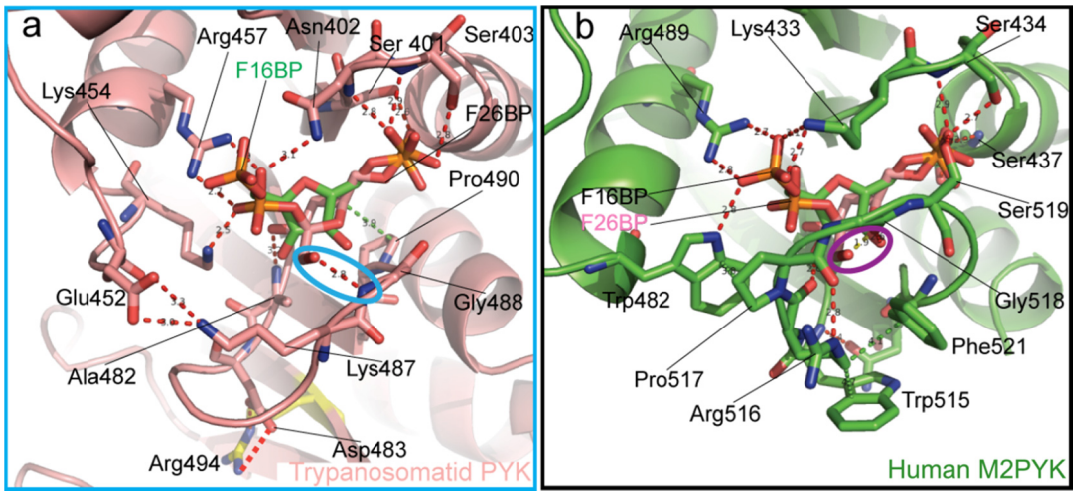


Figure 8.7 F26BP is a better allosteric activator of trypanosomatid PYK than F16BP[†]

The effector sites of *Tc*PYK/OX/F26BP and M2PYK/OX/F16BP (human M2PYK; PDB ID: 3BJF) were superposed to show the different binding modes of effectors. The polypeptide chain is shown in cartoon format, with residues involved in F26BP binding or F16BP binding shown as sticks. (a) An enlarged view of the effector site of *Tc*PYK/OX/F26BP showing the binding mode of F26BP. For comparison, the position of F16BP from the M2PYK/OX/F16BP structure (3BJF) is also shown in green. Effector binding stabilises the effector loop (Ala482-Gly488) and the interaction of the differing 1′-CH_2OH group (blue circle) of F26BP holds the effector loop in a conformation that favours the formation of the allosteric salt bridge (Arg494…Asp483). (b) An enlarged view of the effector site of M2PYK/OX/F16BP showing the binding mode of F16BP. Unlike in *Tc*PYK/OX/F26BP, the effector loop in M2PYK/OX/F16BP moves further from the C-C interface, essentially wrapping around the effector molecule. For comparison the position of F26BP from the *Tc*PYK/OX/F26BP structure is also shown in pink. Interestingly, the 1′-CH_2OH group belonging to F26BP clashes (purple circle) with the human effector loop position. Green dashed lines highlight stacking interactions.

[†]This figure is a modification of one originally generated by my colleague Dr. Hugh Morgan.

8.3.2. The evolution of the PYK allosteric mechanism and the potential to develop allosteric inhibitors

8.3.2.1. The AC-cores of trypanosomatid PYKs can adopt multiple rotation positions

Substrate binding at the active site of PYK can be regulated by effector binding at the effector site, and this allosteric mechanism is mediated by conformational changes. Additionally, PYK allosteric control is distinct from one species to another. Trypanosomatid PYKs perform a symmetrical 8^{o} AC-core rocking motion for the transition between inactive T-state and active R-state.

Interestingly, trypanosomatid PYKs can adopt a third tetrameric conformation in addition to the T-state and R-state, as shown by *Lm*PYK (**Figure 8.8**). In **Figure 8.8**, a third conformation (yellow) of *Lm*PYK complexed with SO_4^{2-} is compared with the T-state conformation (without ligand binding) and R-state conformation (F26BP/OX/ATP/Mg binding) of *Lm*PYK. The SO_4^{2-}-bound structure of *Lm*PYK (yellow) only displayed a 2.9^{o} rigid-body (AC-core) rotation from its T-state (white) compared to that of a 7.9^{o} rigid-body rotation by the R-state (blue). Therefore, the assembly of the AC-cores from trypanosomatid PYKs can be relatively flexible, but their positions are locked by potential ligand binding. Some of these rigid 'locked' conformations (excluding the active R-state conformation) may have rather low affinity for substrates and hence show enzyme inhibition.

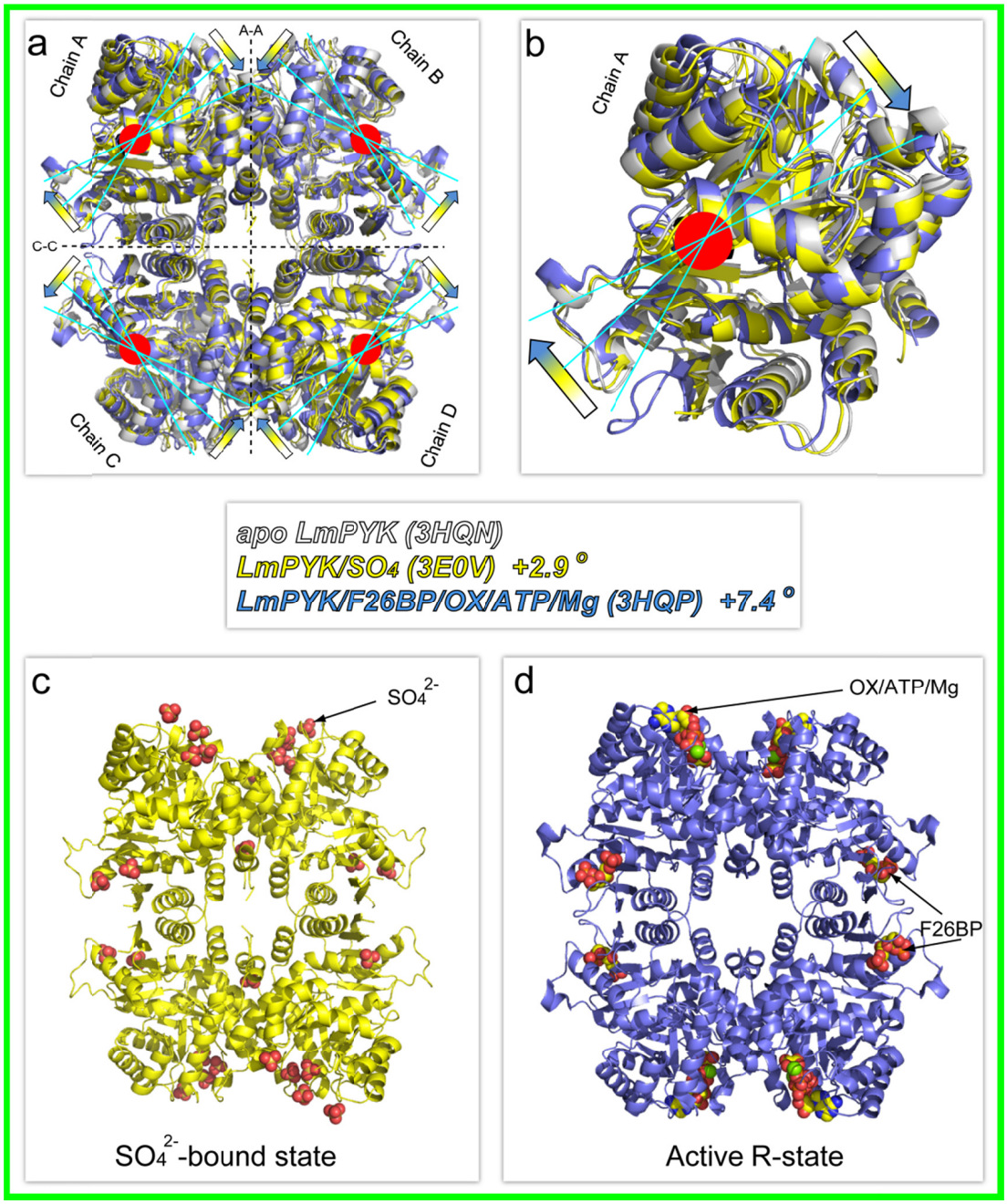


Figure 8.8 Multiple rigid-body (AC-core) positions locked by selected ligands. All B-domains and N-terminal domains are removed, and the interface boundaries are highlighted by dashed lines in panel (a). The superposed polypeptide chains are shown as cartoons. (a) The AC cores (A- and C- domains) of the inactive T-state apoenzyme *Lm*PYK tetramer (white), the SO_4^{2-}-bound state *Lm*PYK/SO_4 tetramer (yellow), and the active R-state *Lm*PYK/F26BP/OX/ATP/Mg tetramer (blue) were superposed. The rigid-body rotations from the inactive state, to the SO_4^{2-}-bound state to the active R-state around the central pivots (indicated in red) are shown by arrows. (b) An enlargement of chain A showing the 8° rotation around the pivot (red). It performs a 2.9° rigid-body rotation from the inactive state (apo *Lm*PYK) to the

SO_4^{2-}-bound state *Lm*PYK/SO_4, while a 7.4° rigid-body rotation is performed from the T-state apo *Lm*PYK to the active R-state *Lm*PYK/F26BP/OX/ATP/Mg. (c) *Lm*PYK/SO_4 (3E0V) tetramer and the bound SO_4^{2-} molecules (shown as spheres). (d) *Lm*PYK/F26BP/OX/ATP/Mg (3HQP) tetramer and the bound F26BP (at the effector site) and OX/ATP/Mg (at the active site) (shown as spheres).

8.3.2.2. The allosteric mechanism of trypanosomatid PYKs is different from that of other allosteric PYKs.

(i) 'Rock-and-lock' rigid-body rotation in Trypanosomatid PYKs

Trypanosomatid PYKs are allosterically regulated by the activator F26BP which binds at the effector site resulting in a stable effector loop and additional salt bridges across the C-C interface. These changes induce the AC-core rotation and lock the enzyme in its thermally more stable R-state which favours substrate binding (**Figure 8.9c** and **8.9d**; Morgan et al. 2010b).

(ii) A monomer-tetramer equilibrium in human M2PYK (*Hs*PYK)

Three out of four mammalian isoenzymes of PYK [RPYK (erythrocyte), LPYK (liver) and M2PYK (embryonic or tumour)] are allosterically activated by F16BP (instead of F26BP as in trypanosomatid PYKs). A comparison of effector recognition between *Tc*PYK and *Hs*M2PYK is shown in **Figure 8.9**. In particular, instead of possessing a rigid-body rocking motion, *Hs*M2PYK exists with a monomer-tetramer equilibrium (Anastasiou et al. 2012) and the binding of the activator F16BP at the effector site shifts the equilibrium to a predominantly tetrameric state where the enzyme attains its optimal conformation for substrate binding and the kinase reaction (**Figure 8.9e** and **8.9f**). By contrast, M1PYK (muscle) is constitutively active but is also regulated by another mechanism, whereby the binding of molecules to the alternative allosteric 'restraining' site inhibits enzyme activity (Williams et al. 2006; Chaneton et al. 2012).

(iii) A hinge-like extra domain in some bacterial PYKs

Another distinct allosteric mechanism is from *Geobacillus stearothermophilus* PYK (*Gs*PYK) which is allosterically activated and thermally stabilized by ribose-5-phosphate (R5P) (Lovell et al. 1998). The extra C-terminal domain (indicated by the green box in **Figure 8.1**) has been suggested to weakly interact with the adjacent monomer (A-domain and C-domain) and may play some role in the structural stability (Lovell et al. 1998; Suzuki et al. 2008). This extra domain may behave as the effector loop of trypanosomatid PYKs and be involved in the allosteric control of *Gs*PYK, as indicated in **Figure 8.9a** and **8.9b**. A similar extra C-terminal domain is also found in *Staphylococcus aureus* (**Figure 8.1**).

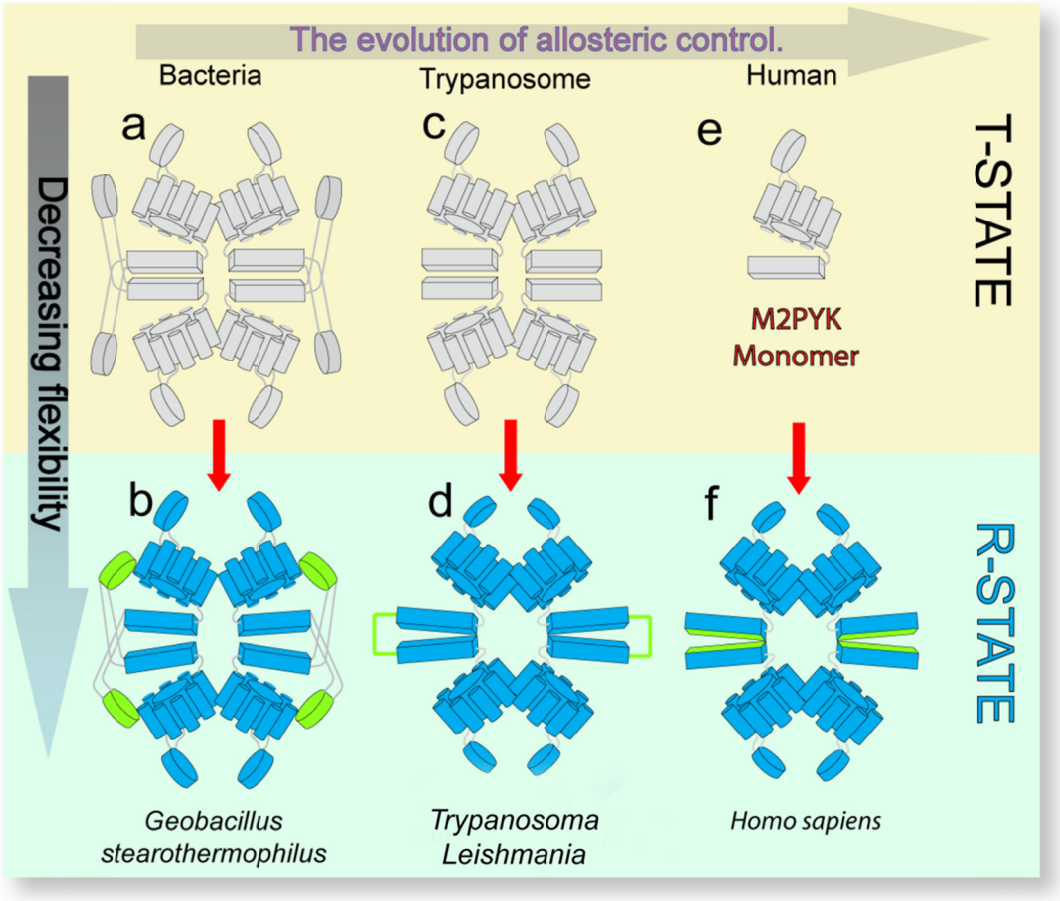


Figure 8.9 The evolution of allosteric control.
Schematic representations of the effector induced T- to R-state allosteric transition are shown. This transition results in the formation of molecular bridges between chains creating a more rigid, highly active R-state conformer. (a) The 'inactive' *Geobacillus stearothermophilus* PYK (*Gs*PYK) T-state tetramer (2E28). An extra C-terminal domain (residues 476-587) corresponding functionally to the effector loop in the trypanosomatid structures is shown in green. This domain movement is likely to mimic the trypanosomatid effector loop, forming molecular bridges between chains (as shown in panel b) to form the thermally stable and highly active R-state tetramer observed in solution. (c) In trypanosomatid PYKs F26BP binding stabilises the effector loop, resulting in the formation of a series of stabilising salt bridge interactions across the C-C interface (as shown in panel d), generating the highly stable R-state tetramer. (e) The 'inactive' form of human M2PYK is monomeric in solution shifting to tetrameric upon F16BP binding. Addition of F16BP, reduces thermal vibration along the C-C interface and locks the effector loops, stabilizing the highly active tetramer conformation.
This figure is a modification of one originally generated by my colleague Dr. Hugh Morgan. The results on human M2PYK presented here are provided by Dr Morgan, and are unpublished.

8.3.2.3. The distinct AC-core rocking motion of Trypanosomatid PYKs could be a potential inhibitor target

As has been discussed in the last section, the allosteric mechanism of trypanosomatid PYKs is different from other PYKs. The distinct rigid-body rocking motion therefore becomes an interesting target for designing allosteric inhibitors. There are three sites on trypanosomatid PYKs which potentially control this rocking motion: 1. effector site (**Figure 8.2e**); 2. C-C interface (**Figure 8.2f**); 3. A-A interface (**Figure 8.2g**).
The effector site can be blocked by selective inhibitors which prevent the binding of F26BP. This type of effector-site specific inhibitor would maintain the enzyme in its inactive state, and prevent substrate binding.

An interface inhibitor is expected to physically impede the movement of AC cores of trypanosomatid PYKs for the T- to R-state transition. Therefore, this proposed inhibitor would lock the enzyme in its inactive state constitutively. Additionally, the interface residues of PYKs are relatively varied between species (**Figure 8.1**) which favours the selectivity of the inhibitor. There are two series of small molecules that have previously been reported to bind at the interface of PYK. One series of nanomolar bis-indole inhibitors (including their derivatives) for methicillin-resistant *Staphylococcus aureus* (MRSA) PYK were reported to selectively bind at the C-C interface and possibly lock the enzyme in its inactive state (**Figure 8.2i**) (Zoraghi et al. 2011; Axerio-Cilies et al. 2012). In addition, a series of nanomolar activators were reported to bind at the A-A interface of human M2PYK (*Hs*M2PYK) which was supposed to be stabilised and locked in an active state (**Figure 8.2j**) (Anastasiou et al. 2012). These two successful examples raise the possibility to design interface inhibitors which are likely to allosterically regulate the enzyme conformation resulting in low substrate affinity.

8.3.3. 'Cocktail'-like PYK inhibitions for drug design

The third method for drug design is to combine the usage of different inhibitors and modulators (**Figure 8.2h**). The combination of modulator and active-site inhibitor (**Figure 8.2b**, **8.2c** and **8.2d**) has already been discussed for the idea of 'cocktail'-like drug design. The inhibitors/modulators binding at different sites of the enzyme have potential to work together leading to the enzyme inhibition with improved efficiency and selectivity.

CHAPTER 9: Summary and Forward Look

9.1. Summary of thesis

This thesis has described biochemical and structural studies of pyruvate kinases from *T. cruzi* and *T. brucei*. The experimental results and well-developed protocols in this thesis provide insight into the allosteric mechanism of trypanosomatid PYKs and may also provide fundamental knowledge and materials for future inhibitor development.

- Pyruvate kinase catalyses the final step of glycolysis to produce ATP, and is a potential drug target for the treatment of a group of diseases caused by trypanosomatid parasites.
- Methods to obtain high quality pure and active enzymes (*Tc*PYK and *Tb*PYK) was a top priority in the project. This thesis has presented successful and effective protocols to over-express and purify *Tc*PYK and *Tb*PYK from *E. coli*. At least 30 mg of pure and active enzyme can be obtained from one-litre *E. coli* cell culture within two days.
- A PYK activity assay using a 96-well plate reader has been optimised to measure enzyme kinetics and to screen potential inhibitors and activators. F26BP is the nanomolar allosteric activator which locks the enzyme in its active R-state, while F16BP has micromolar affinity for *Tb*PYK. Trypanosomatid PYK share similar enzyme kinetics but are distinct from F16BP affinity.
- This thesis described the first crystal structures of *Tc*PYK and *Tb*PYK. Two *Tc*PYK structures and twelve *Tb*PYK structures in complex with various ligands have been determined using the methods of co-crystallisation or crystal soaking. Ponceau S was selected as a ‘silver bullet’ to grow high-quality crystals.

- A concerted 8^{o} rigid-body rotation has been identified to explain the T- and R-state conformational transitions for both *Tc*PYK and *Tb*PYK, which is also the allosteric mechanism for *Lm*PYK. The R-state conformation is in a thermally more stable state due to the additional interactions across the A-A and C-C interfaces.
- Movement of the B-domain is involved in the kinase reaction, and this movement can only be regulated by active-site ligands. The binding of the substrate PEP or its analogues locks the B-domain in a partially closed conformation, while the additional binding of ADP/ATP locks it in a fully closed position.
- Mg^{2+} acts as a co-activator in collaboration with F26BP to maintain the enzyme in its R-state conformation. The position switch of Mg^{2+} between the Mg-3 and Mg-1 positions at the active site is regulated by substrate binding and product release, accompanied by a series of amino acid side-chain reorientations and B-domain movement (**Figure 9.1**).
- PYKs can also catalyse the decarboxylation of oxaloacetate, and the active site for this reaction overlaps with that for the kinase activity. The conserved Lys residue (Lys239 in *Tb*PYK) at the active site and Mg^{2+} are considered to play roles in the decarboxylase activity of PYK.
- The effector F16BP (an analogue of F26BP) was shown to regulate the affinity of substrate analogues for *Tb*PYK which competitively bind at the active site of the enzyme. The allosteric mechanism is an interesting drug target to design allosteric inhibitors to either lock the enzyme in its inactive state, or enhance the affinity of other inhibitors. Additional strategies to design parasite PYK-specific drugs have been proposed in this thesis (**Chapter 8**).

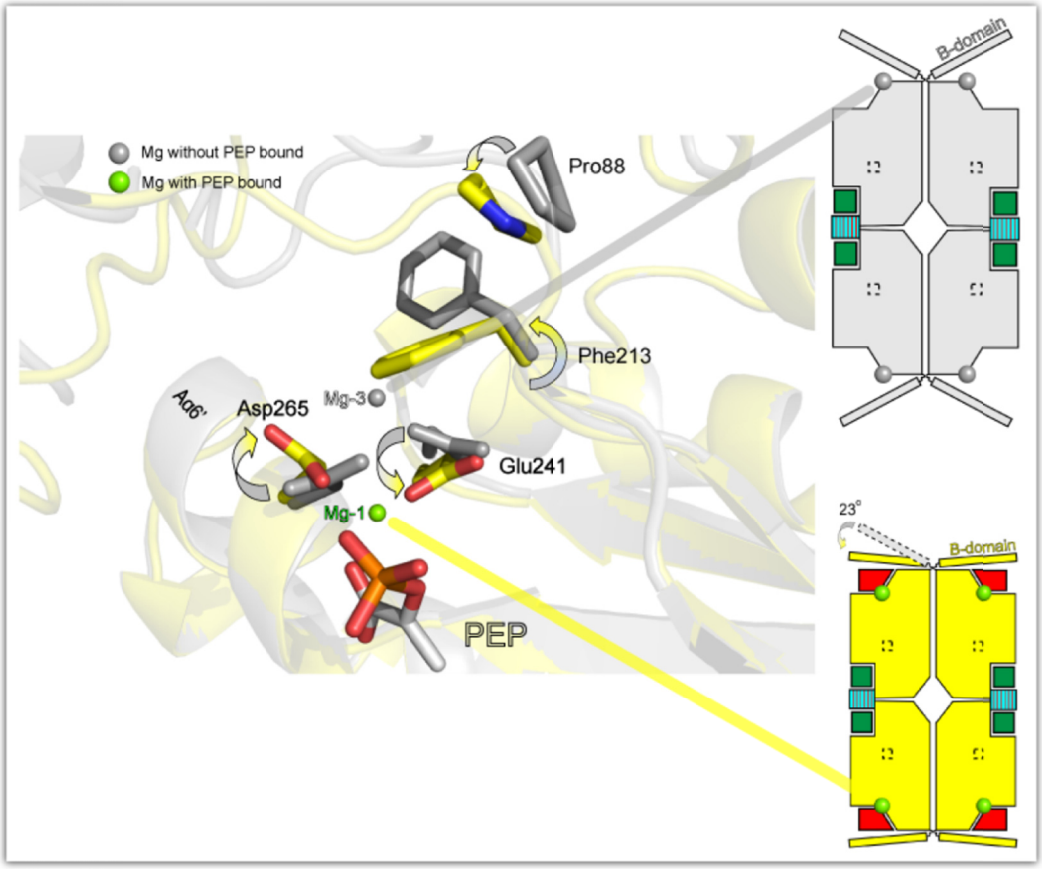


Figure 9.1 Mg^{2+} position switch, side-chain reorientations, and B-domain movement at the active site of *Tb*PYK.
The active site of *Tb*PYK/F26BP/PEP/Mg (yellow, chain A; white, substrate PEP) is shown superposed on *Tb*PYK/F26BP/Mg (grey). Polypeptide chains are shown in cartoon format, while relevant residues and the substrate PEP are shown as sticks. The two Mg^{2+} positions are shown as spheres, with the grey sphere in the 'priming' position (Mg-3 position), and the green sphere in the canonical position (Mg-1 position) for coordination with PEP. The distance between these two Mg^{2+} ions is 3.3 Å. The reorientations of the side chain of Phe213 and the carboxyl side chains of residues Asp265 and Glu241 are indicated by arrows. The small movement of Pro88 as a result of B-domain movement is indicated as well. Schematic representations of *Tb*PYK/F26BP/Mg (grey) tetramer and *Tb*PYK/F26BP/PEP/Mg (yellow) tetramer are also shown (refer to **Figure 5.14** of **Chapter 5** for a more detailed description of the schematic representations).

9.2. Forward look

9.2.1. Do trypanosomatid PYKs exhibit a monomer/dimer-tetramer equilibrium?

Chapter 8 has described human M2PYK that exits in a monomer-tetramer equilibrium (Anastasiou et al. 2012), and the binding of the activator F16BP at the effector site shifts the equilibrium to a predominantly tetrameric state where the enzyme obtains the optimal conformation for substrate binding and kinase reaction. It will be interesting to see whether trypanosomatid PYKs can have this monomer/dimer-tetramer equilibrium in solution as well in the presence of various ligands such as potential inhibitors. One possible experiment is to use SEC-MALS (size-exclusion chromatography coupled with multi-angle light scattering). Different oligomeric states of PYK will be separated by size-exclusion chromatography and the absolute mass value of each separated portion can be measured by the MALS detector.

9.2.2. Is the Mg-3 site a common position for divalent metals in the absence of active-site ligands?

The Mg-1 position has been proved to be a common position for both Mg^{2+} and Mn^{2+} in many PYKs in the presence of active-site ligand (Jurica et al. 1998; Wooll et al. 2001; Williams et al. 2006; Valentini et al. 2002; Fenton et al. 2010). The Mg-3 site has been shown as a Mg^{2+}-binding position at the active site by *Tb*PYK/F26BP/Mg structure in the absence of bound active-site ligand. Some other divalent metals such as Mn^{2+} and Co^{2+} have also been shown to support *Tb*PYK activity, and can replace Mg^{2+} but with different binding kinetics (Callen et al. 1991). It will be interesting to see whether other divalent metals can bind at the Mg-3 position in the absence of active-site ligands. If not, is the absence of affinity for the Mg-3 site related to the

different metal binding kinetics on PYK activity? Crystal soaking with selected divalent metals will be a possible way of testing this.

9.2.3. Enzyme mutagenesis for studying residue functions

It has been discussed in **Chapter 6** that Phe213 and Pro88 in *Tb*PYK might play roles in B-domain movement and substrate binding (**Figure 6.9**). Residues Thr297, Ser331 and Lys239 are proposed to be functional residues involving in the decarboxylase activity of *Tb*PYK in **Chapter 7** (**Figure 7.9**). Nevertheless, more evidence is required to further confirm these suggestions. The mutation of one or more of these residues can be used to study the function of each residue both in structures and enzyme kinetics in the future.

9.2.4. Assay conditions of in-vitro inhibitor screening for Trypanosomatid PYKs

Trypanosomatid PYKs are allosterically regulated by the natural activator F26BP and possess two conformations (T-state and R-state) regulated through AC-domain rotations. Additionally, it has been shown in **Chapter 8** that F16BP (F26BP analogue) can regulate the effect of PEP analogues on *Tb*PYK activity (**Figure 8.6** in **Chapter 8**). Therefore, it is important to carry out inhibitor screening for trypanosomatid PYKs in parallel with and without adding F26BP to the assay. F16BP could be an option (although not ideal) to replace F26BP which is not commercially available. PYK concentration could potentially affect the assay as well if it has a concentration-dependent equilibrium of monomer/dimer-tetramer, particularly for screening interface-binding inhibitors.

References

Albert, M. A., Haanstra, J. R., Hannaert, V., Van Roy, J., Opperdoes, F. R., Bakker, B. M., and Michels, P. A. (2005). Experimental and in silico analysis of glycolytic flux control in bloodstream-form *Trypanosoma brucei*. *J. Biol. Chem.* **280** (31), 28306-28315.

Allen, S. and Muirhead, H. (1996) Refined three-dimensional structure of cat-muscle (M1) pyruvate kinase at a resolution of 2.6 Å. *Acta. Cryst. D* **52** (3), 499-504.

Amato, V. S., Tuon, F. F., Bacha, H. A., Neto, V. A., and Nicodemo, A. C. (2008) Mucosal leishmaniasis. Current scenario and prospects for treatment. *Acta Trop*. **105** (1), 1-9.

Anastasiou, D., Poulogiannis, G., Asara, J. M., Boxer, M. B., Jiang, J. K., Shen, M., Bellinger, G., Sasaki, A. T., Locasale, J. W., Auld, D. S., Thomas, C. J., Vander Heiden, M. G., and Cantley, L. C. (2011) Inhibition of pyruvate kinase M2 by reactive oxygen species contributes to cellular antioxidant responses. *Science* **334** (6060), 1278–1283.

Anastasiou, D., Yu, Y., Israelsen, W. J., Jiang, J. K., Boxer, M. B., Hong, B. S., Tempel, W., Dimov, S., Shen, M., Jha, A., Yang, H., Mattaini, K. R., Metallo, C. M., Fiske, B. P., Courtney, K. D., Malstrom, S., Khan, T. M., Kung, C., Skoumbourdis, A. P., Veith, H., Southall, N., Walsh, M. J., Brimacombe, K. R., Leister, W., Lunt, S. Y., Johnson, Z. R., Yen, K. E., Kunii, K., Davidson, S. M., Christofk, H. R., Austin, C. P., Inglese, J., Harris, M. H., Asara, J. M., Stephanopoulos, G., Salituro, F. G., Jin, S., Dang, L., Auld, D. S., Park, H. W., Cantley, L. C., Thomas, C. J., and Vander Heiden, M. G. (2012). Pyruvate kinase M2 activators promote tetramer formation and suppress tumorigenesis. *Nat. Chem. Biol.* **8** (10), 839-847.

Andrade, L. O. and Andrews, N. W. (2005) The *Trypanosoma cruzi*-host-cell interplay: location, invasion, retention. Nat. Rev. Microbiol. **3** (10), 819-823.

Atanassova, A., Choe, J., Arrowsmith, C., Edwards, A., Sundstrom, M., Bochkarev, A., and Park, H. *unpublished. (PDB = 1ZJH)*.

Aufderheide, A. C., Salo, W., Madden, M., Streitz, J., Buikstra, J., Guhl, F., Arriaza, B., Renier, C., Wittmers, L. E. Jr., Fornaciari, G., and Allison, M. (2004) A 9,000-year record of Chagas disease. *Proc. Natl. Acad. Sci. USA* **101**, 2034-2039.

Axerio-Cilies, P., See, R. H., Zoraghi, R., Worral, L., Lian, T., Stoynov, N., Jiang, J., Kaur, S., Jackson, L., Gong, H., Swayze, R., Amandoron, E., Kumar, N. S., Moreau,

A., Hsing, M., Strynadka, N. C., McMaster, W. R., Finlay, B. B., Foster, L. J., Young, R. N., Reiner, N. E., and Cherkasov, A. (2012) Cheminformatics-driven discovery of selective, nanomolar inhibitors for staphylococcal pyruvate kinase. *ACS Chem. Biol.* **7** (2), 350-359.

Ayi, K., Min-Oo, G., Serghides, L., Crockett, M., Kirby-Allen, M., Quirt, I., Gros, P., and Kain, K. C. (2008) Pyruvate kinase deficiency and malaria. *N. Engl. J. Med.* **358** (17), 1805-1810.

Baker N. A., Sept D., Joseph S., Holst M. J., and McCammon J. A. (2001) *Proc. Natl. Acad. Sci. USA.* **98** (18), 10037-10041.

Baker, N., Glover, L., Munday, J. C., Aguinaga Andrés, D., Barrett, M. P., de Koning, H. P., and Horn, D. (2012). Aquaglyceroporin 2 controls susceptibility to melarsoprol and pentamidine in African trypanosomes. *Proc. Natl. Acad. Sci. USA.* **109** (27), 10996-11001.

Bakszt, R., Wernimont, A., Allali-Hassani, A., Mok, M. W., Hills, T., Hui, R., and Pizarro, J. C. (2010) The crystal structure of Toxoplasma gondii pyruvate kinase 1. *PLoS One* **5** (9), e12736.

Bao, B., Sun, Q., Yao, X., Hong, J., Lee, C. O., Cho, H. Y., and Jung, J. H. (2007) Bisindole alkaloids of the topsentin and hamacanthin classes from a marine sponge Spongosorites sp. *J. Nat. Prod.* **70** (1), 2-8.

Bao, B., Sun, Q., Yao, X., Hong, J., Lee, C. O., Sim, C. J., Im, K. S., and Jung, J. H. (2005) Cytotoxic bisindole alkaloids from a marine sponge Spongosorites sp. *J. Nat. Prod.* **68** (5), 711-715.

Barnard, J. and Pedersen, P. L. (1988). Purification in a single step and kinetic characterization of the pyruvate kinase of *Trypanosoma brucei. Mol. Biochem. Parasitol.* **31** (2), 141-147.

Barrett, M. P., Burchmore, R. J., Stich, A., Lazzari, J. O., Frasch, A. C., Cazzulo, J. J., and Krishna, S. (2003). The trypanosomiases. *Lancet* **362** (9394), 1469-1480.

Barrett, M. P. and Croft, S. L. (2012) Management of trypanosomiasis and leishmaniasis. *Br. Med. Bull.* **104**, 175-196.

Benchimol Barbosa, P. R. (2006) The oral transmission of Chagas' disease: an acute form of infection responsible for regional outbreaks. *Int. J. Cardiol.* **112** (1), 132-133.

Benziman, M., Russo, A., Hochman, S., and Weinhouse, H. (1978) Purification and regulatory properties of the oxaloacetate decarboxylase of *Acetobacter xylinum. J. Bacteriol.* **134** (1), 1-9.

Berens, R. L., Deutsch-King, L. C., and Marr, J. J. (1980) *Leishmania donovani* and *Leishmania braziliensis*: hexokinase, glucose 6-phosphate dehydrogenase and pentose phosphate shunt activity. *Exp. Parasitol.* **49**, 1-8.

Berman, H. M. Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res.* **28** (1), 235-242.

Bern, C., Maguire, J. H., and Alvar, J. (2008) Complexities of assessing the disease burden attributable to leishmaniasis. *PLoS Negl. Trop. Dis.* **2** (10), e313.

Bond, C. S. and Schüttelkopf, A. W. (2009) ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. *Acta Crystallogr D Biol Crystallogr*. **65**, 510-512.

Boxer, M. B., Jiang, J. K., Vander, Heiden, M. G., Shen, M., Skoumbourdis, A. P., Southall, N., Veith, H., Leister, W., Austin, C. P., Park, H. W., Inglese, J., Cantley, L. C., Auld, D. S., and Thomas, C. J. (2010) Evaluation of substituted *N*,*N*′-diarylsulfonamides as activators of the tumor cell specific M2 isoform of pyruvate kinase. *J. Med. Chem.* **53** (3), 1048-1055.

Brun, R., Blum, J., Chappuis, F., and Burri, C. (2010). Human African trypanosomiasis. *Lancet* **375** (9709), 148-159.

Burri, C. and Brun, R. (2003) Human African trypanosomiasis. In Manson's tropical diseases. G.C. Cook and A.I. Zumla, editors. 21st edition. W.B. Saunders/Elsevier. Edinburgh, United Kingdom. 1303-1323.

Buscaglia, C. A., Campo, V. A., Frasch, A. C., and Di Noia, J. M. (2006) *Trypanosoma cruzi* surface mucins: host-dependent coat diversity. *Nat. Rev. Micro.* **4** (3), 229-236.

Cáceres A, J., Portillo, R., Acosta, H., Rosale, D., Quiñones, W., Avilan, L., Salazar, L., Dubourdieu, M., Michels, P. A., and Concepción, J. L. (2003) Molecular and biochemical characterization of hexokinase from *Trypanosoma cruzi. Mol Biochem Parasitol.* **126** (2), 251-262.

Callen, M., Kuntz, D. A., and Opperdoes, F. R. (1991). Characterization of pyruvate kinase of *Trypanosoma brucei* and its role in the regulation of carbohydrate metabolism. *Mol. Biochem. Parasitol.* **47** (1), 19-29.

Callens, M. and Opperdoes, F. R. (1992). Some kinetic properties of pyruvate kinase from *Trypanosoma brucei. Mol. Biochem. Parasitol.* **50** (2), 235-244.

Cannata, J. J, Valle, E., Docampo, R., and Cazzulo, J. J. (1982) Subcellular localization of phosphoenolpyruvate carboxykinase in the trypanosomatids *Trypanosoma cruzi* and *Crithidia fasciculata*. *Mol. Biochem. Parasitol.* **6** (3), 151-160.

Casapullo, A., Bifulco, G., Bruno, I., and Riccio, R. (2000) New bisindole alkaloids of the topsentin and hamacanthin classes from the Mediterranean marine sponge Rhaphisia lacazei. *J. Nat. Prod.* **63** (4), 447-451.

Cazzulo, J. J. (1992) Aerobic fermentation of glucose by trypanosomatids. *FASEB J.* **6** (13), 3153-3161.

Cazzulo J. J., Cazzulo Franke M. C., and Franke de Cazzulo B. M. (1989) On the regulatory properties of the pyruvate kinase from *Trypanosoma cruzi* epimastigotes. *FEMS Microbiol Lett.* **50** (3), 259-263.

Centers for Disease Control and Prevention (CDC) (2006) Chagas disease after organ transplantation - Los Angeles, California, 2006. *MMWR Morb. Mortal. Wkly. Rep.* **55** (29), 798-800.

Chaneton, B., Hillmann, P., Zheng, L., Martin, A. C. L., Maddocks, O. D. K., Chokkathukalam, A., Coyle, J. E., Jankevics, A., Holding, F. P., and Vousden, K. H. (2012). Serine is a natural ligand and allosteric activator of pyruvate kinase M2. *Nature* **491** (7424), 458-462.

Chang, G. G. and Tong, L. (2003) Structure and function of malic enzymes, a new class of oxidative decarboxylases. *Biochemistry* **42** (44), 12721-12733.

Chappuis, F., Sundar, S., Hailu, A., Ghalib, H., Rijal, S., Peeling, R. W., Alvar, J., and Boelaert, M. (2007) Visceral leishmaniasis: what are the needs for diagnosis, treatment and control? *Nat. Rev. Microbiol.* **5** (11), 873-882.

Chatelain, E. and loset, J. R. (2011) Drug discovery and development for neglected diseases: the DNDi model. *Drug Des. Devel. Ther.* **5**, 175-181.

Chirac, P. and Torreele, E. (2006) Global framework on essential health RandD.

Lancet **367** (9522), 1560-1561.

Christofk, H. R., Vander Heiden, M. G., Harris, M. H., Ramanathan, A., Gerszten, R. E., Wei, R., Fleming, M. D., Schreiber, S. L., and Cantley, L. C. (2008a) The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature* **452** (7184), 230-233.

Christofk, H. R., Vander Heiden, M. G., Wu, N., Asara, J. M., and Cantley, L. C. (2008b) Pyruvate kinase M2 is a phosphotyrosine-binding protein. *Nature* **452** (7184), 181–186.

Coley, A. F., Dodson, H. C., Morris, M. T., and Morris, J. C. (2011). Glycolysis in the African trypanosome: Targeting enzymes and their subcellular compartments for therapeutic development. *Mol. Biol. Int.* **2011**, 123702.

Collier, P. (2007) The bottom billion: why the poorest countries are failing and what can be done about it. Oxford: Oxford University Press.

Conteh, L., Engels, T., and Molyneux, D. H. (2010) Socioeconomic aspects of neglected tropical diseases. *Lancet* **375** (9710), 239-247.

Cook, W. J., Senkovich, O., Alleem, K., and Chattopadhyay, D. (2012) Crystal structure of cryptosporidium parvum pyruvate kinase. *PLoS One* **7** (10), e46875.

Creighton, D. J. and Rose, I. A. (1976a) Studies on the mechanism and stereochemical properties of the oxalacetate decarboxylase activity of pyruvate kinase. *J Biol Chem.* **251** (1), 61-68.

Creighton, D. J. and Rose, I. A. (1976b) Oxalacetate decarboxylase activity in muscle is due to pyruvate kinase. *J Biol Chem.* **251** (1), 69-72.

Croft, S. L. (2005) Public-private partnership: from there to here. *Trans R. Soc. Trop. Med. Hyg*. **99** Suppl 1, S9-S14.

Croft, S. L., Barrett, M. P., and Urbina, J. A. (2005) Chemotherapy of trypanosomiases and leishmaniasis. *Trends Parasitol.* **21** (11), 508-512.

Cronin, C. N. and Tipton, K. F. (1985) Purification and regulatory properties of phosphofructokinase from *Trypanosoma* (Trypanozoon) *brucei brucei*. *Biochem. J.* **227** (1), 113-124.

Dahinden, P., Pos, K. M., and Dimroth, P. (2005) Identification of a domain in the alpha-subunit of the oxaloacetate decarboxylase Na^+ pump that accomplishes

complex formation with the gamma-subunit. *FEBS J.* **272** (3), 846-855.

David, C. V. and Craft, N. (2009) Cutaneous and mucocutaneous leishmaniasis. *Dermatol. Ther.* **22** (6), 491-502.

Deane, L. M. (1964) Animal reservoirs of *Trypanosoma cruzi* in Brazil. *Rev. Bras. Malariol. Doencas. Trop.* **16**, 27-48.

DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* DeLano Scientific, San Carolos, CA.

Desjeux P (2004) Leishmaniasis: current situation and new perspectives. *Comp. Immunol. Microbiol. Infect. Dis.* **27** (5), 305-318.

Dimroth, P. (1981) Characterization of a membrane-bound biotin-containing enzyme: oxaloacetate decarboxylase from Klebsiella aerogenes, *Eur. J. Biochem.* **115** (2), 353-358.

Dimroth, P. and Thomer, A. (1983) Subunit composition of oxaloacetate decarboxylase and characterization of the alpha chain as carboxyltransferase, *Eur. J. Biochem.* **137** (1-2), 107-112.

Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G., and Baker, N. A. (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* **35**, W522-525.

Dolinsky, T. J., Nielsen, J.E., McCammon, J. A., Baker, N. A. (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**, W665-667.

Dombrauckas, J. D., Santarsiero, B. D., and Mesecar, A. D. (2005). Structural basis for tumor pyruvate kinase M2 allosteric regulation and catalysis. *Biochemistry* **44** (27), 9417-9429.

Dougherty, T. M. and Cleland, W. W. (1985a) pH studies on the chemical mechanism of rabbit muscle pyruvate kinase. 1. Alternate substrates oxalacetate, glycolate, hydroxylamine, and fluoride. *Biochemistry* **24** (21), 5870-5875.

Dougherty, T. M. and Cleland, W. W. (1985b) pH studies on the chemical mechanism of rabbit muscle pyruvate kinase. 2. Physiological substrates and phosphoenol-alpha-ketobutyrate. *Biochemistry* **24** (21), 5875-5880.

Drew, M. E., Morris, J. C., Wang, Z., Wells, L., Sanchez, M., Landfear, S. M., and

Englund, P. T. (2003). The adenosine analog tubercidin inhibits glycolysis in *Trypanosoma brucei* as revealed by an RNA interference library. *J. Biol. Chem.* **278** (47), 46596-46600.

Ehrlich P. and Shiga K. (1904) Farbentherapeutische Versuche bei Trypanosomenerkrankung. *Berlin Klin. Wochenschrift* **20**, 329-362.

Eigenbrodt, E., Reinacher, M., Scheefers-Borchel, U., Scheefers, H., and Friis, R. (1992). Double role for pyruvate kinase type M2 in the expansion of phosphometabolite pools found in tumor cells. *Crit. Rev. Oncog.* **3** (1-2), 91-115.

Emsley, P. and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta. Cryst. D* **60**, 2126-2132.

Ernest, I., Callens, M., Opperdoes, F. R., and Michels, P. A. (1994). Pyruvate kinase of *Leishmania mexicana mexicana*. Cloning and analysis of the gene, overexpression in *Escherichia coli* and characterization of the enzyme. *Mol. Biochem. Parasitol.***64** (1), 43-54.

Ernest, I., Callens, M., Uttaro, A. D., Chevalier, N., Opperdoes, F. R., Muirhead, H., Michels, P. A. (1998). Pyruvate kinase of *Trypanosoma brucei*: overexpression, purification, and functional characterization of wild-type and mutated enzyme. *Protein Expr. Purif.* **13** (3), 373-382.

Evans, P. (2005) Scaling and assessment of data quality. *Acta. Cryst. D* **62**, 72-82.

Feasey, N., Wansbrough-Jones, M., Mabey, D. C. and Solomon, A. W. (2010) Neglected tropical diseases. *Br. Med. Bull.* **93**, 179-200.

Fenton, A. W., Johnson, T. A., and Holyoak, T. (2010). The pyruvate kinase model system, a cautionary tale for the use of osmolyte perturbations to support conformational equilibria in allostery. *Protein Sci.* **19** (9), 1796-1800.

Fernandes, A. P., Nelson, K., and Beverley, S. M. (1993) Evolution of nuclear ribosomal RNAs in kinetoplastid protozoa: perspectives on the age and origins of parasitism. *Proc. Natl. Acad. Sci. USA* **90**, 11608-11612.

Fèvre, E. M, Wissmann, B. V, Welburn, S. C, and Lutumba, P. (2008) The burden of human African trypanosomiasis. *PLos Negl. Trop. Dis.* **2** (12), e333.

Flynn, I. W. and Bowman, I. B. (1980). Purification and characterization of pyruvate kinase from *Trypanosoma brucei*. *Arch. Biochem. Biophys.* **200** (2), 401-409.

Flynn I. W. and Bowman I. B. (1981) Some kinetic properties of pyruvate kinase from Trypanosoma brucei: influence of pH and fructose-1,6-diphosphate. *Mol Biochem Parasitol*. **4** (1-2), 95-106.

Fothergill-Gilmore, L. A. and Michels, P. A. (1993) Evolution of glycolysis. *Prog. Biophys. Mol. Biol.* **59** (2), 105-235.

Fuad, F. A. A. (2012). Effects of metal ions on the structural and biochemical properties of trypanosomatid phosphoglycerate mutases. PhD thesis. Institute of Structural and Molecular Biology, School of Biological Sciences. University of Edinburgh, Edinburgh.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* **38**, W695-699.

Grab, D. J. and Kennedy, P. G. (2008) Traversal of human and animal trypanosomes across the blood-brain barrier. *J. Neurovirol.* **14** (5), 344-351.

Granjon, T., Maniti, O., Auchli, Y., Dahinden, P., Buchet, R., Marcillat, O., and Dimroth, P. (2010) Structure-function relations in oxaloacetate decarboxylase complex. Fluorescence and infrared approaches to monitor oxomalonate and Na(+) binding effect. *PLoS One* **5** (6), e10935.

Gualdrón-López, M., Brennand, A., Hannaert, V., Quiñones, W., Cáceres, A. J., Nwagwu, M. and Opperdoes, F. R. (1982) Regulation of glycolysis in *Trypanosoma brucei*: hexokinase and phosphofructokinase activity. *Acta Trop.* **39** (1), 61-72.

Gurtler, R. E., Segura, E. L., and Cohen, J. E. (2003) Congenital transmission of *Trypanosoma cruzi* infection in Argentina. *Emerg. Infect Dis.* **9** (1), 29-32.

Hannaert, V., Bringaud, F., Opperdoes, F. R., and Michels, P. A. (2003) Evolution of energy metabolism and its compartmentation in Kinetoplastida. *Kinetoplastid Biol Dis.* **2** (1), 11.

Hart, D. T. and Opperdoes, F. R. (1984) The occurrence of glycosomes (microbodies) in the promastigote stage of four major *Leishmania* species. *Mol. Biochem. Parasitol.* **13** (2), 159-172.

Hawking F. (1978) Suramin: with special reference to onchocerciasis. *Adv. Pharmacol. Chemother.* **15**, 289-322.

Hayward, S. and Berendsen, H. J. C. (1998). Systematic analysis of domain motions in proteins from conformational change: new results on citrate synthase and T4 lysozyme. *Proteins* **30** (2), 144-154.

Hayward, S., Kitao, A., and Berendsen, H. J. (1997). Model-free methods of analyzing domain motions in proteins from simulation: a comparison of normal mode analysis and molecular dynamics simulation of lysozyme. *Proteins* **27** (3), 425-437.

Hitosugi, T. Kang S., Vander Heiden, M. G., Chung, T. W., Elf, S., Lythgoe, K., Dong, S., Lonial, S., Wang, X., Chen, G. Z., Xie, J, Gu, T. L., Polakiewicz, R. D., Roesel, J. L., Boggon, T. 19. J., Khuri, F. R., Gilliland, D. G., Cantley, L. C., Kaufman, J., and Chen, J. (2009) Tyrosine phosphorylation inhibits PKM2 to promote the Warburg effect and tumor growth. *Sci. Signal.* **2** (97), ra73.

Hong, B., Dimov, S., Tempel, W., Auld, D., Thomas, C., Boxer, M., Jianq, J. -K., Skoumbourdis, A., Min, S., Southall, N., Arrowsmith, C. H., Edwards, A. M., Bountra, C., Weigelt, J., Bochkarev, A., Inglese, J., and Park, H. (2009) Activator-Bound Structure of Human Pyruvate Kinase M2. *unpublished. (PDB = 3GQY).*

Hong, B., Dimov, S., Tempel, W., Auld, D., Thomas, C., Boxer, M., Jianq, J. -K., Skoumbourdis, A., Min, S., Southall, N., Arrowsmith, C. H., Edwards, A. M., Bountra, C., Weigelt, J., Bochkarev, A., Inglese, J., and Park, H. (2009) Activator-Bound Structure of Human Pyruvate Kinase M2. *unpublished. (PDB = 3GR4).*

Hong, B., Dimov, S., Tempel, W., Auld, D., Thomas, C., Boxer, M., Jianq, J. -K., Skoumbourdis, A., Min, S., Southall, N., Arrowsmith, C. H., Edwards, A. M., Bountra, C., Weigelt, J., Bochkarev, A., Inglese, J., and Park, H. (2009) Activator-Bound Structure of Human Pyruvate Kinase M2. *unpublished. (PDB = 3H6O).*

Horton, A. A. and Kornberg, H. L. (1964) Oxaloacetate 4-carboxy-lyase from *Pseudomonas ovalis Chester*. *Biochim. Biophys. Acta.* **89**, 381-383.

Hotez, P. J. and Kamath, A. (2009) Neglected tropical diseases in sub-saharan Africa: review of their prevalence, distribution, and disease burden. *PloS Negl. Trop. Dis.* **3** (8), e412.

Hotez, P. J., Molyneux, D. H., Fenwick, A., Kumaresan, J., Sachs, S. E., Sachs, J. D., and Savioli, L. (2007) Control of neglected tropical diseases. *N. Engl. J. Med.* **357** (10), 1018-1027.

Hotez, P. J. and Pecoul, B. (2010) “Manifesto” for advancing the control and elimination of neglected tropical diseases. *PloS Negl. Trop. Dis.* **4** (5), e718.

Jacobs, R. T., Plattner, J. J., Nare, B., Wring, S. A., Chen, D., Freund, Y., Gaukel, E. G, Orr, M. D., Perales, J. B., Jenks, M., Noe, R. A., Sligar, J. M., Zhang, Y. K., Bacchi, C. J, Yarlett, N., and Don, R. (2011) Benzoxaboroles: a new class of potential drugs for human African trypanosomiasis. *Future Med. Chem.* **3** (10), 1259-1278.

Jetten, M. S., and A. J. Sinskey. (1995) Purification and properties of oxaloacetate decarboxylase from *Corynebacterium glutamicum*. *Antonie Van Leeuwenhoek* **67** (2), 221-227.

Jiang, J. K., Boxer, M. B., Vander Heiden, M. G., Shen M., Skoumbourdis, A. P., Southall, N., Veith, H., Leister, W., Austin, C. P., Park, H. W., Inglese, J., Cantley, L. C., Auld, D. S., and Thomas, C. J. (2010) Evaluation of thieno[3,2-b]pyrrole[3,2-d]pyridazinones as activators of the tumor cell specific M2 isoform of pyruvate kinase. *Bioorg. Med. Chem. Lett.* **20** (11), 3387-3393.

Jockel, P., Di Berardino, M., and Dimroth, P. (1999) Membrane topology of the beta-subunit of the oxaloacetate decarboxylase Na^+ pump from Klebsiella pneumoniae, *Biochemistry* **38** (41), 13461-13472.

Joosten, R. P., te Beek, T. A., Krieger, E., Hekkelman, M. L., Hooft, R. W., Schneider, R., Sander, C., and Vriend, G. (2011) A series of PDB related databases for everyday needs. *Nucleic Acids Res.* **39**, D411-419.

Juan, S. M., Cazzulo, J. J., and Segura, E. C. (1976) The pyruvate kinase of *Trypanosoma cruzi*. *Acta Physiol. Lat. Am.* **26** (5), 424-426.

Jurica, M. S., Mesecar, A., Heath, P. J., Shi, W., Nowak, T., and Stoddard, B. L. (1998). The allosteric regulation of pyruvate kinase by fructose-1,6-bisphosphate. *Structure* **6** (2), 195-210.

Jursinic, S. B. and Robinson, J. L. (1978) The active site of rabbit muscle pyruvate kinase. Evidence for a site common to the oxalacetate decarboxylase and pyruvate kinase reactions. *Biochim. Biophys. Acta.* **523** (2), 358-367.

Kabsch, W. and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers.* **22**, 2577-2637.

Kennedy, P. G. (2008a) The continuing problem of human African trypanosomiasis

(sleeping sickness). *Ann. Neurol.* **64** (2), 116-126.

Kennedy, P. G. (2008b) Diagnosing central nervous system trypanosomiasis: two stage or not to stage? *Trans. R. Soc. Trop. Med. Hyg.* **102** (4), 306-307.

Klaffl, S. and Eikmanns, B. J. (2010) Genetic and functional analysis of the soluble oxaloacetate decarboxylase from *Corynebacterium glutamicum. J. Bacteriol.* **192** (10), 2604-2612.

Kornberg, A., Ochoa, S., and Mehler, A. H. (1948) Spectrophotometric studies on the decarboxylation of beta-keto acids. *J. Biol. Chem.* **174** (1), 159-172.

Krebs, H. A. (1942) The effect of inorganic salts on the ketone decomposition of oxaloacetic acid. *Biochem. J.* **36** (3-4), 303-305.

Krissinel, E. and Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta. Cryst. D* **60**, 2256-2268.

Krissinel, E. and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372** (3), 774-797.

Kung, C., Hixon, J., Choe, S., Marks, K., Gross, S., Murphy, E., DeLaBarre, B., Cianchetta, G., Sethumadhavan, S., Wang, X., Yan, S., Gao, Y., Fang, C., Wei, W., Jiang, F., Wang, S., Qian, K., Saunders, J., Driggers, E., Woo, H. K., Kunii, K., Murray, S., Yang, H., Yen, K., Liu, W., Cantley, L. C., Vander, Heiden, M. G., Su, S. M., Jin, S., Salituro, F. G., and Dang, L. (2012) Small molecule activation of PKM2 in cancer cells induces serine auxotrophy. *Chem. Biol.* **19** (9), 1187-1198.

Labrou, N. E. and Clonis, Y. D. (1999) Oxaloacetate decarboxylase from *Pseudomonas stutzeri*: purification and characterization. *Arch. Biochem. Biophys.* **365** (1), 17-24.

Larsen, T. M., Benning, M. M., Rayment, I., and Reed G. H. (1998). Structure of the bis(Mg^{2+})-ATP-oxalate complex of the rabbit muscle pyruvate kinase at 2.1 A resolution: ATP binding over a barrel. *Biochemistry* **37** (18), 6247-6255.

Larsen, T. M., Benning, M. M., Wesenberg, G. E., Rayment, I., and Reed, G. H. (1997). Ligand-induced domain movement in pyruvate kinase: structure of the enzyme from rabbit muscle with Mg^{2+}, K^{+}, and L-phospholactate at 2.7 Å resolution. *Arch. Biochem. Biophys.* **345** (2), 199-206.

Larsen, T. M., Laughlin, L. T., Holden, H. M., Rayment, I., and Reed, G. H. (1994).

Structure of rabbit muscle pyruvate kinase complexed with Mn^{2+}, K^{+}, and pyruvate. *Biochemistry* **33** (20), 6301-6309.

Larson, S. B., Day, J. S., Cudney, R., and McPherson, A. (2007) A novel strategy for the crystallization of proteins: X-ray diffraction validation. *Acta Cryst. D* **63**, 310-318.

Laussermair, E., Schwarz, E., Oesterhelt, D., Reinke, H., Beyreuther, K., and Dimroth, P. (1989) The sodium ion translocating oxaloacetate decarboxylase of Klebsiella pneumoniae. Sequence of the integral membrane-bound subunits beta and gamma, *J. Biol. Chem.* **264** (25), 14710-14715.

López, C., Chevalier, N., Hannaert, V., Rigden, D. J., Michels, P. A., and Ramirez, J. L. (2002) *Leishmania donovani* phosphofructokinase. Gene characterization, biochemical properties and structure-modeling studies. *Eur. J. Biochem.* **269** (16), 3978-3989.

Lovell, S. C., Mullick, A. H., and Muirhead, H. (1998). Cooperativity in *Bacillus stearothermophilus* pyruvate kinase. *J. Mol. Biol.* **276** (4), 839-851.

Lv, L. Li, D., Zhao, D., Lin, R., Chu, Y., Zhang, H., Zha, Z., Liu, Y., Li, Z., Xu, Y., Wang, G., Huang, Y., Xiong, Y., Guan, K. L., and Lei, Q. Y. (2011) Acetylation targets the M2 isoform of pyruvate kinase for degradation through chaperone-mediated autophagy and promotes tumor growth. *Mol. Cell* **42** (6), 719-730.

Manderson, L. (2012) Neglected diseases of poverty. (2012) *Med. Anthropol.* **31** (4), 283-286.

Mathers, C. D, Ezzati, M., and Lopez, A. D. (2007) Measuring the burden of neglected tropical diseases: the global burden of disease framework. *PLoS Negl. Trop. Dis*. **1** (2), e114.

Mattevi, A., Valentini, G., Rizzi, M., Speranza, M. L., Bolognesi, M., Coda, A. (1995). Crystal structure of *Escherichia coli* pyruvate kinase type I: molecular basis of the allosteric transition. *Structure* **3** (7), 729-741.

Matthews B.W. (1968) *J. Mol. Biol.* **33**(2), 491-497.

Mazurek, S. (2010). Pyruvate kinase type M2: A key regulator of the metabolic budget system in tumor cells. *Int. J. Biochem. Cell. Biol.* **43** (7), 969-980.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C.

and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658-674.

McPherson, A. and Cudney, B. Searching for silver bullets: an alternative strategy for crystallizing macromolecules. (2006). *J. Struct. Biol.* **156** (3), 387-406.

Mesecar, A. D. and Nowak, T. (1997a). Metal-ion-mediated allosteric triggering of yeast pyruvate kinase. 1. A multidimensional kinetic linked-function analysis. *Biochemistry* **36** (22), 6792-6802.

Mesecar, A. D. and Nowak, T. (1997b). Metal-ion-mediated allosteric triggering of yeast pyruvate kinase. 2. A multidimensional thermodynamic linked-function analysis. *Biochemistry* **36** (22), 6803-6813.

Minodier, P. and Parola, P. (2007) Cutaneous leishmaniasis treatment. *Travel Med. Infect. Dis.* **5** (3), 150-158.

Molyneux, D. (2007) *Control of Human Parasitic Diseases,* Elsevier Books, Oxford.

Moncayo, A., Ortiz, and Yanine, M. I. (2006) An update on Chagas disease (human American trypanosomiasis). *Ann Trop Med Parasitol.* **100** (8), 663-677.

Moncayo, A. and Silveira, A. C. (2009) Current epidemiological trends for Chagas disease in Latin America and future challenges in epidemiology, surveillance and health policy. *Mem. Inst. Oswaldo. Cruz.* **104** Suppl 1, 17-30.

Moran, M. (2005) A breakthrough in RandD for neglected diseases: new ways to get the drugs we need. *PLoS Med*. **2** (9), e302.

Moran, M. et al. (2012) *G-FINDER 2012.* Neglected disease research and development: a five year rivew. Sydney, Policy Cures.

Moran, M., Guzman, J., Ropars, A. L. McDonald, A., Jameson, N., Omune, B., Ryan, S., and Wu, L. (2009) Neglected disease research and development: how much are we really spending? *PLoS Med*. **6** (2), e30.

Morgan, H. P., McNae, I. W., Hsin, K. Y., Michels, P. A., Fothergill-Gilmore, L., A., and Walkinshaw, M. D. (2010a) An improved strategy for the crystallization of *Leishmania mexicana* pyruvate kinase. *Acta Cryst. F* **66**, 215-218.

Morgan, H. P., McNae, I. W., Nowicki, M. W., Hannaert, V., Michels, P. A., Fothergill-Gilmore, L. A., and Walkinshaw, M. D. (2010b). Allosteric mechanism of pyruvate kinase from *Leishmania mexicana* uses a rock and lock model. *J. Biol. Chem.* **285** (17), 12892-12898.

Morgan, H. P., McNae, I. W., Nowicki, M. W., Zhong, W., Michels, P. A., Auld, D. S., Fothergill-Gilmore, L., A., and Walkinshaw, M. D. (2011). The trypanocidal drug suramin and other trypan blue mimetics are inhibitors of pyruvate kinases and bind to the adenosine site. *J. Biol. Chem.* **286** (36), 31232-31240.

Morgan, H. P., O'Reilly, F., Palmer, R., McNae, I. W., Nowicki, M. W., Wear, M. A., Fothergill-Gilmore, L. A., and Walkinshaw, M. D. (2012b) *unpublished. (PDB = 3SRF).*

Morgan, H. P., O'Reilly, F., Palmer, R., McNae, I. W., Nowicki, M. W., Wear, M. A., Fothergill-Gilmore, L. A., and Walkinshaw, M. D. (2012c) *unpublished. (PDB = 3SRD).*

Morgan, H. P., O'Reilly, F., Palmer, R., McNae, I. W., Nowicki, M. W., Wear, M. A., Fothergill-Gilmore, L. A., and Walkinshaw, M. D. (2012d) *unpublished. (PDB = 3SRH).*

Morgan, H. P., Walsh, M. J., Blackburn, E. A., Wear, M. A., Boxer, M. B., Shen, M., Veith, H., McNae, I. W., Nowicki, M. W., Michels, P. A., Auld, D. S., Fothergill-Gilmore, L. A., and Walkinshaw, M. D. (2012a) A new family of covalent inhibitors block nucleotide binding to the active site of pyruvate kinase. *Biochem. J.* **448** (1), 67-72.

Morrison, L. J., Marcello, L., and McCulloch, R. (2009) Antigenic variation in the African trypanosome: molecular mechanisms and phenotypic complexity. *Cell Microbiol.* **11** (12), 1724-1734.

Mubayi A, Castillo-Chavez C, Chowell G, Kribs-Zaleta C, Ali Siddiqui N, Kumar N, and Das, P. (2009) Transmission dynamics and underreporting of Kala-azar in the Indian state of Bihar. *J. Theor. Biol.* **262** (1), 177-185.

Murai, K., Morishita, M., Nakatani, R., Kubo, O., Fujioka, H., and Kita, Y. (2007) Concise total synthesis of (-)-spongotine A. *J. Org. Chem.* **72** (23), 8947-8949.

Murray, H. W., Berman, J. D., Davies, C. R., and Saravia, N. G. (2005). Advances in leishmaniasis. *Lancet* **366** (9496), 1561-1577.

Murshudov, G. N., Vagin, A. A., and Dodson E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta. Cryst. D* **53**, 240-255.

Nagajyothi, F., Machado, F. S., Burleigh, B. A., Jelicks, L. A., Scherer, P. E., Mukherjee, S., Lisanti, M. P, Weiss, L. M, Garg, N. J., and Tanowitz, H. B. (2012)

Mechanisms of *Trypanosoma cruzi* persistence in Chagas disease. *Cell Microbiol.* **14** (5), 634-643.

Narayanan, B. C., Niu, W., Han, Y., Mariano, P. S., Dunaway-Mariano, D., and Herzberg, O. (2008) Structure and function of PA4872 from *Pseudomonas aeruginosa*, a novel class of oxaloacetate decarboxylase from the PEP mutase/isocitrate lyase superfamily. *Biochemistry* **47** (1), 167-182.

Ng, S. K., Wong, M., and Hamilton, I. R. (1982) Properties of oxaloacetate decarboxylase from *Veillonella parvula. J. Bacteriol.* **150** (3), 1252-1258.

Nowicki, M. W., Tulloch, L. B., Worralll, L., McNae, I. W., Hannaert, V., Michels, P. A. M., Fothergill-Gilmore, L. A., Walkinshaw, M. D., and Turner, N. J. (2008) Design, synthesis and trypanocidal activity of lead compounds based on inhibitors of parasite glycolysis. *Bioorg. Med. Chem.* **16** (9), 5050-5061.

Nwaka, S. and Ridley, R. G. (2003) Virtual drug discovery and development for neglected diseases through public-private partnerships. *Nat. Rev. Drug. Disc*. **2** (11), 919-928.

O'Brien R., Chuang, D. T., Taylor, B. L., and Utter, M. F. (1977) Novel enzymic machinery for the metabolism of oxalacetate, phosphoenolpyruvate, and pyruvate in *Pseudomonas citronellolis*. *J. Biol. Chem.* **252** (4), 1257-1263.

Opperdoes, F. R. (1987) Compartmentation of carbohydrate metabolism in trypanosomes. *Annu. Rev. Microbiol.* **41**, 127-151.

Opperdoes, F. R. (1990) The glycosome of trypanosomes and *Leishmania*. *Biochem. Soc. Trans.* **18** (5), 729-731.

Opperdoes, F. R and Borst, P. (1977) Localization of nine glycolytic enzymes in a microbody-like organelle in *Trypanosoma brucei*: the glycosome. *FEBS Lett.* **80** (2), 360-364.

Pendergrass, D. C., Williams, R., Blair, J. B., and Fenton, A. W. (2006) Mining for allosteric information: natural mutations and positional sequence conservation in pyruvate kinase. *IUBMB Life* **58** (1), 31-38.

Pereira, K. S., Schmidt, F. L., Guaraldo, A. M., Franco, R. M., Dias, V. L., and Passos, L. A. (2009). Chagas' disease as a foodborne illness. *J. Food Prot.* **72** (2), 441-446.

Pérez-Molina, J. A., Norman, F., and López-Vélez, R. (2012). Chagas disease in

non-endemic countries: epidemiology, clinical presentation and treatment. *Curr. Infect. Dis. Rep.* **14** (3), 263-274.

Plaut, G. W. and Lardy, H. A. (1949) The oxalacetate decarboxylase of *Azotobacter vinelandii*. *J. Biol. Chem.* **180** (1), 13-27.

Potterton, E., Briggs, P., Turkenburg, M., and Dodson, E. (2003) A graphical user interface to the CCP4 program suite. *Acta. Cryst. D* **59**, 1131-1137.

Potterton, E., McNicholas, S., Krissinel, E., Cowtan, K., and Noble, M. (2002) The CCP4 molecular-graphics project. *Acta Cryst. D* **58**, 1955-1957.

Quijano-Hernandez, I. and Dumonteil, E. (2011) Advances and challenges towards a vaccine against Chagas disease. *Hum. Vaccin.* **7** (11), 1184-1191.

Racagni G. E. and Machado de Domenech, E. E. (1983) Characterization of *Trypanosoma cruzi* hexokinase. *Mol. Biochem. Parasitol.* **9** (2), 181-188.

Ran, T., Gao, Y., Marsh. M., Zhu, W., Wang, M., Mao, X., Xu, L., Xu, D., and Wang, W. (2012) Crystal structures of Cg1458 reveal a catalytic lid domain and a common catalytic mechanism for FAH family. *Biochem. J.* **449** (1), 51-60.

Ran, T., Wang, Y., Xu, D., and Wang, W. (2011) Expression, purification, crystallization and preliminary crystallographic analysis of Cg1458: a novel oxaloacetate decarboxylase from *Corynebacterium glutamicum*. *Acta Crystallogr. Sect. F.* **67**, 968-970.

Rassi, A. Jr., Rassi, A., and Marcondes de Rezende, J. (2012) American trypanosomiasis (Chagas disease). *Infect. Dis. Clin. North Am.* **26** (2), 275-291.

Rassi, A. Jr., Rassi, A., and Marin-Neto, J. Á. (2009). Chagas heart disease: pathophysiologic mechanisms, prognostic factors and risk stratification. *Mem.Inst.Oswaldo Cruz* **104**, 152-158.

Rassi, A. Jr., Rassi, A., and Marin-Neto, J. Á. (2010) Chagas disease. *Lancet* **375** (9723), 1388-1402.

Rigden, D. J., Phillips, S. E., Michels, P. A., and Fothergill-Gilmore, L. A. (1999). The structure of pyruvate kinase from *Leishmania mexicana* reveals details of the allosteric transition and unusual effector specificity. *J. Mol. Biol.* **291** (3), 615-635.

Ríos, J. F., Arboleda, M., Montoya, A. N., and Alarcón, E. P. (2011) Probable outbreak of oral transmission of Chagas disease in Turbo, Antioquia. *Biomedica.* **31**

(2), 185-195.

Robinson, J. L and Rose, I. A. (1972) The proton transfer reactions of muscle pyruvate kinase. *J. Biol. Chem.* **247** (4), 1096-1105.

Romano, A. H., and Conway, T. (1996). Evolution of carbohydrate metabolic pathways. *Res. Microbiol.* **147** (6-7), 448-455.

Rudenko, G. (2011) African trypanosomes: the genome and adaptations for immune evasion. *Essays Biochem.* **51**, 47-62.

Sakai, H. (2004) Possible structure and function of the extra C-terminal sequence of pyruvate kinase from *Bacillus stearothermophilus*. *J. Biochem.* **136** (4), 471-476.

Sakai, H. and Ohta, T. (1993). Molecular cloning and nucleotide sequence of the gene for pyruvate kinase of *Bacillus stearothermophilus* and the production of the enzyme in *Escherichia coli*. *Eur. J. Biochem.* **211** (3), 851-859.

Schaftingen, E., Opperdoes, F. R. and Hers, H. G. (1987). Effects of various metabolic conditions and of the trivalent arsenical melarsen oxide on the intracellular levels of fructose 2, 6 bisphosphate and of glycolytic intermediates in *Trypanosoma brucei*. *Eur. J. Biochem.* **166** (3), 653-661.

Schwarz, E., Oesterhelt, D., Reinke, H., Beyreuther, K., and Dimroth, P. (1988) The sodium ion translocating oxalacetate decarboxylase of Klebsiella pneumoniae. Sequence of the biotin-containing alpha-subunit and relationship to other biotin-containing enzymes, *J. Biol.Chem.* **263** (20), 9640-9645.

Shaw, J. J. (1994) Taxonomy of the genus *Leishmania*: present and future trends and their implications. *Mem. Inst. Oswaldo. Cruz.* **89** (3), 471-478.

Shin, J., Seo, Y., Cho, K. W., Rho, J. R., and Sim, C. J. (1999) New Bis(Indole) alkaloids of the topsentin class from the sponge spongosorites genitrix. *J. Nat. Prod.* **62** (4), 647-649.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol.* **7**, 539.

Simarro, P. P., Jannin, J., and Cattand, P. (2008) Eliminating human African trypanosomiasis: where do we stand and what comes next? *PLoS Med.* **5** (2), e55.

Solomons, J. T. G., Johnson, U. Schoenheit, P., and Davies, C. (2012) *unpublished. (PDB = 3QTG).*

Sols, A. (1981) Multimodulation of enzyme activity. *Curr. Top. Cell. Regul.* **19**, 77-101.

Speck, J. F. (1949) The effect of cations on the decarboxylation of oxalacetic acid. *J. Biol. Chem.* **178** (1), 315-324.

Steverding, D. (2008). The history of African trypanosomiasis. *Parasit. Vectors* 1, 3.

Stuart, D. I., Levine, M., Muirhead, H., and Stammers, D. K. (1979) Crystal structure of cat muscle pyruvate kinase at a resolution of 2.6 A.*J. Mol. Biol.* **134** (1), 109-142.

Stuart, K., Brun, R., Croft, S., Fairlamb, A., Gürtler, R. E., McKerrow, J., Reed, S., and Tarlenton, R. (2008) Kinetoplastids: related protozoan pathogens, different diseases. *J. Clin. Invest.* **118** (4), 1301-1310.

Sundar, S. and Chatterje, M. (2006) Visceral leishmaniasis - current therapeutic modalities. *Indian J. Med. Res*. **123** (3)*,* 345-352.

Studer, R., Dahinden, P., Wang, W. W., Auchli, Y., Li, X. D., and Dimroth, P. (2007) Crystal structure of the carboxyltransferase domain of the oxaloacetate decarboxylase Na^+ pump from *Vibrio cholerae*, *J. Mol. Biol.* **367** (2), 547-557.

Suzuki, K., Ito, S., Shimizu-lbuka, A. and Sakai, H. (2008) Crystal structure of pyruvate kinase from *Geobacillus stearothermophilus*. *J. Biochem.* **144** (3), 305-312.

Tanne, J. H. (2012) How the search for new drugs for neglected diseases is paying off. *BMJ*. (Overseas and Retired Doctors Edition) **344**, 18.

Taylor, J. E. and Rudenko, G. (2006) Switching trypanosome coats: what's in the wardrobe? *Trends Genet.* **22** (11), 614-620.

Taylor, M. B, Berghausen, H., Heyworth, P., Messenger, N., Rees, L. J., and Gutteridge, W. E. (1980) Subcellular localization of some glycolytic enzymes in parasitic flagellated protozoa. *Int. J. Biochem.* **11** (2), 117-120.

Tielens, A. G. and Van Hellemond, J. J. (1998) Differences in energy metabolism between trypanosomatidae. *Parasitol Today.* **14** (7), 265-272.

Tsai, C. S. (1967) Spontaneous decarboxylation of oxalacetic acid. *Can J Chem.* **45** (8), 873-880.

Tulloch, L. B., Morgan, H. P., Hannaert, V., Michels, P. A., Fothergill-Gilmore, L. A., and Walkinshaw, M. D. (2008). Sulphate removal induces a major conformational change in *Leishmania mexicana* pyruvate kinase in the crystalline state. *J. Mol. Biol.* **383** (3), 615-626.

Urbina, J. and Crespo, A. (1984) Regulation of energy metabolism in *Trypanosoma* (Schizotrypanum) *cruzi* epimastigotes. I. Hexokinase and phosphofructokinase. *Mol. Biochem. Parasitol*. **11**, 225-239.

Valentini, G., Chiarelli, L., Fortin, R., Speranza, M. L., Galizzi, A., and Mattevi, A. (2000) The allosteric regulation of pyruvate kinase. *Biol. Chem.* **275** (24), 18145-18152.

Valentini, G., Chiarelli, L. R., Fortin, R., Dolzan, M., Galizzi, A., Abraham, D. J., Wang, C., Bianchi, P., Zanella, A., and Mattevi, A. (2002). Structure and function of human erythrocyte pyruvate kinase. Molecular basis of nonspherocytic hemolytic anemia. *J. Biol. Chem.* **277** (26), 23807-23814.

Vander Heiden, M. G., Christofk, H. R., Schuman, E., Subtelny, A. O., Sharfi, H., Harlow, E. E., Xian, J., and Cantley, L. C. (2010) Identification of small molecule inhibitors of pyruvate kinase M2. *Bioccem. Pharmacol.* **79** (8), 1118-1124.

van Schaftingen, E., Opperdoes, F. R., and Hers, H. G. (1985) Stimulation of *Trypanosoma brucei* pyruvate kinase by fructose 2,6-bisphosphate. *Eur. J. Biochem.* **153** (2), 403-406.

van Schaftingen, E., Opperdoes, F. R., and Hers, H. G. (1987) Effects of various metabolic conditions and of the trivalent arsenical melarsen oxide on the intracellular levels of fructose 2,6-bisphosphate and of glycolytic intermediates in *Trypanosoma brucei*. *Eur. J. Biochem.* **166** (3), 653-661.

Verlinde, C. L., Hannaert, V., Blonski, C., Willson, M., Périé J. J., Fothergill-Gilmore, L. A., Opperdoes, F. R., Gelb, M. H., Hol, W. G., and Michels, P. A. (2001). Glycolysis as a target for the design of new anti-trypanosome drugs. *Drug Resist. Updat.* **4** (1), 50-65.

Vollmer, S. H., Walner, M. B., Tarbell, K. V., and Colman, R. F. (1994) Guanosine 5'-O-[S-(4-bromo-2,3-dioxobutyl)]thiophosphate and adenosine 5'-O-[S-(4-bromo-2,3-dioxobutyl)]thiophosphate. New nucleotide affinity labels which react with rabbit muscle pyruvate kinase. *J. Biol. Chem.* **269** (11), 8082-8090.

von Geldern, T., Harhay, M. O., Scandale, I., and Don, R. (2011). Kinetoplastid parasites. *Top. Med. Chem.* **7**, 181-241.

Wernimont, A. K., Hutchinson, A., Hassanali, A., Kozieradzki, I., Cossar, D., Bochkarev, A., Arrowsmith, C. H., Edwards, A. M., Bountra, C., Weigelt, J., Hui, R., Hills, T., and Pizarro, J.C. (2012) *unpublished. (PDB = 3MA8).*

WHO. (2002) Control of Chagas disease. Second report of the WHO Expert Committee. Technical report series no 905. Geneva: World Health Organization.

WHO. (2004) The World Health Report 2004. Changing history. Annex Table 3: Burden of Disease in DALYs by cause, sex and mortality stratum in WHO regions, estimates for 2002, Geneva: World Health Organization.

WHO. (2006a) Human African trypanosomiasis (sleeping sickness): epidemiological update. *Wkly Epidemiol. Rec.* **81** (8), 71-80.

WHO. (2006b) Neglected tropical diseases: hidden successes, emerging opportunities. Geneva: World Health Organization.

Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., and Hochstrasser, D. F. (1999) Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.* **112**, 531-552.

Wilkinson, S. R. and Kelly, J. M. (2009) Trypanocidal drugs: mechanisms, resistance and new targets. *Expert Rev. Mol. Med.* **11**, e31.

Williams, R., Holyoak, T., McDonald, G., Gui, C., and Fenton, A. W. (2006). Differentiating a ligand's chemical requirements for allosteric interactions from those for protein binding. Phenylalanine inhibition of pyruvate kinase. *Biochemistry* **45** (17), 5421-5429.

Wooll, J. O., Friesen, R. H. E., White, M. A., Watowich, S. J., Fox, R. O., Lee, J. C., and Czerwinski, E. W. (2001). Structural and functional linkages between subunit interfaces in mammalian pyruvate kinase. *J. Mol. Biol.* **312** (3), 525-540.

World Health Organization/Industry Drug Development Working Group. (2001) Working paper on priority infectious diseases requiring additional RandD. http://www.who.int/intellectualproperty/documents/en/IFPMA. pdf.

Young, C., Losikoff, P., Chawla, A., Glasser, L., and Forman, E. (2007) Transfusion-acquired *Trypanosoma cruzi* infection. *Transfusion.* **47** (3), 540-544.

Zeledón, R. and Rabinovich, J. E. (1981) Chagas' disease: an ecological appraisal with special emphasis on its insect vectors. *Annu. Rev. Entomol.* **26**, 101-133.

Zijlstra, E. E., Musa, A. M., Khalil, E. A. G., El Hassan, I. M., and El-Hassan, A. M. (2003) Post -kala-azar dermal leishmaniasis. *Lancet infect.Dis*. **3** (2), 87-98.

Zoraghi, R., See, R. H., Axerio-Cilies, P., Kumar, N. S., Gong, H., Moreau, A., Hsing, M, Kaur, S., Swayze, R. D., Worrall, L., Amandoron, E., Lian, T., Jackson, L., Jiang, J., Thorson, L., Labriere, C., Foster, L., Brunham, R. C., McMaster, W. R., Finlay, B. B., Strynadka, N. C., Cherkasov, A., Young, R. N., and Reiner, N. E. (2011a) Identification of pyruvate kinase in methicillin-resistant *Staphylococcus aureus* as a novel antimicrobial drug target. *Antimicrob Agents Chemother.* **55** (5), 2042-2053.

Zoraghi, R., Worrall, L., See R. H., Strangman, W., Popplewell, W. L., Gong, H., Samaai, T., Swayze, R. D., Kaur, S., Vuckovic, M., Finlay, B. B., Brunham, R. C., McMaster, W. R., Davies-Coleman, M. T., Strynadka, N. C., Andersen, R. J., and Reiner, N. E. (2011b) Methicillin-resistant *Staphylococcus aureus* (MRSA) pyruvate kinase as a target for bis-indole alkaloids with antibacterial activities. *J. Biol. Chem.* **286** (52), 44716-44725.