METABOLOMIC ANALYSIS OF *TRYPANOSOMA CONGOLENSE* AND *LEISHMANIA MEXICANA* TREATED WITH SELECTED ANTIPARASITIC DRUGS AND *IN SILICO* MODELLING OF POTENTIAL DRUG TARGETS

ONDARI LAURAH NYASITA

A thesis submitted in partial fulfilment of the requirements for the Degree in Master of Science in Bioinformatics of Pwani University

DECEMBER 2022

DECLARATION

This thesis is my original work and has not been presented in any other University or any other award.

Name: Laurah Nyasita Ondari   **Reg no**: SG30/PU/36050/20

Signature   Date: 13th July 2023

We confirm that the work reported in this thesis was carried out by the candidate under our supervision.

1. **Prof. Michael P. Barrett**

Institute of Infection and Immunity

University of Glasgow

Signature   Date: 13th July, 2023

2. **Dr. Daniel Masiga**

Animal and Human Health Theme

International Centre of Insect Physiology and Ecology (ICIPE)

Signature   Date: 13th July, 2023

3. **Prof. Suhaila Omar Hashim**

Department of Biochemistry and Biotechnology

Pwani University, Kilifi

Signature   Date:

DEDICATION

This work is dedicated to my parents Mary and Tom Ondari, and to Valarie, Byron, and

Kevin who have been the wind in my wings.

ACKNOWLEDGEMENT

iv

ABSTRACT

Animal African Trypanosomiasis (AAT) is a devastating parasitic disease caused by a variety of haemoflagellate extracellular protozoa of the genus *Trypanosoma*. It is endemic in 37 of the 55 countries in Africa and is transmitted by the obligate blood feeder of the genus *Glossina*. It affects both domestic and wild animals causing huge economic losses and severe health problems. Chemotherapeutic and prophylactic agents such as Isometamidium chloride (ISM) are used to prevent the progression of AAT in endemic areas. These agents face limitations such as resistance, toxicity, low efficacy, and unknown mechanisms of action. Therefore, it is crucial to develop new, safe and effective trypanocides or optimize those in current use by understanding their mechanism of action (MOA) through untargeted metabolomics. In this study, the MOA of ISM was first investigated through statistical and metabolomic analysis and molecular docking. Statistical analysis was done through Principal Component Analysis (PCA), one-way Analysis of Variance (ANOVA), and Tukey's HSD *post hoc* analysis, cluster, and fold change analysis. Significant perturbations were observed in glycolysis, acetate: succinate CoA transferase succinyl CoA synthetase (ASCT/SCS), and energy metabolites indicating that ISM may have an inhibitory effect on the glucose transporter. Molecular docking assays showed ISM interacting with the glucose binding site of the transporter. Isometamidium chloride interacted with amino acid residues known to be conserved in all sugar transporter members and to play an important role in transport of hexoses and formation of the exofacial substrate binding site. Secondly, to increase the drug arsenal against AAT, new or existing drugs used against biologically and biochemically similar kinetoplastids have been probed for their novel activity against *Trypanosoma* species. In this study, the MOA of novel anti-trypanosomatid compounds dubbed JYH and VMS were elucidated through the aforementioned analyses. The novel compounds are toxic to

kinetoplastid protozoa with activity recorded against *Leishmania mexicana.* For one of the compounds (VMS), statistical and metabolomic analysis showed possible perturbations to *T. congolense* metabolism, particularly in energy metabolism, aromatic amino acid metabolism, and cellular redox. However, from the analyses, it was not possible to readily assign candidate target proteins in *T. congolense*. In *L. mexicana*, inhibition of the Coenzyme A biosynthesis pathway at dephospho-CoA kinase (DPCK) or phosphopantetheine adenylyltransferase (PPAT) enzymes was evident for the two novel compounds. Molecular docking assays showed a possible inhibition of dephospho-CoA kinase by the drugs interacting with the CoA binding domain where the native substrate dephospho-CoA binds. On phosphopantetheine adenylyl transferase the drugs seemed to interact with the active site where adenosine triphosphate (ATP) binds. This study identified potential drug targets for ISM, JYH and VMS. These findings provide new insights into the mechanisms of action of these drugs and suggest avenues for future drug development efforts.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

AAT	Animal African Trypanosomiasis
ADT	AutoDock tools
ANOVA	Analysis of variance
ASCT-SCS	Acetate: succinate CoA transferase succinyl CoA synthetase
ATP	Adenosine triphosphate
BSF	Bloodstream forms
CDD	Conserved Domain Database
DNA	Deoxyribonucleic acid
DPCK	Dephospho-CoA kinase
DZ	Diminazene aceturate
FAO	Food and Agriculture Organization
GUI	Graphical user interface
HTS	High throughput screening
IDEOM	Identification and evaluation of metabolomics data from LC-MS
IL-12	Interleukin-12
IL-6	Interleukin-6
IM	Intramuscular
ISM	Isometamidium chloride
JYH	JYH-G-52-1, a novel anti-trypanosomatid compound from Kip Guy & Scott Landfear (USA)
kDNA	Kinetoplast DNA
LC/MS	Liquid chromatography/mass spectrometry
MF	Metacyclic forms
MOA	Mechanism of action

MS	Mass spectrometry
NMR	Nuclear magnetic resonance
ODC	Ornithine decarboxylase
PCA	Principal component analysis
PCF	Procyclic forms
PGK	Phosphoglycerate kinase
PLS	Partial least squares
PLS-DA	Partial least squares discriminant analysis
PPAT	Phosphopantetheine adenylyltransferase
RCSB-PDB	Research Collaboratory for Structural Bioinformatics Protein Data Bank
RNA	Ribonucleic acid
ROC	Receiver operating characteristic curve
SBDD	Structure-based drug design
SC	Subcutaneous
TCA	Tricarboxylic acid
TNF	Tumor necrosis factor
TOPO II	Kinetoplastid topoisomerase type II
VMS	VMS-7-25, a novel anti-trypanosomatid compound from Kip Guy & Scott Landfear (USA)
VSG	Variant surface glycoprotein

CHAPTER ONE

1.0 INTRODUCTION

1.1 Background

African Animal Trypanosomiasis (AAT) is a serious and fatal parasitic disease that is endemic in 37 of the 55 countries found on the African continent (Yaro et al., 2016). Millions of animals in the sub-Saharan region are grossly affected leading to losses that equate to approximately 4.5 billion USD per annum in food, dung, and drafting power (Yaro et al., 2016). The disease is transmitted by an obligate blood feeder of the genus *Glossina* commonly referred to as the tsetse fly. Tsetse flies are exclusively found within the African continent, and are distributed within the tropics between latitudes 12°N and 25°S (Dorn et al., 2011). A combination of biological and mechanical factors is responsible for the confinement of AAT within the tropics such as equable climates and infestation by the tsetse fly vector (Shereni et al., 2021). Animal African Trypanosomiasis is characterised by fever, weakness and lethargy, weight loss, anaemia, staring coat, and discharge from the eyes among other symptoms depending on the extracellular protozoan responsible for the infection (Spickler, 2003).

According to the Food and Agriculture Organization (FAO), various protozoan species of the genus *Trypanosoma* are responsible for AAT including *Trypanosoma congolense, Trypanosoma vivax, Trypanosoma simiae,* and *Trypanosoma brucei* (*African Animal Trypanosomiasis; Selected Articles from the World Animal Review*, n.d.). *T. congolense* is known to cause a majority of infections in domestic animals, causing fatal chronic disease that leads to death if left untreated (Coustou et al., 2010). These pathogenic African trypanosomes, including *T. congolense*, have complex life cycles that involve the tsetse vector and mammalian host (Peacock et al., 2012). The procyclic form (PCF), confined to

the tsetse fly and the bloodstream form (BSF) confined to the mammalian host are the two important replicating stages (Colasante et al., 2006). The best-studied trypanosome is *Trypanosoma brucei* where the two life forms differ significantly in their metabolic requirements. For example, BSF are completely reliant on glycolysis and substrate-level phosphorylation for their energy requirements while PCF favour proline as their main energy source (Eyford et al., 2011; Steketee et al., 2021). For *T. congolense* however, the BSF appear to be less dependent on glucose and their metabolism has key differences when compared to *T. brucei,* as shown in a thorough metabolome overview of *T. congolense* blood stream form by Steketee et al., (2021).

Although there are other methods for controlling progression of AAT in endemic areas, several interventions target the BSF form of the parasite. They include prophylactic and chemotherapeutic trypanocides that are administered parenterally by either intramuscular (IM) or subcutaneous (SC) routes (Richards et al., 2021). The most common trypanocides available in African markets include homidium chloride (Novidium), homidium bromide (Ethidium), diminazene aceturate (Berenil), quinapyramine sulfate (Antrycide), and isometamidium chloride (Samorin, trypamidium) (Richards et al., 2021).

Diminazene aceturate (DZ) was discovered in 1955 (Kuriakose & Uzonna, 2014). It is the first-line treatment for AAT (Giordani et al., 2016) and is administered IM/SC. There is reported resistance to diminazene aceturate following high accessibility, limited knowledge, and improper use (Joshua et al., 1995; Moti et al., 2015). Isometamidium chloride is an amphiphilic cationic phenanthridine used in veterinary medicine for prophylaxis and chemotherapy. Its trypanocidal action was first described by Wragg et al., (1958). The mechanism of action of ISM is poorly understood and resistance to it is thought to be due to

recurrent exposure to the drug as well as frequent under-dosing (Delespaux & de Koning, 2007).

Due to the aforementioned challenges as a result of toxicity and resistance, there is a pressing need to discover and develop new prophylactic and chemotherapeutic agents against AAT as well as optimize the ones in current use. Drugs used against neglected diseases caused by taxonomically related kinetoplastids have been used in an effort to increase the drug arsenal against AAT (Alkhaldi et al., 2019; Croft et al., 1996; Papagiannaros et al., 2005). Such drugs include those used against Leishmaniasis. Leishmaniasis is a parasitic disease that is endemic in 98 countries in Asia, Africa, Central America and southern Europe (Barrett & Croft, 2012). The disease is caused by infection with Leishmania parasites, which are spread by the bite of infected female phlebotomine sand flies (Barrett & Croft, 2012). Leishmaniasis manifests in two major forms; visceral and cutaneous Leishmaniasis.

The process of optimizing and developing new drugs for AAT is complex and difficult. The process is enhanced through analysis of the mechanism of action (MOA) of existing drugs (Creek & Barrett, 2014) as well as advancing knowledge of how *T. congolense* develops resistance. One approach to achieve this is through an untargeted overview of cellular metabolism following drug exposure through metabolomics.

Metabolomics is a relatively new 'omics' technology that refers to the study of small biomolecules that are the intermediates or end products of metabolic processes (Hendriks et al., 2011; Patti, 2011). Metabolomics can be targeted (hypothesis-driven) or untargeted (hypothesis-generating). Untargeted metabolomics is particularly advantageous when finding drug-mediated responses in biological systems as it can classify metabolic identities and biochemical pathways involved in trypanocidal effects (Creek & Barrett, 2014). This approach has been used in previous studies to give insight on the MOA of eflornithine

(Vincent et al., 2010), a trypanocide used against *T. brucei* that inhibits the ornithine decarboxylase enzyme (Mbekeani et al., 2019) as well as the antileishmanial alkylphospholipid, miltefosine (Armitage et al., 2018).

This study is aimed at profiling the cellular metabolome of *T. congolense* parasites upon exposure to selected antiprotozoal drugs; including isometamidium (ISM) and novel drug candidates (here dubbed JYH and VMS) (Hammill et al., 2021; Ortiz et al., 2017) with metabolomics based on liquid chromatography-mass spectrometry (LC-MS). The metabolomes of *T. congolense* and *Leishmania mexicana,* two related kinetoplast parasite, were studied. A comparison of the metabolome outputs of *T. congolense* and *L. mexicana* treated with one of the novel compound was conducted with the aim of seeking differences and similarities in the MOA. Furthermore, protein structure predictions for the target candidates were conducted and evaluated, *in silico,* to determine docked poses of the drugs to the identified targets.

1.2 Problem statement

Trypanosoma congolense is the major causal pathogen of the economically and socially important disease AAT (Coustou et al., 2010). Animal African Trypanosomiasis causes high morbidity and mortality, as well as huge economic losses in Africa (Kasozi et al., 2022; Richards et al., 2021; Spickler, 2003). A majority of the efforts to decrease disease burden and eliminate the disease rely on the use of chemotherapeutic drugs, especially isometamidium chloride which is used for both chemoprophylaxis and cure (Kasozi et al., 2022; Stevenson et al., 1995; Tihon et al., 2017). Efforts to eliminate this disease, through the use of trypanocides, have been hindered by resistance, toxicity, poor efficacy, unsuitable pharmacokinetics, and unaffordable costs (Creek & Barrett, 2014). The aforementioned limitations, coupled with lack of proper and well-funded pharmaceutical research, as well as

limited knowledge, or lack thereof, of the MOA of existing drugs, contribute to the persistent progression of AAT. There is, therefore, an urgent need for new and improved drugs.

Compared to its counterpart, *T. brucei*, little is known about *T. congolense* biochemistry (Steketee et al., 2021). Understanding its metabolism, especially in a host environment that contains trypanocides, can contribute to understanding drug MOA, the parasite's mechanism of drug resistance and identifying unknown drug targets. The discovery of the MOA sheds light and allows optimization of the pharmacokinetics of current chemotherapeutic compounds thus allowing monitoring of efficacy, resistance and toxicity.

1.3 Justification

Many routinely used antiprotozoal drugs act by uncertain mechanisms of action, making appropriate clinical use and monitoring for efficacy, toxicity, and resistance difficult (Creek & Barrett, 2014). The process of drug discovery, development, and appropriate use of trypanocides is substantially aided by understanding the mechanism of action of drugs.

The neglected diseases Leishmaniasis and Trypanosomiasis are caused by taxonomically, structurally, and biochemically related kinetoplastid haemoflagellates (Barrett & Croft, 2012; Stuart et al., 2008). Owing to the structural and biochemical relatedness, previous studies have probed for novel activity of antileishmanials in *Trypanosoma spp* (Alkhaldi et al., 2019; Bouton et al., 2021; Croft et al., 1996; Papagiannaros et al., 2005). This has been done in an effort to increase the drug arsenal against AAT. This study investigated how a novel compound known to be active against Leishmania parasites perturbs the metabolome of *T. congolense* and *L. mexicana*.

Elucidating the mechanism of action of drugs used against AAT is crucial for improving treatment outcomes, reducing drug resistance, identifying new drug targets, and advancing

the understanding of the parasite's biology (Giordani et al., 2016; Yang et al., 2014). Previous studies have indicated that understanding the mechanism of action of drugs can help in developing new and more effective drugs, leading to better treatment outcomes (Giordani et al., 2016). Additionally, *T. congolense* and other AAT causing parasites have developed resistance to drugs used for treatment, and understanding how drugs work against the parasite can help in developing strategies to prevent or slow down the development of resistance. Identifying new pharmacological targets is essential, and studying the mechanism of action of drugs used against AAT can provide insights into the biology of the parasite, leading to new strategies for disease control. Therefore, elucidating the mechanism of action of drugs used against AAT is crucial for the development of effective treatments and disease control.

The University of Glasgow has generated metabolomics data of *T. congolense* parasites treated with isometamidium chloride, and *T. congolense* and *L. mexicana* treated with novel antiprotozoal agents, VMS and JYH, for this investigation. In this study, changes in metabolite levels relative to untreated parasites were analysed in order to elucidate the drugs' mechanisms of action.

1.4 Objectives

1.4.1 Overall objective
To analyze the metabolic effects of isometamidium chloride and two novel drug candidates, VMS and JYH, on *Trypanosoma congolense* and *Leishmania mexicana* and identify potential drug targets *in silico.*

1.4.2 Specific objectives
1. To investigate the effects of isometamidium chloride on the metabolome of *Trypanosoma congolense* for identification of potential drug targets.

2. To compare the effects of VMS on the metabolome of *Trypanosoma congolense* and *Leishmania mexicana* for identification of potential drug targets.

3. To infer the interaction mechanisms of isometamidium chloride, VMS and JYH agents against the identified drug targets through *in silico* modeling.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 The life cycle of *Trypanosoma congolense* and *Leishmania mexicana*

Trypanosoma congolense is the major causal agent of the economically and socially important disease called Animal African Trypanosomiasis (AAT) also known as Nagana (Coustou et al., 2010). Although knowledge about this parasite is comparatively limited compared to its counterpart *T. brucei*, evidence to date suggests that it also undergoes a complex developmental cycle spanning two hosts; the tsetse fly vector and the mammalian host (Bringaud et al., 2006; Peacock et al., 2012). Most of the developmental stages are found in the insect vector. Bloodstream forms (BSF) that actively divide and multiply in the blood of an infected mammalian host are ingested by the tsetse vector during a blood meal. In the insect vector, they differentiate into procyclic forms (PCF) in the insect's midgut then migrate to the proboscis where they attach to the labrum and proliferate into epimastigotes. The parasites finally transform into metacyclic forms (MF) which are non-dividing and have a variant surface glycoprotein (VSG) coat. Metacyclic forms are infective when transferred to a mammalian host during the next blood meal (Hirumi & Hirumi, 1984). *T. congolense*'s life cycle is illustrated in **Figure 1** below.
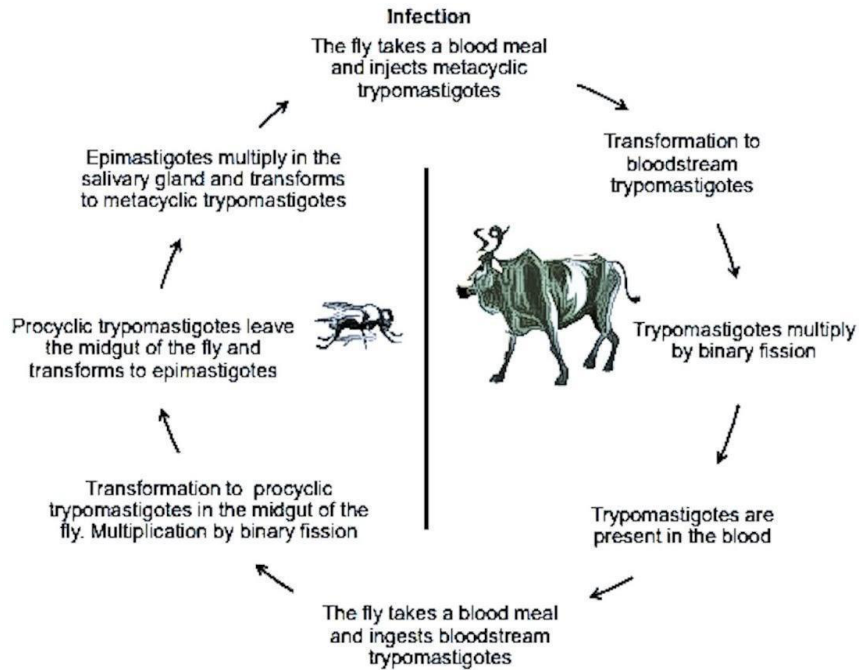
Figure 1: The developmental life cycle of *Trypanosoma brucei* showing its life stages both in the tsetse fly vector and the mammalian host (Stein et al., 2014).

Leishmania mexicana is a kinetoplastid hameoflagellate that is taxonomically, structurally, and biochemically related to *T. congolense* (Barrett & Croft, 2012; Stuart et al., 2008). *L. mexicana* causes the cutaneous form of the neglected disease Leishmaniasis. According to Damianou et al., (2020), Leishmaniasis creates significant public health issues and accounts for 20,000 – 40,000 annual deaths in 98 tropical and subtropical nations. *Leishmania spp.* and *Trypanosoma spp.* belong to the phylum *Euglenozoa*, family *Trypanosomatidae* and class *Kinetoplastida* (Kaufer et al., 2017; Kostygov et al., 2021). The developmental cycle of *Leishmania spp.* involves two primary morphological forms, intracellular amastigotes in the mammalian host and motile promastigotes in the sand fly vector. Infection to humans occurs when infective metacyclic promastigotes are injected into the body when an infected female phlebotomine sandfly draws a blood meal. The promastigotes are phagocytized by

macrophages and other types of mononuclear phagocytic cells. The promastigotes transform into amastigotes which multiply by simple division in the mononuclear phagocytic cells. In turn sand flies become infected when they ingest these amastigotes during a blood meal. The amastigotes transform into promastigotes and develop in the midgut. This life cycle is illustrated in **Figure 2** below.

The most well-known types of promastigotes, known as the mammal-infective phases of metacyclic promastigotes, have been described in sandfly infections. Gossage et al., (2003) evidenced two unique, consecutive growth cycles with four distinct life cycle phases occur during development in *Lutzomyia longipalpis* sand fly larvae for *L. mexicana*. Procyclic promastigotes start the first growth cycle. These promastigotes then divide in the bloodmeal of the abdominal midgut and give rise to nectomonad promastigotes, which do not divide. Nectomonad forms control the infection's anterior migration; they later change into leptomonad promastigotes, which start a new growth cycle in the anterior midgut. In order to be transmitted to a mammalian host, leptomonad promastigotes later differentiate into non-diverging metacyclic promastigotes.
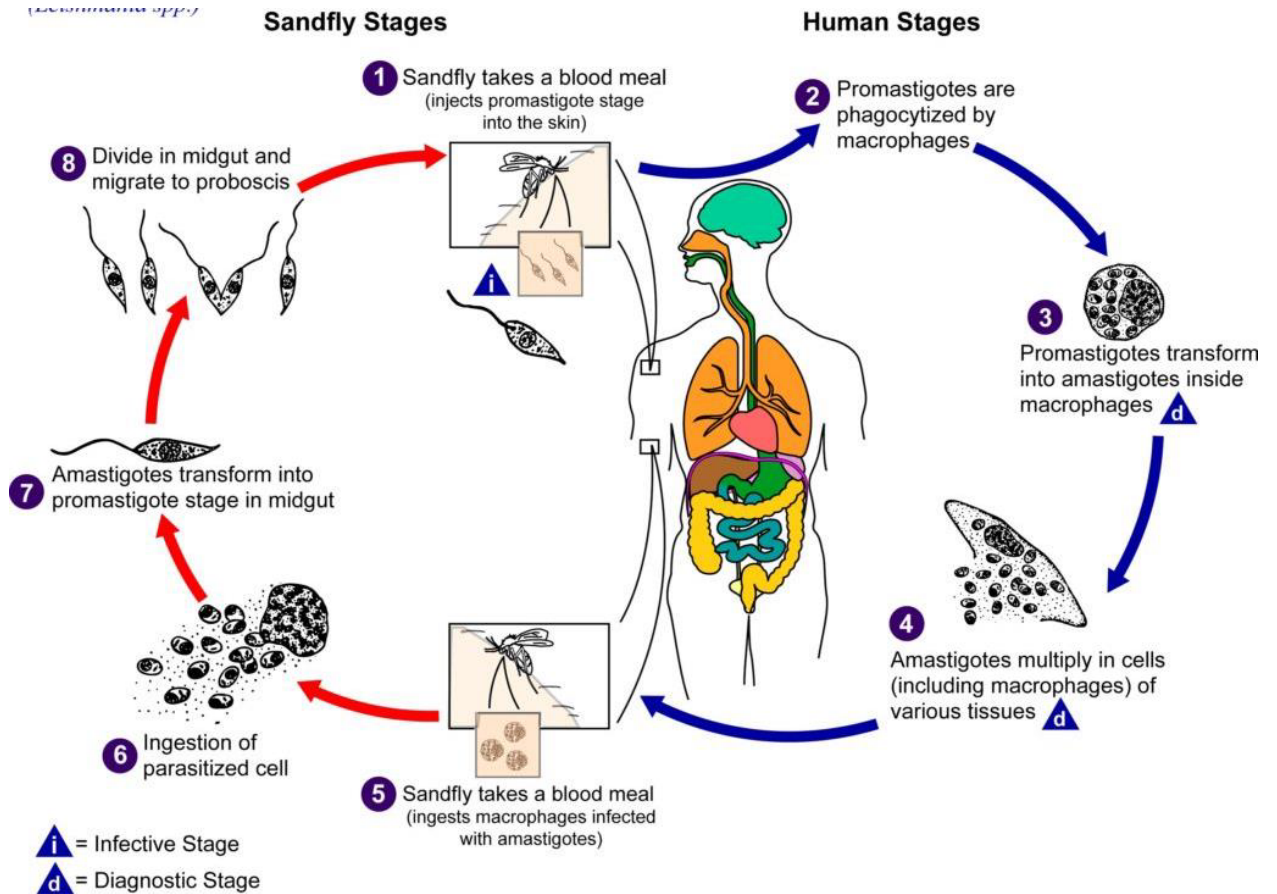
Figure 2: The developmental cycle of *Leishmania mexicana* in the human host and the sandfly vector. Adapted from (CDC - Leishmaniasis - Biology, 2020)

2.2 The *Trypanosoma congolense* and *Leishmania mexicana* metabolomes

The metabolome refers to the collection of small molecules known as metabolites in a cell, organ or organism (Wishart et al., 2007). According to the metabolic profiles of the best-studied trypanosomatids including *T. brucei*, *T. cruzi* and *Leishmania spp.*, the metabolic complexities of these expressed forms differ significantly due to the different environments found in hosts (Bringaud et al., 2006). For example, *T. brucei* and *T. cruzi* trypomastigotes depend heavily on glucose catabolism for their energy requirements because their vertebrate host environment is rich in glucose. However, in the insect vector haemolymph where glucose is scarce, their energy is prominently from L-proline or L-glutamine (Steketee et al.,

2021). Observations from a comparative metabolomics and transcriptomics study of *T. congolense* and *T. brucei* conducted by Steketee et al., (2021) showed that *T. congolense*'s metabolic consumption and output differed from that of *T. brucei*. In the glycolytic pathway, contrary to *T. brucei* where the major output is pyruvate*, T. congolense* did not produce high levels of this metabolite but instead showed increased levels of succinate, acetate and malate. This is in line with previous observations seen by Agosin & von Brand, (1954). Although these tricarboxylic acid (TCA) cycle intermediates were found, further results from the Steketee et al., (2021) study showed that in addition to glycolysis and the TCA cycle, the reversal of the glycosomal succinate shunt is also involved in their generation. In addition, the study suggested ATP generation could also be from the acetate: succinate CoA transferase succinyl-CoA synthetase (ASCT-SCS) cycle for *T. congolense* as the uptake of D-glucose in this parasite was decreased compared to *T. brucei*. Further, increased mitochondrial activity was observed that resulted in high levels of acetate (Steketee et al., 2021). Regarding transcriptomics, *T. congolense* showed increased levels in transcripts associated with gluconeogenesis, the succinate shunt and the acetate generation pathway. Moreover, the study proposed that *T. congolense* relies on alternative sources to obtain ribose instead of pentose phosphate pathway (Steketee et al., 2021). Finally, it showed reduced levels of fatty acids derived from threonine and glucose suggesting that *T. congolense* scavenges fatty acids as opposed to synthesizing them.

Similarly, metabolome studies of Leishmania parasites have revealed distinct metabolic profiles in different life stages. An integrated metabolomics and proteomics investigation by Akpunarlieva et al., (2017) sought to study molecular variations of wild-type and mutant models of Leishmania. The study identified a number of variations, including a lack of glucose transport and several phenotypic changes in the mutant model. Additionally, the

findings showed that, in response to oxidative stress, metabolic pathways of glycoconjugate production and redox homeostasis involved in glucose metabolism underwent a series of changes. These changes demonstrated compatibility with the loss of sugar uptake capacity and explained the mutant's low virulence among the hosts. Another study by Westrop et al., (2015) found that *in vitro* cultures of logarithmic and stationary phase promastigotes of Leishmania parasites exhibit distinct metabolic profiles, suggesting increased differentiation into metacyclic forms. These differences included changes in nucleoside and nucleobase levels, as well as alterations in membrane glycerol phospholipid structure that affect membrane fluidity. The study which made a comparison between promastigotes and axenic amastigotes (cultured amastigotes) of *L. mexicana* revealed that promastigotes consume more glucose and amino acids, releasing partially catabolized products into the medium (Westrop et al., 2015). Metabolomic analysis using labeled substrates showed the release of specific metabolites by promastigotes, whereas axenic amastigotes exhibited a more conservative metabolic response characterized by decreased glucose and amino acid uptake and increased fatty acid catabolism. These metabolic differences were also observed in amastigotes derived from animal lesions, indicating their natural occurrence. Furthermore, a comparison of different strains of *L. donovani* with varying susceptibility to sodium stibogluconate identified metabolic differences related to oxidative stress protection and membrane fluidity (Westrop et al., 2015). However, analysis of *L. infantum* promastigotes treated with miltefosine did not show significant alterations in membrane phospholipids but indicated increased turnover of internal membrane lipids (Westrop et al., 2015).

These differences in the metabolism with respect to the parasite's environment are advantageous in studying how they survive and thrive (Steketee et al., 2021). In this study, this underlying knowledge of metabolic differences was applied to investigate how the

metabolome of *T. congolense* and *L. mexicana* is altered when they encounter chemotherapeutic compounds as a way of: 1) explaining how these drugs work, 2) determining potential targets of these drugs.

2.3 Use of trypanocides to control animal African Trypanosomiasis

Animal African Trypanosomiasis jeopardizes the lives of about 50 million cattle and causes approximately 3 million cattle deaths every year (FAO, 2019). To prevent the risk to development and food security caused by this disease, appropriate control measures need to be taken (Richards et al., 2021). A systematic review on Trypanosomiasis and tsetse fly control options conducted by Meyer et al., (2016) stated that prior to the 1950s, control primarily consisted of practices that had significant environmental consequences, such as bush clearing, dichlorodiphenyltrichloroethane (DDT) ground spraying, and elimination of wildlife that exhibited undesirable qualities both in terms of traits and economic production through a process referred to as wildlife culling. Since the 1980s, however, more environmentally and politically acceptable approaches, such as selective bush clearing, sequential aerial spraying (SAS), insecticide-treated traps and targets (ITT), insecticide-treated cattle (ITC) used as live baits, and the sterile insect technique (SIT), were developed (Meyer et al., 2016).

Thus far, interventions to keep the disease at bay have included: 1) area-wide elimination and control of the tsetse fly vector (Schofield & Kabayo, 2008), 2) use of trypanocidal compounds as prophylactic and chemotherapeutic agents (Hargrove et al., 2012; Prayag et al., 2020), and 3) pyrethroid insecticide spraying of cattle (Muhanguzi et al., 2015). These control options are summarized in **Figure 3** below. A review by Richards et al., (2021) stated that control against AAT relies heavily on anti-trypanosomal chemotherapy. Furthermore,

FAO stipulates that 35 million doses of trypanocides are administered to cattle annually to prevent the progression of this disease in endemic areas (FAO, 2019).
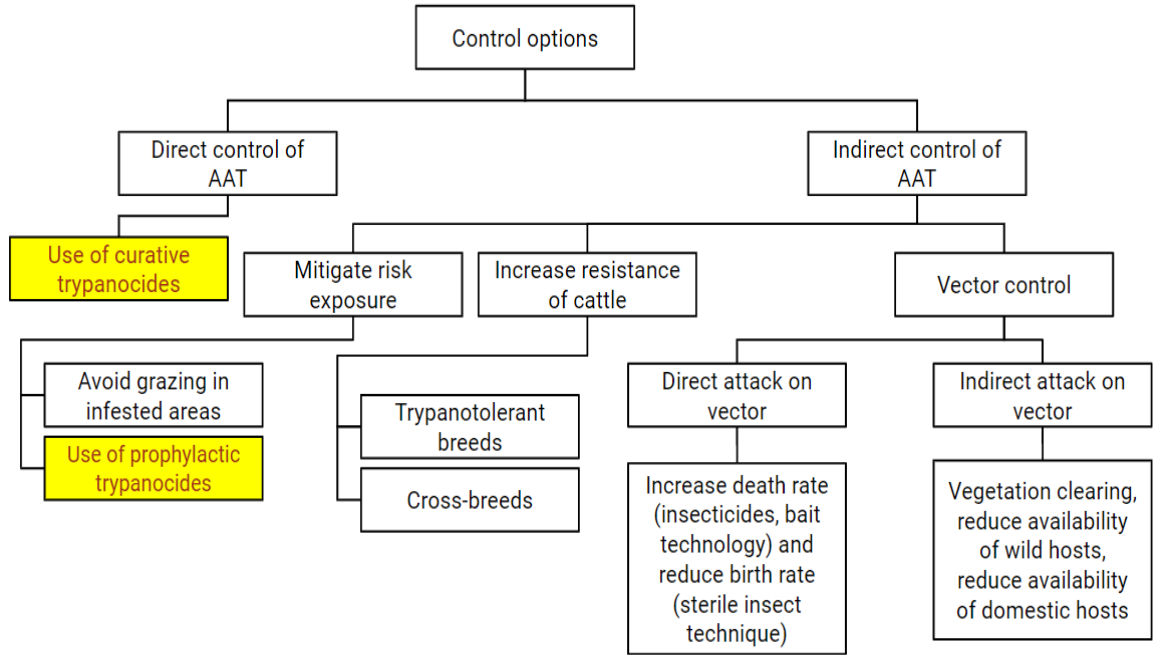
Figure 3: A summary diagram of the control options available for controlling AAT. Methods highlighted in yellow emphasize the use of trypanocides (adapted from Meyer et al., (2016)).

Six veterinary trypanocidal compounds are licensed for use against AAT. These trypanocides include homidium bromide (ethidium), homidium chloride (Novidium), diaminazene aceturate (Berenil), quinapyramine sulfate (Antrycide), isometamidium chloride (Samorin) and suramin sodium. Isometamidium chloride (ISM), Diminazine aceturate (DZ) and homidium salts are frequently used in Africa (Giordani et al., 2016).

The phenanthridine ISM and the diamidine DZ are the most commonly used drugs against AAT caused by *T. congolense* (and *T. vivax*) (FAO, 2019). Diminazine aceturate is the first line treatment for AAT. Its mechanism of action has been proposed to involve host effects,

including altering pathways associated with cytokine production reducing the production of interleukin-6 (IL-6), interleukin-12 (IL-12), and tumor necrosis factor (TNF) in macrophages (Kuriakose et al., 2012; Kuriakose & Uzonna, 2014). Additionally, binding to the parasite's kinetoplast (mitochondrial DNA) has also been proposed (Brack et al., 1972). Resistance to DZ has been observed in areas where there is continuous use of the trypanocide. In a study conducted by De Koning et al., (2004), it was observed that DZ resistance is due to its loss as it is transported across the plasma membrane, as previously suggested by Barrett et al., (1995) in *Trypanosoma equiperdum.* Further, a study by Carruthers et al., (2021) observed that DZ resistance in *T. congolense* seems to be associated with reduced mitochondrial membrane potential.

Isometamidium chloride is the most recent drug developed for control of AAT and has been in use since it was first introduced six decades ago (Richards et al., 2021). It is the only recommended prophylactic trypanocide at higher doses. The typical dose for treatment is 0.25-1.0 mg/Kg while that for prophylaxis is 0.5-1 mg/Kg (Giordani et al., 2016). Despite the popularity of ISM, Onyeyili & Egwu, (1995) credited its narrow use to low safety margins and severe reactions at inoculation sites. Furthermore, research conducted by Tihon et al., (2017) indicates that 17 out of the 37 countries where AAT is endemic have reported resistance towards ISM. Further contraindications for its use have been attributed to toxicity (Kaminsky et al., 1997). The mechanism of action (MOA) of ISM is not properly known. However, previous studies have ascribed its ability to inhibit kinetoplast topoisomerase type II (TOPO II) which is involved in the linking and unlinking of the dense minicircle DNA network essential for encoding guide ribonucleic acid (RNAs) (Kaminsky et al., 1997). The cleaving of deoxyribonucleic acid (DNA) has also been proposed (Delespaux & de Koning, 2007). The homidium salts (homidium chloride and homidium bromide) have been halted

for use in certain countries where AAT is endemic such as Nigeria due to widespread resistance (Kasozi et al., 2022). However, they are used in areas where resistance to ISM and DZ is observed (Onyeyili & Egwu, 1995).

Due to the low-profit market in the African continent, high costs in design, development, and licensing of new drugs, limited funding for vaccine research, and little interest in neglected diseases it is vital to optimize the use of current trypanocides (Geerts et al., 2001). The mechanism of action of most of the drugs used to control AAT is not fully understood (Creek & Barrett, 2014) which hinders surveillance on mechanisms that render them toxic, and their modes of resistance.

Besides the known trypanocides, other bioactive compounds against related pathogens have been assessed for cross-species potency in *Trypanosoma* species (Alkhaldi et al., 2019; Croft et al., 1996; Ortiz et al., 2017). The neglected diseases Leishmaniasis and Trypanosomiasis are caused by taxonomically, structurally, and biochemically related kinetoplastid haemoflagellates (Barrett & Croft, 2012; Stuart et al., 2008). Owing to the structural and biochemical relatedness, previous studies have probed for novel activity of antileishmanials in *Trypanosoma spp* (Alkhaldi et al., 2019). Two such compounds (JYH and VMS) are evaluated in this study.

The *p*-chloronitrobenzamide scaffold JYH was identified through phenotypic screening and showed strong potency against all *Trypanosoma* species *in vitro* (Hwang et al., 2013) and efficacious against *T. congolense* and *T. brucei in vivo* (Ortiz et al., 2017). Ortiz et al., (2017) previously assessed JYH's potency against *L. mexicana*. It's mechanism of action prior to this study was unknown despite its evidence of great potency, acceptable metabolic stability in both human and rat microsomal models, modest solubility, and reasonable permeability

(Hwang et al., 2013). Additionally, the presence of microtubule-interacting and vimentin-like proteins in Leishmania parasites led to the evaluation of 3-arylquinolines, which were previously studied as oncogene leads, in Leishmania parasites and resulted in the modified arylquinolone VMS (Hammill et al., 2021). From the study by Hammill et al., (2021), VMS was identified as an early lead against Leishmaniasis since it possessed good intracellular amastigote activity (EC50 = 120 nM) and a 30-fold selectivity index. In the study by Hammill et al., (2021), it was also shown that VMS retained activity against several patient-derived antimony-resistant strains of the parasite *L. donovani* that causes visceral Leishmaniasis (Hammill et al., 2021). In this study, the MOA of VMS against *T. congolense* and *L. mexicana* was compared and the MOA of both JYH and VMS was evaluated in *L. mexicana*.

2.4 Metabolomics-based strategies for drug target identification

Metabolomics is a relatively new 'omics' technology that involves the identification and quantification of the cellular metabolome of a particular system (Vincent et al., 2012). Unlike other omics technologies, namely proteomics and genomics, which are frequently complicated by modifications such as epigenetic and post-translational modifications, metabolomics provides unique molecular profiles of activities that occur at the cellular level hence allowing easier correlation with observable physical metabolic outputs (Patti, 2011). Metabolomics is usually conducted with nuclear magnetic resonance (NMR) or mass spectrometry (MS) to analyze highly abundant biomolecules or for wide coverage metabolome analyses respectively (Vincent et al., 2016). Mass spectrometry is advantageous over NMR in terms of sensitivity and accuracy when detecting the mass of hundreds to thousands of metabolites in a single sample (Vincent & Barrett, 2015). There are two general approaches in metabolomics; the hypothesis-driven approach known as targeted

metabolomics, and the hypothesis-generating approach also known as untargeted metabolomics (Creek & Barrett, 2014).

Targeted metabolomics quantifies a list of targeted metabolites discriminately while untargeted metabolomics aims to quantify as many small biomolecules as possible in an unbiased manner (Patti, 2011). Untargeted metabolomics is often conducted with liquid chromatography-mass spectrometry (LC/MS) because of: 1) its ability for global profiling (Vincent & Barrett, 2015), 2) its high sensitivity (Gertsman & Barshop, 2018; Patti, 2011), and 3) its ability to proceed without chemical derivatization (Cui et al., 2018). Untargeted metabolomics has a wide range of applications including in drug discovery to elucidate the MOA of drugs that act by unknown means (Creek & Barrett, 2014; Li et al., 2016), identifying disease biomarkers, finding mechanisms of toxicity, and determining mechanisms of resistance (Vincent et al., 2016).

Pharmaceutical drug discovery programs have relied heavily on high throughput screening (HTS) (Halouska et al., 2011) and phenotypic screening (Vincent & Barrett, 2015). High throughput target-based screening involves screening for lead compounds in large chemical libraries to find those that interact with specific targets while eliminating compounds that do not show significant interactions (Jahnke, 2007). Phenotypic screening, on the other hand, aims to find compounds that kill parasites in cell-based systems, irrespective of the knowledge of the target. While these methods are widely applied, they suffer various challenges which mean additional refinement of compounds is necessary to create useful drugs. High throughput screening against targets has high failure rates because compounds that inhibit particular targets often fail to cross membranes to enter parasites or might be metabolically unstable (Halouska et al., 2011). Phenotypic screening, which by definition is finding compounds that kill parasites, is unable to pinpoint the toxic and synergistic effects

of lead compounds since drug MOAs are not known (Vincent & Barrett, 2015). Untargeted metabolomics can therefore help elucidate the mechanisms by which drugs cause cellular responses, making it an attractive and advantageous route.

Untargeted metabolomics has been previously studied for anti-protozoan MOA research. Vincent et al., (2012) conducted a study that aimed to ascertain the mechanism of action of the ornithine decarboxylase inhibitor, eflornithine. The study showed how metabolomics can be applied to such research, especially if the drug's mechanism of action involves enzyme inhibition. In the study, an increase in ornithine decarboxylase's (ODC) substrate, L-ornithine was observed with a concomitant decrease in its product, putrescine, showing that the drug acted by a single means that was inhibition of ODC (Vincent & Barrett, 2015). The same study also analysed the *T. brucei* metabolome to study the synergistic effects of nifurtimox and eflornithine and showed that the two drugs were mildly antagonistic *in vitro*. Another study by Armitage et al., (2018) designed to investigate the effect of miltefosine on the metabolome and lipidome of treated Leishmania amastigotes, hypothesized the importance of sphingolipids and ergosterol for the drug's sensitivity although not able to pinpoint a single target.

2.5 Metabolomics data preprocessing and pretreatment

Metabolomics experiments are typically designed to assess the abundances of endogenous biomolecules in biological systems (Bartel et al., 2013). These abundances are unique molecular signatures that infer the physiological state of the biological system (Saccenti et al., 2013). The resulting metabolic profiles usually consist of huge amounts of data that pose a challenge when making relevant biological inferences from the data (Worley & Powers, 2013). Two broad statistical analysis approaches, namely univariate and multivariate methods, are used to extract meaning from metabolomics datasets depending on experimental

design (Saccenti et al., 2013; Worley & Powers, 2013). Univariate methods focus on the observation and analysis of individual variables (Saccenti et al., 2013) while multivariate methods consider multiple variables simultaneously along with underlying relationships between the variables (Vinaixa et al., 2012). Biological systems and processes, as well as LC/MS, are inherently multivariate in nature (Vinaixa et al., 2012) rendering many metabolomics analyses multivariate, although picking individual metabolites whose abundance is affected in different situations is fundamental. Univariate methods only focus on mean and variance (Saccenti et al., 2013) and include t-tests or analysis of variance (ANOVA) (Saccenti et al., 2013) for data that comes from a normal distribution and Kruskal-Wallis, Mann-Whitney, Wilcoxon signed-rank and Friedman tests for data that is not normally distributed (Vinaixa et al., 2012). Multivariate analyses, in contrast, besides focusing on simultaneous analysis of all variables, also focus on covariance and correlations (Saccenti et al., 2013). In metabolomics, two routinely applied multivariate analyses are dimensionality reduction models such as principal component analysis (PCA) and partial least squares (PLS).

Before metabolomics datasets are analyzed they have to undergo pre-processing and pretreatment (Hendriks et al., 2011; van den Berg et al., 2006). Data preprocessing refers to the generation of clean data by removing noise. Specifically, in untargeted metabolomics through MS, it refers to dealing with artifacts and other noise in the data, retention time shifts, and baseline drifting (Hendriks et al., 2011; van den Berg et al., 2006). Data preprocessing of the datasets used in this study was done using Identification and Evaluation of Metabolomics data from LC/MS (IDEOM). The IDEOM platform is a user-friendly Excel based template with many macros that enable pre-processing of raw LC/MS data to annotated and hyperlinked metabolite lists (Creek et al., 2012). The template does not require any

installation and can be loaded directly onto the latest version of Microsoft Excel (Creek et al., 2012). Many other freely available applications exist for alignment, deconvolution, detection, integration and identification (Blekherman et al., 2011). Many of these applications are, however, built on statistical software and require some expertise to navigate them (Creek et al., 2012). The IDEOM template on the other hand has a GUI with R-based msconvert, XCMS and MzMatch for preprocessing, and additional automated filtering and annotation to remove artefacts in LC/MS (Creek et al., 2012) which is a major problem in LC/MS data (Blekherman et al., 2011). The IDEOM template also conducts automated identification of putative metabolites based on their retention time and mass and then annotates the metabolites with confidence levels. **Figure 4** is a representation of how IDEOM pre-processes the data and outputs annotated and identified putative metabolites with confidence levels to allow quick and easy data analysis and visualisation that can be used to make biological inferences.
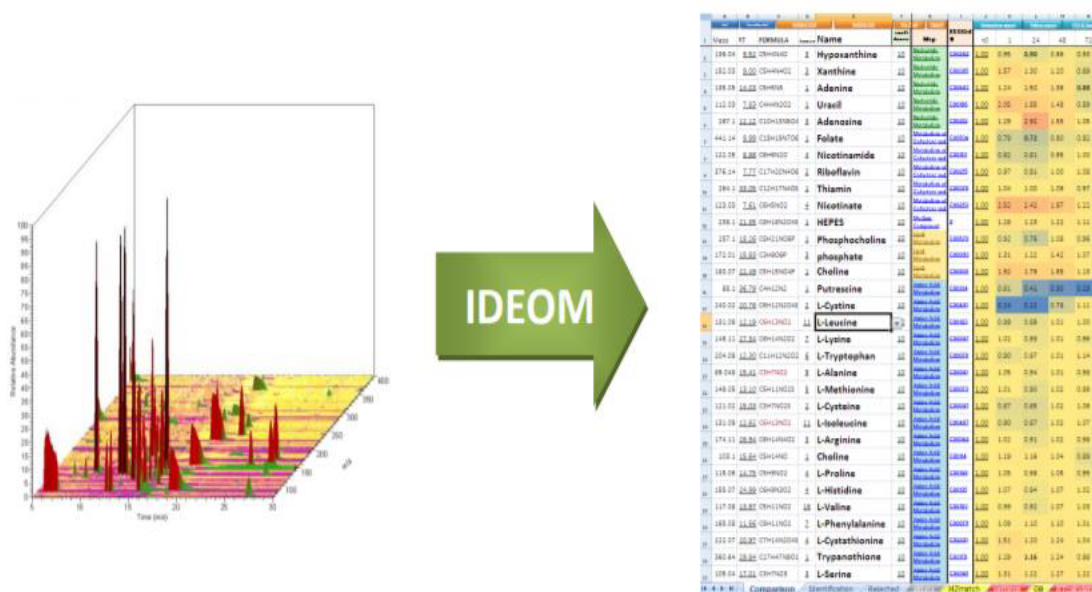
Figure 4: A representation of LC/MS data alignment, identification and annotation of putative metabolites using IDEOM. On the left is raw peaks obtained after LC/MS while on the right is a comparison sheet obtained after data preprocessing using IDEOM (Creek et al., 2012; *MzMatch/PeakML: Metabolomics Data Analysis*, n.d.).

The output obtained from IDEOM consists of several Excel sheets including: 1) "Alldata" that contains information about every peak set in the peak list, 2) "Rejected" that contains metabolites whose confidence level is below 5 (these peaks represent noise and artefacts), 3) "Identification" that contains metabolites whose confidence level is above 5, 4) "AllBasepeaks" that contains all base peaks from MzMatch, and 5) "Comparison" that contains the output from the comparison of peak intensities and has many evaluation and visualisation functions.

The comparison sheet contains information on the mass of the putative metabolite, its retention time, its putative formula, number of isomers in the database with the exact mass and retention time, the name of the putative metabolite, the confidence level with which it was identified and annotated, the metabolic pathway that it maps to as well as information

about the samples used to generate the data as shown in **Figure 5**. The information contains

hyperlinks to their source databases (Creek et al., 2012).
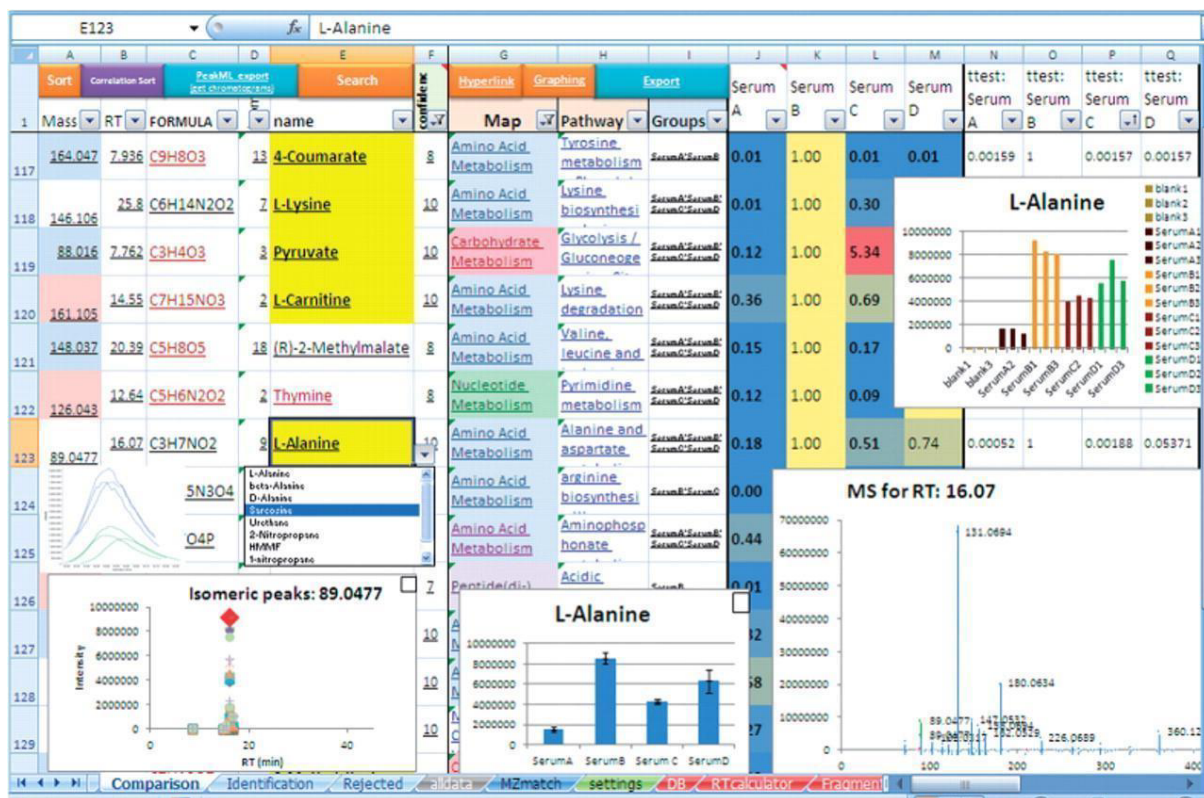
Figure 5: A screenshot of the comparison sheet in IDEOM containing lists of hyperlinked lists of important metabolite information (Creek et al., 2012).

After data pre-processing and metabolite identification, data is pretreated for downstream

data analysis. Data pretreatment aims to showcase the biological relevance of clean data by

emphasizing the differences in the relative peak areas of metabolites in different samples.

This is done by relating these differences to physiological states of the biological systems in

which they occur (van den Berg et al., 2006). Methods for pretreating data include scaling,

centering and normalization (Karaman, 2017; van den Berg et al., 2006). Centering ensures

all metabolite concentrations are fluctuations, not around a central mean, but around zero

therefore large differences between highly and lowly abundant metabolite concentrations are

adjusted (van den Berg et al., 2006). This adjustment minimizes the high offsets in the data

leaving biologically relevant variations only (Karaman, 2017; van den Berg et al., 2006). Centering is vital in metabolomics as seen in studies conducted by Ducruix et al., (2008), Narduzzi et al., (2020), among many others. Scaling involves division of metabolite relative peak areas observed by a scaling factor which can be mean or standard deviation (van den Berg et al., 2006). Scaling aims to convert the differences observed in metabolite concentrations to values that are relative to the scaling factor (Karaman, 2017). Different scaling techniques exist including autoscaling, Pareto scaling, range scaling, vast scaling and level scaling. All scaling techniques with the exception of level scaling, which uses mean as a scaling factor, use standard deviation (Karaman, 2017; van den Berg et al., 2006). Autoscaling is the best known and most used scaling technique in metabolomics (Gromski et al., 2014).

Transformations are often applied to metabolomics datasets to convert multiplicative effects to additive relations, therefore correcting for heteroscedasticity in the data (Karaman, 2017; van den Berg et al., 2006). Two common transformation methods used in metabolomics include log and power transformations (van den Berg et al., 2006). Log transformation corrects data heteroscedasticity. However its disadvantages include: 1) inability to deal with zero characters as their transformed values begin to approach negative infinity and 2) over emphasis of low concentration metabolites (van den Berg et al., 2006). Power transformations on the other hand are able to deal with zero values but their inability to turn multiplicative effects into additive effects make them undesirable (Karaman, 2017; van den Berg et al., 2006).

2.6 Metabolomics statistical analysis software

For downstream metabolomics data analysis, many metabolomics analyses are now performed locally using commercial statistical software tools such as MatLab, MS-Excel,

SigmPlot, and SIMCA-P (Xia et al., 2009). The metabolomics community, in particular, makes extensive use of SIMCA-P (Umetrics) due to its great graphic capabilities and comprehensive analytic choices (Xia et al., 2009). SIMCA-P is, however, quite expensive offering only one month of free trial. Software used for metabolomics analysis that do not require installation include MeltDB and MetaboAnalyst. MeltDB is a web-based software platform that focuses on the storage, administration, analysis, and annotation of MS-based metabolomics data (Neuweger et al., 2008; Xia et al., 2009). MeltDB allows user access control and a more flexible ontology-based metabolomics experiment annotation. It also provides the ability to integrate and parameterize preprocessing procedures and methodologies that can be uploaded to a compute cluster (Neuweger et al., 2008). Unfortunately, a number of popular browsers (Firefox, Netscape) appear to have security certificate difficulties with this server, and access requires a user login and password (Xia et al., 2009). Furthermore, the statistical capabilities of MeltDB such as t-tests, volcano plots, PCA, and heat maps are limited to GC/LC-MS data only (Xia et al., 2009).

MetaboAnalyst version 5.0 is a user-friendly web application frequently used in analysis of metabolomics datasets (Xia & Wishart, 2016) with over 300,000 users according to the metaboanalyst.ca website. The application can handle data from both NMR and MS platforms either in their raw form containing spectral bins or in the preprocessed form containing metabolite concentrations (Xia & Wishart, 2016). The platform allows for both univariate and multivariate data analysis depending on one's experimental design and study goals. It also has the capability of recording R commands used in data analysis and generation of various outputs such as heatmaps, PCA plots, PLS plots and receiver operating characteristic (ROC) curves. Furthermore, MetaboAnalyst is also available as a package for use in R programming, that allows users flexibility in commands for their desired output.

Compared to other software tools, MetaboAnalyst supports a wider range of data types, and more complex data analysis methodologies, data annotation tools, and automated report generating utilities. In addition, it includes a number of other methods not available in SIMCA-P, such as volcano plots, Significance Analysis of Microarrays (and Metabolites) (SAM), k-means, self-organizing maps (SOM), random forest, and support vector machine (SVM). While MetaboAnalyst lacks the graphical versatility of SIMCA-P, it is more accessible (through the web), free, and simple to use even on R. MetaboAnalyst also includes its own metabolite and route identification tools, which are not available in any other statistical software package. However, because of its reliance on the Human Metabolome Database (HMDB) infrastructure, MetaboAnalyst's coverage of plant and microbial metabolism is limited.

Several studies have shown the use of MetaboAnalyst and MetaboAnalystR for metabolomics data analysis (Cadena-Zamudio et al., 2022; Costanzo et al., 2022). In this study, MetaboAnalystR package was used for statistical data analysis.

Figure 6 below shows the general metabolomics workflow from data generation to statistical analysis.
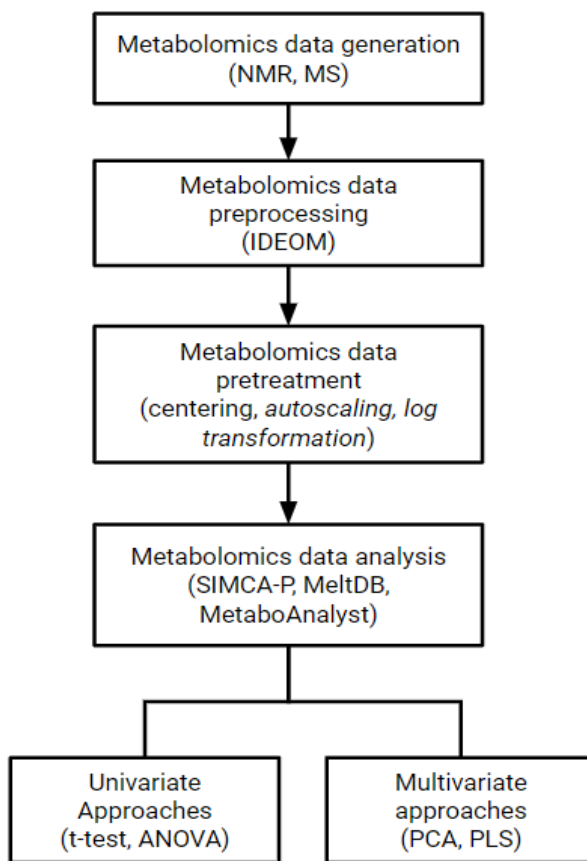
Figure 6: A general metabolomics workflow. Text within parentheses show the different methods available for the process. Metabolomics data is generated through NMR or MS, it is then preprocessed and pretreated before it is analysed using different statistical tools.

2.6 Molecular docking

The metabolomics analysis aims to identify changes in abundance of metabolites in drug treated cells – which, it is hoped, will point to metabolic enzymes inhibited by the drugs. Upon identification of such enzymes, *in silico* predictions of drug binding potential can be performed. Molecular docking is a prominent *in silico* structure-based drug design (SBDD) method that predicts the interaction between small molecules known as ligands and their biological macromolecules also known as target (Pinzi & Rastelli, 2019). The main aim of molecular docking is to find ligands that have particular electrostatic and stereochemical

properties to acquire the greatest binding affinity to the biological receptor (Ferreira et al., 2015). The process typically begins with obtaining the molecular pose of the ligand within its target and then establishing if the ligand and the receptor are complementary to each other through the use of a scoring function (Pinzi & Rastelli, 2019). In the context of drug design, the ligands are usually drugs while the targets could be proteins (enzymes, cell surface receptors, ion gated channels, membrane transport proteins) or nucleic acids (McMasters, 2018; Santos et al., 2016; Singh et al., 2006). Prior to docking it is necessary to first generate a three-dimensional structure of the target. This can be achieved through a number of methods such as crystallography on purified targets then exposed to X-rays, nuclear magnetic resonance (NMR) or through computational structure prediction using various software packages. Increasingly, molecular modelling is used to generate structures *in silico* and most recently an Artificial Intelligence based algorithm, called Alpha-fold 2.0, was released by DeepMind and produces robust structures from protein amino acid sequences (Jumper et al., 2021).

A number of different software packages have been developed to allow docking. AutoDock is a popular example alongside DOCK, GLIDE, GOLD, FlexX (Pakpahan et al., 2013), and LUDI (Morris et al., 2008) among others. AutoDock is a freely available suite of C programs that perform computational docking and virtual screening (Forli et al., 2016). The C programs work in tandem and they include: 1) AutoDock Tools (ADT) – graphical user interface (GUI) for preparation of receptor and ligand coordinates and thereafter docking and analysis, 2) AutoDock Vina – a relatively new program to AutoDock that performs both screening and docking with increased speed and accuracy (Trott & Olson, 2010), 3) AutoDock - for docking simulations between the ligand and receptor, 4) Raccoon2 - for virtual screening 5) AutoLigand - for prediction of suitable binding pockets for the ligand (Forli et al., 2016) and

6) AutoGrid that predetermines a grid box onto where the ligand binds on the receptor based on the interaction of energies (Morris et al., 1996). The current distribution of the original AutoDock suite (AutoDock 4) is available in two main distributions, AutoDock and AutoDock Vina and they both contain ADT and AutoGrid.

Docking programs generally culminate in two components: 1) ligand poses in a binding pocket, and 2) ranked poses based on binding affinities (Tanchuk et al., 2016). The programs use a scoring function to predict the affinity of a ligand (or many ligands) binding to a receptor's binding site (Pakpahan et al., 2013). Most scoring functions are designed to score interactions between a flexible ligand and a rigid receptor meaning that the receptor does not undergo any changes during the docking process. However, for protein targets, it is widely known that once a ligand binds, they undergo conformational changes to facilitate this binding and therefore increase the number of interactions (Anderson et al., 2001). Ensemble docking is therefore desired for protein targets to generate ensemble candidates upon binding of the ligand (Amaro et al., 2018). AutoDock and AutoDock Vina use quite different scoring functions (Tanchuk et al., 2016). The empirical scoring function of AutoDock Vina is a machine learning algorithm compared to AutoDock's that is physics-based (Quiroga & Villarreal, 2016). The AutoDock Vina scoring function calculates the contribution of several individual forces such as gauss1, gauss2, hydrogen bonding, hydrophobic and repulsion forces to the target-ligand binding (Quiroga & Villarreal, 2016). The Scoring function is as follows:

$$\Delta G_{binding} = \Delta G_{vdW} + \Delta G_{elec} + \Delta G_{hbond} + \Delta G_{desolv} + \Delta G_{tors}$$

Where ΔG_{vdW} is the Leonard-Jones potential, ΔG_{elec} is the coulombic with Solmajer-dielectric, ΔG_{hbond} is the Goodford directionality, ΔG_{desolv} is the Stouten pairwise atomic

solvation parameters and ΔG_{tors} is the number proportional to rotatable bonds (Pakpahan et al., 2013).

The AutoDock Vina scoring function slightly outperforms AutoDock 4 (Tanchuk et al., 2016). AutoDock and AutoDock Vina were used in this study for the identification and scoring of binding poses of the drugs to the identified drug targets. Prior studies have used molecular docking to validate leishmanicidal and trypanocidal compound activities against their molecular targets (Ibezim et al., 2018; Ogungbe & Setzer, 2009; Rock et al., 2021) . Molecular docking studies were therefore performed to infer interaction mechanisms between the drug compounds and their macromolecular targets.

CHAPTER THREE

3.0 MATERIALS AND METHODS

3.1 Experimental design

Materials used in this study comprise of three global metabolomics datasets profiling the chemical perturbations of *Trypanosoma congolense* and *Leishmania mexicana* parasites which were treated with anti-parasitic molecules. The three datasets were as follows: Isometamidium chloride treated *T. congolense*, *T. congolense* treated with VMS-7-25 and *L. mexicana* treated with VMS-7-25 and JYH-G-52-1.

Data generation which involved cell cultures of the parasites exposed to the anti-parasitic drugs and data preprocessing was previously performed at Glasgow Polyomics (2020) prior to this study. Metabolomics data was generated as described below:

For *T. congolense* treated with ISM, BSF *T. congolense* parasites were grouped into three groups representing three types of sample groups (Untreated, Low dose, High dose). The three sample groups were cultured in SCM-6 minimal media at different drug concentrations as follows: Untreated – 0 nM ISM, Low dose – 0.6 nM ISM, High dose – 6.0 nM ISM. The samples were made in 4 replicates of 60 ml each (total of 12 replicates) that were incubated at 34° C and monitored in different concentrations of ISM as stated above for 24 hrs.

For *T. congolense* treated with VMS-7-25 from Kip Guy & Scott Landfear (USA), the parasites were treated with 7.6-8 µl of 19.5 µM VMS-7-25 (Hammill et al., 2021) for 15 min at 34°C. Samples were grouped into control and VMS with 4 replicates in each sample group. For *L. mexicana* treated with JYH-G-52-1 and VMS-7-25 from Kip Guy & Scott Landfear (USA), 38 µl of 25mM JYH-G-52-1(Ortiz et al., 2017) and 1.09 µl of 100 mM VMS-7-25 and 37.91 µl of DMSO was added to 10 ml of culture. Cells were incubated at 25 °C for 10

min with JYH-G-52-1 and 15min with VMS-7-25. Samples were grouped into control and treated with 4 replicates in each sample group.

Metabolites were then detected using Orbitrap mass spectrometry and data in the form of mass peaks as well as retention times was obtained (Kamleh et al., 2008). Metabolomics data was generated and automatic metabolite annotation and identification was performed using Identification and Evaluation of Metabolomics data (IDEOM) software (Creek et al., 2012).

3.2 Study materials

The materials used in this study included four datasets and three compounds. A summary of samples that were used in this study as well as the total putative metabolites identified in the samples and their dataset assignment are detailed in **Table 1** below.

Table 1: Summary of samples of the datasets

Dataset	Number of samples used	Number of putative metabolites identified
ISM *T. congolense* IL3000	12	776
VMS-7-25 *T. congolense*	8	565
VMS-7-25 *L. mexicana*	8	830
JYH-G-52-1 *L. mexicana*	8	830

The structure of the compounds evaluated i.e. isometamidium chloride, JYH-G-52-1 and VMS-7-25 are shown in **Figure 7.**
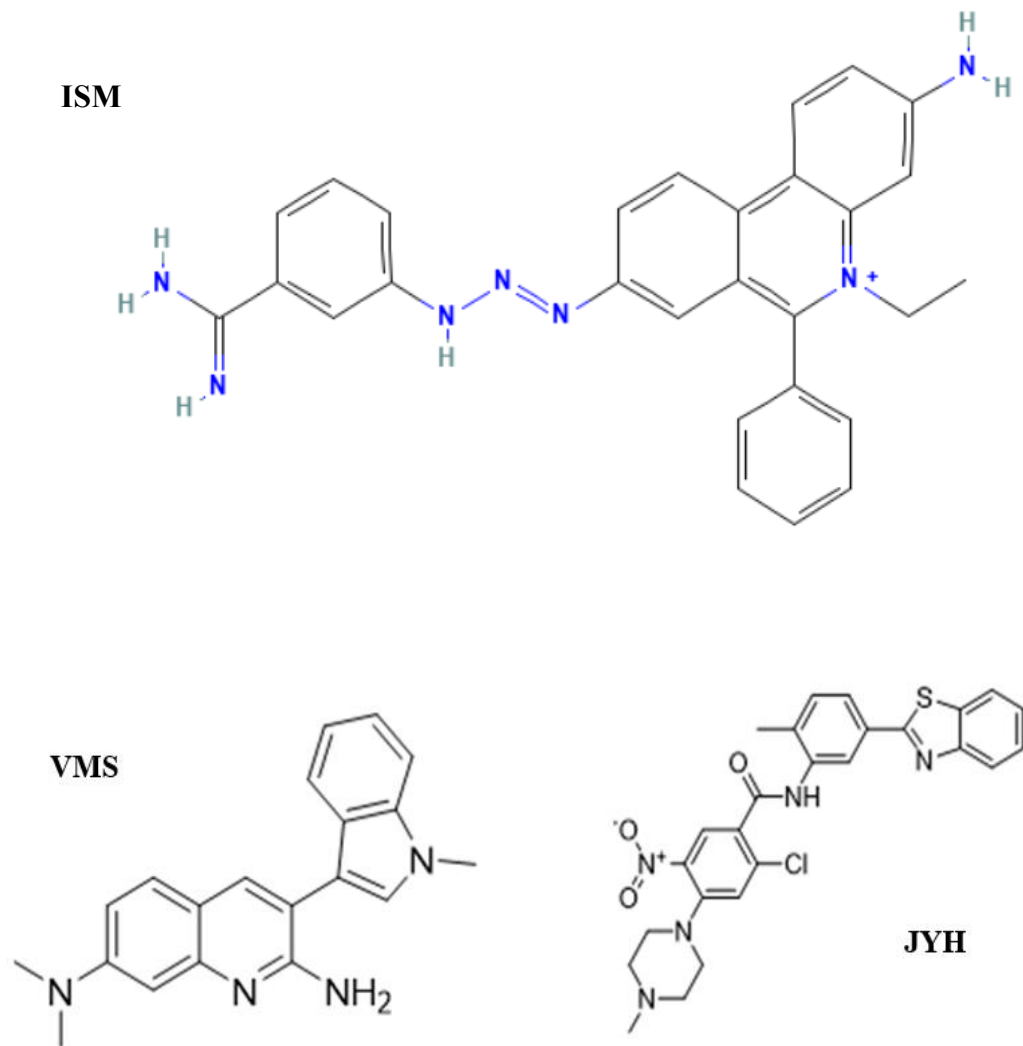
Figure 7: The two-dimensional structure of isometamidium chloride (ISM), VMS and JYH.

3.3 Data preprocessing and quality control

The three datasets were preprocessed before downstream data analysis using IDEOM (Creek et al., 2012). Preprocessing included data filtering to remove baseline noise, as well as centering fluctuations around zero. This step was also performed at Glasgow Polyomics (2020) prior to statistical and metabolomic analysis.

3.4 Statistical and metabolomic analysis

Data was analysed using the MetaboAnalystR package version 3.2.0 in R (Chong & Xia, 2018). Baseline noise was filtered using the interquartile range robust measure of scale because of its efficiency, robustness to estimate both symmetric and asymmetric data as well as its relatively higher breakdown point of 0.25. Data was then normalized through scaling using autoscaling to adjust for fold changes in different metabolites and transformation using log transformation (log_{10}) to correct for heteroscedasticity and multiplicative effects in the data (van den Berg et al., 2006). Dimension reduction was performed using unsupervised PCA and supervised PLS-DA and to investigate discriminate clustering between the samples as well as metabolites that contributed to the discriminate clustering. A one-way ANOVA with Tukey's HSD *post hoc* test was conducted to compare the effects of the different drug concentrations on the metabolites' levels for the ISM treated dataset. For the novel compound (JYH-G-52-1 and VMS-7-25) datasets, fold change analysis and t-test were used to analyse the metabolic effects of the drugs on the parasite. Cluster analysis was then performed based on Euclidean distance using k-means and hierarchical clustering to further identify discriminate clustering of dysregulated metabolites. From the above analyses, dysregulated metabolites were identified and further metabolomic analysis conducted to identify the drug targets.

3.5 Retrieval and analysis of primary protein sequence

The primary sequence of the annotated putative drug targets was retrieved from the TriTryp database (https://tritrypdb.org/tritrypdb/app) in FASTA format. A similarity search using Blastp was performed against the NCBI database (https://blast.ncbi.nlm.nih.gov/) for sequence confirmation and further functional and molecular analyses on the sequences were conducted using Conserved Domain Database (CDD), InterPro and Pfam (https://pfam.xfam.org/).

3.6 Three dimensional structure prediction

The secondary structures of the retrieved sequences were predicted using Psipred (http://bioinf.cs.ucl.ac.uk/psipred/). The retrieved sequences were used to model the 3-Dimensional structure of the predicted drug targets using Alphafold2 (https://alphafold.ebi.ac.uk/) (Jumper et al., 2021). The obtained models were refined using GalaxyRefine server (https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE) (Heo et al., 2013). The validity of the predicted structures was evaluated using Protein Structure Analysis web server (https://prosa.services.came.sbg.ac.at/prosa.php) as well as ERRAT, PROCHECK, and VERIFY 3D from Structural Analysis and verification server (https://saves.mbi.ucla.edu/). Structures with high quality factor in ERRAT (> 90%), minimal stereochemical errors in PROCHECK analysis and 3D model compatibility with the protein primary sequence of > 80% on VERIFY 3D proceeded to molecular docking analysis.

3.7 Ligand Retrieval and Optimization

The 3D structure of isometamidium chloride in structural data format (SDF) was retrieved from the PubChem database (https://pubchem.ncbi.nlm.nih.gov/compound/Isometamidium-chloride). The 3D SDF file was converted to MOL2 format using OPEN BABEL graphical user interface (GUI) software. The structures of JYH-G-52-1 and VMS-7-25 were drawn using MarvinSketch from Marvin suite programs on Chem axon (https://chemaxon.com/). The resulting MDL molfiles were optimized through energy minimization on UCSF chimera. The 1000 steps of Steepest descent minimization were performed first to relieve highly unfavorable clashes, followed by 500 steps of conjugate gradient minimization, to achieve a conformation with the lowest energy minimum after severe clashes had been relieved (Das et al., 2020; Pettersen et al., 2004). The optimization was necessary to achieve conformations with the lowest energy that illustrate stability since 2D structures of ligands are not energetically stable.

3.8 Preparation of the ligand and receptor for molecular docking

The predicted macromolecules were prepared using AutoDock 1.5.7. Polar hydrogens were added and Gasteiger charges were computed for the receptors. AD4 atom types were assigned to the receptor defining aromatic and aliphatic amino acids as well as hydrogen bond donors and acceptors. The resulting structures were saved as PDBQT files. The ligands were also inputted into AutoDock, torsion tree detected and the resultant structures saved as PDBQT files.

3.9 Grid parameters and prediction of binding pockets

Three-dimensional grid boxes for the receptors were created by the AutoGrid algorithm to evaluate the binding energies on the macromolecule coordinates. The grid maps representing the intact ligand in the actual docking target sites were calculated with AutoGrid (part of the AutoDock package). The DoG (difference of Gaussian) site scorer from ProteinPlus (https://proteins.plus/) was used to predict the receptors' binding pockets (Volkamer et al., 2010). The DoG site scorer utilises 3D difference of Gaussian to filter grid representations of the protein to recognise patterns which resemble spherical structures (Volkamer et al., 2010). This filter approximates the second derivative of the Gaussian function (Laplacian of gaussian) to identify invaginations of the protein surface which are suitable to accommodate ligands (Volkamer et al., 2010). The druggability score was calculated based on the linear contribution of the 3D descriptors describing volume, hydrophobicity and enclosure. The higher the drug score the more druggable the pocket is. The hydrophobicity ratio is dependent on the most occurring residues in the pocket while enclosure is dependent on if the pocket is closed or open and near the protein's surface. The top 5 ranked binding pockets were investigated.

3.10 Molecular docking assay

Rigid receptor-flexible ligand docking was performed using AutoDock Vina version 1.2.0 using the computed grid parameters from AutoDock Tools and pocket centers computed from Protein plus. The configuration file was created as follows:

receptor = name_of_receptor.pdbqt

ligand = name_of_ligand.pdbqt

out = name_of_output_file.pdbqt

log = name_of_log_file.txt

center_x = x coordinate computed from ProteinPlus

center_y = y coordinate computed from ProteinPlus

center_z = z coordinate computed from ProteinPlus

size_x = number of points in x dimension computed from AutoGrid

size_y = number of points in y dimension computed from AutoGrid

size_z = number of points in z dimension computed from AutoGrid

exhaustiveness = 16

num_modes = 20

The configuration file was then used to perform docking by running the following command on the command line:

C:\Program Files (x86)\The Scripps Research Institute\Vina\vina.exe - - config config.txt

The output of the command was ligand binding poses compiled in the name_of_output_file.pdbqt file.

3.11 Post-docking interaction analysis and visualisation

Obtained ligand poses within the receptor with the lowest binding energy were analysed and visualized using the Discovery Studio client v21.1.0.20298, Dassault Systèmes Biovia Corp, to illustrate receptor-ligand 3D interaction profiles with their corresponding 2D diagrams. Further analysis and visualisation was performed using UCSF Chimera version 1.15 and

PyMOL software (The PyMOL Molecular Graphics System, Version 2.5.2 Schrödinger, LLC, San Diego, CA, USA).

3.12 Screening of pharmacodynamics and pharmacokinetic parameters for JYH and VMS

Candidate drug compounds fail to proceed in drug development pipelines due to unsuitable pharmacokinetics and toxicity (Xiong et al., 2021). Absorption, distribution, metabolism, excretion (ADME) and toxicity have been longstanding parameters that should be evaluated early drug design. Molecular physicochemical descriptors, drug-likeness and ADME properties of the novel compounds JYH-G-52-1 and VMS-7-25 were analysed using the freely accessible web tools SwissADME (Daina et al., 2017) and ADMETlab2.0 (Xiong et al., 2021). JYH and VMS were loaded onto SWISS ADME (http://www.swissadme.ch/) in their SMILES (Simplified Molecular Input Line Entry System) format and their descriptors computed automatically using the tool.

CHAPTER FOUR

4.0 RESULTS

4.1 *Trypanosoma congolense* treated with isometamidium chloride

4.1.1 Data preprocessing and transformation

Data provided as an IDEOM file from Glasgow Polyomics was exported to R for further analysis using the MetaboAnalystR package version 3.2.0. The data contained 12 samples and 775 processed peaks for each sample as tabulated in **Table 2** below. The samples were grouped into three sample groups; control (TcISM_0001, TcISM_0002, TcISM_0003, TcISM_0004), low dose (TcISM_0005, TcISM_0006, TcISM_0007, TcISM_0008) and high dose (TcISM_0009, TcISM_0010, TcISM_0011, TcISM_0012). Positive features indicated peaks that had recorded peak intensity values while processed features indicated all the features that were used in subsequent analysis after zero and missing values were corrected with 1/5th of the smallest value in the sample group.

Table 2: Summary of data processing results for *T. congolense* treated with isometamidium chloride

	Sample	*Features (positive)*	*Missing/Zero*	*Features (processed)*
Control	*TcISM_0001*	765	10	775
	TcISM_0002	765	10	775
	TcISM_0003	763	12	775
	TcISM_0004	760	15	775
Low dose ISM	*TcISM_0005*	773	2	775
	TcISM_0006	768	7	775
	TcISM_0007	772	3	775
	TcISM_0008	772	3	775
High dose ISM	*TcISM_0009*	768	7	775
	TcISM_0010	766	9	775
	TcISM_0011	752	23	775
	TcISM_0012	766	9	775

4.1.2 Global metabolomics profiles of the control and isometamidium chloride treated
Trypanosoma congolense **samples**

Multivariate analyses PCA and PLS-DA were used in this instance. From PCA analysis, the

first two PCs (PC 1 and PC 2) were chosen explaining 49.8% and 13% of the variance in the

data respectively, both totaling to 62.8% (**Figure 8**). Preliminary PCA clustering showed that

the drug caused significant perturbation to the *T. congolense* metabolome with the control

samples clustering discriminately from the low and high dose indicating that the groups are

distinct.
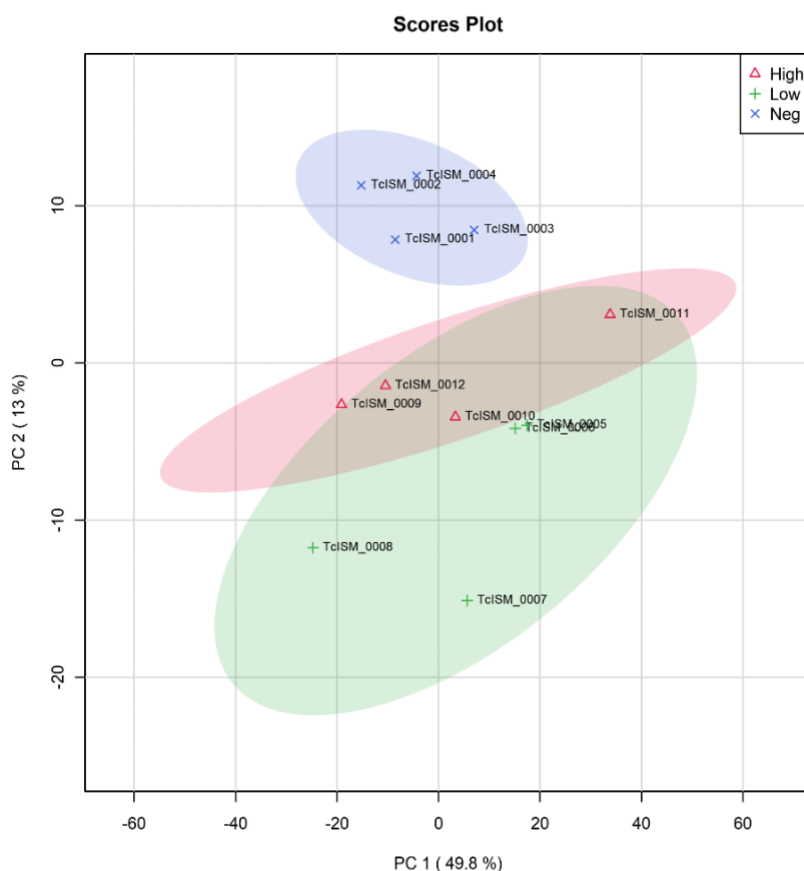
Figure 8: Principal component analysis (PCA) score plots of the isometamidium chloride *T. congolense* treated dataset.

Further feature selection was done using PLS-DA, a supervised multivariate statistical analysis technique. The PLS-DA scores graph in **Figure 9**, clearly distinguishes between the control and treatment groups. Through internal cross validation, which determines the accuracy and the model's predictive power and goodness of fit parameter (R^2), the effectiveness of these PLS-DA models was assessed (Q^2). High R^2 and Q^2 values were observed, and the difference between R^2 and Q^2 was 0.20969. Together, these findings demonstrated significant metabolic perturbations caused by isometamidium chloride on the metabolome of the *T. congolense* parasite.
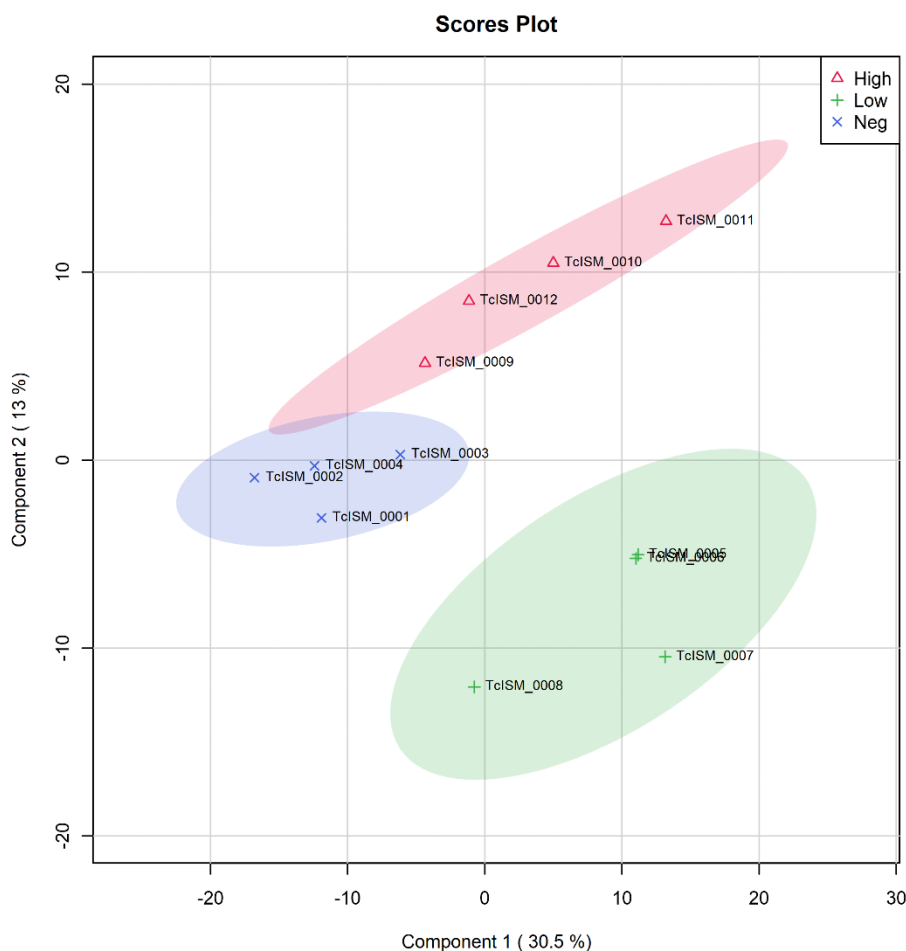
Figure 9: Partial least squares discriminant analysis (PLS-DA) scores plot between component 1 (30.5%) and component 2 (13%) (accuracy = 0.91667, R2 = 0.9995, Q2 = 0.78981) scores plot with their explained variance.

4.1.3 Identification of statistically significant metabolites

Dysregulated metabolites in the low (0.6 nM) and high (6.0 nM) isometamidium chloride

treatment groups were identified by comparing the metabolite relative peak intensities in the

two groups upon exposure to the drug. One-way ANOVA with Tukey's *post hoc* analysis

revealed 23 metabolites that showed significant differences (adjusted FDR < 0.05, P ≤ 0.05)

when the treatment groups were compared to the control group (**Figure 10, Table 3**).
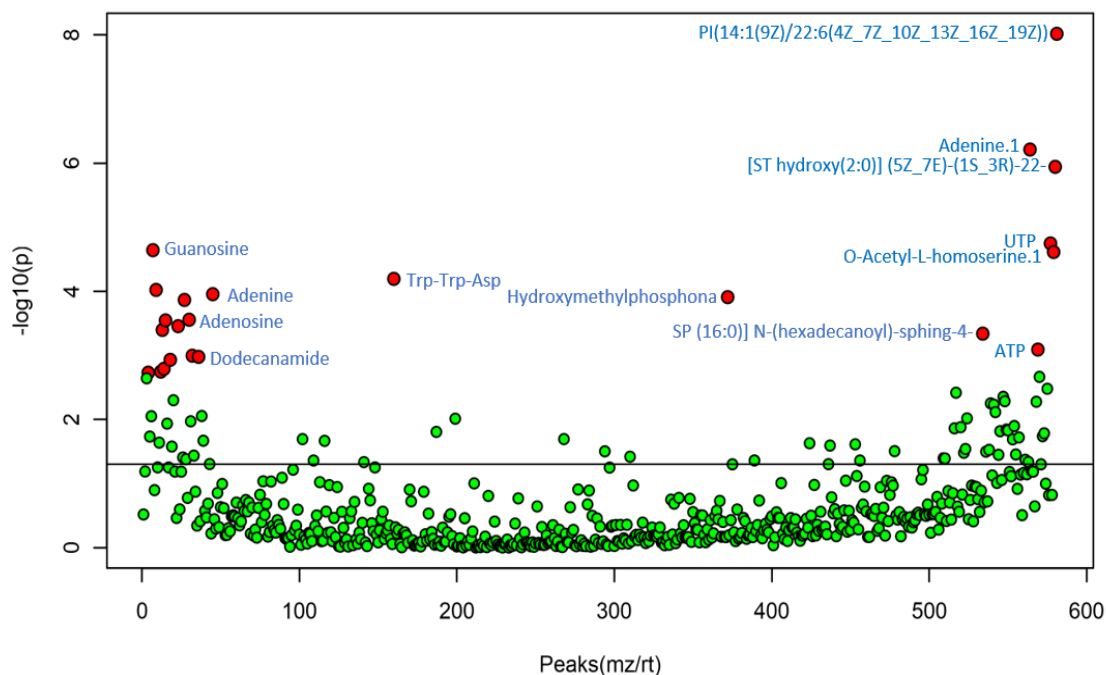
Figure 10: Significant features selected by one-way ANOVA analysis (adjusted FDR < 0.05, P ≤ 0.05). Statistically significant features are shown as red dots. The details of the significant features are shown in Table 3.

Table 3: Important features identified by one-way ANOVA and Tukey's HSD *post-hoc* analysis. The metabolites are color coded depending on whether they increased or decreased following exposure to ISM

Metabolite	f.value	p.value	-log10(p)	FDR
PI(14:1(9Z)/22:6(4Z_7Z_10Z_13Z_16Z_19Z))	267.82	9.5857E-9	8.0184	7.4289E-6
Adenine.1	103.58	6.1319E-7	6.2124	2.3761E-4
[ST hydroxy(2:0)] (5Z_7E)-(1S_3R)-22-(3-hydroxyphenyl)-23_24-dinor-9_10-seco-5_7_10(19)-cholatriene-1_3-diol	89.653	1.1408E-6	5.9428	2.947E-4
UTP	46.577	1.7882E-5	4.7476	0.0031431
ATP	17.39	8.0978E-4	3.0916	0.034866
[SP (16:0)] N-(hexadecanoyl)-sphing-4-enine-1-phosphate	20.353	4.5733E-4	3.3398	0.020849
CDP-choline	14.27	0.001618	2.7911	0.034004
Palmiticamide	13.859	0.001787	2.7479	0.046926
Uracil	13.702	0.001858	2.731	0.046926
Guanosine	43.909	2.2767E-5	4.6427	0.0031431
O-Acetyl-L-homoserine.1	43.198	2.4334E-5	4.6138	0.0031431
Trp-Trp-Asp	34.003	6.379E-5	4.1952	0.0070624
AMP	30.743	9.4982E-5	4.0224	0.0092014
Adenine	29.56	1.1075E-4	3.9557	0.0095368
Hydroxymethylphosphonate	28.767	1.2314E-4	3.9096	0.0095434
Asn-Met-Asn-Gln	28.034	1.3613E-4	3.866	0.0095912
Adenosine	23.273	2.7745E-4	3.5568	0.015716
Deoxyadenosine	23.131	2.8389E-4	3.5468	0.015716
3-Ethylmalate	21.89	3.4911E-4	3.457	0.018037
dAMP	21.108	3.9975E-4	3.3982	0.019363
Cytidine	16.344	0.0010093	2.996	0.040784
Dodecanamide	16.151	0.0010525	2.9778	0.040784
4-deoxypyridoxine	15.669	0.0011705	2.9316	0.043199

KEY

	Metabolites with decreasing peak intensity
	Metabolites with increasing peak intensity

A volcano plot analysis (fold change $0.5 \geq (x) \geq 2$, t-test threshold $(y) > 0.1$) between the control samples and the high dose samples revealed similar changes as evidenced in **Figure**

11. Unphosphorylated nucleobases exemplified by Uracil and Guanosine had x ≥ 2 fold changes while nucleotide triphosphates exemplified by UTP had x ≤ 0.5 fold changes.
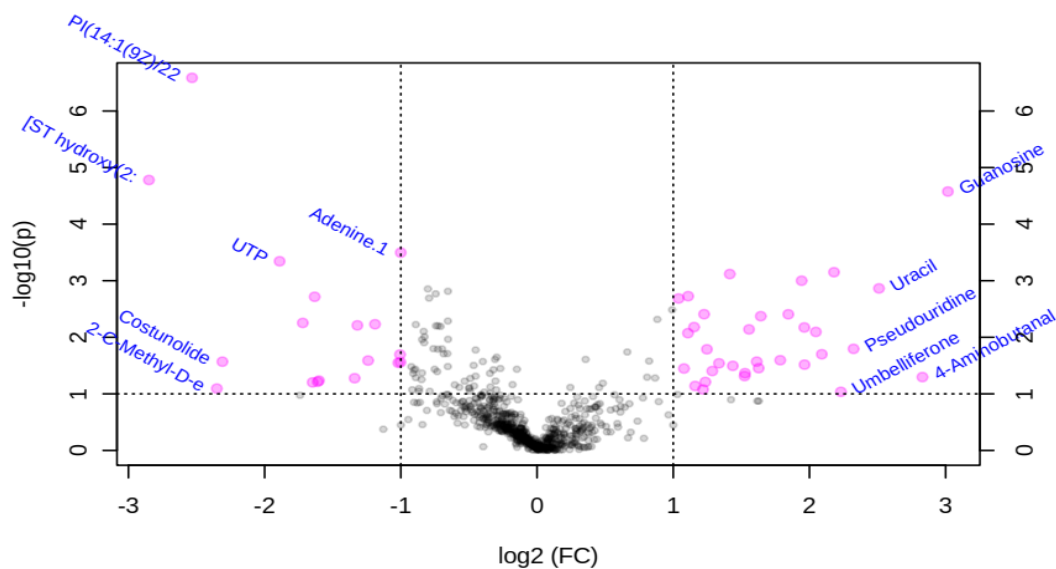
Figure 11: Features selected by volcano plot with fold change threshold (x) 2 and t-tests threshold (y) 0.1. The pink circles represent features above the threshold.

The clustering heat map of the dysregulated metabolites illustrated in **Figure 12** showed distinct clustering between the control and treatment groups wherein UTP, ATP, 2-methylcholine and PI (14:1(9Z)/22:6(4Z_7Z_10Z_13Z_16Z_19Z)) underwent consistent decrease in relative peak intensity as the treatment dosage increased. In contrast, cytidine, AMP, and guanosine underwent increase in relative peak intensity.
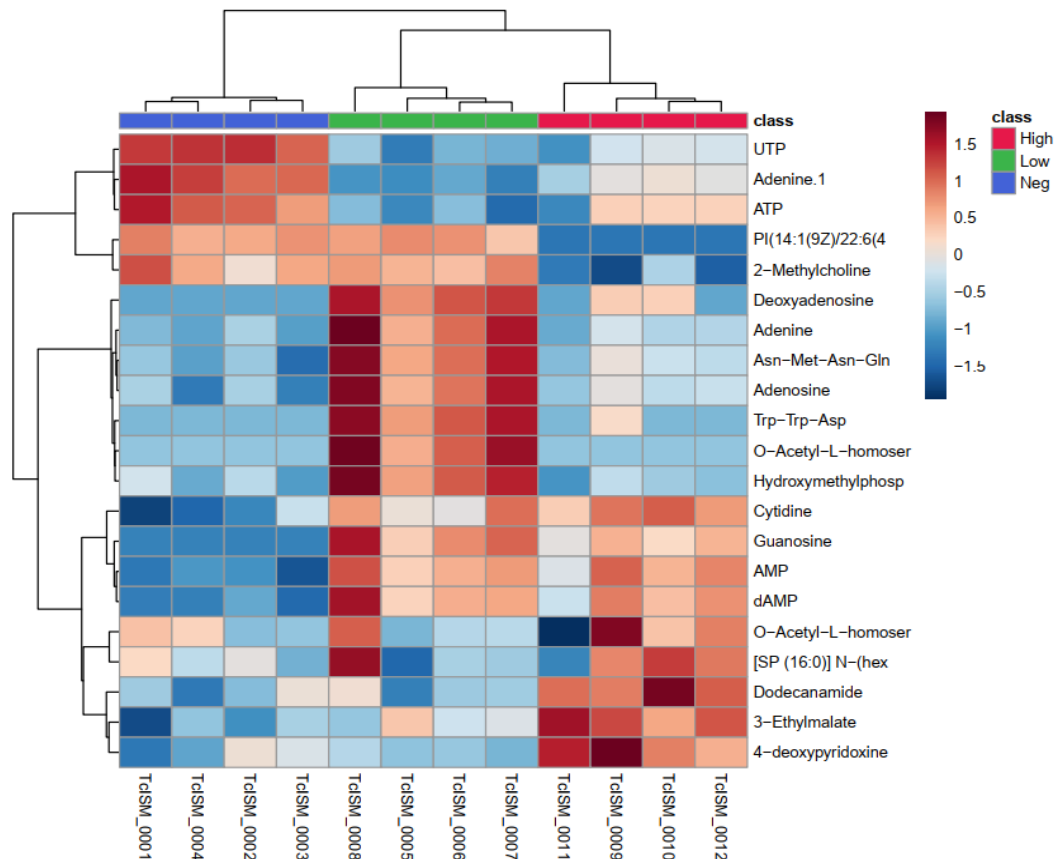
Figure 12: Clustering heatmap of the dysregulated metabolites showing distinct clustering between the control and treatment groups. Distinct clusters indicate 1) metabolites that underwent consistent decrease upon exposure to isometamidium chloride at both low and high dose, 2) metabolites that spiked on exposure to low dose and then decreased at high dose and 3) metabolites that steadily increased over the treatment dosages.

The one-way ANOVA with Tukey's HSD *post hoc* and clustering statistical analyses identified significant alterations in metabolites of the nucleotide metabolism pathway. Unphosphorylated nucleobases and nucleoside monophosphates underwent significant increase in relative peak intensity while a concomitant decrease in nucleoside diphosphates and nucleoside triphosphates was observed. In addition to ATP and ADP, UTP, UDP, GTP and GDP underwent the following sample relative intensity changes 1.00/0.18/0.27, 1.00/0.82/0.92, 1.00/0.27/0.61, 1.00/0.77/0.94 in 0.0nM, 0.6nM and 6.0nM ISM treatment groups respectively. In stark contrast, AMP, dAMP, CMP, GMP, adenosine, guanine, cytidine, cytosine, uracil, and guanosine showed significant increase with the following

sample relative intensity changes recorded 1.00/4.80/4.53, 1.00/4.97/3.84, 1.00/3.40/3.06, 1.00/3.64/2.88, 1.00/36.69/2.24, 1.00/1.99/2.22, 1.00/1.89/2.16 and 1.00/1.93/1.84 in 0.0nM, 0.6nM and 6.0nM ISM treatment groups respectively. This observation points to a hypothetical nucleoside kinase system (**Figure 13**) (Miranda et al., 2022) similar to that proposed in other organisms (Appanna et al., 2016) that could be observed in these parasites. In the study by Appanna et al., (2016), *Pseudomonas fluorescens* appeared to invoke the participation of pyruvate orthophosphate dikinase (PPDK) and phosphoenol pyruvate synthase (PEPS) to fuel production of adenosine triphosphate (ATP) in stressed cells in an oxygen independent manner. Since PPDK and PEPS necessitate utilization of adenosine monophosphate (AMP) as a substrate, this substrate was readily obtained from adenylate kinase (AK) which is in high abundance in oxidatively stressed cells (Appanna et al., 2016). Adenylate kinase converts adenosine diphosphate (ADP) to ATP and AMP. ADP is in turn in constant supply due to increased activity of nucleotide diphosphate kinase (NDPK) – which mediate the conversion of ATP and NDP to ADP and NTP – and acetate kinase (AK) – which transforms acetate to acetate phosphate and ADP (Appanna et al., 2016). This process is illustrated in **Figure 13** below. It is then possible that when the parasites use up ATP, they resort to other NTPs, transferring their high energy phosphates to ADP to sustain ATP for energy. This process is mediated by NDPK (Miranda et al., 2022). In that case when the NTPs are used up, the parasites resort to NDPs. This leads to an accumulation of NMPs (with a lower energy phosphate) with further knock-on impacts on unphosphorylated nucleosides and nucleobases which rise in abundance.
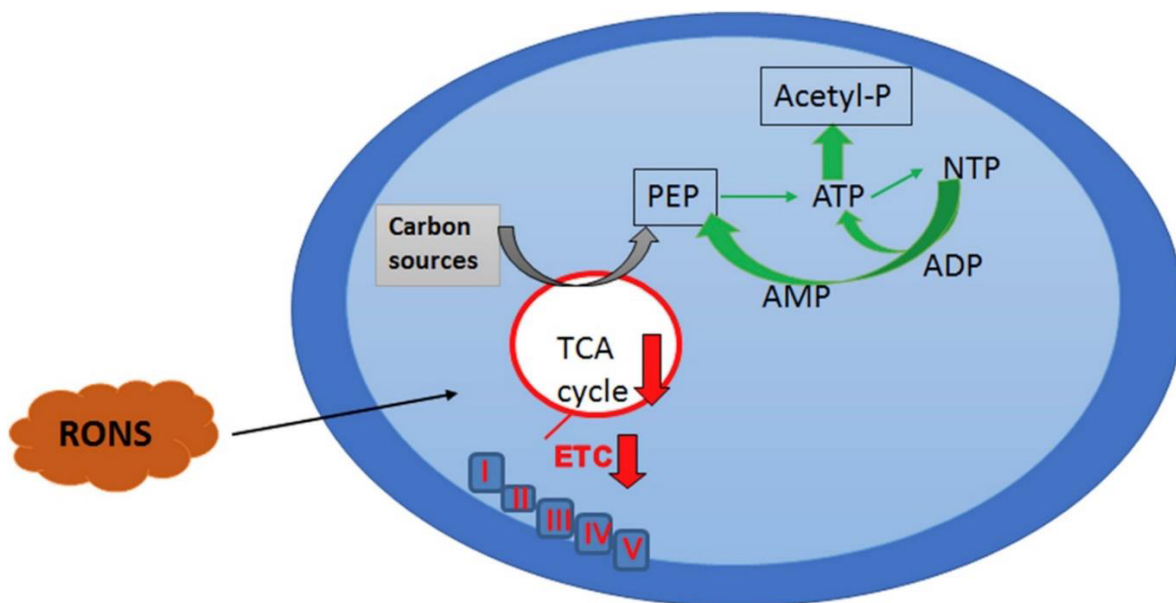
Figure 13: The nucleotide phosphorelay system in *Pseudomonas fluorescens*. Adapted from Appanna et al., (2016).

Additionally, metabolomic analysis revealed seven metabolites within carbohydrate metabolism showing significant alterations upon exposure to isometamidium chloride wherein six of them showed significant decrease while one showed an increase. Specifically, metabolites of glycolysis e.g. D-glucose, D-glucose-6-phosphate, glycerol, D-glyceraldehyde-3-phosphate and D-fructose-1,6-bisphosphate all decreased. There was also a concomitant decrease in metabolites related to the TCA and ASCT/SCS cycles such as malate, fumarate, cis aconitate and citrate which showed the following relative peak intensity observations 1.00/0.67/0.78, 1.00/0.69/0.77, 1.00/0.77/0.71, 1.00/0.65/0.53 in the control, low and high groups respectively. No significant changes were observed in succinate. Adenosine diphosphate and ATP of the energy metabolism pathway exhibited consistent decrease in the low and high doses. The decrease in glycolysis, including a decrease in intracellular glucose, TCA and ASCT/SCS cycles and energy metabolites indicates that isometamidium chloride may have an inhibitory effect on the glucose transporter as shown

in **Figure 14**. Glucose enters the parasite through the *T. congolense* hexose transporter 1 (TcoHT1) and is metabolized further in the glycosome to its end products. If inhibition of glycolysis came after the transporter then glucose levels would increase inside the cell as has been seen with some inhibitors of hexokinase (Barrett, personal communication).
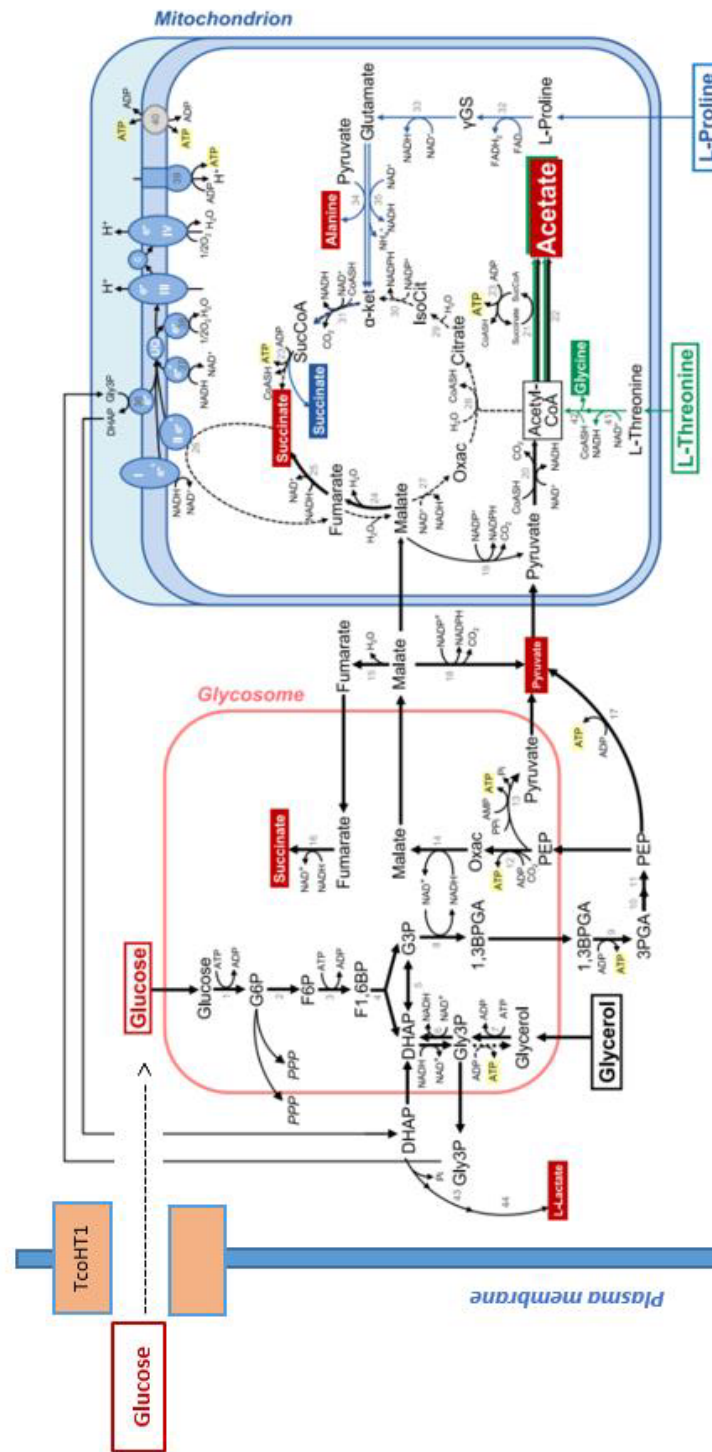
Figure 14: Classic glycolysis in PCF *T. brucei* that is hypothesized to resemble BSF *T. congolense* (adapted from Michels et al., (2021)).

4.1.4 Glucose transporter sequence identification and structure prediction

The protein sequence of the annotated putative glucose/hexose transporter identified by gene ID TcIL3000_10_7310 was obtained from TriTrypDB database (https://tritrypdb.org) in FASTA format. The putative sequences were validated using CDD (conserved domain database) ((M. Yang et al., 2020), https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) and Integrative protein signature database InterPro (Hunter et al., 2009). The CDD results predicted that the sequence belonged to the major facilitator superfamily (MFS), cd06174 while InterPro predicted GO terms GO:005585 (transmembrane transport) and GO:0022857 (transmembrane transporter activity) for the sequence's biological and molecular functions respectively. The retrieved sequence was used to model the structure of the *T. congolense* glucose transporter using Alphafold2 (https://alphafold.ebi.ac.uk/). The resultant structure is as shown in **Figure 15**. The structure passed the VERIFY 3D check with 85.52% of the residues averaging a 3D-1D score of ≥ 0.2. No errors were detected in the residue-by-residue geometry and overall structure geometry. Further, 93.4% of the residues were within the allowed regions and 0.0% of the residues were in the disallowed regions. According to ProSA, the structure had a Z-score of -7.38 which is within the range of energy distribution for proteins within the same size. The PDB format of the predicted structure was downloaded to be used in subsequent docking steps.
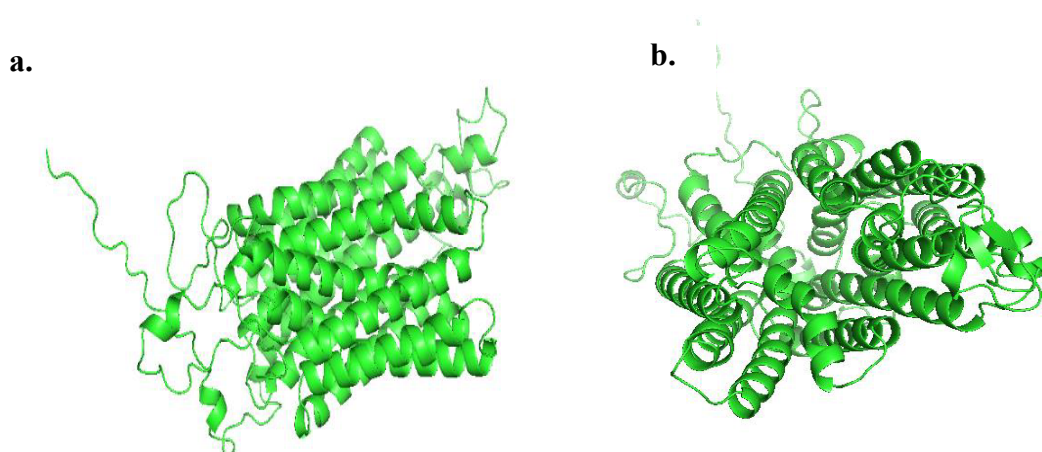
Figure 15: The predicted structure of the *T. congolense* glucose transporter (TcoHT1) visualized on PyMol. (**a**) side view of the transporter and (**b**) axial view.

4.1.5 Molecular docking analysis

Potential binding pockets of the predicted *T. congolense* glucose transporter were predicted by the DoGSite scorer function on ProteinPlus. The predicted binding pockets are as shown in **Table 4** indicating their volume, surface area, drug score and simple score.

Table 4: Details of the top five predicted binding pockets as identified by ProteinPlus based on drug score.

Name	Volume	Surface	Drug score	Simple score
P_0	1076.11	861.28	0.82	0.65
P_1	567.03	949.12	0.8	0.43
P_2	445.39	719.43	0.71	0.24
P_3	181.09	218.7	0.64	0.0
P_4	262.19	382.66	0.63	0.03

The largest pocket was selected with a hydrophobicity ratio of 0.40 and xyz coordinates (4.32, -0.10, -1.66). The binding pocket contained residues predicted to be on the active site of the protein according to CDD. These values were used to create the Vina configuration file describing the xyz centres, a grid size of 60 x 60 x 60 A^0, exhaustiveness = 8 and num_modes = 20.

The output files from AutoDock Vina were analysed based on the absolute docking energies of isometamidium chloride and glucose on the predicted glucose transporter when docked. The docking energies were compared between isometamidium chloride and glucose. The binding affinities are as shown in **Table 5** below.

Table 5: Binding energies between isometamidium chloride and glucose and the predicted *T. congolense* hexose transporter 1

Mode	ISM affinity (kcal/mol)	Glucose affinity (kcal/mol)
1	-10.1	-5.7
2	-10.0	-5.6
3	-10.0	-5.5
4	-9.7	-5.4
5	-9.7	-5.4
6	-9.7	-5.3
7	-9.5	-5.2
8	-9.5	-5.1
9	-9.2	-5.1
10	-9.1	-4.9

4.1.6 Docking visualisation

Figure 16 and **Figure 17** illustrates the binding results of isometamidium chloride and glucose to the predicted glucose transporter at the predicted binding pockets. The protein is represented as a cartoon and colored in rainbow from N-terminal to C-terminal. The ligands are represented in stick format while the bonds between the ligand and protein are depicted in dashed lines that are color-coded depending on the type of bond formed. Glucose interacted with Ile 53, Gln 216, Ile 219, Leu 313, Gln 314, Ile 318, Asn 319, Phe 411, Glu 412, Phe 420, Gln 443, Phe 444 and Asn 447 while isometamidium chloride interacted with Ile 53, Gly 54, Val 56, Ala 57, Ala 124, Met 127, Ile 128, Gln 216, Ile 219, Thr 220, Ile 223, Leu 313, Gln 314, Ile 318, Asn 319, Met 322, Asn 347, Phe 411, Glu 412, Gly 416, Phe 419, Phe 420, Gln 443 and Asn 447. Further, **Figure 18** shows the residues interacting with

glucose and ISM respectively while **Figure 19** shows the positioning of the docked glucose and ISM on the predicted transporter.

a.

b.
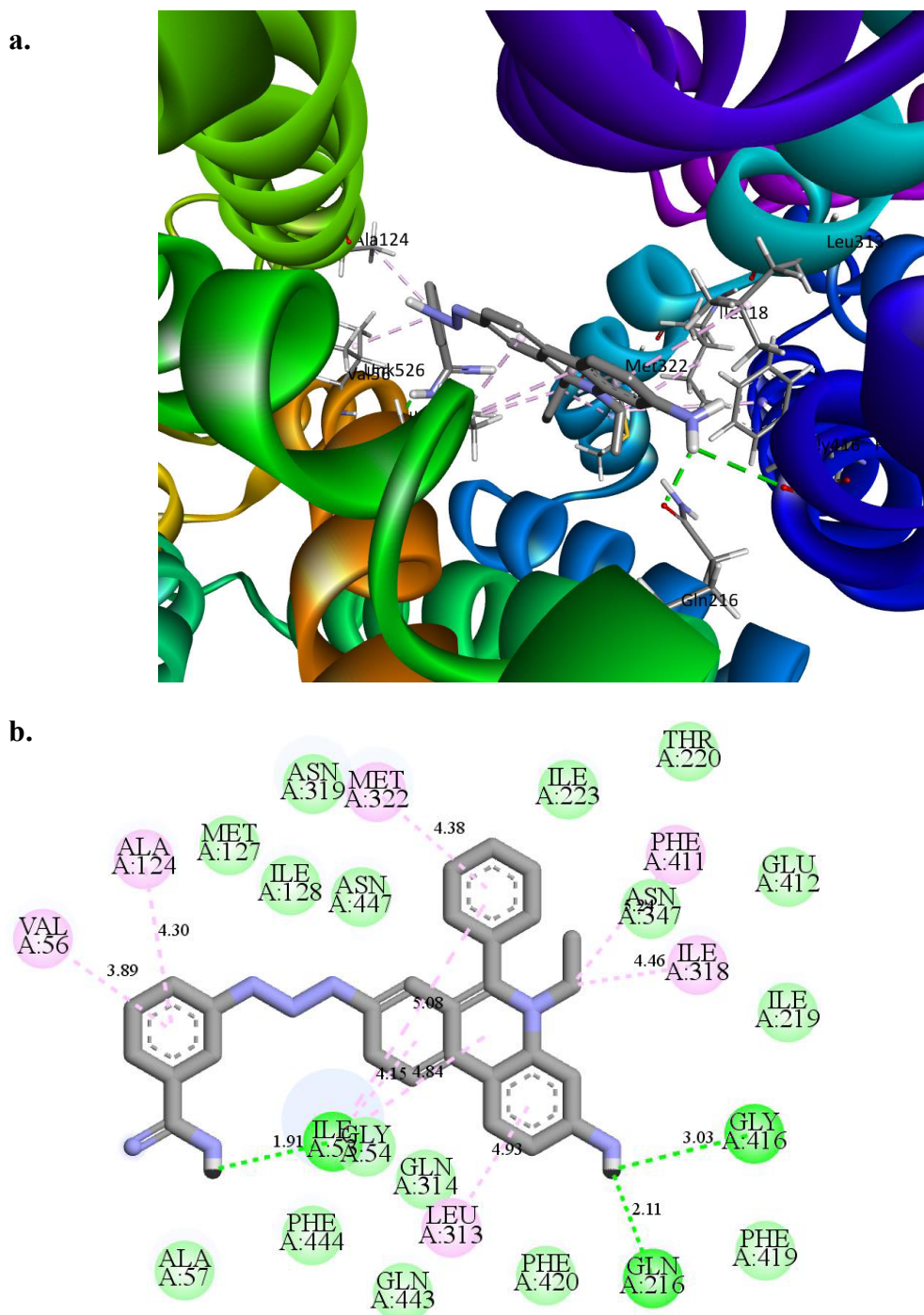
Figure 16: Interaction of isometamidium with the predicted *T. congolense* glucose transporter. Interactions within the 3D molecule are shown in (**a**) while 2D interactions are shown in (**b**).
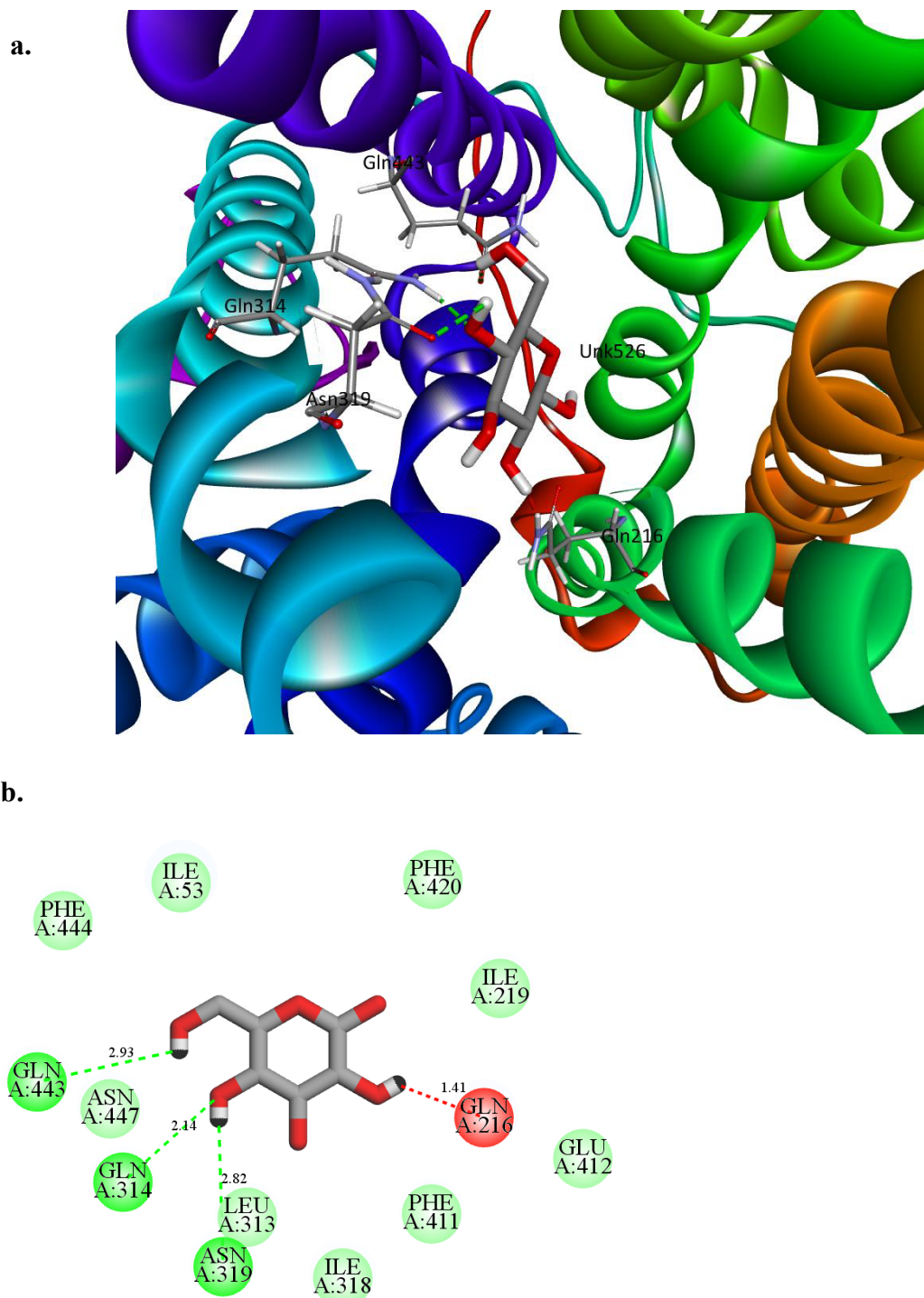
Figure 17:Interaction of glucose with the predicted *T. congolense* glucose transporter. Interactions within the 3D molecule are shown in (**a**) while 2D interactions are shown in (**b**).
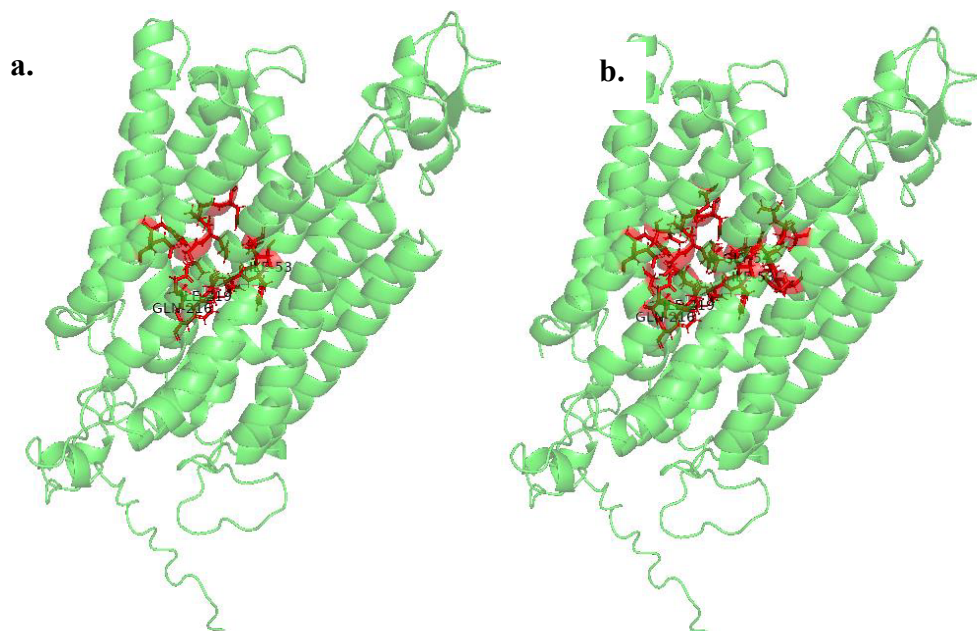
Figure 18: Structure of the predicted *T. congolense* glucose transporter 1 (TcoHT1) showing residues that interact with (**a**) glucose and (**b**) isometamidium chloride. The interacting residues are shown as red sticks and are labelled while the rest of the protein is shown as a green cartoon
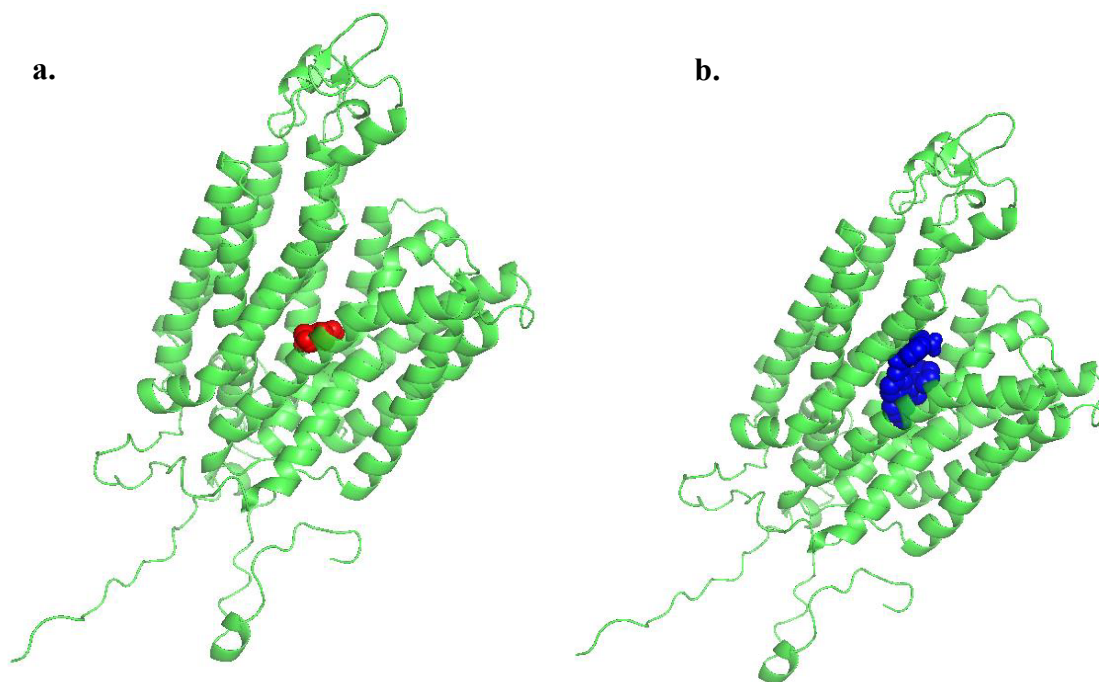
Figure 19: Structure of the predicted *T. congolense* glucose transporter 1 (TcoHT1) showing (**a**) glucose and (**b**) ISM docked. Glucose is shown in red spheres and ISM is shown in blue spheres while the rest of the protein is shown as a green cartoon.

This modelling data indicates the possibility that ISM can bind the TcoHT1 transporter at a site overlapping with the exofacial glucose binding site (**Figure 18, Figure 19**). The binding energy (-10.1 kcal/mol) is compatible with relatively avid binding that could lead to inhibition of glucose uptake by the transporter. This in turn could explain the metabolomics data which indicated inhibition of glucose uptake into *T. congolense* with knock-on effects on the production of ATP, triggering a phosphorelay system that then consumed phosphate from other nucleotides in attempting to restore ATP homeostasis.

4.2 *Trypanosoma congolense* treated with VMS

4.2.1 Data preprocessing and transformation

The untargeted metabolomics data was analysed using MetaboanalystR package version 3.2.0 in R. Data details from the *T. congolense* dataset is presented in **Table 6** below. The samples were in two sample groups; control (FG0663, FG0667, FG0671, FG0675) and VMS (FG0662, FG0666, FG0670, FG0674). Positive features indicated peaks that had recorded peak intensity values while processed features indicated all the features that were used in subsequent analysis after zero and missing values were corrected with $1/5^{th}$ of the smallest value in the sample group. Features were preprocessed and normalized as explained in Section **3.3**.

Table 6: Summary of data processing results of the *T. congolense* treated with VMS

	Sample	Features (positive)	Missing/zero	Features (processed)
Control	FG0663	563	2	565
	FG0667	561	4	565
	FG0671	560	5	565
	FG0675	563	2	565
VMS treated	FG0662	561	4	565
	FG0666	558	7	565
	FG0670	560	5	565
	FG0674	561	4	565

4.2.2 Global metabolomics profiles of the control and VMS treatment *Trypanosoma congolense* samples

Multivariate analysis using unsupervised PCA was performed on eight samples to get an overview of the metabolic variations between the control and treatment groups. The first two principal components (PC1 and PC2) were used, explaining 43% and 19.3% variance respectively. The PCA scores plot effectively distinguished between the control and treatment groups, as illustrated in **Figure 20**. The metabolic characteristics contributing to the metabolic differences between the control and treatment samples were then determined using the PCA loadings plot. The top 20 loadings having positive influence on PC1 and the bottom 20 loading having negative influence on PC1 are as shown in **Table 7** below. Features that had positive influence on PC 1 drove the clustering of the treated groups while those that had negative influence drove clustering of the control groups.
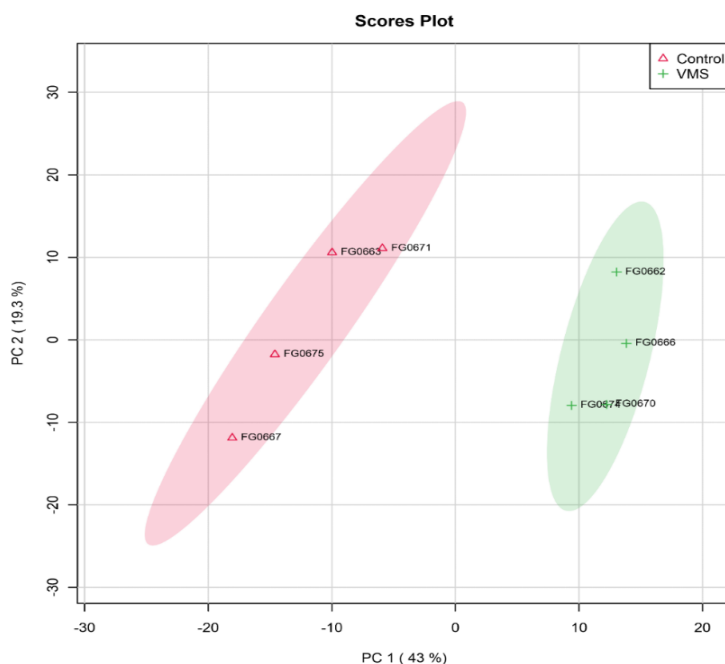
Figure 20: PCA Scores plot between PC 1 covering 43% variance and PC 2 covering 19.3% variance. The control groups are marked by red triangles while the treatment groups are marked by green crosses.

Further feature selection was done using supervised PLS-DA. The PLS-DA scores graph in **Figure 21**, clearly distinguishes between the control and treatment groups. In PLS-DA analyses, metabolic features with VIP scores ≥ 1.5 were extracted as significant metabolic characteristics for further study. Through internal cross validation, which determines the model's predictive power and goodness of fit parameter (R^2), the effectiveness of these PLS-DA models was assessed (Q^2). High R^2 and Q^2 values were observed, and the difference between R^2 and Q^2 was less than 0.2, which excludes model overfitting. Together, these findings demonstrated significant metabolic perturbations caused by VMS on the metabolome of the *T. congolense* parasite.
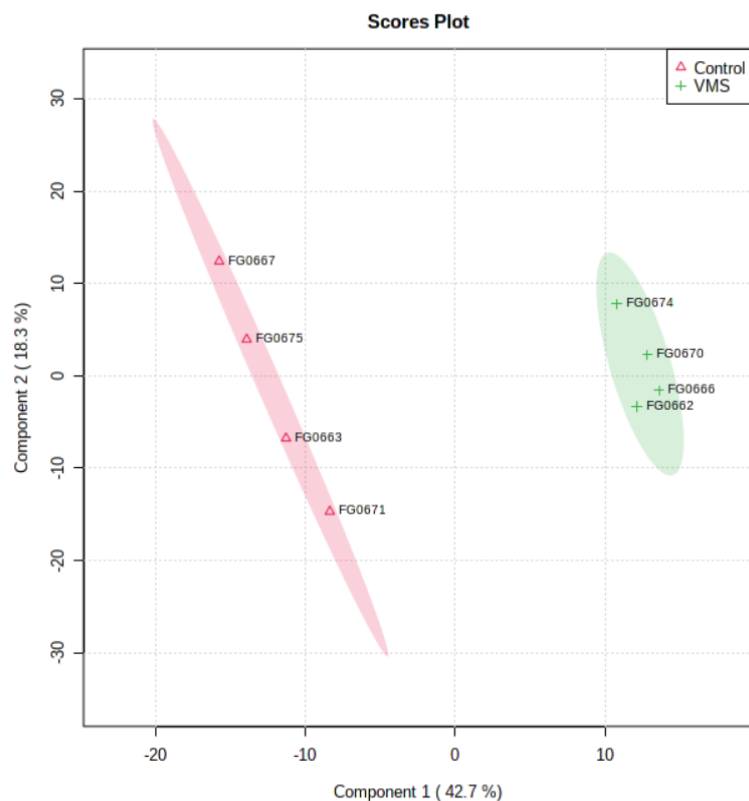
Figure 21: PLS-DA scores plot between component 1 (42.7%) and component 2 (18.3%) , R^2 = 1.0, Q^2 = 0.89258. The control groups are marked by red triangles while the treated groups are marked by green crosses.

Table 7: Top 20 and bottom 20 loadings on PC1 identified by PCA loadings plot.

	Name	Loadings 1	Loadings 2
1	Eseramine	0.072483	-0.010578
2	Leu-Pro-Tyr	0.072317	-0.011547
3	NADP+	0.072151	-9.17E-05
4	Sodiumlaurylsulfate	0.072046	0.002054
5	Acetoacetate	0.071958	-0.016713
6	(R)-2-Hydroxyglutarate	0.071734	1.10E-04
7	2-Oxoglutarate	0.071478	-0.022448
8	3-(4-Hydroxyphenyl)lactate	0.071404	-0.022577
9	(S)-2-Aceto-2-hydroxybutanoate	0.070923	-0.022751
10	O-Butanoylcarnitine	0.070627	0.010882
11	NG_NG-Dimethyl-L-arginine	0.070618	0.0254
12	3-Methyl-2-oxobutanoic acid	0.070408	-0.029751
13	[PC (14:0)] 1-tetradecanoyl-sn-glycero-3-phosphocholine	0.070217	0.018909
14	N-Butyryl-L-homoserine lactone	0.069977	-0.013797
15	(S)-3-Methyl-2-oxopentanoic acid	0.069882	-0.025071
16	Hypoxanthine	0.06942	0.0051487
17	Ascorbate	0.069305	-0.024627
18	Diacetyl	0.069072	-0.031215
19	3-Methylguanine	0.068811	-0.020526
20	Uracil	0.068329	-0.035128
21	L-Tryptophan	-0.072939	0.0011074
22	Ala-Gly-Ser	-0.072784	0.0014926
23	Thr-Ser	-0.072634	0.0048685
24	beta-Alanyl-L-arginine	-0.072583	0.0038874
25	Lys-Pro-Pro	-0.072522	0.0027299
26	Gly-Ser	-0.072438	0.014259
27	Phenylacetonitrile	-0.072398	0.0020062
28	Indole	-0.072384	-0.0040705
29	L-Ala-L-Glu	-0.072376	0.0065254
30	L-Leucine	-0.072304	-0.0071934
31	Ala-Ser	-0.072253	0.014253
32	Piperidine	-0.072206	-0.0081839
33	N2-Succinyl-L-arginine	-0.072204	0.0055231
34	Thr-His	-0.072128	0.01643
35	Quinoline	-0.072125	0.015942
36	Indoleacrylicacid	-0.072124	0.0020418
37	2-Naphthylamine	-0.072066	0.0061523
38	L-Phenylalanine	-0.071999	0.0085359
39	Ala-Gly-His	-0.07199	0.013891
40	Gly-Pro-Arg	-0.071961	0.015943

Key

| | Top 20 loadings on PC 1 |
| Bottom 20 loadings on PC 1 |
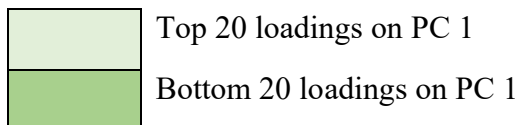
4.2.3 Identification of statistically significant metabolites

Further identification of statistically significant metabolite features between the two sample groups was performed using fold change analysis (FC $\geq$ 2 or $\leq$ 0.5) and two sample t-tests (P $\leq$ 0.05). Fold change analysis identified 94 significant features, 73 undergoing significant depletion and 21 undergoing significant increases (**Table 8**) while *t*-tests identified 138 statistically significant metabolites (**Figure 23**). A large portion of the decreased-abundance metabolites were small peptides, metabolites of the aromatic amino acids metabolism pathways such as L-tryptophan, L-phenylalanine and indole, and pseudouridine and its derivatives. The top 50 features identified by fold change analysis are summarized in **Table 8** below in descending order relative to fold change and visualized in the volcano plot in **Figure 22**.
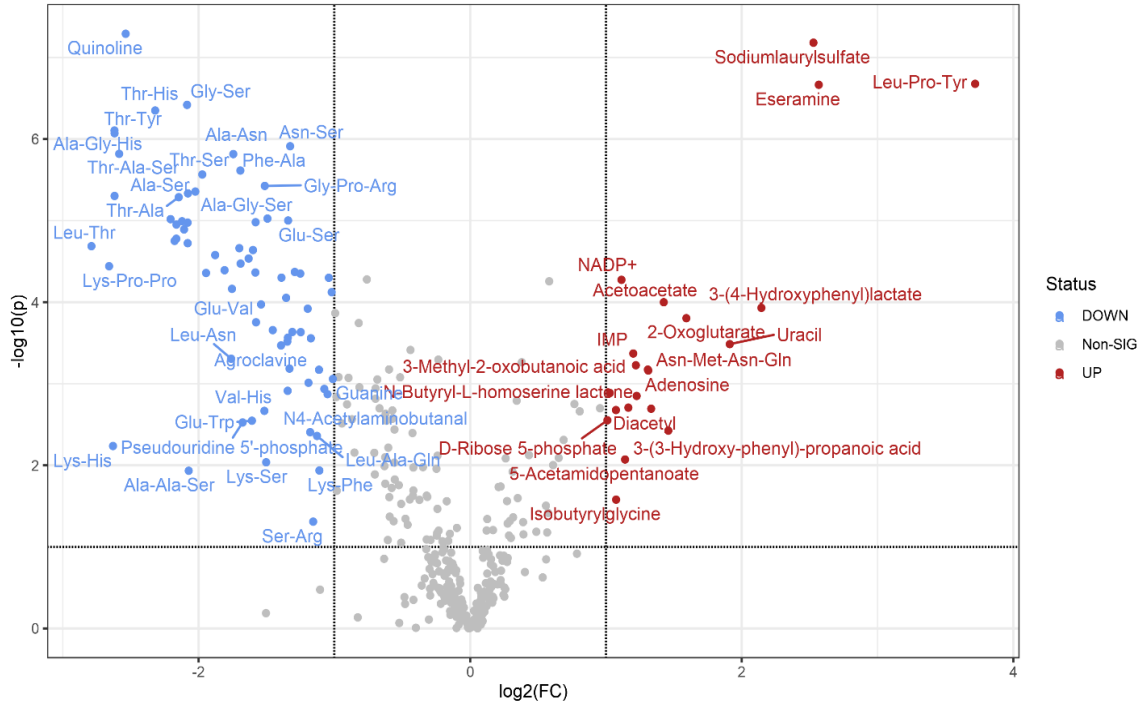
Figure 22: Important features selected by volcano plot with a fold change (x) 2 and t-test threshold (y) 0.1. Red circles represent features above the threshold while the blue circles represent features below the threshold. Both fold change and P values are log transformed.

Table 8: Top 50 features identified by fold change analysis

	Peaks(mz/rt)	FC	log2(FC)	raw.pval	-log10(p)
1	Leu-Pro-Tyr	13.17	3.7192	2.10E-07	6.6772
2	Eseramine	5.9281	2.5676	2.16E-07	6.6656
3	Sodiumlaurylsulfate	5.769	2.5283	6.59E-08	7.1814
4	NADP+	2.1651	1.1145	5.31E-05	4.2746
5	Proclavaminic acid	0.49321	-1.0197	7.53E-05	4.1231
6	Dethiobiotin	0.48604	-1.0409	5.02E-05	4.2996
7	Ala-Pro-Ser	0.44623	-1.1641	1.04E-05	4.9817
8	(S)-AMPA	0.42027	-1.2506	4.46E-05	4.351
9	Asn-Asp-Pro	0.4085	-1.2916	4.26E-05	4.3708
10	Asn-Ser	0.39898	-1.3256	1.22E-06	5.9125
11	Glu-Ser	0.39486	-1.3406	9.96E-06	5.0019
12	pentane-1 3 4 5-tetracarboxylate	0.3909	-1.3551	8.84E-05	4.0537
13	Pseudouridine	0.38167	-1.3896	5.00E-05	4.3012
14	Coformycin	0.35538	-1.4926	9.44E-06	5.0253
15	Gly-Pro-Arg	0.35055	-1.5123	3.75E-06	5.4256
16	Leu-Asp-Ser	0.3349	-1.5782	4.46E-06	5.3506
17	Glu-Thr	0.33452	-1.5798	1.04E-05	4.981
18	L-Ala-L-Glu	0.33411	-1.5816	4.32E-05	4.3641
19	Phe-Tyr	0.33003	-1.5993	2.30E-05	4.6391
20	4-Nitrophenyl-3-ketovalidamine	0.32279	-1.6313	2.91E-05	4.5354
21	Leu-Lys-Ala	0.30982	-1.6905	3.36E-05	4.4732
22	Phe-Ala	0.30944	-1.6923	2.43E-06	5.6142
23	Asn-His	0.30783	-1.6998	2.18E-05	4.6624
24	Ala-Asn	0.2985	-1.7442	1.53E-06	5.8162
25	Phenylacetonitrile	0.29649	-1.754	6.85E-05	4.164
26	L-Phenylalanine	0.28549	-1.8085	4.05E-05	4.3923
27	Ala-Pro.1	0.27208	-1.8779	2.64E-05	4.5779
28	N-Acetyl-beta-D-glucosaminylamine	0.25975	-1.9448	4.38E-05	4.3584
29	Thr-Ser	0.25477	-1.9727	2.72E-06	5.566
30	Ala-Gly-Ser	0.24591	-2.0238	4.40E-06	5.3565
31	His-His	0.23701	-2.077	2.87E-06	5.5419
32	Ala-Ser	0.23678	-2.0784	4.63E-06	5.3345
33	Phe-Val	0.23648	-2.0802	1.89E-05	4.7237
34	Isoquinoline	0.23641	-2.0806	1.06E-05	4.9757
35	Gly-Ser	0.23576	-2.0846	3.81E-07	6.4195
36	Skatole	0.23217	-2.1067	1.28E-05	4.8921
37	L-Tryptophan	0.22979	-2.1216	1.02E-05	4.9924
38	Thr-Ala	0.22594	-2.146	5.15E-06	5.2878
39	beta-Alanyl-L-arginine	0.22326	-2.1632	1.12E-05	4.9521
40	Indoleacrylicacid	0.22315	-2.1639	1.67E-05	4.7783

	Peaks(mz/rt)	FC	log2(FC)	raw.pval	-log10(p)
41	Indole	0.22122	-2.1765	1.77E-05	4.7516
42	2-Naphthylamine	0.2167	-2.2062	9.59E-06	5.0183
43	Thr-His	0.2003	-2.3197	4.46E-07	6.3504
44	Quinoline	0.17233	-2.5368	5.12E-08	7.2909
45	Thr-Ala-Ser	0.16669	-2.5848	1.52E-06	5.8185
46	Ala-Gly-His	0.16283	-2.6186	8.49E-07	6.0712
47	N2-Succinyl-L-arginine	0.16281	-2.6187	4.98E-06	5.3027
48	Thr-Tyr	0.16276	-2.6192	7.85E-07	6.1054
49	Lys-Pro-Pro	0.15834	-2.6589	3.62E-05	4.4412
50	Leu-Thr	0.1447	-2.7889	2.06E-05	4.687
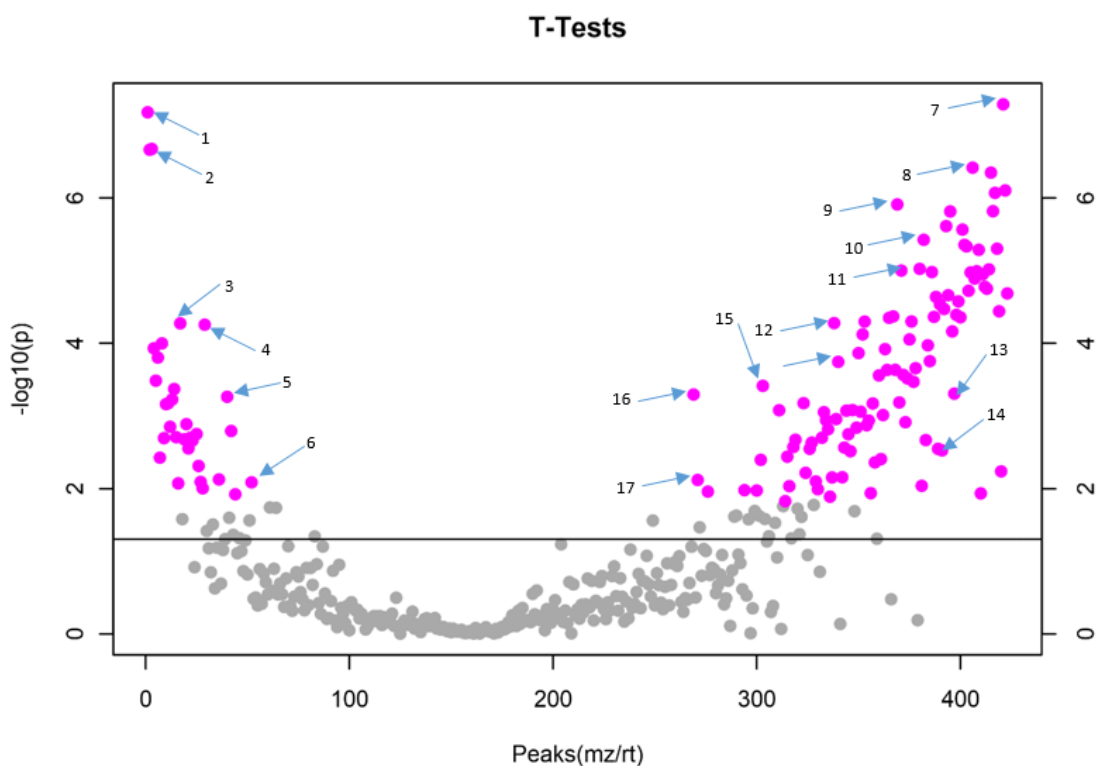
Figure 23: Important features selected by t-tests with a threshold of 0.1. Purple dots represent features above the threshold. The P values have been transformed to -log10 so that the more significant features (with smaller P values) are plotted higher on the graph. The features are labelled as follows: 1. Sodiumlaurylsulphate, 2. Leu-Pro-Tyr, 3. NADP$^+$, 4. (R)-2-hydroxyglutarate, 5. NG_NG_Dimethyl-L-arginine, 6. [PC(14:0)]1-tetradecanoyl-sn-glycero-3-phosphocholine, 7. Quinoline, 8. Gly-Ser, 9. Asn-Ser, 10. Gly-Pro-Arg, 11. Glu-Ser, 12. N(6)-[(Indol-3-yl)acetyl]-L-lysine, 13. Leu-Asn, 14. Pseudouridine-5'-phosphate, 15. L-glutamine, 16. Gly-His, 17. L-1-pyrolline-3-hydroxy-5-carboxylate

Further metabolomic investigations revealed induction of the pentose phosphate pathway marked by increase in metabolites of this pathway specifically glucose-6-phosphate, ribose-5-phosphate (or other pentose phosphates), fructose-6-phosphate and glyceraldehyde-3-phosphate (**Figure 24**).
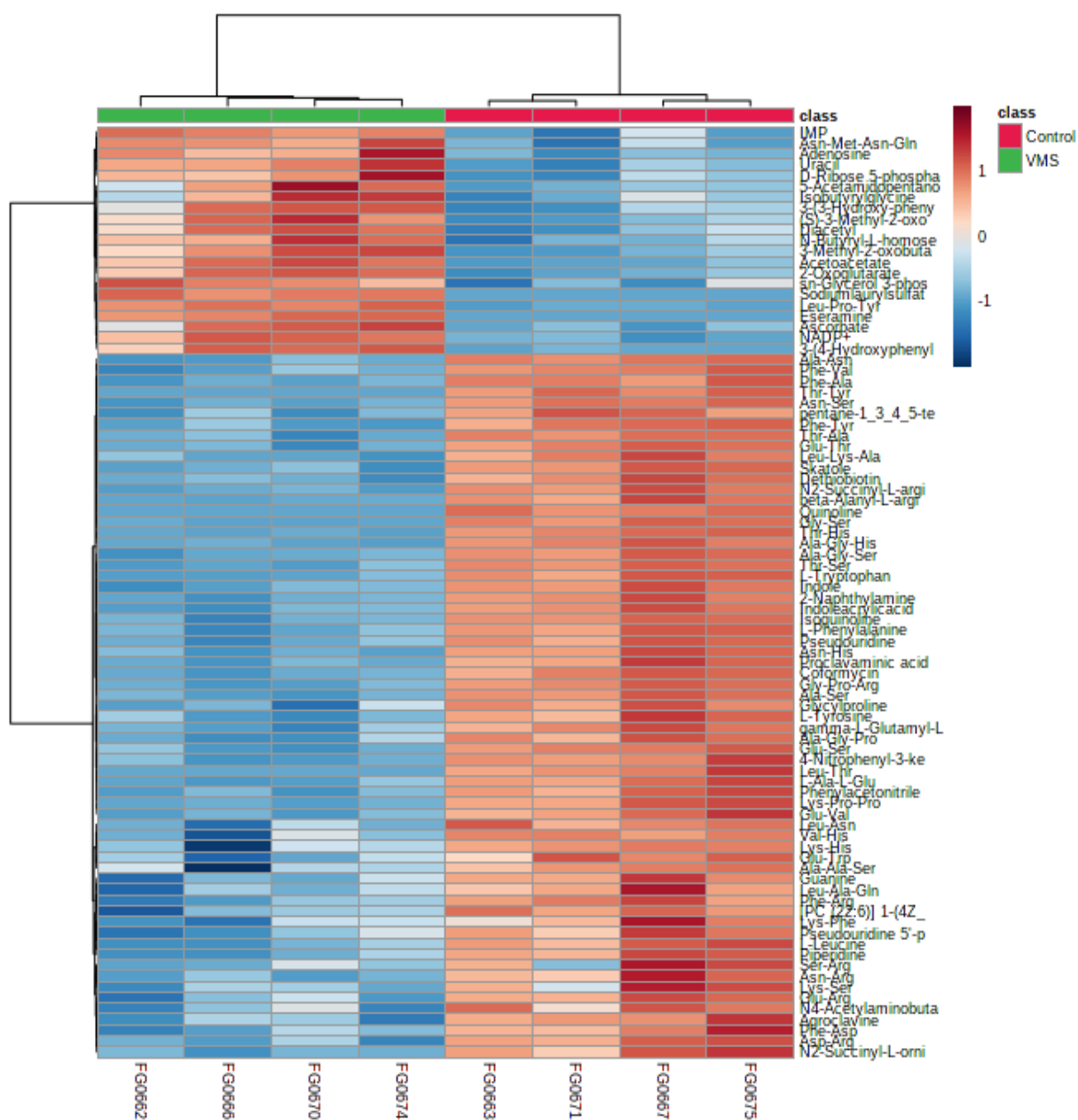
Figure 24: Cluster heatmap of the top 50 metabolites identified by volcano plot. The heatmap visualizes the reordered data matrix and the trees separately as dendrograms in the margins.

It is possible that treatment with VMS could have induced the parasite's pentose phosphate pathway to increase circulating NADPH and ribose-5-phosphate. Two points of increase are especially important i.e. glucose-6-phosphate and 6-phosphogluconate. Although, marked increase was not observed with 6-phosphogluconate, there was consistent increase in downstream metabolites of this pathway. NADPH was not identified in the analysis, however, its oxidized form, $NADP^+$, was seen to be undergoing significant increase (**Figure 23** and **Figure 24**). Increases in $NADP^+$ occur when NADPH is consumed, for example in protecting against oxidative stress. $NADP^+$ also triggers the pentose phosphate pathway which could explain the increased flux through this pathway. However, there might also have been regulation in processes that mediate conversion of NAD^+ to $NADP^+$ therefore modulating the intracellular level of NADPH. A possible reason for the increase in $NADP^+$ would be activation of NAD^+ kinases to convert NAD^+ to $NADP^+$ which was evidenced in the data by an observed increase in $NADP^+$ with a decrease in NAD^+ or an inhibition of NADPases that would yield the same observation (Singh et al., 2007). The activation of NAD^+ kinases to synthesise $NADP^+$ would have a positive influence on the metabolic modules involved in the production of NADPH thereby providing the reductive environment essential for neutralizing oxidative stress. Furthermore, metabolites in the glutathione biosynthesis pathway were depleted after treatment, specifically γ-glutamyl cysteine and glutathione indicating possible disruption in the parasite's anti-oxidant arsenal. This was further evidenced by slight decrease in ovothiol and increase in trypanothione disulfide, possibly as a result of oxidation of dihydrotrypanothione. Ascorbate, a non-enzymatic mechanism of defense against free radicals, accumulated (**Figure 24**) possibly due to inhibition of ascorbate peroxidase that is also involved in protection against oxidative stress. In the aromatic amino acids metabolism pathway, aromatic amino acids (phenylalanine,

tyrosine and tryptophan) all underwent depletion while their catabolic products (e.g. phenylpyruvate, 3-(4-hydroxyphenyl)lactate)) increased. These pathways involve NADH-dependent dehydrogenases that utilize NADH and in turn produce NAD^+ (Cazzulo et al., 1999; Marchese et al., 2018).

The metabolomics data revealed significant perturbations to *T. congolense* metabolism, particularly areas of energy metabolism, aromatic amino acid metabolism and cellular redox. However, it was not possible to readily assign candidate target proteins for VMS in *T. congolense*.

4.3 *Leishmania mexicana* treated with JYH and VMS

4.3.1 Data preprocessing and transformation

Data preprocessing and transformation was carried out as described in aforementioned sections using MetaboanalystR package version 3.2.0 in R and the data details are as shown in **Table 9** below. The JYH and VMS treatment groups were analysed concurrently as the dataset contained both treatment groups and a control group. The samples were in three sample groups; control (LM2101, LM2102, LM2103, LM2104), VMS (LM2105, LM2108, LM2107) and JYH (LM2109, LM2110, LM2111, LM2112). Positive features indicated peaks that had recorded peak intensity values while processed features indicated all the features that were used in subsequent analysis after zero and missing values were corrected with 1/5th of the smallest value in the sample group. Features were preprocessed and normalized as described in Section **3.4**.

Table 9: Summary of data processing results of the *L. mexicana* dataset

	Sample	*Features (positive)*	*Missing/zero*	*Features (processed)*
Control	LM2101	801	28	829
	LM2102	805	24	829
	LM2103	798	31	829
	LM2104	802	27	829
JYH	LM2109	783	46	829
	LM2110	791	38	829
	LM2111	770	59	829
	LM2112	778	51	829
VMS	LM2105	801	28	829
	LM2107	808	21	829
	LM2108	791	38	829

4.3.2 Global metabolomics profiles of the control and JYH and VMS treatment samples

Multivariate analysis was conducted using unsupervised PCA and supervised PLS-DA to assess drug perturbations on the metabolome of the parasite. The first two components, PC 1 and PC 2, were chosen explaining 47% and 20.5% of the variance respectively. Distinct clustering was observed in the two treatment groups and the control group (**Figure 25**).
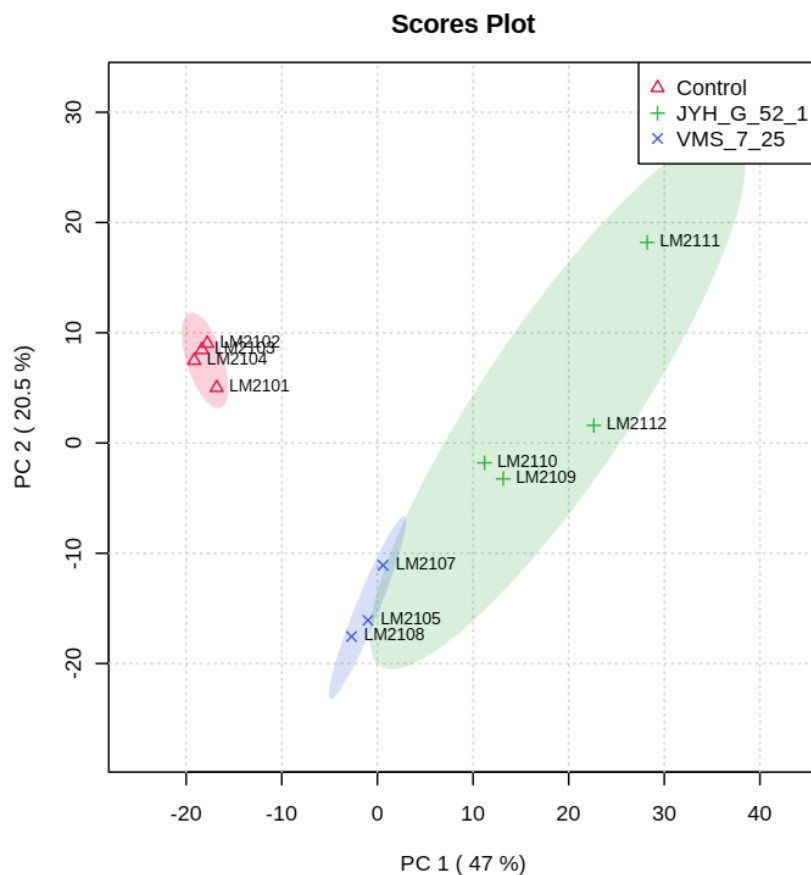
Figure 25: PCA scores plot between PC1 covering 47% variance and PC2 covering 20.5% variance. Control groups are marked by red triangles, the JYH treatment groups by green crosses and VMS treatment groups by blue crosses.

Further summary of the data was conducted using PLS-DA analysis (accuracy = 1.0, R^2 = 0.99619, Q^2 = 0.88574). The difference between R^2 and Q^2 was < 0.2 excluding model overfitting and consistent with PCA showed that the drugs had effects on the metabolome of *L. mexicana*.
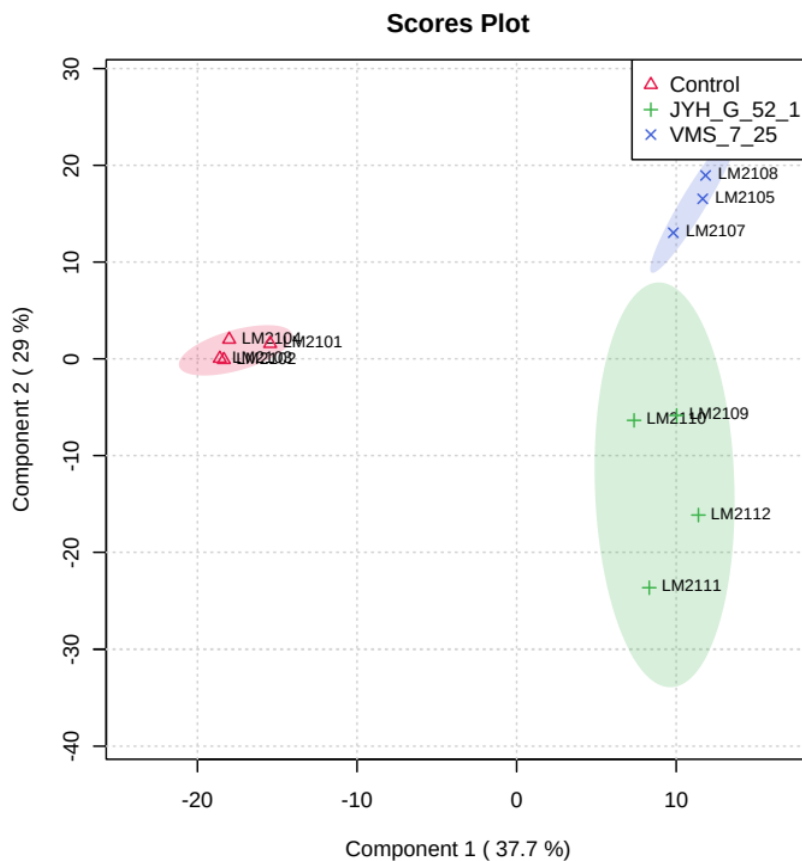
Figure 26: PLS-DA scores plots between component 1 (37.7%) and component 2 (29%). Control groups are marked by red triangles, the JYH treatment groups by green crosses and VMS treatment groups by blue crosses.

4.3.3 Identification of statistically significant metabolites

Significantly altered metabolites were identified using one-way ANOVA with Tukey's *post hoc* analysis. **Table 10** shows 50 of the identified metabolites with their corresponding F value and P value. Most of the metabolites identified were phospholipids and lysophospholipids.

Table 10: Significantly altered metabolites as identified by one-way ANOVA with Tukey's *post hoc* analysis

	Peaks(mz/rt)	f.value	p.value	-log10(p)	FDR	Tuk
1	[PC acety] 1-acetyl-sn-glycero-3-phosphocholine	2350.9	8.32E-12	11.08	5.17E-09	VMS
2	CMP-3-deoxy-D-manno-octulosonate	1440.8	5.88E-11	10.231	1.82E-08	JYH
3	Tetradecanoyl-CoA	1184.9	1.28E-10	9.8924	2.65E-08	JYH
4	Sinalbin	773.4	7.01E-10	9.1543	1.09E-07	JYH
5	Deoxyinosine	623.55	1.65E-09	8.7823	2.05E-07	JYH
6	Xanthine	587.25	2.09E-09	8.6788	2.17E-07	JYH
7	[PC (18:3)] 1-(9Z 12Z 15Z-octadecatrienoyl)-sn-glycero-3-phosphocholine	525.82	3.25E-09	8.4883	2.48E-07	JYH
8	[SP] Sphinganine-1-phosphate	520.09	3.39E-09	8.4694	2.48E-07	JYH
9	2-Aminoethylphosphocholate	512.83	3.59E-09	8.4451	2.48E-07	JYH
10	1-(beta-D-Ribofuranosyl)-1 4-dihydronicotinamide	492.41	4.22E-09	8.3751	2.50E-07	JYH
11	[FA (6:0)] R-hexanoyl CoA	486.2	4.43E-09	8.3533	2.50E-07	JYH
12	[PC (22:6)] 1-(4Z 7Z 10Z 13Z 16Z 19Z-docosahexaenoyl)-sn-glycero-3-phosphocholine	392.04	1.04E-08	7.9827	5.39E-07	JYH
13	[PC (18:2)] 1-(9Z 12Z-octadecadienoyl)-sn-glycero-3-phosphocholine	374.76	1.24E-08	7.9052	5.52E-07	JYH
14	S-Acetyldihydrolipoamide	374.64	1.25E-08	7.9046	5.52E-07	JYH
15	L-Tyrosine methyl ester	352.94	1.58E-08	7.8021	6.53E-07	JYH
16	Alpha-Linoleoylcholine	328.23	2.10E-08	7.6775	8.01E-07	VMS
17	425.5869874	324.69	2.19E-08	7.6589	8.01E-07	JYH
18	[PC (18:0)] 1-octadecanoyl-sn-glycero-3-phosphocholine	312.12	2.56E-08	7.5912	8.84E-07	JYH
19	[PE (16:1/16:1)] 1 2-di-(9Z-hexadecenoyl)-sn-glycero-3-phosphoethanolamine	300.85	2.96E-08	7.5281	9.39E-07	JYH
20	[PC acetyl(12:2)] 1-dodecyl-2-acetyl-sn-glycero-3-phosphocholine	299.32	3.02E-08	7.5194	9.39E-07	JYH
21	LPA(0:0/18:2(9Z_12Z))	277.3	4.09E-08	7.3884	1.21E-06	JYH
22	Phe-Thr	265.35	4.86E-08	7.313	1.37E-06	VMS
23	LysoPC(20:3(5Z_8Z_11Z))	255.67	5.63E-08	7.2494	1.52E-06	JYH
24	LysoPE(0:0/18:1(11Z))	244.39	6.73E-08	7.1723	1.74E-06	JYH
25	Leucyl-leucine	240.81	7.13E-08	7.147	1.77E-06	JYH
26	Hypoxanthine	234.01	7.98E-08	7.0982	1.85E-06	JYH
27	LysoPC(22:5(4Z_7Z_10Z_13Z_16Z))	233.63	8.03E-08	7.0954	1.85E-06	JYH

	Peaks(mz/rt)	f.value	p.value	-log10(p)	FDR	Tuk
28	1-Oleoylglycerophosphocholine	230.36	8.49E-08	7.0713	1.85E-06	JYH
29	(2E)-Octadecenoyl-CoA	227.73	8.88E-08	7.0517	1.85E-06	JYH
30	coronatine	227.26	8.95E-08	7.0482	1.85E-06	VMS
31	2-Deoxy-D-ribose-5-phosphate	214.91	1.11E-07	6.9528	2.23E-06	JYH
32	LysoPC(20:2(11Z_14Z))	180.95	2.19E-07	6.66	4.18E-06	JYH
33	5-6-Dihydrouridine	180.27	2.22E-07	6.6536	4.18E-06	JYH
34	Val-Tyr	177.52	2.36E-07	6.6274	4.31E-06	JYH
35	sn-glycero-3-Phosphocholine	164.73	3.16E-07	6.5006	5.60E-06	JYH
36	Uracil	161.56	3.41E-07	6.4676	5.88E-06	JYH
37	1-18:3-lysophosphatidylethanolamine	159.43	3.59E-07	6.4451	6.02E-06	JYH
38	Cys-Met-Phe-His	154.62	4.04E-07	6.3932	6.61E-06	JYH
39	Nicotinate	153.2	4.19E-07	6.3776	6.67E-06	JYH
40	L-Aspartate	150.45	4.50E-07	6.3469	6.98E-06	JYH
41	CMP	144.89	5.21E-07	6.2832	7.89E-06	JYH
42	Trimetaphosphate	137.86	6.32E-07	6.1992	9.13E-06	JYH
43	LysoPC(22:4(7Z_10Z_13Z_16Z))	137.37	6.41E-07	6.1931	9.13E-06	JYH
44	Adenosine	137.04	6.47E-07	6.1892	9.13E-06	JYH
45	Westiellamide	135.71	6.72E-07	6.1726	9.21E-06	JYH
46	5'-Deoxyadenosine	135.19	6.82E-07	6.1662	9.21E-06	JYH
47	Leu-Thr-Pro	134.16	7.03E-07	6.1532	9.28E-06	VMS
48	[PC (14:0)] 1-tetradecanoyl-sn-glycero-3-phosphocholine	130.21	7.89E-07	6.1029	1.02E-05	JYH
49	histidine methyl ester	123.59	9.66E-07	6.015	1.22E-05	JYH
50	[PC (16:0)] 1-hexadecanoyl-sn-glycero-3-phosphocholine	121.91	1.02E-06	5.992	1.27E-05	JYH

Further metabolic signatures were identified using volcano plot analysis (**Figure 27)** and fold change analysis (**Table 11**). Most of the metabolites that were being depleted in the JYH treatment group seemed to belong to lipid pathways exemplified by CoA, tetradecanoyl-CoA, hexanoyl-CoA as well as various phospholipids.
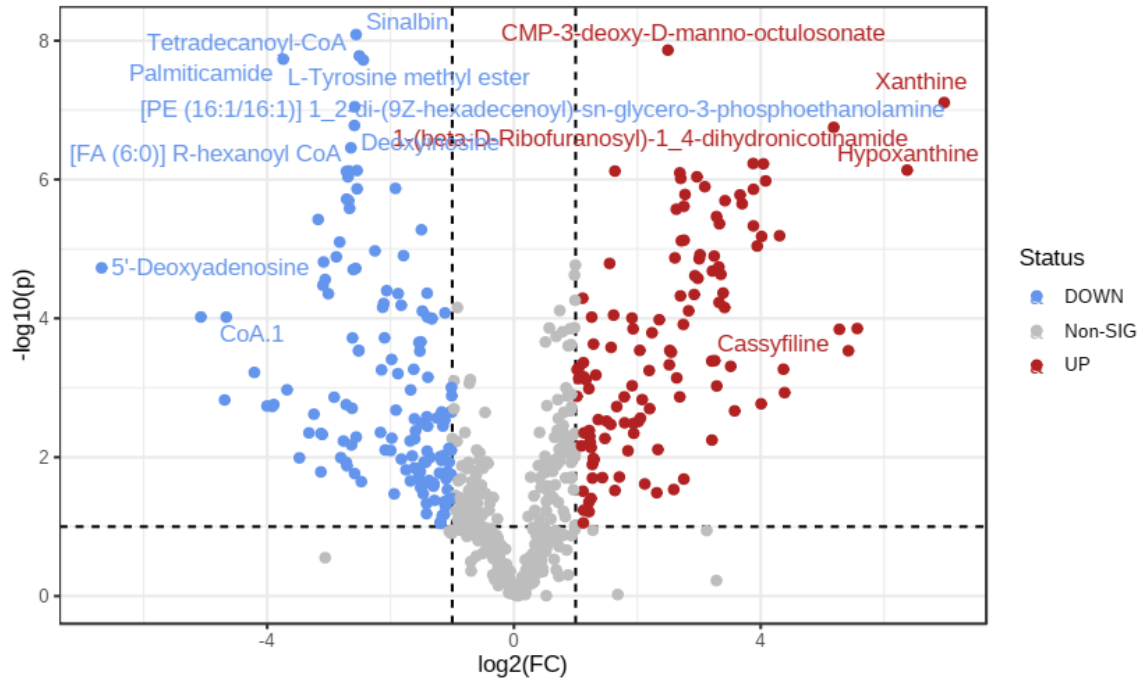
Figure 27: Volcano plot showing metabolites altered upon exposure to JYH. Metabolites marked by red dots increased while those marked in blue dots decreased.

Table 11: Metabolic signatures identified by fold change analysis between the control and JYH treatment group.

	Peaks(mz/rt)	Fold Change	log2(FC)
1	Xanthine	125.7	6.9738
2	Hypoxanthine	82.879	6.3729
3	Propanoyl phosphate	47.308	5.564
4	[PI (18:1)] 1-(9Z-octadeconyl)-sn-glycero-3-phosphoethanolamine	42.839	5.4209
5	Cassyfiline	38.7	5.2743
6	1-(beta-D-Ribofuranosyl)-1 4-dihydronicotinamide	36.33	5.1831
7	[Fv] Okanin 4'-(2" 4"-diacetyl-6"-p-coumarylglucoside)	20.916	4.3865
8	[PE (18:1)] 1-(9Z-octadecenoyl)-sn-glycero-3-phosphoethanolamine	20.648	4.3679
9	LysoPE(18:2(9Z 12Z)/0:0)	19.77	4.3053
10	[PC (14:0)] 1-tetradecanoyl-sn-glycero-3-phosphocholine	16.908	4.0796
11	2-Aminoethylphosphocholate	16.483	4.0429
12	Adenosine	16.159	4.0142
13	[PC (16:1)] 1-(9Z-hexadecenoyl)-sn-glycero-3-phosphocholine	16.071	4.0064
14	Nicotinate	15.374	3.9425
15	[PC (22:6)] 1-(4Z 7Z 10Z 13Z 16Z 19Z-docosahexaenoyl)-sn-glycero-3-phosphocholine	14.743	3.882
16	LysoPC(22:5(4Z 7Z 10Z 13Z 16Z))	14.733	3.881
17	[PC (18:3)] 1-(9Z 12Z 15Z-octadecatrienoyl)-sn-glycero-3-phosphocholine	14.706	3.8783
18	LPA(0:0/18:2(9Z_12Z))	12.988	3.6991
19	[PC (18:2)] 1-(9Z 12Z-octadecadienoyl)-sn-glycero-3-phosphocholine	12.68	3.6645
20	LPA(0:0/18:1(9Z))	11.936	3.5773
21	[Fv] Okanin 4'-(3" 4"-diacetyl-6"-p-coumarylglucoside)	11.415	3.5129
22	Uracil	10.71	3.4209
23	Asn-Leu-Trp-Tyr	10.661	3.4142
24	Pantetheine	10.482	3.3898
25	LysoPE(0:0/18:2(9Z_12Z))	10.224	3.354
26	LysoPC(20:3(5Z_8Z_11Z))	10.057	3.3301
27	Pantetheine 4'-phosphate	10.023	3.3253
28	LysoPC(22:4(7Z 10Z 13Z 16Z))	9.9865	3.32
29	L-methioninamide	9.7588	3.2867
30	1-Oleoylglycerophosphocholine	9.7353	3.2832

	Peaks(mz/rt)	Fold Change	log2(FC)
31	Methylenediurea	9.7205	3.281
32	2-methylphosphinoyl-2-hydroxyacetate	9.5406	3.2541
33	LysoPC(20:2(11Z_14Z))	9.4859	3.2458
34	[PC (16:0)] 1-hexadecanoyl-sn-glycero-3-phosphocholine	9.3002	3.2173
35	Thiamin monophosphate	9.2445	3.2086
36	5-Guanidino-2-oxopentanoate	9.2305	3.2064
37	L-Aspartate	0.11083	-3.1736
38	Aspartyl-L-proline	0.10566	-3.2424
39	PC(18:3(9Z 12Z 15Z)/22:6(4Z 7Z 10Z 13Z 16Z 19Z))	0.099958	-3.3225
40	[PC (18:2/18:2)] 1 2-di-(9Z 12Z-octadecadienoyl)-sn-glycero-3-phosphocholine	0.089742	-3.4781
41	Acetyl-CoA	0.078115	-3.6783
42	Palmiticamide	0.07488	-3.7393
43	UTP	0.067334	-3.8925
44	UDP-glucose	0.066268	-3.9156
45	13 14-Dihydro- lipoxin A4	0.062376	-4.0029
46	prostaglandin H1	0.054124	-4.2076
47	CoA.1	0.039487	-4.6625
48	3-Methylbutanoyl-CoA	0.038751	-4.6896
49	3-heptaprenyl-4-hydroxy-5-methoxybenzoate	0.02967	-5.0748
50	5'-Deoxyadenosine	0.009729	-6.6835

Key

	Increasing metabolites
	Decreasing metabolites

4.3.4 Pathway enrichment analysis

To further understand significantly altered metabolites at the pathway level, MetaboAnalyst

5.0 (https://www.metaboanalyst.ca) was used to carry out pathway enrichment analysis using

the pathway analysis feature. The *T. brucei* (KEGG) pathway library was used for prediction

analysis. For the over representation analysis and the pathway topology analysis,

respectively, the hypergeometric test and relative-betweenness centrality were chosen to

determine whether the pathway was significant. Findings from pathway enrichment analysis are projected in **Figure 28** below.
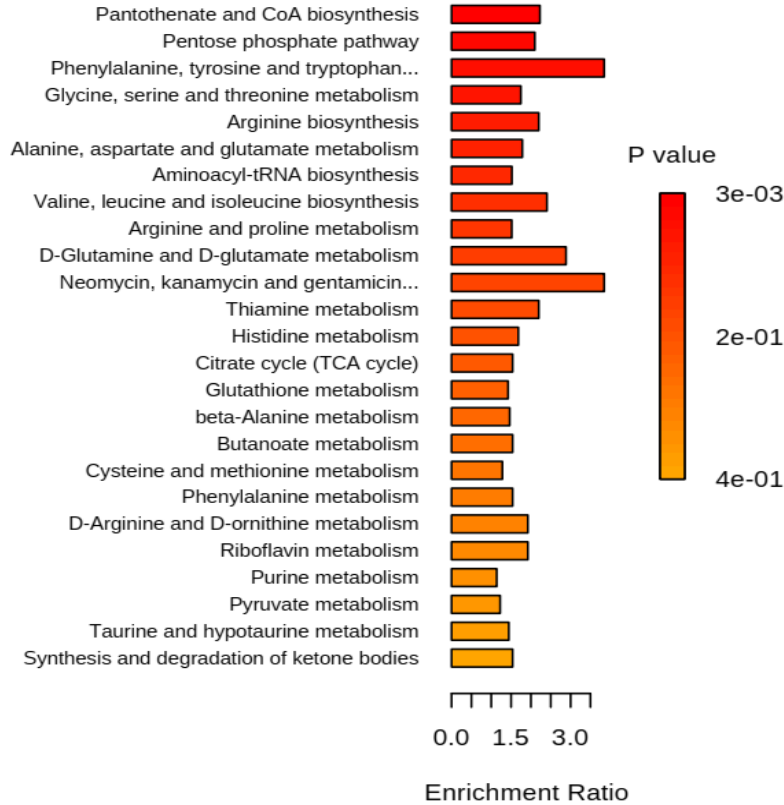
Figure 28: The top 25 metabolic pathways identified with pathway significance P ≤ 0.05. Pantothenate and CoA biosynthesis, pentose phosphate pathway, and phenylalanine, tyrosine and tryptophan metabolism were identified to be significantly altered metabolic pathways.

Analysis of the perturbed pathways revealed large increases in lysophopholipids (single acyl chain) and decreases in selected phospholipids (two acyl chains). There is a dynamic interplay between lysophospholipids and phospholipids with single chains cleaved from phospholipids to yield lysophospholipids which regain a second chain after transfer of an acyl group from an acyl-CoA. The TCA cycle intermediates were also of diminished abundance with knock-on effects on ATP since ATP synthesis depends upon the TCA cycle feeding the electron transport chain. The TCA cycle depends upon Acetyl CoA which was

also diminished. Sphingolipids which depend on palmitoyl CoA were also diminished. Coenzyme A itself was diminished, however, key metabolites in its biosynthesis via the pantothenate pathway increased. The combined data indicated that the drastic impact of JYH and VMS on the *L. mexicana* metabolome could have been due to loss of coenzyme A synthesis which has knock-on effects across the metabolome. An increase in pantothenate pathway intermediates is consistent with inhibition of either phosphopantetheine adenylyltransferase (PPAT) or dephospho-CoA kinase (DPCK) (**Figure 29**). Since dephospho-CoA was not found in the dataset it was not possible to distinguish the two, hence from the analysis either enzyme could be targeted.
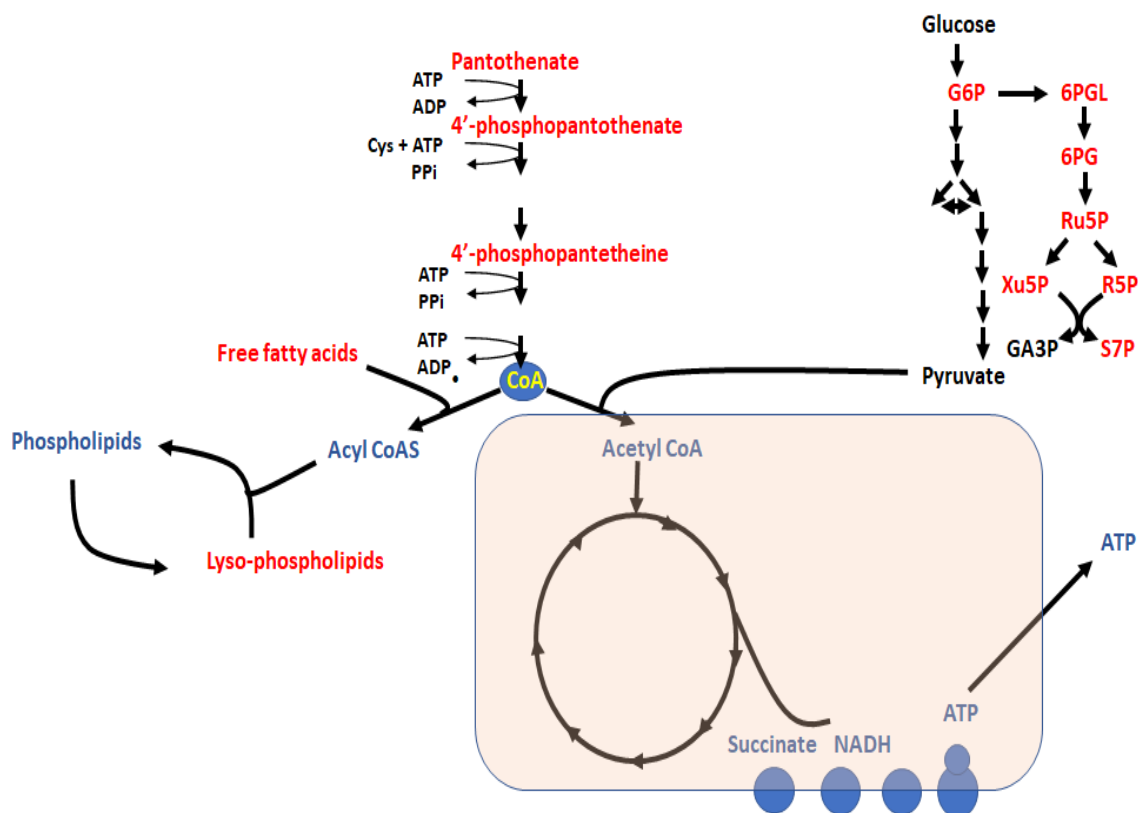
Figure 29: The Coenzyme A biosynthesis pathway, glycolysis, pentose phosphate pathway and the TCA cycle. Metabolites in red increased after treatment of *L. mexicana* with VMS and JYH while those in blue decreased.

4.3.5 Target validation

To perform further target validation, screening VMS and JYH targets was done using the online reverse docking tool, PharmMapper (http://lilab.ecust.edu.cn/pharmmapper/). PharmMapper uses pharmacophore mapping to find probable targets for specified small compounds. Both VMS and JYH were used as query structures in a PharmMapper search to locate potential targets. The 16159 druggable pharmacophore models and 52431 ligandable pharmacophore models were employed in this instance as the search databases, with a limit of 300 reserved matching targets. Dephospho-CoA kinase (Rank 153) as verified in the DrugBank database was identified by PharmMapper.

4.3.6 Receptor sequence identification and structure prediction

Leishmania mexicana Dephospho-CoA kinase and PPAT sequences were retrieved from the TriTryp database (LmxM.22.1530, putative dephospho-CoA kinase and LmxM.31.2070 cytidyltransferase). The sequence used to generate the PPAT structure was annotated as cytidyltransferase. To confirm the possibility of the sequence being PPAT, annotated putative PPAT sequence from *Leishmania major* (LmjF.32.2070) was blasted against the *L. mexicana* on Blastp. Blastp identified the sequence LmxM.31.2070 (GenBank accession number: XP_003878055.1) with 87.72% sequence similarity. The putative sequences were validated using CDD which predicted that these proteins belong to the nucleoside/nucleotide kinases (cd02019) and nucleotidyltransferases (cl00015) superfamilies respectively. Similar findings were identified from InterPro searches, where the protein sequences for the InterPro family signatures IPR001977 (dephospho-CoA kinase) and IPR001980 (phosphopantetheine adenylyltransferase) were found. Additionally, GO:0015937 (CoA biosynthetic process) was predicted by InterPro for the biological process and molecular function of DPCK and PPAT. Using the primary sequence of the proteins their three dimensional structures were predicted

using AlphaFold2 with subsequent refinement using GalaxyRefine. The results were visualized in Pymol and are as shown in **Figure 30** below.
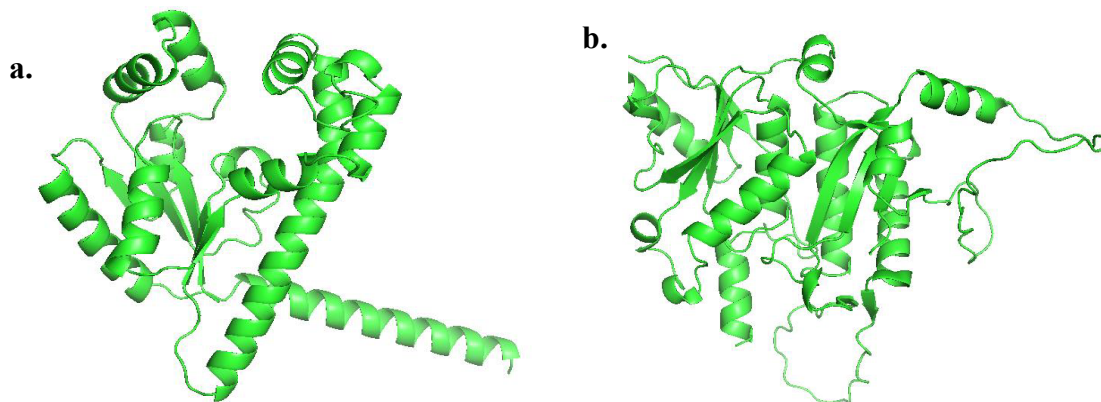
Figure 30: Predicted structure of *L. mexicana* (**a**) DPCK (**b**) PPAT

The structures' validity was evaluated and the results are as shown in **Table 12** below.

Table 12: DPCK and PPAT structure validity evaluations

	DPCK	*PPAT*
ERRAT	95.339	99.2674
VERIFY 3D (% of residues that have averaged 3D-1D score >= 0.2)	84.84	86.83
% residues in allowed regions	95.9	96.3
% residues in additional allowed regions	4.1	4.0
% residues in generously allowed regions	-	0.7
ProSA z-score	-7.29	-9.23

The PDB format of the predicted structures were downloaded to be used in subsequent docking steps.

4.3.7 Molecular docking analysis

Parameters stated in **Table 13** below were used to carry out molecular docking. The centres were determined from binding pocket predictions obtained from ProteinPlus and active sites confirmed using CDD analysis.

Table 13: Docking parameters used in AutoDock Vina for docking JYH and VMS on DPCK and PPAT

Receptor		DPCK	PPAT
Grid centre (xyz coordinates)		-0.021, 0.163, 1.964	-10.027, -3.34, -1.564
Number of modes generated		20	20
Exhaustiveness		16	16
Grid size (xyz)	JYH	100	100
	VMS	100	100

Output from AutoDock Vina were analysed based on the absolute docking energies of JYH, VMS and the corresponding endogenous substrate. The binding affinities (kcal/mol) of ten generated binding poses of the compounds and endogenous substrates on the protein macromolecules are as shown in **Table 14** and **15** below.

Table 14: Binding affinity of dephospho-CoA, JYH and VMS on the predicted *L. mexicana* DPCK

Mode	Dephospho-CoA affinity (kcal/mol)	VMS affinity (kcal/mol)	JYH affinity (kcal/mol)
1	-6.2	-8.1	-7.9
2	-5.9	-7.4	-7.4
3	-5.9	-7.2	-7.3
4	-5.7	-7.1	-7.0
5	-5.5	-6.9	-7.0
6	-5.5	-6.8	-7.0
7	-5.5	-6.6	-7.0
8	-5.5	-6.5	-6.8
9	-5.4	-6.3	-6.8
10	-5.3	-6.3	-6.8

Table 15: Binding affinity of ATP, JYH and VMS on the predicted *L. mexicana* PPAT

Mode	ATP affinity (kcal/mol)	VMS affinity (kcal/mol)	JYH affinity (kcal/mol)
1	-7.7	-8.2	-9.7
2	-7.0	-8.2	-9.4
3	-6.9	-8.1	-8.8
4	-6.9	-8.1	-8.8
5	-6.9	-8.1	-8.7
6	-6.8	-8.1	-8.2
7	-6.8	-7.8	-8.1
8	-6.7	-7.7	-7.8
9	-6.7	-7.7	-7.7
10	-6.6	-7.5	-7.5

4.3.8 Docking visualisation

The output from AutoDock Vina were visualized on Biovia discovery studio as shown in **Figures 31 - 34** below. The figures show interactions between the compounds and the predicted *L. mexicana* DPCK and PPAT proteins. The protein is depicted as a cartoon and colored according to secondary structures while the ligand is depicted in stick format where (**a**) shows 3D visualisation of interactions between residues in the binding pocket and the ligand compounds. Interacting residues are also depicted in stick format and bonds are shown as color-coded dashed lines (**b**) shows a 2D illustration of the ligand interactions with their corresponding distances. The ligand is shown as a stick figure while the interacting residues are shown as color-coded balls, while (**c**) shows the positioning of ligand within the protein.

Key

🟪	Alkyl bonds
🟩	Conventional hydrogen bond
🟦	Hydrophobic bond
🟪	Pi-sigma bond
🟥	Pi-Pi T shaped
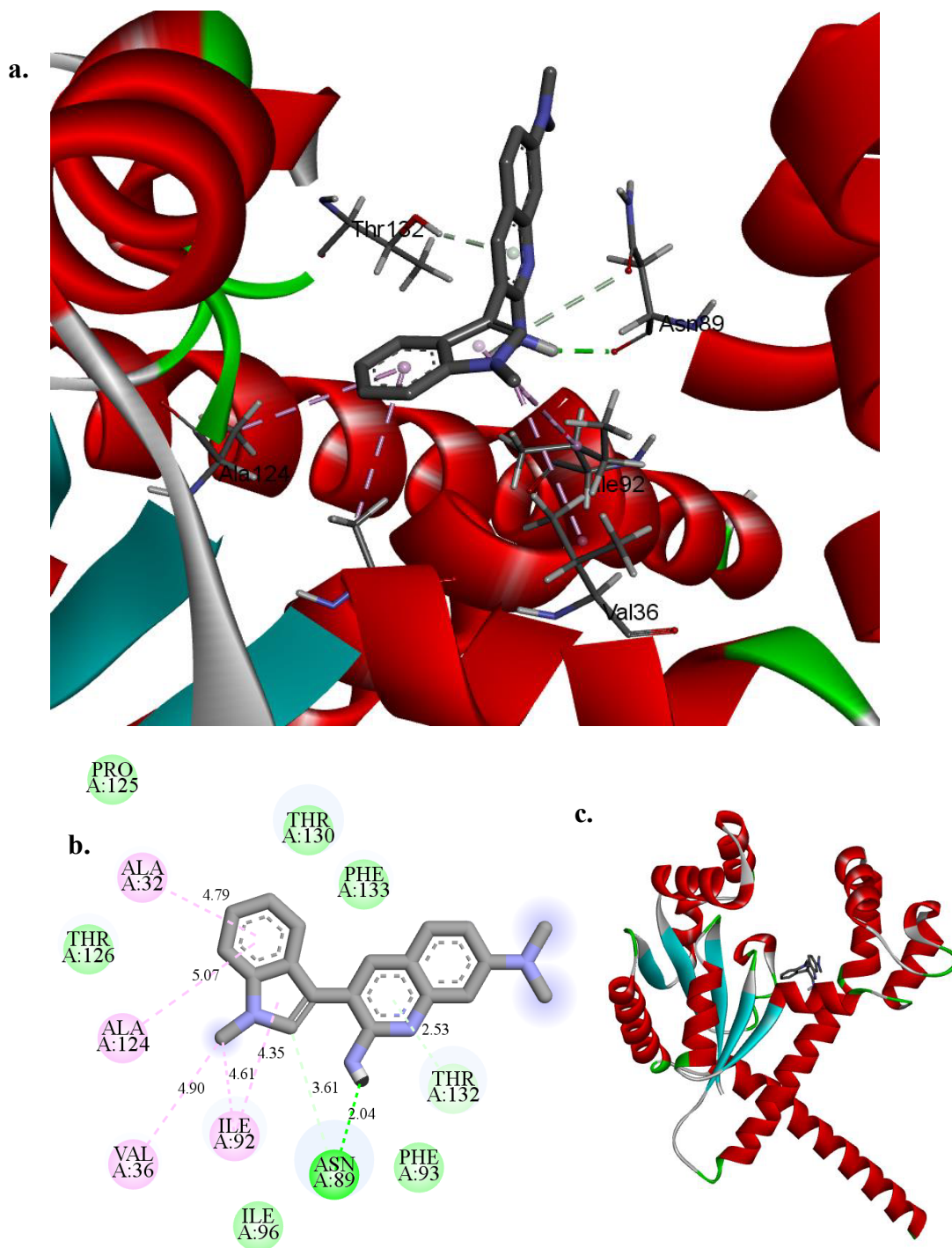
Figure 31: Interactions between VMS ligand and the predicted *L. mexicana* DPCK protein. VMS interacted with Ala 32, Val 36, Asn 89, Ile 92, Phe 93, Ile 96, Ala 124, Pro 125, Thr 126, Thr 130, Thr 132 and Phe 133
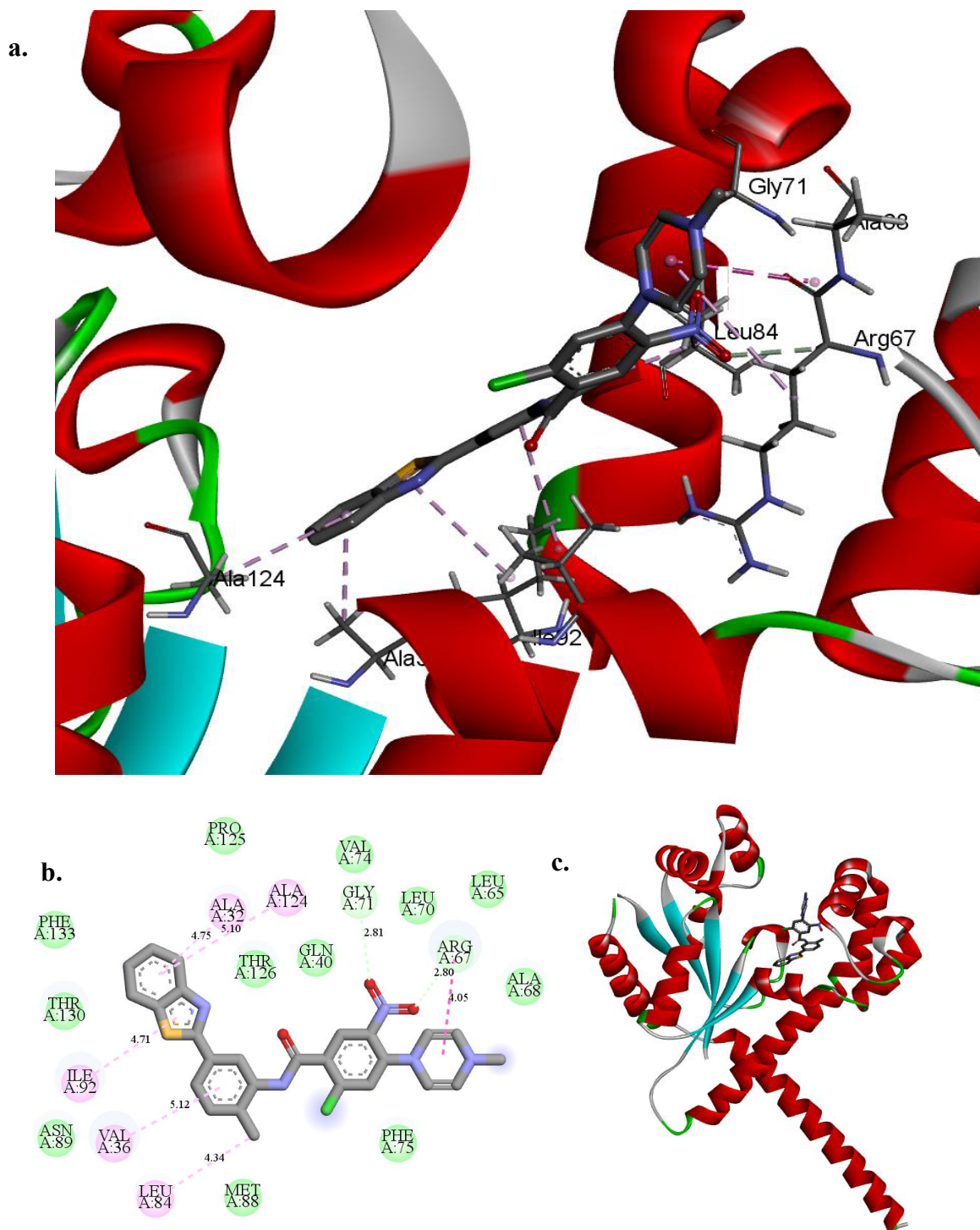
Figure 32: Interactions between JYH ligand and the predicted *L. mexicana* DPCK protein. JYH interacted with Ala 32, Val 36, Gln 40, Leu 65, Arg 67, Ala 68, Leu 70, Gly 71, Val 74, Phe 75, Leu 84, Met 88, Asn 89, Ile 92, Ala 124, Pro 125, Thr 126, Thr 130 and Phe 133.
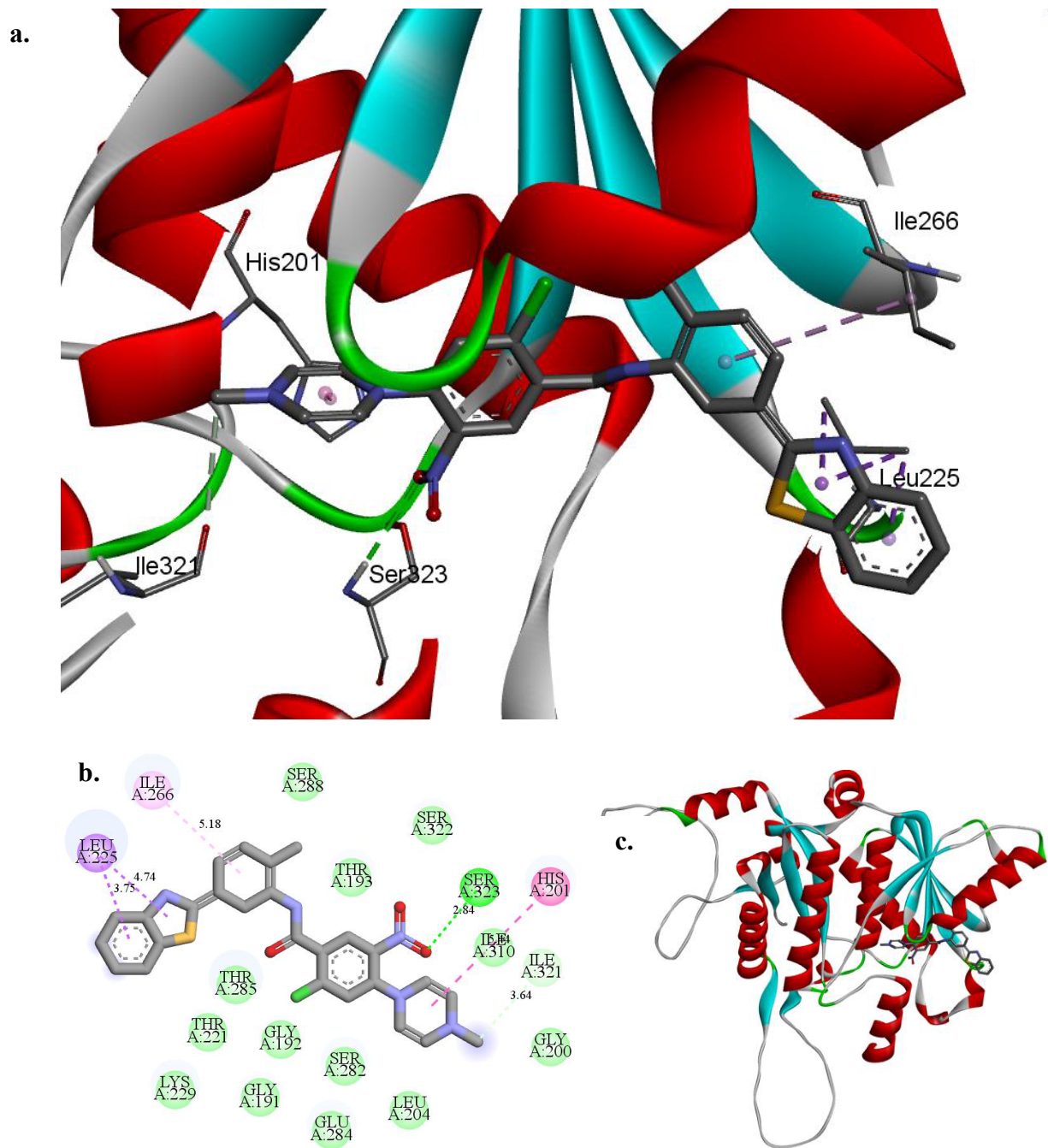
Figure 33: Interactions between JYH ligand and the predicted *L. mexicana* PPAT protein. JYH interacted with Gly 191, Gly 192, Thr 193, Gly 200, His 201, Leu 204, Thr 221, Leu 225, Lys 229, Ile 266, Glu 284, Ser 282, Thr 285, Ser 288, Ile 310, Ile 321, Ser 322, and Ser 323
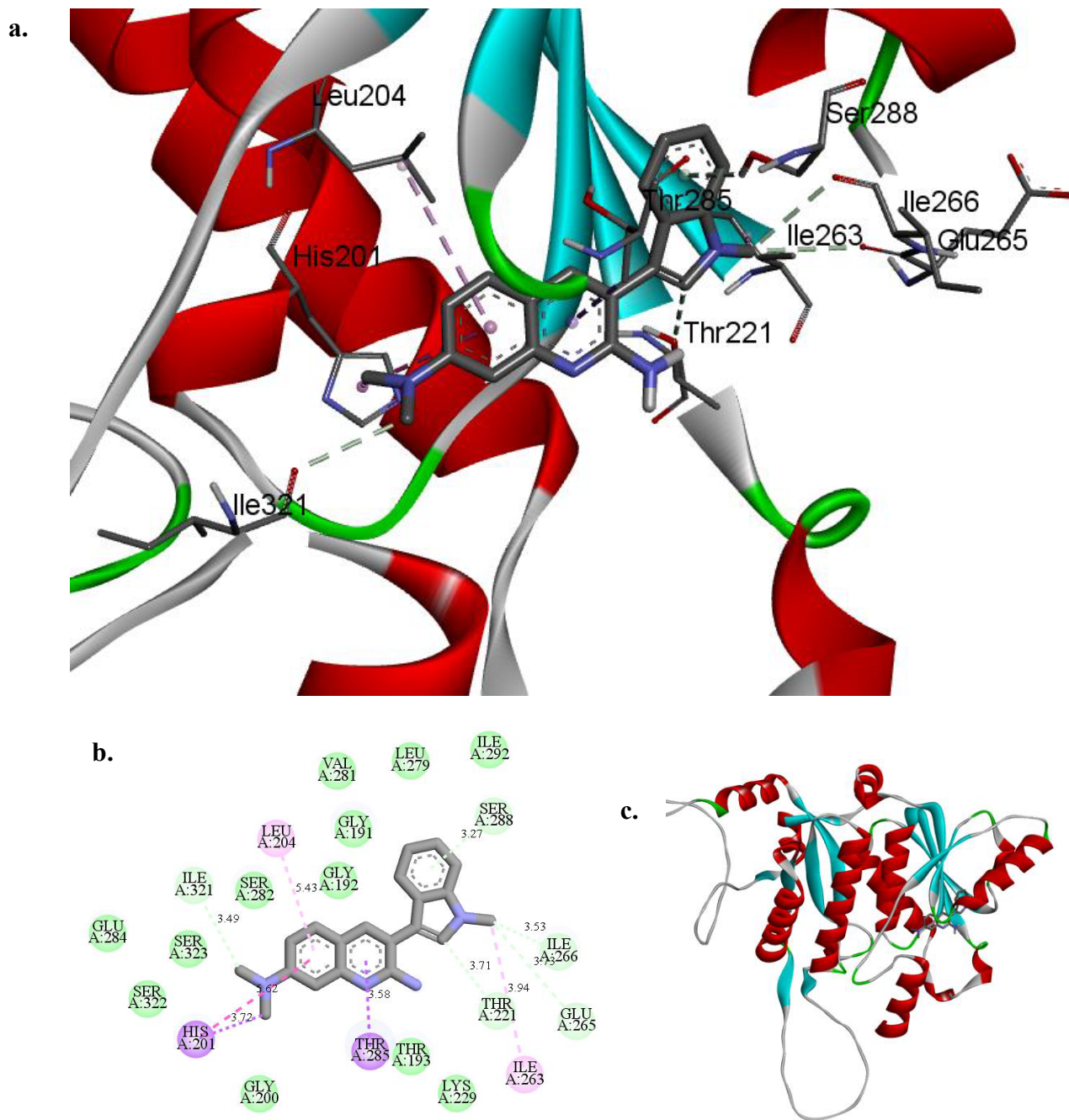
Figure 34: Interactions between VMS ligand and the predicted *L. mexicana* PPAT protein. VMS interacted with Gly 191, Ile 192, Thr 193, Gly 200, His 201, Leu 204, Thr 221, Lys 229, Ile 263, Glu 265, Ile 266, Leu 279, Val 281, Ser 282, Glu 284, Thr 285, Ser 288, 292, Ser 322 and Ser 323.

4.4 Comparison of *Leishmania mexicana* and *Trypanosoma congolense* dephospho-CoA kinase and phosphopantetheine adenylyltransferase

4.4.1 Pairwise protein sequence alignment

From Section **4.2** and Section **4.3**, different mechanisms of action were identified for VMS – inhibition of either DPCK or PPAT in *L. mexicana* and oxidative stress in *T. congolense*. Having identified these different mechanisms of action for VMS, pairwise protein sequence alignment was performed on the DPCK and PPAT sequences of the respective parasites (**Figure 35 - 36**) using EMBOSS and BLASTp. This was done to identify any changes in the active sites or substrate binding sites that might have dictated the different mechanisms of action – binding to DPCK or PPAT in *L. mexicana* and not in *T. congolense*. The sequences of *T. congolense* DPCK (Gene bank accession number: CCC90800.1) and PPAT (Gene bank accession number: CCC96024.1) were obtained from NCBI database.

The pairwise protein sequence alignment of the *T. congolense* DPCK enzyme against that of *L. mexicana* (**Figure 35**) showed amino acid variations in the CoA binding domain in 11 positions which include, 4 non-conservative mutations (P90F, I96T, T126L and T132I), 5 conservative mutations (L65I, V74I, T130S, K131D and F133Y) and 2 semi-conservative mutations (A32C and G85R). The pairwise protein sequence alignment of the *T. congolense* PPAT enzyme against that of *L. mexicana* (**Figure 36**) showed amino acid variations at four positions on the active site namely E265D, I266R, I321V and S88A while the rest of the amino acids that interacted with VMS were conserved between the two sequences.

```
L.mexicana     1 MILIGLTGGIACGKSSVSRILRDEFHIEVIDADLVVRELQTPNSACTRRI    50
                 |:|:||||||||||||:||.||:...||.|:|.|.:||.||.|.|||.|||
```

```
T.congolense      1 MLLVGLTGGIACGKSTVSTILQGRHHITVVDCDKLVRNLQQPFSACARRI      50

L.mexicana       51 AARWPLCVHPETGELNRAELGKVVFSDARARRELGKVMNPAIFKAILKRI     100
                    |.|||.||:|.:||::||.||.::|.|..|||:|.::|..||.|.:|.:
T.congolense     51 ARRWPQCVNPLSGEIDRAALGGIIFGDPIARRDLARIMNFPIFCATMKLL     100

L.mexicana      101 AAAWWRDLWRSGAVSSPSIVVLDAPTLFETKTFTYFISASVVVSCSEQRQ     150
                    ...||..|.|......|.:||||||.|:|...:|:.:...:||:|.|::|
T.congolense    101 LGLWWESLCRQLKGGEPLLVVLDAPLLYESNIYTWIVDCVMVVACREEQQ     150

L.mexicana      151 IERLRSRDGFSREAALQRIGSQMPLEAKCRLADYIIENDCADDLDALRGG     200
                    :||:..|:|.:||.|:||.:.|||:..||:.||.:|.|:|  .|..|...
T.congolense    151 VERIMKRNGLNREQAVQRVSAQMPISEKCKRADQVIFNEC--PLSELEQL     198

L.mexicana      201 VCACVAWMSRQSNKRLTYMFGTVAVGAVGVAAAVGYACYRLLLA       244
                    |...|.||.:||.|:::|.:....|...:|.||...|..:|..:
T.congolense    199 VDDAVLWMRQQSGKQVTRILLATATAGIGFAAVTAYIVFRFFV-       241
```

Figure 35: Protein sequence alignment of *T. congolense* DPCK (CCC90800.1) and *L. mexicana* DPCK (LmxM.22.1530.1). The sequences had 46% sequence identity. The nucleotide binding GGXXXGK[T/S] motif is underlined. Amino acids predicted to interact with VMS on *L. mexicana* DPCK are highlighted.

```
L.mexicana        1 MTPLVLLSTSRGKDANAKSLVAYLTRSVLTPIATKAATAADSASTASESS      50
                    |.|| .|:||.|.:.|.|.|.|.::|  |:|            .||....|||
```

```
T.congolense      1 MHPL-RLNTSAGIENNRKLLFSHL--SLL-----------GSAKPVGESS      36

L.mexicana       51 KYRCSSTELQCGYGGGVDVYLAIHNQHRSTFLEHSVYLYSAALEVCPQLS     100
                    :            .:.|.|.:.:..|.|.|:|...||.|:|:..|:.:
T.congolense     37 E-------------PLHVQLLVDDGRRDTALQHIKALYDASLQHHPETA      72

L.mexicana      101 ISIVPVIGTAEAVACDTTNKEKTYAPAASSSPSARSKQQEQQLDGMQGGS     150
                    ::::|.       |:...:.|.|          ..|:...|
T.congolense     73 VTVIPF-----ALGRLSGNTE--------GGPNTEKK-------------      96

L.mexicana      151 SDVIELYDDAAAVLREWKLVDKAFMRDAAGFQPHYKYVAVGGTFDHFHSG     200
                                    .|..|...:.|...|.|.|.||.||:|||||.|:|
T.congolense     97 --------------NEMSLFTPSLMTHGAAFDPLYKSVALGGTFDRLHAG     132

L.mexicana      201 HKVLLSTAALNAMQKLRVGVTDASLLTQKRFAESLQSIELRMENVAQFLH     250
                    ||:||||||.|.|....:|:|||...||:.|...|:.:.:..::|.|.|::|:.
T.congolense    133 HKLLLSTALLYATHFVRIGVTLPPLLSTKAHADLIEPFDVRTEAVSRFVR     182

L.mexicana      251 KMRPDLELELAPISEISGGTKSIPDVEALVVSPETAKSLGIINEMRAANG     300
                    .:||||:::::|.|.:.|||....|.:|||||||.||..:|..|||.| .:.
T.congolense    183 LLRPDLDVDIAGIEDRSGGADQDPALEALVVSSETVGALSFINEAR-VSA     231

L.mexicana      301 GLAPMVGITIPQVESPTGES--ISSTALRECQTQAD-      334
                    |:.|:..:.:|.|.|..||.  :|||.||....:..
T.congolense    232 GMKPLEAVVVPYVGSREGEEGRVSSTDLRARGREGAR      268
```

Figure 36: Protein sequence alignment of *T. congolense* PPAT (CCC96024.1) and *L. mexicana* PPAT (LmxM.31.2070.1). The sequences had 33.97% sequence identity and 77% query cover. The HXGH motif that contacts the adenine ring of ATP is underlined. Amino acids predicted to interact with VMS on *L. mexicana* PPAT are highlighted.

4.4.2 Molecular docking analysis of *Trypanosoma congolense* dephospho-CoA kinase

The structure of *T. congolense* DPCK was predicted and refined as described in Section **4.3.6**.

Blind docking was performed using AutoDock Tools 1.5.7 and AutoDock Vina and the

output is shown in **Table 16** and **Figure 37** below.

Compared to *L. mexicana*, the visualisation on **Figure 37** shows that VMS would interact with a different binding pocket in the *T. congolense* enzyme, which could indicate how the apparent selectivity comes about based on the metabolomics data.

Table 16: Binding affinity of VMS on the predicted *T. congolense* DPCK

Mode	DPCK VMS affinity (kcal/mol)
1	-6.5
2	-6.4
3	-6.3
4	-6.2
5	-6.1
6	-6.0
7	-5.8
8	-5.8
9	-5.6
10	-5.5
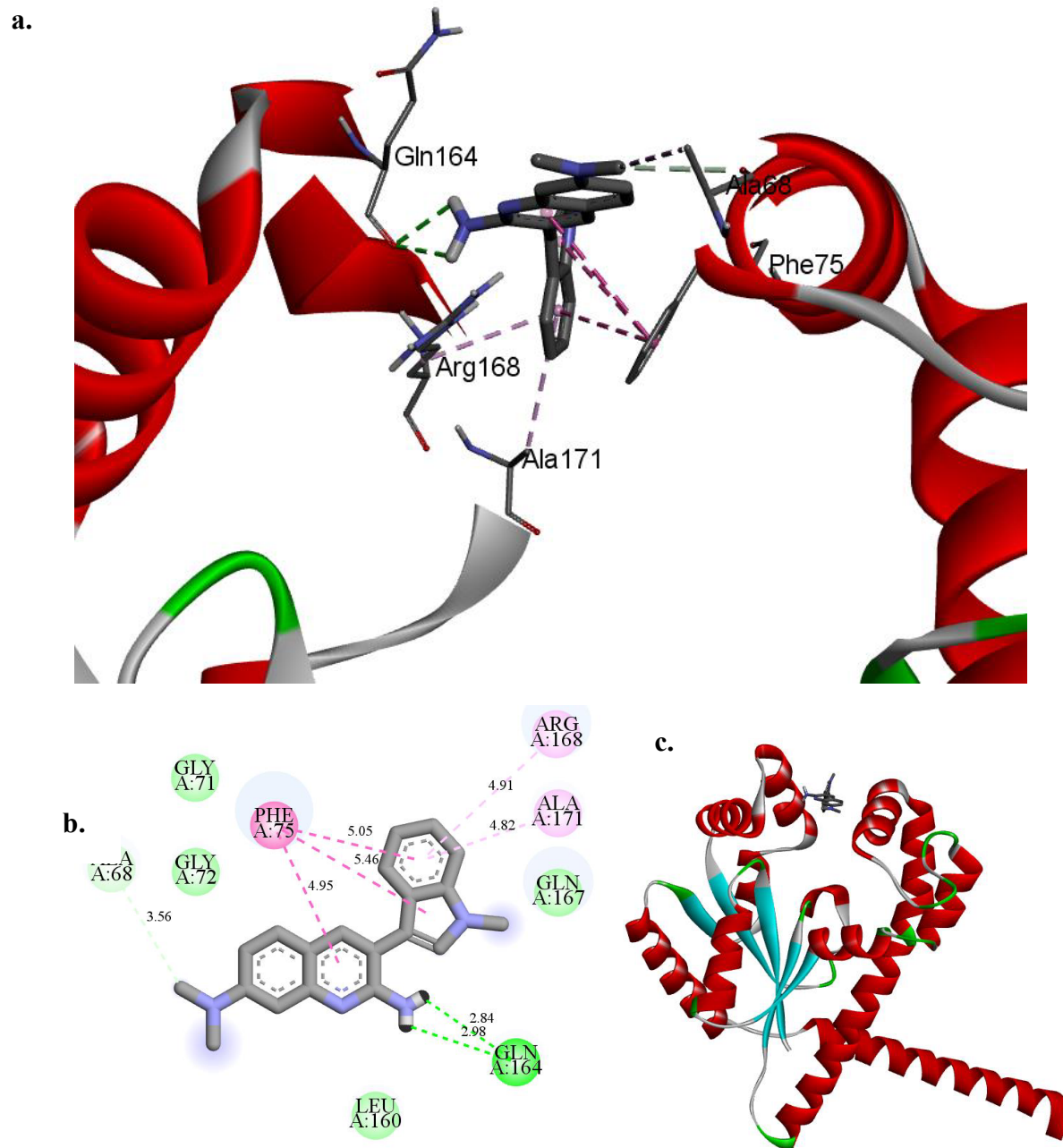
Figure 37: Interactions of VMS with the *T. congolense* DPCK enzyme. VMS interacted with Ala 68 via carbon-hydrogen bond, Phe 75 via pi-pi stacked bond, Ala 168 and Ala 171 via pi-pi T shaped bond, Gln 164 via hydrogen bond (2.84 A, 2.98 A) and Gly 71, Gly 72, Gln 167 and Leu 160 via Van der waals forces. (**a**) shows 3D visualisation of interactions between residues in the binding pocket and the ligand compounds. Interacting residues are also depicted in stick format and bonds are shown as color-coded dashed lines (**b**) shows a 2D diagram of the ligand interactions with their corresponding distances. The ligand is shown as a stick figure while the interacting residues are shown as color-coded balls, while (**c**) shows the positioning of ligand within the protein.

4.4 Pharmacokinetics and pharmacodynamics properties of JYH and VMS

In Sections **4.2** and **4.3** above, consideration is made as to how the JYH and VMS compounds

might inhibit CoA synthesis in Leishmania, and probably different targets in *T. congolense*.

Bioinformatics can also provide key data as to the possible suitability of compounds to

progress as drugs based on pharmacokinetic parameters. The results of computing the

physicochemical descriptors, ADME parameters, pharmacokinetic characteristics, drug-like

nature, and medicinal chemistry friendliness of JYH and VMS are provided in the **Table 17**

below.

Table 17: Pharmacokinetics and pharmacodynamics properties of JYH and VMS

	JYH	VMS
Physicochemical properties		

Canonical SMILES	CN1CCN(CC1)c1cc(Cl)c(cc1[N+](=O)[O-])C(=O)Nc1cc(ccc1C)c1nc2c(s1)cccc2	Nc1nc2cc(ccc2cc1c1cn(c2c1cccc2)C)N(=C)=C
Chemical formula	C25H23ClN5O3S	C20H18N4
Molecular weight	522.02	314.38
Number of heavy atoms	36	24
Number of aromatic heavy atoms	21	19
Fraction Csp3	0.23	0.05
Number of rotatable bonds	6	2
Number of H-bond acceptors	5	1
Number of H-bond donors	1	1
Molar refractivity	154.33	104.67
TPSA	122.53	55.52
Pharmacokinetics		
GI absorption	Low	High
BBB permeant	No	Yes
Pgp substrate	No	Yes
CYP1A2 inhibitor	No	Yes
CYP2C19 inhibitor	Yes	Yes
CYP2C9 inhibitor	Yes	No
CYP2D6 inhibitor	No	Yes
CYP3A4 inhibitor	Yes	Yes
log Kp (cm/s)	-5.49	-4.45
Drug-likeness		
Lipinski violations	1	0
Ghose violations	2	0
Veber violations	0	0
Egan violations	0	0
Muegge violations	1	1
Bioavailability Score	0.55	0.55
Medicinal chemistry		
PAINS alerts	0	0
Brenk alerts	2	1
Leadlikeness violations	2	1
Synthetic accessibility	3.78	2.9

Key

TPSA   Topological Polar Surface Area

BBB   Blood-Brain barrier

GI   Gastrointestinal

Pgp substrate   P-glycoprotein substrate

CYP   Cytochrome P450

Log Kp   Skin permeation coefficient

SMILES   Simplified Molecular Input Line Entry System

The SwissADME platform was used to evaluate the probability of JYH and VMS to possess drug-likeness. Drug-likeness is a qualitative evaluation of therapeutic candidate compounds to qualify as drugs with respect to oral availability (Halder & Elma, 2020). The Lipinski rule of 5 directs that the best drug candidates should demonstrate a molecular weight of ≤ 500, number of hydrogen bond donors ≤ 5, number of hydrogen bond acceptors ≤ 10, lipophilicity expressed as (logP) ≤ 5 and a molar refractivity of between 40 and 130 (Lipinski, 2000; Sarkar et al., 2019). In this respect, JYH had one violation while VMS had no violations. According to the Ghose filter (Ghose et al., 1999), the molecular weight, molar refractivity, and total number of atoms must all fall within certain ranges: computed log P must be between -0.4 and 5.6, molecular weight must be between 160 and 480, and the molar refractivity must be between 40 and 130. JYH violated two Ghose drug-likeliness rules (in terms of molecular weight which was 522.02 and molar refractivity that was 154.33). The Muegge filter instructs that therapeutic compounds should have molecular weight beween 200 and 600, log P values should range between -2 and 5, topological polar surface area (TPSA) ≤ 150, the number of rings ≤ 7, the number of carbons > 4, the number of heteroatoms > 1, the number of rotatable bonds ≤ 15, number of hydrogen bond acceptors ≤ 10 and hydrogen bond donors ≤ 5 (Muegge et al., 2001). Both JYH and VMS violated Muegge rules in terms of log P values where both compounds exhibited log P values > 5.

CHAPTER FIVE

5.0 DISCUSSION

Understanding how a drug impacts upon the metabolome of a parasite offers a path to understanding its mechanism(s) of action. These drug-metabolome interactions have been clarified by recent progress in discovering the mechanisms of action for anti-kinetoplastid medicines already in clinical use (Armitage et al., 2018; Vincent & Barrett, 2015) or in development pipelines. For isometamidium chloride, the mechanism of action has generally

been unclear but through the years, hypotheses of it affecting the kinetoplastid DNA, where the drug accumulates, have been proposed (Delespaux & de Koning, 2007). This study was set out to elucidate the mechanism of action of isometamidium chloride and two newly discovered compounds through hypothesis-generating untargeted metabolomics by examining the drug's interactions with *T. congolense* and *L. mexicana* parasites *in vitro* to uncover impacts on their metabolome.

5.1 Isometamidium chloride analysis

Results from statistical and metabolomic analysis revealed significant perturbations in metabolites belonging to nucleotide and energy metabolism pathways with decreases in tri-phosphorylated and di-phosphorylated nucleosides such as ATP, UTP, UDP, GDP, ADP. The analyses also showed increases in mono-phosphorylated and un-phosphorylated nucleosides such as AMP, dAMP, guanosine, adenosine, adenine and cytidine (**Figure 10, Figure 12, Table 3**). Adenosine triphosphate is critical for sustaining cell division and growth in all life forms and its reduction or depletion would arrest growth. In some *Trypanosoma* species such as *T. brucei* there is complete reliance on glucose for the parasite's energy needs with one molecule of glucose being catabolized to yield two molecules of ATP under aerobic conditions and one molecule of ATP under anaerobic conditions. Previous work by Steketee et al., (2021), pointed to *T. congolense* requiring lower concentrations (2mM compared to *T. brucei's* 5mM) of glucose to maintain growth, morphology and motility although glucose was vital for the parasite's energy needs and survival as complete deprivation or inhibition of the glycolytic flux led to cell death within 48h. In *Trypanosoma* parasites, glucose is shuttled into the cell via a transmembrane glucose transporter (**Figure 14**) of which *T. congolense* possesses two isoforms; TcoHT1 in bloodstream and metacyclic forms and TcoHT2 in procyclic forms. These two isoforms are 92.4% similar (Vedrenne et

al., 2000). Glucose then enters a membrane bound organelle, the glycosome, probably via a relatively non-specific pore at the organelle's membrane. The significant decrease of ATP as revealed by statistical analysis is indicative of interference with the glycolytic pathway that is well known to yield this metabolite. As evidenced by further analysis, there were indeed noteworthy decreases in D-glucose and other metabolites of the glycolytic pathway, followed by a decrease in metabolites of the TCA cycle and the acetate: succinate CoA-transferase succinyl-CoA synthetase (ASCT/SCS) cycle. In the ASCT/SCS cycle, pyruvate – the end product of glycolysis – is catabolized to succinate, alanine and acetate. Further breakdown of ASCT/SCS cycle metabolites has been hypothesized to play a role in mitochondrial ATP production in the related parasite, *T. brucei* (Michels et al., 2021). Data from a previous study suggested that *T. congolense* does not rely on oxidative phosphorylation for ATP production (Steketee et al., 2021). Since the main entry point of glycolytic flux is through the TcoHT1 hexose transporter, it is possible that isometamidium chloride may impact the uptake of glucose into the parasite by inhibiting its transporter.

According to a previous study by Girgis-Takla & James (1974), observations that isometamidium chloride uptake into *T. brucei* parasite increased with complete glucose privation of the parasites. This observation was attributed to cellular damage that provided access of the parasite's kinetoplast to isometamidium chloride. Findings from this study suggest it is possible that the privation could possibly allow more uptake of isometamidium chloride. If the drug enters via this particular transporter, lack of glucose to compete for binding to the transporter can allow more access for isometamidium chloride binding at the transporter's active site.

The observed pattern, where decrease in triphosphate and diphosphate nucleosides with a corresponding increase in their monophosphates could be due to a hypothetical nucleoside

phosphorelay system (**Figure 13**) (Appanna et al., 2016; Miranda et al., 2022). With the inhibition of the transporter – hence reduced glucose available for ATP formation through glycolysis, TCA cycle and respiratory chain – it is possible that the parasites resort to other NTPs to transfer phosphate ultimately to ADP creating ATP for their energy needs. In that case when the NTPs are used up, the parasites could have resorted to NDPs for energy and subsequently NMPs. Since these measurements were made over a short period of time, this cycle is evidenced by the decrease in NTPs and NDPs and increase in NMPs, non-phosphorylated nucleosides and nucleobases.

The largest binding pockets predicted by DoGsite scorer, when aligned to the *Homo sapiens* glucose transporter 1 (GLUT1, SLC2A1), revealed common residues defining the substrate binding site (Galochkina et al., 2019) (**Appendix 1**).

The top ranked ligand pose of isometamidium chloride bound to TcoHT1 exhibited a binding affinity of -10.1 kcal/mol while the top ranked ligand pose of glucose exhibited -5.7kcal/mol (**Table 5**). The degree of interaction between a ligand and a receptor is dependent on binding affinity. The stronger the ligand-receptor interaction, which shows a more accurate the molecular docking prediction, is exhibited by a more negative binding affinity (Pantsar & Poso, 2018). Owing to the lower energy of binding displayed by isometamidium chloride it is possible that isometamidium chloride has a higher competitive affinity for TcoHT1 compared to glucose and occupies the same binding site that the endogenous substrate glucose occupies (**Figure 18** and **Figure 19**). Isometamidium chloride formed three hydrogen bonds with TcoHT1 and additional hydrophobic interactions (**Figure 16**). Both isometamidium chloride and glucose interacted with Gln 216 (**Figure 16** and **17**) through hydrogen bonding. This particular Gln 216 (Gln 161 in GLUT1) is conserved in 28 members of the sugar transporter superfamily and in all 20 members that transport hexose and pentose

sugars i.e. glucose, galactose, xylose, arabinose and fructose (Galochkina et al., 2019; Mueckler et al., 1994). The conservation of this residue strongly suggests that it plays an important role in transport of hexose sugars and formation of the exofacial binding site (Mueckler et al., 1994). Substitution of this residue with leucine or asparagine in a previous study indicates reduced transport activity by 50 and 10 fold respectively (Mueckler et al., 1994). As isometamidium chloride and glucose both interact with this residue it is possible that once isometamidium chloride occupies the substrate binding site of TcoHT1 and forms a strong hydrogen bond it blocks glucose access and transport into the cell. Trp388 (Phe 420), Trp412 (Phe 444), and Phe379 (Phe 411) play crucial roles in GLUT1's ability to transport glucose (Almahmoud et al., 2019). Particularly, Trp388 (Phe 420) plays an important role in the access of GLUT1 between the outward open conformation and the inward open conformations and mutations such as W388L reduced the activity of glucose influx. A study by Almahmoud et al., (2019) showed most GLUT1 inhibitors interacted with Asn 34 (Ala 57), Gln 37 (Leu 60), Gln 161 (Gln 216), Gln 172 (Ala 227), Gln 282 (Leu 313), Gln 283 (Gln 314), Asn 288 (Asn 319), Tyr 292 (Met 322), Asn 317 (-), Phe 379 (Phe 411), Glu 380 (Gln 412) in the outward open conformation and a further Trp 388 (Phe 420), Trp 412 (Phe 444) in the inward open conformation.

Among the other kinetoplastids the *T. vivax* glucose transporter can form similar isometamidium chloride interactions as *T. congolense* with a hydrogen bond at Gln 215 (equivalent to Gln 216 in *T. congolense*) (**Appendix 2**) suggesting that the drug may work in a similar fashion in this parasite. The hydrogen bond between the Gln 215 and the *T. vivax* glucose transporter is, however, larger in distance (2.62 A˚) compared to the bond with TcoHT1 that is (2.11 A˚) showing less tight binding at this residue. While in *T. brucei*, isometamidium chloride seems to sit at the same glucose substrate binding site, it does not

form a hydrogen bond with Gln 215 (Gln 216), instead predicted to bind to Gln 244 (equivalent to Gln 443 in *T. congolense* and Asn 411 in GLUT1) (**Appendix 3**). This residue is also involved in glucose transport according to GLUT1 studies (Galochkina et al., 2019) and its substitution in other glucose transporters reduces the binding affinity of glucose to the transporters.

It is possible that isometamidium chloride is not entirely selective to the parasite glucose transporters as the drug seems to bind to some residues of the glucose binding site in animals when docking simulations were performed (**Appendix 4**). In the *Capra hircus* GLUT1 model (GenBank accession number: NP_001301152.1), both isometamidium chloride and glucose interact with Gln 161 (Gln 161 in human GLUT1 and Gln 216 in *T. congolense*). The docking simulation showed ISM exhibiting a binding affinity of -12.5 kcal/mol in the *Capra hircus* model. In the *Bos taurus* model (GenBank accession number: NP_777027.1), although isometamidium chloride seemed to interact with some residues that define the substrate binding site, it did not form interactions with the residues known to bind glucose e.g. Gln 216, Gln 282, Gln 283, Asn 288, Phe 291, Asn 317, Phe 379, Glu 380, Trp 388 (Galochkina et al., 2019; Ung et al., 2016). Therefore, in spite of ISM's ability to bind to both host and trypanosome glucose transporters, anti-parasite specificity could be due to the parasites having a greater dependency on glucose for energy production (Steketee et al., 2021). This could also be because of the drug having additional modes of action, such as binding to the kinetoplast and preventing its replication (Delespaux & de Koning, 2007).

With respect to resistance, a study on the genomic analysis of isometamidium chloride resistance did not identify changes in the glucose transporter gene. The study attributed resistance to a GAA insertion in a gene coding for an ATP-binding cassette (ABC) transporter (energy-dependent transport of different substrates across biological membranes)

in a variety of ISM resistant parasites (Tihon et al., 2017). This suggests that while the uptake of the drug may be through the transmembrane glucose transporter, drug efflux mechanisms are involved in cases of resistance.

5.2 JYH and VMS analysis

For the *Leishmania mexicana* dataset, statistical analyses using one-way ANOVA with Tukey's *post hoc* test (**Table 9**), fold change analysis (**Table 10**) as well as pathway enrichment analysis (**Figure 28**) revealed significant perturbations in metabolites belonging to lipid metabolism pathways exemplified by phospholipids, lysophospholipids, sphingolipids and metabolites with CoA moieties. Acetyl-CoA, CoA and 3-methylbutanoyl-CoA were seen to be depleted. Further metabolomic analysis from the raw dataset showed an interplay in increases in lysophospholipids with a decrease in phospholipids. This interplay is shown in the Land's cycle where acyl CoA moieties are transferred to lysophospholipids mediated by lysophospholipid acyltransferases to yield phospholipids. In turn, the acyl CoA moieties are cleaved from phospholipids by phospholipase A 2 to yield lysophospholipids (O'Donnell, 2022). Since CoA was being depleted it is possible that this led to the depletion of acyl-CoAs in the parasite which would then not be transferred to lysophospholipids to yield phospholipids. Precursors of CoA biosynthesis were also seen to be increasing while CoA itself was being depleted showing alteration in this pathway at the point of DPCK or PPAT, since the metabolite dephospho-CoA was not detected in the dataset, as shown in **Figure 29**. CoA is a cofactor that is essential for cell survival and growth as it is involved in many metabolic processes, including the biosynthesis of phospholipids, synthesis and degradation of fatty acids, and the operation of the TCA cycle (Leonardi & Jackowski, 2007), Trypanocyc (http://vm-trypanocyc.toulouse.inra.fr/). There are various registered drugs and potential inhibitors against Leishmaniasis that target lipid and fatty acid

metabolisms such as amphotericin B, sodium stibogluconate, miltefosine, imipramine, myriocin and tamoxifen (de Aquino et al., 2021) and the compounds from Kip Guy and Scott Landfear, coded as JYH and VMS are potential additions to this list.

Dephospho-CoA Kinase (DPCK) catalyses the last step of Coenzyme A biosynthesis while phosphopantetheine adenylyltransferase (PPAT) catalyses the penultimate step. Unlike in mammalian cells where these two steps are catalyzed by a 62-kD bifunctional enzyme, in Leishmania parasites the steps are carried out by distinct enzymes (Opperdoes & Michels, 2008). Other protozoa, e.g. *Entamoeba histolytica* are known to contain more than one DPCK related enzyme (Nurkanto et al., 2018) and in *L. mexicana* the genes that encode these enzymes are annotated as LmxM.22.1530 and LmxM.18.0290 in the TriTryp database (https://tritrypdb.org). The structure predicted by AlphaFold2 (**Figure 30**) was characteristic of other DPCK crystal structures from *Escherichia coli* (O'toole et al., 2003), *H. influenzae* (Obmolova et al., 2001) and predicted structures of *Entamoeba histolytica* (Nurkanto et al., 2018). The DPCK enzyme and its homologs contain the highly conserved P-loop also known as the Walker A sequence motif. The motif is encoded by the sequence GXXXXGKT/S (where X is any residue) responsible for nucleotide binding. The enzyme also contains a LID domain that forms a lid over the active site during catalysis to prevent transfer of the phosphate group to water and a CoA binding domain known to bind CoA during feedback inhibition (**Figure 30**) (O'toole et al., 2003). In this study, JYH and VMS seemed to interact with the CoA binding domain within the substrate binding site (**Figure 31**) which in previous studies has been identified to be the deep cleft between the LID domain and the CoA domains opposite the P-loop (Nurkanto et al., 2018; O'toole et al., 2003; Obmolova et al., 2001). In similar studies with the *E. coli* DPCK key amino acids identified to be involved in binding of CoA are Thr 8, Asp 33, His 89, Pro 113, Leu 114 and Gln 159 (O'toole et al., 2003) while

in *E. histolytica* amino acids involved in CoA binding include Thr-10, Asp 35, Try 90, Ala 117, Leu 118 and Glu 163 (Nurkanto et al., 2018). These residues align to Thr 7, Asp 33, Asn 89, Pro 125, Thr 126 and Gln 172. Collectively JYH and VMS interacted with Ala 32, Asp 33, Val 36, Gln 40, Leu 65, Arg 67, Ala 68, Leu 70, Gly 71, Val 74, Phe 75, Leu 84, Met 88, Asn 89, Ile 92, Ala 124, Pro 125, Thr 126, Thr 130, Thr 132 and Phe 133 (**Figure 31** and **32**).

A previous study by Obmolova et al., (2001) that involved modelling dephospho-CoA into the *Haemophilus influenza* DPCK revealed the dephospho-CoA adenine base interacts with the hydrophobic environment created by the side chains of Ile 92, Val 112, Leu 114, Leu 120, and Met 96, in the cleft known to bind dephospho-CoA in other studies. The study by Obmolova et al., (2001) also indicated the possible formation of hydrogen bonds between the base and the His 89 and Asn 118 residues at the bottom of the binding site with the 3-hydroxyl group of the ribose unit of dephospho-CoA possibly interacting with Asp 33 via hydrogen bonding. The JYH benzothiazole group from this study's model (**Figure 32**) interacted with Ile 92, Ala 124, and Ala 32 which align with Ile 92, Val 112 and Asp 33 respectively in *H. influenza* indicating interactions with the same residues that the indigenous substrate dephospho-CoA interacts with. In the VMS model, similar interactions were formed between the compound and Ala 32 (4.79 A˚), Ala 124 (5.07 A˚) and Ile 92 (4.37 A˚) via Pi-alkyl, with Val 36 (4.90 A˚) and Ile 92 (4.61 A˚) via alkyl bonds, with Thr 132 via a pi-donor hydrogen bond (2.59 A˚), with Asn 89 via a conventional hydrogen bond (2.04 A˚) and with Phe 93, Pro 125, Ile 96, Thr 126, Thr 130, and Phe 133 via hydrophobic interactions (**Figure 31**). Molecular docking analysis with the native substrate dephospho-CoA showed interactions with Phe 133, Asp 33, Arg 67, Gln 40, Met 88, Phe 75, Thr 130, Arg 81, Asn 89, Thr 126, Pro 125, Thr 132, Val 36, Ile 92 (**Appendix 5**). The model 6ARI

(https://www.rcsb.org/structure/6ARI) depicting an inhibitor of the *E. coli* DPCK shows the N-(methylsulfonyl)-3-{[(thiophen-2-yl)sulfanyl]methyl}benzamide (BQV) inhibitor interacting with the same substrate binding site in the CoA binding domain. Furthermore, JYH and VMS had lower binding energies to DPCK (-7.9 kcal/mol and -7.4 kcal/mol respectively) compared to dephospho-CoA (-6.2 kcal/mol). This indicates higher affinity of the compounds to DPCK than dephospho-CoA (**Table 13**) as well as formation of a more stable complex when JYH and VMS bind to DPCK. It is therefore possible that the metabolic effects observed on the metabolome of *L. mexicana* parasites upon exposure to the two drugs might have been due to inhibition of the DPCK enzyme via competitive inhibition with dephospho-CoA. This is especially true because in further wet lab analyses when Coenzyme A was added to cells exposed to VMS they were rescued. Dephospho-Coenzyme A addition, however, could not rescue the cells. This is consistent with DPCK as the target since dephospho-CoA cannot be converted to CoA if DPCK is inhibited (Barrett, personal communication).

Molecular docking analysis was also done using the predicted PPAT 3D structure (**Figure 30**). The predicted structure was similar to PPAT crystal structures from *E. coli* 1H1T (Izard, 2003), 1QJC (Izard, 2002), *Archaeoglobus fulgidus* 3DO8 (https://www.rcsb.org/), and *Staphylococcus aureus* 4NAU (De Jonge et al., 2013) with a characteristic Rossmann fold that forms the canonical nucleotide binding site and the T/HXGH (where X is any residue) motif that plays an important role in stabilizing the pentaco-ordinate transition state of ATP (De Jonge et al., 2013). In this study, VMS was predicted to interact with Gly 191, Gly 192, Thr 193, Gly 200, Lys 229, Thr 285, Ser 322 and Ile 292 via hydrophobic interactions with Leu 204 (5.41 A˚) and His 201 (4.19 A˚) via alkyl interactions and with Thr 193 (2.58 A˚) via hydrogen bonding (**Figure 34**). The JYH compound interacted with Gly 191, Gly 192,

Thr 193, Gly 200, Leu 204, Lys 229, Leu 279, Val 281, Ser 282, Glu 284, Ile 292, Ser 322, and Ser 323 via hydrophobic interactions and His 201 via pi-pi stacked bond, with Thr 285 and His 201 via Pi-sigma bond, with Ile 263 via alkyl bonding, with Leu 204 via pi-alkyl bonding, with Ser 288 via pi-donor hydrogen bonding and with Thr 221, Ile 266 and Glu 265 via carbon hydrogen bonding (**Figure 33**). These interactions were similar to the ones formed by the native substrate ATP (**Appendix 6**). From NCBI CDD and InterPro analysis the active site on the *L. mexicana* PPAT was predicted to involve Gly 192, Thr 193, Thr 221 and Val 281 and the HXGH motif lied on His 198, Ile 199, Gly 200 and His 201 showing that the drugs interacted with both the active site and the ATP binding site represented by the HXGH motif. In PPATs fully conserved residues thought to bind CoA in CoaD found in *E. coli* are His 18, Lys 42, Arg 51, Arg 91, Asp 95 and Glu 99 (Geerlof et al., 1999) which align with His 201, Asp 222, Glu 243, Glu 247, Glu 257 respectively in *L. mexicana*. Besides His 201, these interactions were not observed in the JYH and VMS interactions therefore pointing to them interacting with the ATP binding active site. In a previous study by Izard, (2002) ATP bound to *E.coli* PPAT was shown to interact with Tyr 7, Gly 9, Thr 10, Phe 11, Gly 17, His 18, Ile 21, Gly 89, Arg 91, Glu 99, Pro 120, Trp 124, Ile 127, Ser 128, Ser 129 and Ser 130 which align to Val 190, Gly 192, Thr 193, Phe 194, Gly 200, His 201, Leu 204, Arg 241, Glu 243, Pro 283 and Lys 287. The VMS and JYH collectively interacted with 6 of these residues. The sequence used did not contain the characteristic KMSKS motif also known to form part of the enzyme active site (Ser 128-130) although this is mostly identified in class I amino-acyl tRNA synthetases where the motif is important in stabilization of the aminoacyl-adenylate. The top ranked predicted binding pose of JYH displayed a binding affinity of -9.7 kcal/mol and VMS had -8.2 kcal/mol. Compared to the native ATP whose top ranked ligand pose had a binding affinity of -6.7 kcal/mol (**Appendix 6**), JYH and VMS had higher binding

affinity for the PPAT enzyme. Findings from this study therefore indicate a possible binding of the antileishmanial compounds to the active site of the PPAT enzyme causing competitive inhibition with ATP. Further studies, however, will be required to prove this hypothesis. It is of note that the fact that modelling predicted potential interactions between VMS and JYH with both DPCK and PPAT, which could indicate that docking algorithms are relatively relaxed in how they enable ligands to bind to proteins. This emphasizes the necessity to follow up such predictions with experimentation e.g. purifying the Leishmanial proteins and testing whether activity is inhibited by drugs and crystallising the proteins for structural analysis including with drugs bound.

The mechanism of action of VMS appeared to differ between *L. mexicana* and *T. congolense* parasites (**Section 4.2**) based on metabolomics experiments. While a possible inhibition of either DPCK or PPAT was seen in *L. mexicana*, in *T. congolense* the compound seemed to alter energy metabolism, aromatic amino acid metabolism as well as cellular redox systems with no impact on CoA synthesis evident. Protein sequence alignment of the enzymes from the respective parasites (**Figure 35** and **36)** showed amino acid variations in the amino acids that form the substrate binding sites and active sites. In a pairwise DPCK sequence alignment, although 18 amino acids that form the substrate binding site and active site were conserved, amino acid variations occurred in CoA binding domain at 11 positions which include, 4 non-conservative mutations (P90F, I96T, T126L and T132I, 5 conservative mutations (L65I, V74I, T130S, K131D and F133Y) and 2 semi-conservative mutations (A32C and G85R). A multiple sequence alignment of other *Leishmania* and *Trypanosoma* species using Clustal Omega alignment program (https://www.ebi.ac.uk/Tools/msa/clustalo/) (**Appendix 7**) showed a semi-conservative A32S substitution in *T. vivax* and *T. b. brucei* as opposed to the A32C substitution. At positions 126 and 132 the hydrophilic threonines were substituted with

hydrophobic leucine and isoleucine respectively in the trypanosome species. Furthermore, a structure comparison and surface binding analysis of the *L. mexicana* and *T. congolense* enzymes showed differences in their conformations (**Appendix 8** and **Appendix 9**).

The results of this study could explain the different binding interactions observed in the molecular docking analysis of the compound into the *T. congolense* DPCK enzyme that were dissimilar to those seen in the *L. mexicana* (**Figure 37**). The VMS compound seemed to dock into a separate site on the enzyme that was neither the nucleotide binding site where ATP binds nor the CoA domain where dephospho-CoA binds. Given the possibility that docking is not always accurate, it is possible that the compounds do not bind effectively to the *T. congolense* enzyme if at all. Future work including enzyme assays in the presence of drug, crystallization and structural analysis and also site directed mutagenesis – for example, replacing the Cys 32 in the *T. congolense* enzyme with Ala and *vice versa* for the Leishmania enzyme then testing ability of the compound(s) to bind need to be conducted. *In silico* site directed mutagenesis performed at A32C, T126L and T132I on the predicted *L. mexicana* DPCK using UCSF Chimera version 1.15 however did not show alterations to how the compounds bound to the enzyme.

The DPCK and PPAT enzymes are suitable drug targets because of their low percent identity with the human counterpart. A protein sequence alignment between the *Homo sapiens* bifunctional CoA synthetase (AAM19996.1) and *L. mexicana* DPCK (XP_003875607.1) and *L. mexicana* PPAT (XP_003878055.1) showed a sequence similarity of 26.29% and 42.61% respectively. This < 50% alignment similarity makes the two proteins pharmacologically attractive because they present lower risk for adverse effects (dos Santos Vasconcelos & Rezende, 2021).

Upon assessing the pharmacological suitability of the VMS and JYH, pharmacokinetics and pharmacodynamics analysis showed possible suitability of VMS, but not of JYH (**Table 17**). In particular, a compound likely to be active and permeable should have < 5 hydrogen bond donors, < 10 hydrogen bond acceptors, < 500 molecular weight and a calculated logP value of < 5 (Benet et al., 2016). The JYH compound had violations in terms of molecular weight and may therefore require modification to create a more drug-like derivative.

CHAPTER SIX

CONCLUSION AND RECOMMENDATIONS

6.1 Conclusion

From this study, untargeted metabolomics and *in silico* modeling hypothesized the potential drug targets of the widely used isometamidium chloride against Animal African Trypanosomiasis (AAT) and the two recently discovered antileishmanials discovered by the groups of Kip Guy and Scott Landfear known as JYH and VMS. Statistical and metabolomic analysis of the drugs' perturbations on the parasites' metabolomes identified alteration in the glycolysis and energy metabolism pathways when isometamidium chloride was administered to *Trypanosoma congolense* and in CoA biosynthesis when the new drugs were administered to *Leishmania mexicana*. Further metabolomic analysis identified possible competitive inhibition of the *T. congolense* glucose transporter in the case of isometamidium chloride which was further evidenced in the molecular docking assay where isometamidium chloride seemed to interact with the same substrate binding site that glucose interacts with. Isometamidium chloride also exhibited lower binding energy indicating higher affinity of isometamidium chloride for the TcoHT1 compared to glucose. For JYH and VMS, statistical and metabolomic analysis indicated possible inhibition of either dephospho-CoA kinase (DPCK) or phosphopantetheine adenylyltransferase (PPAT). Molecular docking analysis

showed possible competitive inhibition with dephospho-CoA, with JYH and VMS binding to the CoA binding domain of *L. mexicana* DPCK. The compounds also exhibited lower binding energies and higher binding affinities for the enzyme compared to dephospho-CoA. In PPAT both drugs seemed to interact with the enzyme active site that binds ATP. However, JYH had higher binding affinity for the enzyme while VMS had lower binding affinity compared to the native substrate ATP. The mechanism of action of VMS in *L. mexicana* differed with that of *T. congolense* where alterations were observed in energy metabolism, aromatic amino acid metabolism and cellular redox. Additional research is needed to find out which of the two, if not both, enzymes is targeted by the two compounds. This study has identified new targets in *L. mexicana* that have potential for drug development. This will hopefully lead to the discovery of other highly effective compounds with low resistance potential that can be used to treat Leishmaniasis for which current therapies often require lengthy hospitalization of at least 21 days (Hammill et al., 2021). Even when the oral miltefosine is used in combination with sodium stibogluconate, the length of hospitalization is only marginally reduced (Omollo et al., 2011). Therefore, the need for novel drugs is urgent.

6.2 Recommendations

From the findings in this study, it is recommended that:

i. Targeted metabolomics studies be conducted to confirm the results published in this study. The studies should target the affected metabolic pathways identified i.e. glycolysis and energy metabolism pathways for isometamidium chloride used against *T. congolense* and Coenzyme A biosynthesis pathways for JYH and VMS compounds used against *L. mexicana*.

ii.   The JYH structure should be modified because the compound violates drug-likeness and lead-likeness rules (**Table 17**). This is especially necessary if the drug is ever to enter *in vivo* and eventually human clinical trials.

iii.   The protein sequence used to predict the structure of PPAT is annotated as a cytidylyltransferase. To ascertain the possibility that it is indeed the *L. mexicana* PPAT, the protein should be purified after overexpression assay with substrates known to bind onto cytidylyltransferases and in the presence of ATP or CTP. Activity with the compounds should then be evaluated.

iv.   The 3-Dimensional structures of the proteins used in this study be determined using experimental methods such as X-ray crystallography or Nuclear Magnetic Resonance (NMR) studies to fully characterize their catalytic mechanism. This is because high-resolution experimentally determined structures maintain their active sites. Furthermore, JYH and VMS should be co-crystallised with the enzymes to determine whether they address the proteins as predicted.

v.   Enzyme assays for *L. mexicana* DPCK and PPAT should be carried out and JYH and VMS compounds tested for inhibitory activity.

vi.   Mutagenesis of key residues in the possible target enzymes should be performed to determine impact on activity and the ability of the compounds to inhibit activity.

REFERENCES

African animal trypanosomiasis; selected articles from the World Animal Review. (n.d.). Retrieved October 12, 2021, from https://www.fao.org/3/ah809e/ah809e02.htm

Agosin, M., & von Brand, T. (1954). Studies on the carbohydrate metabolism of *Trypanosoma congolense*. *Experimental Parasitology*, *3*(6), 517–524. https://doi.org/10.1016/0014-4894(54)90047-8

Akpunarlieva, S., Weidt, S., Lamasudin, D., Naula, C., Henderson, D., Barrett, M., Burgess, K., & Burchmore, R. (2017). Integration of proteomics and metabolomics to elucidate metabolic adaptation in Leishmania. *Journal of Proteomics*, *155*, 85–98. https://doi.org/10.1016/J.JPROT.2016.12.009

Alkhaldi, A. A. M., de Koning, H. P., & Bukhari, S. N. A. (2019). Antileishmanial and antitrypanosomal activity of symmetrical dibenzyl-substituted α,β-unsaturated carbonyl-based compounds. *Drug Design, Development and Therapy*, *13*, 1179–1185. https://doi.org/10.2147/DDDT.S204733

Almahmoud, S., Wang, X., Vennerstrom, J. L., & Zhong, H. A. (2019). Conformational Studies of Glucose Transporter 1 (GLUT1) as an Anticancer Drug Target. *Molecules 2019, Vol. 24, Page 2159*, *24*(11), 2159.

https://doi.org/10.3390/MOLECULES24112159

Amaro, R. E., Baudry, J., Chodera, J., Demir, Ö., McCammon, J. A., Miao, Y., & Smith, J. C. (2018). Ensemble Docking in Drug Discovery. *Biophysical Journal*, *114*(10), 2271–2278. https://doi.org/10.1016/J.BPJ.2018.02.038

Anderson, A. C., O'Neil, R. H., Surti, T. S., & Stroud, R. M. (2001). Approaches to solving the rigid receptor problem by identifying a minimal set of flexible residues during ligand docking. *Chemistry & Biology*, *8*(5), 445–457. https://doi.org/10.1016/S1074-5521(01)00023-0

Appanna, V. P., Alhasawi, A. A., Auger, C., Thomas, S. C., & Appanna, V. D. (2016). Phospho-transfer networks and ATP homeostasis in response to an ineffective electron transport chain in *Pseudomonas fluorescens*. *Archives of Biochemistry and Biophysics*, *606*, 26–33. https://doi.org/10.1016/J.ABB.2016.07.011

Armitage, E. G., Alqaisi, A. Q. I., Godzien, J., Peña, I., Mbekeani, A. J., Alonso-Herranz, V., López-Gonzálvez, Á., Martín, J., Gabarro, R., Denny, P. W., Barrett, M. P., & Barbas, C. (2018). Complex interplay between sphingolipid and sterol metabolism revealed by perturbations to the leishmania metabolome caused by miltefosine. *Antimicrobial Agents and Chemotherapy*, *62*(5), 1–12. https://doi.org/10.1128/AAC.02095-17

Ballante, F. (2018). Protein-ligand docking in drug design: Performance assessment and binding-pose selection. *Methods in Molecular Biology*, *1824*, 67–88. https://doi.org/10.1007/978-1-4939-8630-9_5/COVER

Barrett, M. P., & Croft, S. L. (2012). Management of trypanosomiasis and leishmaniasis. *British Medical Bulletin*, *104*(1), 175–196. https://doi.org/10.1093/BMB/LDS031

Barrett, M. P., Zhang, Z. Q., Denise, H., Giroud, C., & Baltz, T. (1995). A diamidine-resistant *Trypanosoma equiperdum* clone contains a P2 purine transporter with reduced substrate affinity. *Molecular and Biochemical Parasitology*, *73*(1–2), 223–229. https://doi.org/10.1016/0166-6851(95)00120-P

Bartel, J., Krumsiek, J., & Theis, F. J. (2013). Statistical methods for the analysis of high-throughput metabolomics data. *Computational and Structural Biotechnology Journal*, *4*(5), e201301009. https://doi.org/10.5936/CSBJ.201301009

Benet, L. Z., Hosey, C. M., Ursu, O., & Oprea, T. I. (2016). BDDCS, the Rule of 5 and drugability. *Advanced Drug Delivery Reviews*, *101*, 89–98. https://doi.org/10.1016/J.ADDR.2016.05.007

Blekherman, G., Laubenbacher, R., Cortes, D. F., Mendes, P., Torti, F. M., Akman, S., Torti, S. V., & Shulaev, V. (2011). Bioinformatics tools for cancer metabolomics. *Metabolomics*, *7*(3), 329–343. https://doi.org/10.1007/S11306-010-0270-3/TABLES/2

Bouton, J., Ferreira, D. A. F. L., Cardoso, S. C., Mazzarella, M. A., De Nazaré, C. S. M., Maes, L., Karalic, I., Caljon, G., & Van Calenbergh, S. (2021). Revisiting Pyrazolo[3,4-d]pyrimidine Nucleosides as Anti- *Trypanosoma cruzi* and Antileishmanial Agents. *Journal of Medicinal Chemistry*, *64*(7), 4206–4238. https://doi.org/10.1021/ACS.JMEDCHEM.1C00135/SUPPL_FILE/JM1C00135_SI_002.CSV

Brack, C., Delain, E., Riou, G., & Festy, B. (1972). Molecular organization of the kinetoplast DNA of *Trypanosoma cruzi* treated with berenil, a DNA interacting drug. *Journal of Ultrastructure Research*, *39*(5–6), 568–579. https://doi.org/10.1016/S0022-5320(72)90122-0

Bringaud, F., Rivière, L., & Coustou, V. (2006). Energy metabolism of trypanosomatids: Adaptation to available carbon sources. *Molecular and Biochemical Parasitology*, *149*(1), 1–9. https://doi.org/10.1016/J.MOLBIOPARA.2006.03.017

Cadena-Zamudio, J. D., Monribot-Villanueva, J. L., Pérez-Torres, C. A., Alatorre-Cobos, F., Jiménez-Moraila, B., Guerrero-Analco, J. A., & Ibarra-Laclette, E. (2022). The use of ecological analytical tools as an unconventional approach for untargeted metabolomics data analysis: the case of *Cecropia obtusifolia* and its adaptive responses to nitrate starvation. *Functional & Integrative Genomics 2022*, *1*, 1–27. https://doi.org/10.1007/S10142-022-00904-1

Carruthers, L. V., Munday, J. C., Ebiloma, G. U., Steketee, P., Jayaraman, S., Campagnaro, G. D., Ungogo, M. A., Lemgruber, L., Donachie, A. M., Rowan, T. G., Peter, R., Morrison, L. J., Barrett, M. P., & De Koning, H. P. (2021). Diminazene resistance in *Trypanosoma congolense* is not caused by reduced transport capacity but associated with reduced mitochondrial membrane potential. *Molecular Microbiology*, *116*(2), 564–588. https://doi.org/10.1111/MMI.14733

Cazzulo, F. M. C., Vernal, J., Cazzulo, J. J., & Nowicki, C. (1999). The NAD-linked aromatic α-hydroxy acid dehydrogenase from *Trypanosoma cruzi*. *European Journal of Biochemistry*, *266*(3), 903–910. https://doi.org/10.1046/J.1432-1327.1999.00926.X

CDC - Leishmaniasis - Biology. (2020). Retrieved May 2, 2023, from https://www.cdc.gov/parasites/leishmaniasis/biology.html

Chong, J., & Xia, J. (2018). MetaboAnalystR: an R package for flexible and reproducible analysis of metabolomics data. *Bioinformatics*, *34*(24), 4313–4314. https://doi.org/10.1093/BIOINFORMATICS/BTY528

Colasante, C., Ellis, M., Ruppert, T., & Voncken, F. (2006). Comparative proteomics of glycosomes from bloodstream form and procyclic culture form *Trypanosoma brucei brucei*. *Proteomics*, *6*(11), 3275–3293. https://doi.org/10.1002/PMIC.200500668

Costanzo, M., Caterino, M., Salvatori, I., Manganelli, V., Ferri, A., Misasi, R., & Ruoppolo, M. (2022). Proteome data of neuroblastoma cells overexpressing Neuroglobin. *Data in Brief*, *41*, 107843. https://doi.org/10.1016/J.DIB.2022.107843

Coustou, V., Guegan, F., Plazolles, N., & Baltz, T. (2010). Complete In Vitro Life Cycle of *Trypanosoma congolense*: Development of Genetic Tools. *PLOS Neglected Tropical Diseases*, *4*(3), e618. https://doi.org/10.1371/JOURNAL.PNTD.0000618

Creek, D. J., & Barrett, M. P. (2014). Determination of antiprotozoal drug mechanisms by metabolomics approaches. *Parasitology*, *141*(1), 83–92. https://doi.org/10.1017/S0031182013000814

Creek, D. J., Jankevics, A., Burgess, K. E. V., Breitling, R., & Barrett, M. P. (2012). IDEOM: an Excel interface for analysis of LC–MS-based metabolomics data. *Bioinformatics*, *28*(7), 1048–1049. https://doi.org/10.1093/BIOINFORMATICS/BTS069

Croft, S. L., Snowdon, D., & Yardley, V. (1996). The activities of four anticancer alkyllysophospholipids against *Leishmania donovani*, *Trypanosoma cruzi* and *Trypanosoma brucei*. *Journal of Antimicrobial Chemotherapy*, *38*(6), 1041–1047. https://doi.org/10.1093/JAC/38.6.1041

Cui, L., Lu, H., & Lee, Y. H. (2018). Challenges and emergent solutions for LC-MS/MS based untargeted metabolomics in diseases. *Mass Spectrometry Reviews*, *37*(6), 772–792. https://doi.org/10.1002/MAS.21562

Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports 2017 7:1*, *7*(1), 1–13. https://doi.org/10.1038/srep42717

Damianou, A., Burge, R. J., Catta-Preta, C. M. C., Geoghegan, V., Romina Nievas, Y., Newling, K., Brown, E., Burchmore, R., Rodenko, B., & Mottram, J. C. (2020). Essential roles for deubiquitination in Leishmania life cycle progression. *PLOS Pathogens*, *16*(6), e1008455. https://doi.org/10.1371/JOURNAL.PPAT.1008455

Das, N., Jena, P. K., & Pradhan, S. K. (2020). Arabinosyltransferase C enzyme of *Mycobacterium tuberculosis*, a potential drug target: An insight from molecular docking study. *Heliyon*, *6*(2). https://doi.org/10.1016/j.heliyon.2019.e02693

de Aquino, G. P., Gomes, M. A. M., Salinas, R. K., & Laranjeira-Silva, M. F. (2021). Lipid and fatty acid metabolism in trypanosomatids. *Microbial Cell*, *8*(11), 262. https://doi.org/10.15698/MIC2021.11.764

De Jonge, B. L. M., Walkup, G. K., Lahiri, S. D., Huynh, H., Neckermann, G., Utley, L., Nash, T. J., Brock, J., San Martin, M., Kutschke, A., Johnstone, M., Laganas, V., Hajec, L., Gu, R. F., Ni, H., Chen, B., Hutchings, K., Holt, E., McKinney, D., … Thresher, J. (2013). Discovery of inhibitors of 4'-phosphopantetheine adenylyltransferase (PPAT) to validate PPAT as a target for antibacterial therapy. *Antimicrobial Agents and Chemotherapy*, *57*(12), 6005–6015. https://doi.org/10.1128/AAC.01661-13/SUPPL_FILE/ZAC012132367SO1.PDF

De Koning, H. P., Anderson, L. F., Stewart, M., Burchmore, R. J. S., Wallace, L. J. M., & Barrett, M. P. (2004). The Trypanocide Diminazene Aceturate Is Accumulated Predominantly through the TbAT1 Purine Transporter: Additional Insights on

Diamidine Resistance in African Trypanosomes. *Antimicrobial Agents and Chemotherapy*, *48*(5), 1515–1519. https://doi.org/10.1128/AAC.48.5.1515-1519.2004

Delespaux, V., & de Koning, H. P. (2007). Drugs and drug resistance in African trypanosomiasis. *Drug Resistance Updates*, *10*(1–2), 30–50. https://doi.org/10.1016/j.drup.2007.02.004

Dorn, P. L., Noireau, F. C., Krafsur, E. S., Lanzaro, G. C., & Cornel, A. J. (2011). Genetics of major insect vectors. *Genetics and Evolution of Infectious Diseases*, 411–472. https://doi.org/10.1016/B978-0-12-384890-1.00015-7

dos Santos Vasconcelos, C. R., & Rezende, A. M. (2021). Systematic in silico Evaluation of *Leishmania spp.* Proteomes for Drug Discovery. *Frontiers in Chemistry*, *9*, 228. https://doi.org/10.3389/fchem.2021.607139

Ducruix, C., Vailhen, D., Werner, E., Fievet, J. B., Bourguignon, J., Tabet, J. C., Ezan, E., & Junot, C. (2008). Metabolomic investigation of the response of the model plant *Arabidopsis thaliana* to cadmium exposure: Evaluation of data pretreatment methods for further statistical analyses. *Chemometrics and Intelligent Laboratory Systems*, *91*(1), 67–77. https://doi.org/10.1016/J.CHEMOLAB.2007.08.002

Eyford, B. A., Sakurai, T., Smith, D., Loveless, B., Hertz-Fowler, C., Donelson, J. E., Inoue, N., & Pearson, T. W. (2011). Differential protein expression throughout the life cycle of *Trypanosoma congolense*, a major parasite of cattle in Africa. *Molecular and Biochemical Parasitology*, *177*(2), 116–125. https://doi.org/10.1016/J.MOLBIOPARA.2011.02.009

FAO. (2019). *The disease. Programme Against African Trypanosomosis (PAAT). Food and Agriculture Organization of the United Nations (FAO).* https://www.fao.org/paat/the-

programme/the-disease/en/

Ferreira, L. G., Dos Santos, R. N., Oliva, G., & Andricopulo, A. D. (2015). Molecular Docking and Structure-Based Drug Design Strategies. *Molecules 2015, Vol. 20, Pages 13384-13421*, *20*(7), 13384–13421. https://doi.org/10.3390/MOLECULES200713384

Ferreira, R. S., Pons, J. L., & Labesse, G. (2019). Insights into Substrate and Inhibitor Selectivity among Human GLUT Transporters through Comparative Modeling and Molecular Docking. *ACS Omega*, *4*(3), 4748–4760. https://doi.org/10.1021/ACSOMEGA.8B03447/SUPPL_FILE/AO8B03447_SI_001.PDF

Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., & Olson, A. J. (2016). Computational protein–ligand docking and virtual drug screening with the AutoDock suite. *Nature Protocols 2016 11:5*, *11*(5), 905–919. https://doi.org/10.1038/nprot.2016.051

Galochkina, T., Ng Fuk Chong, M., Challali, L., Abbar, S., & Etchebest, C. (2019). New insights into GluT1 mechanics during glucose transfer. *Scientific Reports 2019 9:1*, *9*(1), 1–14. https://doi.org/10.1038/s41598-018-37367-z

Geerlof, A., Lewendon, A., & Shaw, W. V. (1999). Purification and characterization of phosphopantetheine adenylyltransferase from *Escherichia coli*. *Journal of Biological Chemistry*, *274*(38), 27105–27111. https://doi.org/10.1074/jbc.274.38.27105

Geerts, S., Holmes, P. H., Eisler, M. C., & Diall, O. (2001). African bovine trypanosomiasis: the problem of drug resistance. *Trends in Parasitology*, *17*(1), 25–28. https://doi.org/10.1016/S1471-4922(00)01827-4

Gertsman, I., & Barshop, B. A. (2018). Promises and pitfalls of untargeted metabolomics. *Journal of Inherited Metabolic Disease 2018 41:3*, *41*(3), 355–366. https://doi.org/10.1007/S10545-017-0130-7

Ghose, A. K., Viswanadhan, V. N., & Wendoloski, J. J. (1999). A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *Journal of Combinatorial Chemistry*, *1*(1), 55–68. https://doi.org/10.1021/CC9800071/ASSET/IMAGES/MEDIUM/CC9800071N00001. GIF

Giordani, F., Morrison, L. J., Rowan, T. G., De Koning, H. P., & Barrett, M. P. (2016). The animal trypanosomiases and their chemotherapy: A review. *Parasitology*, *143*(14), 1862–1889. https://doi.org/10.1017/S0031182016001268

Girgis-Takla, P., & James, D. M. (1974). In Vitro Uptake of Isometamidium and Diminazene by *Trypanosoma brucei*. *Antimicrobial Agents and Chemotherapy*, *6*(3), 372–374. https://doi.org/10.1128/AAC.6.3.372

Gossage, S. M., Rogers, M. E., & Bates, P. A. (2003). Two separate growth phases during the development of Leishmania in sand flies: implications for understanding the life cycle. *International Journal for Parasitology*, *33*(10), 1027. https://doi.org/10.1016/S0020-7519(03)00142-5

Gromski, P. S., Xu, Y., Hollywood, K. A., Turner, M. L., & Goodacre, R. (2014). The influence of scaling metabolomics data on model classification accuracy. *Metabolomics 2014 11:3*, *11*(3), 684–695. https://doi.org/10.1007/S11306-014-0738-7

Halder, S. K., & Elma, F. (2020). In silico identification of novel chemical compounds with

anti-TB potential for the inhibition of InhA and EthR from *Mycobacterium tuberculosis*. *BioRxiv*, 2020.12.04.411967. https://doi.org/10.1101/2020.12.04.411967

Halouska, S., Fenton, R. J., Barletta, R. G., & Powers, R. (2011). Predicting the in Vivo Mechanism of Action for Drug Leads Using NMR Metabolomics. *ACS Chemical Biology*, *7*(1), 166–171. https://doi.org/10.1021/CB200348M

Hammill, J. T., Sviripa, V. M., Kril, L. M., Ortiz, D., Fargo, C. M., Kim, H. S., Chen, Y., Rector, J., Rice, A. L., Domagalska, M. A., Begley, K. L., Liu, C., Rangnekar, V. M., Dujardin, J. C., Watt, D. S., Landfear, S. M., & Guy, R. K. (2021). Amino-Substituted 3-Aryl- And 3-Heteroarylquinolines as Potential Antileishmanial Agents. *Journal of Medicinal Chemistry*, *64*(16), 12152–12162. https://doi.org/10.1021/ACS.JMEDCHEM.1C00813/SUPPL_FILE/JM1C00813_SI_0 01.CSV

Hargrove, J. W., Ouifki, R., Kajunguri, D., Vale, G. A., & Torr, S. J. (2012). Modeling the Control of Trypanosomiasis Using Trypanocides or Insecticide-Treated Livestock. *PLOS Neglected Tropical Diseases*, *6*(5), e1615. https://doi.org/10.1371/JOURNAL.PNTD.0001615

Hendriks, M. M. W. B., Eeuwijk, F. A. va., Jellema, R. H., Westerhuis, J. A., Reijmers, T. H., Hoefsloot, H. C. J., & Smilde, A. K. (2011). Data-processing strategies for metabolomics studies. *TrAC Trends in Analytical Chemistry*, *30*(10), 1685–1698. https://doi.org/10.1016/J.TRAC.2011.04.019

Heo, L., Park, H., & Seok, C. (2013). GalaxyRefine: protein structure refinement driven by side-chain repacking. *Nucleic Acids Research*, *41*(W1), W384–W388. https://doi.org/10.1093/NAR/GKT458

Hirumi, H., & Hirumi, K. (1984). Continuous cultivation of animal-infective bloodstream forms of an East African *Trypanosoma congolense* stock. *Annals of Tropical Medicine and Parasitology*, *78*(3), 327–330. https://doi.org/10.1080/00034983.1984.11811824

Hunter, S., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L., Duquenne, L., Finn, R. D., Gough, J., Haft, D., Hulo, N., Kahn, D., Kelly, E., Laugraud, A., Letunic, I., Lonsdale, D., … Yeats, C. (2009). InterPro: the integrative protein signature database. *Nucleic Acids Research*, *37*(suppl_1), D211–D215. https://doi.org/10.1093/NAR/GKN785

Hwang, J. Y., Smithson, D., Zhu, F., Holbrook, G., Connelly, M. C., Kaiser, M., Brun, R., & Guy, R. K. (2013). Optimization of chloronitrobenzamides (CNBs) as therapeutic leads for human African trypanosomiasis (HAT). *Journal of Medicinal Chemistry*, *56*(7), 2850–2860. https://doi.org/10.1021/JM301687P/SUPPL_FILE/JM301687P_SI_001.PDF

Ibezim, A., Obi, C. B., Oforkansi, N. M., Mbah, C. J., & Nwodo, N. J. (2018). Discovery of Trypanocidal Bioactive Leads by Docking Study, Molecular Dynamic Simulation and In Vivo Screening. *ChemistrySelect*, *3*(8), 2386–2389. https://doi.org/10.1002/SLCT.201702972

Ishak, N. A., Tahir, N. I., Mohd Sa'id, S. N., Gopal, K., Othman, A., & Ramli, U. S. (2021). Comparative analysis of statistical tools for oil palm phytochemical research. *Heliyon*, *7*(2), e06048. https://doi.org/10.1016/J.HELIYON.2021.E06048

Izard, T. (2002). The crystal structures of phosphopantetheine adenylyltransferase with bound substrates reveal the enzyme's catalytic mechanism. *Journal of Molecular Biology*, *315*(4), 487–495. https://doi.org/10.1006/JMBI.2001.5272

Izard, T. (2003). A novel adenylate binding site confers phosphopantetheine adenylyltransferase interactions with coenzyme A. *Journal of Bacteriology*, *185*(14), 4074–4080. https://doi.org/10.1128/JB.185.14.4074-4080.2003/ASSET/DA7F324E-516D-41E5-9A53-D99DFA88456A/ASSETS/GRAPHIC/JB1430116003.JPEG

Jahnke, W. (2007). Fragment-Based Approaches. *Comprehensive Medicinal Chemistry II*, *3*, 939–957. https://doi.org/10.1016/B0-08-045044-X/00114-0

Joshua, R. A., Obwolo, M. J., Bwangamoi, O., & Mandebvu, E. (1995). Resistance to diminazine aceturate by *Trypanosoma congolense* from cattle in the Zambezi Valley of Zimbabwe. *Veterinary Parasitology*, *60*(1–2), 1–6. https://doi.org/10.1016/0304-4017(94)00780-G

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature 2021 596:7873*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kaminsky, R., Schmid, C., & Lun, Z. R. (1997). Susceptibility of dyskinetoplastic *Trypanosoma evansi* and *T. equiperdum* to isometamidium chloride. *Parasitology Research*, *83*(8). https://doi.org/10.1007/s004360050346

Kamleh, A., Barrett, M. P., Wildridge, D., Burchmore, R. J. S., Scheltema, R. A., & Watson, D. G. (2008). Metabolomic profiling using Orbitrap Fourier transform mass spectrometry with hydrophilic interaction chromatography: a method with wide applicability to analysis of biomolecules. *Rapid Communications in Mass Spectrometry*,

22(12), 1912–1918. https://doi.org/10.1002/RCM.3564

Karaman, I. (2017). Preprocessing and Pretreatment of Metabolomics Data for Statistical Analysis. *Advances in Experimental Medicine and Biology*, *965*, 145–161. https://doi.org/10.1007/978-3-319-47656-8_6

Kasozi, K. I., MacLeod, E. T., Ntulume, I., & Welburn, S. C. (2022). An Update on African Trypanocide Pharmaceutics and Resistance. *Frontiers in Veterinary Science*, *9*, 828111. https://doi.org/10.3389/FVETS.2022.828111

Kaufer, A., Ellis, J., Stark, D., & Barratt, J. (2017). The evolution of trypanosomatid taxonomy. *Parasites & Vectors 2017 10:1*, *10*(1), 1–17. https://doi.org/10.1186/S13071-017-2204-7

Kostygov, A. Y., Karnkowska, A., Votýpka, J., Tashyreva, D., Maciszewski, K., Yurchenko, V., & Lukeš, J. (2021). Euglenozoa: taxonomy, diversity and ecology, symbioses and viruses. *Open Biology*, *11*(3). https://doi.org/10.1098/RSOB.200407

Kuriakose, S., Muleme, H. M., Onyilagha, C., Singh, R., Jia, P., & Uzonna, J. E. (2012). Diminazene Aceturate (Berenil) Modulates the Host Cellular and Inflammatory Responses to *Trypanosoma congolense* Infection. *PLOS ONE*, *7*(11), e48696. https://doi.org/10.1371/JOURNAL.PONE.0048696

Kuriakose, S., & Uzonna, J. E. (2014). Diminazene aceturate (Berenil), a new use for an old compound? *International Immunopharmacology*, *21*(2), 342–345. https://doi.org/10.1016/J.INTIMP.2014.05.027

Leonardi, R., & Jackowski, S. (2007). Biosynthesis of Pantothenic Acid and Coenzyme A. *EcoSal Plus*, *2*(2). https://doi.org/10.1128/ECOSALPLUS.3.6.3.4/ASSET/0787962E-

C619-46F8-BC7B-A476DBCBA345/ASSETS/GRAPHIC/3.6.3.4_FIG_002.GIF

Li, B., Tang, J., Yang, Q., Cui, X., Li, S., Chen, S., Cao, Q., Xue, W., Chen, N., & Zhu, F. (2016). Performance Evaluation and Online Realization of Data-driven Normalization Methods Used in LC/MS based Untargeted Metabolomics Analysis. *Scientific Reports 2016 6:1*, *6*(1), 1–13. https://doi.org/10.1038/srep38881

Lipinski, C. A. (2000). Drug-like properties and the causes of poor solubility and poor permeability. *Journal of Pharmacological and Toxicological Methods*, *44*(1), 235–249. https://doi.org/10.1016/S1056-8719(00)00107-6

Marchese, L., Nascimento, J. D. F., Damasceno, F. S., Bringaud, F., Michels, P. A. M., & Silber, A. M. (2018). The Uptake and Metabolism of Amino Acids, and Their Unique Role in the Biology of Pathogenic Trypanosomatids. *Pathogens 2018, Vol. 7, Page 36*, *7*(2), 36. https://doi.org/10.3390/PATHOGENS7020036

Mbekeani, A. J., Jones, R. S., Bassas Llorens, M., Elliot, J., Regnault, C., Barrett, M. P., Steele, J., Kebede, B., Wrigley, S. K., Evans, L., & Denny, P. W. (2019). Mining for natural product antileishmanials in a fungal extract library. *International Journal for Parasitology: Drugs and Drug Resistance*, *11*(May), 118–128. https://doi.org/10.1016/j.ijpddr.2019.05.003

McMasters, D. R. (2018). Knowledge-Based Approaches to Off-Target Screening. In *Methods in Enzymology* (Vol. 610, pp. 311–323). Academic Press Inc. https://doi.org/10.1016/bs.mie.2018.09.023

Meyer, A., Holt, H. R., Selby, R., & Guitian, J. (2016). Past and Ongoing Tsetse and Animal Trypanosomiasis Control Operations in Five African Countries: A Systematic Review. *PLOS Neglected Tropical Diseases*, *10*(12), e0005247.

https://doi.org/10.1371/JOURNAL.PNTD.0005247

Michels, P. A. M., Villafraz, O., Pineda, E., Alencar, M. B., Cáceres, A. J., Silber, A. M., & Bringaud, F. (2021). Carbohydrate metabolism in trypanosomatids: New insights revealing novel complexity, diversity and species-unique features. *Experimental Parasitology*, *224*. https://doi.org/10.1016/J.EXPPARA.2021.108102

Miranda, M. R., Sayé, M., Reigada, C., Galceran, F., Rengifo, M., Maciel, B. J., Digirolamo, F. A., & Pereira, C. A. (2022). Revisiting trypanosomatid nucleoside diphosphate kinases. *Memórias Do Instituto Oswaldo Cruz*, *116*, e210339. https://doi.org/10.1590/0074-02760210339

Morris, G. M., Goodsell, D. S., Huey, R., & Olson, A. J. (1996). Distributed automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4. *Journal of Computer-Aided Molecular Design 1996 10:4*, *10*(4), 293–304. https://doi.org/10.1007/BF00124499

Morris, G. M., Huey, R., & Olson, A. J. (2008). Using AutoDock for Ligand-Receptor Docking. *Current Protocols in Bioinformatics*, *24*(1), 8.14.1-8.14.40. https://doi.org/10.1002/0471250953.BI0814S24

Moti, Y., De Deken, R., Thys, E., Van Den Abbeele, J., Duchateau, L., & Delespaux, V. (2015). PCR and microsatellite analysis of diminazene aceturate resistance of bovine trypanosomes correlated to knowledge, attitude and practice of livestock keepers in South-Western Ethiopia. *Acta Tropica*, *146*, 45–52. https://doi.org/10.1016/J.ACTATROPICA.2015.02.015

Mueckler, M., Weng, W., & Kruse, M. (1994). Glutamine 161 of Glut1 glucose transporter is critical for transport activity and exofacial ligand binding. *Journal of Biological*

Chemistry, *269*(32), 20533–20538. https://doi.org/10.1016/S0021-9258(17)32026-4

Muegge, I., Heald, S. L., & Brittelli, D. (2001). Simple selection criteria for drug-like chemical matter. *Journal of Medicinal Chemistry*, *44*(12), 1841–1846. https://doi.org/10.1021/JM015507E/ASSET/IMAGES/MEDIUM/JM015507EN00001 .GIF

Muhanguzi, D., Okello, W. O., Kabasa, J. D., Waiswa, C., Welburn, S. C., & Shaw, A. P. M. (2015). Cost analysis of options for management of African Animal Trypanosomiasis using interventions targeted at cattle in Tororo District; south-eastern Uganda. *Parasites & Vectors 2015 8:1*, *8*(1), 1–9. https://doi.org/10.1186/S13071-015-0998-8

mzMatch/PeakML: metabolomics data analysis. (n.d.). Retrieved December 7, 2021, from http://mzmatch.sourceforge.net/ideom.php

Narduzzi, L., Dervilly, G., Marchand, A., Audran, M., Bizec, B. Le, & Buisson, C. (2020). Applying metabolomics to detect growth hormone administration in athletes: Proof of concept. *Drug Testing and Analysis*, *12*(7), 887–899. https://doi.org/10.1002/DTA.2798

Neuweger, H., Albaum, S. P., Dondrup, M., Persicke, M., Watt, T., Niehaus, K., Stoye, J., & Goesmann, A. (2008). MeltDB: a software platform for the analysis and integration of metabolomics experiment data. *Bioinformatics (Oxford, England)*, *24*(23), 2726–2732. https://doi.org/10.1093/BIOINFORMATICS/BTN452

Nurkanto, A., Jeelani, G., Yamamoto, T., Hishiki, T., Naito, Y., Suematsu, M., Hashimoto, T., & Nozaki, T. (2018). Biochemical, metabolomic, and genetic analyses of dephospho coenzyme a kinase involved in coenzyme a biosynthesis in the human enteric parasite *Entamoeba histolytica*. *Frontiers in Microbiology*, *9*(NOV), 2902. https://doi.org/10.3389/FMICB.2018.02902/BIBTEX

O'Donnell, V. B. (2022). New appreciation for an old pathway: the Lands Cycle moves into new arenas in health and disease. *Biochemical Society Transactions*, *50*(1), 1–11. https://doi.org/10.1042/BST20210579

O'toole, N., Joã˜, J., Barbosa, J. A. R. G., Li, Y., Hung, L.-W., Matte, A., & Cygler, M. (2003). Crystal structure of a trimeric form of dephosphocoenzyme A kinase from *Escherichia coli. Protein Science*, *12*(2), 327–336. https://doi.org/10.1110/PS.0227803

Obmolova, G., Teplyakov, A., Bonander, N., Eisenstein, E., Howard, A. J., & Gilliland, G. L. (2001). Crystal Structure of Dephospho-Coenzyme A Kinase from *Haemophilus influenzae. Journal of Structural Biology*, *136*(2), 119–125. https://doi.org/10.1006/JSBI.2001.4428

Ogungbe, I. V., & Setzer, W. N. (2009). Comparative molecular docking of antitrypanosomal natural products into multiple *Trypanosoma brucei* drug targets. *Molecules*, *14*(4). https://doi.org/10.3390/molecules14041513

Omollo, R., Alexander, N., Edwards, T., Khalil, E. A. G., Younis, B. M., Abuzaid, A. A., Wasunna, M., Njoroge, N., Kinoti, D., Kirigi, G., Dorlo, T. P. C., Ellis, S., Balasegaram, M., & Musa, A. M. (2011). Safety and Efficacy of miltefosine alone and in combination with sodium stibogluconate and liposomal amphotericin B for the treatment of primary visceral leishmaniasis in East Africa: study protocol for a randomized controlled trial. *Trials*, *12*, 166. https://doi.org/10.1186/1745-6215-12-166

Onyeyili, P., & Egwu, G. (1995). Chemotherapy of African trypanosomiasis: a historical perspective. *Protozoological Abstracts*, *Vol.19 No.*(0309–1287). https://doi.org/10.1079/cabireviews/19960801151

Opperdoes, F. R., & Michels, P. A. M. (2008). The metabolic repertoire of Leishmania and

implications   for   drug   discovery.   *Leishmania*,   123–158. https://doi.org/https://doi.org/10.21775/9781910190272

Ortiz, D., Armand Guiguemde, W., Hammill, J. T., Carrillo, A. K., Chen, Y., Connelly, M., Stalheim, K., Elya, C., Johnson, A., Min, J., Shelat, A., Smithson, D. C., Yang, L., Zhu, F., Kiplin Guy, R., & Landfear, S. M. (2017). Discovery of novel, orally bioavailable, antileishmanial compounds using phenotypic screening. *PLOS Neglected Tropical Diseases*, *11*(12), e0006157. https://doi.org/10.1371/JOURNAL.PNTD.0006157

Pakpahan, M. T., Rusmerryani, M., Kawaguchi, K., Saito, H., & Nagao, H. (2013). Evaluation of scoring functions for protein-ligand docking. *AIP Conference Proceedings*, *1518*(1), 645. https://doi.org/10.1063/1.4794652

Pantsar, T., & Poso, A. (2018). Binding Affinity via Docking: Fact and Fiction. *Molecules 2018, Vol. 23, Page 1899*, *23*(8), 1899. https://doi.org/10.3390/MOLECULES23081899

Papagiannaros, A., Bories, C., Demetzos, C., & Loiseau, P. M. (2005). Antileishmanial and trypanocidal activities of new miltefosine liposomal formulations. *Biomedicine & Pharmacotherapy*, *59*(10), 545–550. https://doi.org/10.1016/J.BIOPHA.2005.06.011

Patti, G. J. (2011). Separation strategies for untargeted metabolomics. *Journal of Separation Science*, *34*(24), 3460–3469. https://doi.org/10.1002/JSSC.201100532

Peacock, L., Cook, S., Ferris, V., Bailey, M., & Gibson, W. (2012). The life cycle of *Trypanosoma (Nannomonas) congolense* in the tsetse fly. *Parasites & Vectors 2012 5:1*, *5*(1), 1–13. https://doi.org/10.1186/1756-3305-5-109

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C.,

& Ferrin, T. E. (2004). UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/JCC.20084

Pinzi, L., & Rastelli, G. (2019). Molecular docking: Shifting paradigms in drug discovery. In *International Journal of Molecular Sciences* (Vol. 20, Issue 18). MDPI AG. https://doi.org/10.3390/ijms20184331

Prayag, K., Surve, D. H., Paul, A. T., Kumar, S., & Jindal, A. B. (2020). Nanotechnological interventions for treatment of trypanosomiasis in humans and animals. *Drug Delivery and Translational Research 2020 10:4*, *10*(4), 945–961. https://doi.org/10.1007/S13346-020-00764-X

Quiroga, R., & Villarreal, M. A. (2016). Vinardo: A Scoring Function Based on Autodock Vina Improves Scoring, Docking, and Virtual Screening. *PLOS ONE*, *11*(5), e0155183. https://doi.org/10.1371/JOURNAL.PONE.0155183

Ravindranath, P. A., Forli, S., Goodsell, D. S., Olson, A. J., & Sanner, M. F. (2015). AutoDockFR: Advances in Protein-Ligand Docking with Explicitly Specified Binding Site Flexibility. *PLOS Computational Biology*, *11*(12), e1004586. https://doi.org/10.1371/JOURNAL.PCBI.1004586

Richards, S., Morrison, L. J., Torr, S. J., Barrett, M. P., Manangwa, O., Mramba, F., & Auty, H. (2021). Pharma to farmer: field challenges of optimizing trypanocide use in African animal trypanosomiasis. *Trends in Parasitology*, *37*(9), 831–843. https://doi.org/10.1016/J.PT.2021.04.007

Rock, J., Garcia, D., Espino, O., Shetu, S. A., Chan-Bacab, M. J., Moo-Puc, R., Patel, N. B., Rivera, G., & Bandyopadhyay, D. (2021). Benzopyrazine-Based Small Molecule

Inhibitors As Trypanocidal and Leishmanicidal Agents: Green Synthesis, In Vitro, and In Silico Evaluations. *Frontiers in Chemistry*, *9*, 757. https://doi.org/10.3389/FCHEM.2021.725892/BIBTEX

Saccenti, E., Hoefsloot, H. C. J., Smilde, A. K., Westerhuis, J. A., & Hendriks, M. M. W. B. (2013). Reflections on univariate and multivariate analysis of metabolomics data. *Metabolomics 2013 10:3*, *10*(3), 361–374. https://doi.org/10.1007/S11306-013-0598-6

Santos, R., Ursu, O., Gaulton, A., Bento, A. P., Donadi, R. S., Bologa, C. G., Karlsson, A., Al-Lazikani, B., Hersey, A., Oprea, T. I., & Overington, J. P. (2016). A comprehensive map of molecular drug targets. *Nature Reviews Drug Discovery 2016 16:1*, *16*(1), 19–34. https://doi.org/10.1038/nrd.2016.230

Sarkar, B., Islam, S. S., Ullah, M. A., Hossain, S., Prottoy, M. N. I., Araf, Y., & Taniya, M. A. (2019). Computational Assessment and Pharmacological Property Breakdown of Eight Patented and Candidate Drugs against Four Intended Targets in Alzheimer's Disease. *Advances in Bioscience and Biotechnology*, *10*(11), 405–430. https://doi.org/10.4236/ABB.2019.1011030

Schofield, C. J., & Kabayo, J. P. (2008). Trypanosomiasis vector control in Africa and Latin America. *Parasites & Vectors 2008 1:1*, *1*(1), 1–7. https://doi.org/10.1186/1756-3305-1-24

Shereni, W., Neves, L., Argilés, R., Nyakupinda, L., & Cecchi, G. (2021). An atlas of tsetse and animal African trypanosomiasis in Zimbabwe. *Parasites and Vectors*, *14*(1), 1–10. https://doi.org/10.1186/S13071-020-04555-8/FIGURES/4

Singh, R., Mailloux, R. J., Puiseux-Dao, S., & Appanna, V. D. (2007). Oxidative stress evokes a metabolic adaptation that favors increased NADPH synthesis and decreased

NADH production in *Pseudomonas fluorescens*. *Journal of Bacteriology*, *189*(18), 6665–6675. https://doi.org/10.1128/JB.00555-07

Singh, S., Malik, B. K., & Sharma, D. K. (2006). Molecular drug targets and structure based drug design: A holistic approach. *Bioinformation*, *1*(8), 314–320. https://doi.org/10.6026/97320630001314

Spickler, A. R. D. (2003). *African Trypanosomiasis*. www.cfsph.iastate.edu

Stein, J., Mogk, S., Mudogo, C. N., Sommer, B. P., Scholze, M., Meiwes, A., Huber, M., Gray, A., & Duszenko, M. (2014). Drug development against sleeping sickness: old wine in new bottles? *Current Medicinal Chemistry*, *21*(15), 1713–1727. https://doi.org/10.2174/09298673206661311119121636

Steketee, P. C., Dickie, E. A., Iremonger, J., Crouch, K., Paxton, E., Jayaraman, S., Alfituri, O. A., Awuah-Mensah, G., Ritchie, R., Schnaufer, A., Rowan, T., de Koning, H. P., Gadelha, C., Wickstead, B., Barrett, M. P., & Morrison, L. J. (2021). Divergent metabolism between *Trypanosoma congolense* and *Trypanosoma brucei* results in differential sensitivity to metabolic inhibition. *PLoS Pathogens*, *17*(7). https://doi.org/10.1371/JOURNAL.PPAT.1009734

Stevenson, P., Sones, K. R., Gicheru, M. M., & Mwangi, E. K. (1995). Comparison of isometamidium chloride and homidium bromide as prophylactic drugs for trypanosomiasis in cattle at Nguruman, Kenya. *Acta Tropica*, *59*(2), 77–84. https://doi.org/10.1016/0001-706X(94)00080-K

Stuart, K., Brun, R., Croft, S., Fairlamb, A., Gürtler, R. E., McKerrow, J., Reed, S., & Tarleton, R. (2008). Kinetoplastids: related protozoan pathogens, different diseases. *The Journal of Clinical Investigation*, *118*(4), 1301–1310. https://doi.org/10.1172/JCI33945

Tanchuk, V. Y., Tanin, V. O., Vovk, A. I., & Poda, G. (2016). A New, Improved Hybrid Scoring Function for Molecular Docking and Scoring Based on AutoDock and AutoDock Vina. *Chemical Biology & Drug Design*, *87*(4), 618–625. https://doi.org/10.1111/CBDD.12697

Tihon, E., Imamura, H., Van den Broeck, F., Vermeiren, L., Dujardin, J. C., & Van Den Abbeele, J. (2017). Genomic analysis of Isometamidium Chloride resistance in *Trypanosoma congolense. International Journal for Parasitology: Drugs and Drug Resistance*, *7*(3), 350–361. https://doi.org/10.1016/J.IJPDDR.2017.10.002

Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, *31*(2), 455–461. https://doi.org/10.1002/JCC.21334

Ung, P. M. U., Song, W., Cheng, L., Zhao, X., Hu, H., Chen, L., & Schlessinger, A. (2016). Inhibitor Discovery for the Human GLUT1 from Homology Modeling and Virtual Screening. *ACS Chemical Biology*, *11*(7), 1908. https://doi.org/10.1021/ACSCHEMBIO.6B00304

van den Berg, R. A., Hoefsloot, H. C. J., Westerhuis, J. A., Smilde, A. K., & van der Werf, M. J. (2006). Centering, scaling, and transformations: Improving the biological information content of metabolomics data. *BMC Genomics*, *7*(1), 1–15. https://doi.org/10.1186/1471-2164-7-142

Vedrenne, C., Bringaud, F., Barrett, M. P., Tetaud, E., & Baltz, T. (2000). The structure–function relationship of functionally distinct but structurally similar hexose transporters from *Trypanosoma congolense. European Journal of Biochemistry*, *267*(15), 4850–

4860. https://doi.org/10.1046/J.1432-1327.2000.01543.X

Vinaixa, M., Samino, S., Saez, I., Duran, J., Guinovart, J. J., & Yanes, O. (2012). A Guideline to Univariate Statistical Analysis for LC/MS-Based Untargeted Metabolomics-Derived Data. *Metabolites 2012, Vol. 2, Pages 775-795*, *2*(4), 775–795. https://doi.org/10.3390/METABO2040775

Vincent, I. M., & Barrett, M. P. (2015). Metabolomic-based strategies for anti-parasite drug discovery. *Journal of Biomolecular Screening*, *20*(1), 44–55. https://doi.org/10.1177/1087057114551519

Vincent, I. M., Creek, D. J., Burgess, K., Woods, D. J., Burchmore, R. J. S., & Barrett, M. P. (2012). Untargeted Metabolomics Reveals a Lack Of Synergy between Nifurtimox and Eflornithine against *Trypanosoma brucei*. *PLOS Neglected Tropical Diseases*, *6*(5), e1618. https://doi.org/10.1371/JOURNAL.PNTD.0001618

Vincent, I. M., Creek, D., Watson, D. G., Kamleh, M. A., Woods, D. J., Wong, P. E., Burchmore, R. J. S., & Barrett, M. P. (2010). A molecular mechanism for eflornithine resistance in African trypanosomes. *PLoS Pathogens*, *6*(11). https://doi.org/10.1371/JOURNAL.PPAT.1001204

Vincent, I. M., Ehmann, D. E., Mills, S. D., Perros, M., & Barrett, M. P. (2016). Untargeted Metabolomics To Ascertain Antibiotic Modes of Action. *Antimicrobial Agents and Chemotherapy*, *60*(4), 2281–2291. https://doi.org/10.1128/AAC.02109-15

Volkamer, A., Griewel, A., Grombacher, T., & Rarey, M. (2010). Analyzing the Topology of Active Sites: On the Prediction of Pockets and Subpockets. *Journal of Chemical Information and Modeling*, *50*(11), 2041–2052. https://doi.org/10.1021/CI100241Y

Westrop, G. D., Williams, R. A. M., Wang, L., Zhang, T., Watson, D. G., Silva, A. M., & Coombs, G. H. (2015). Metabolomic Analyses of Leishmania Reveal Multiple Species Differences and Large Differences in Amino Acid Metabolism. *PLoS ONE*, *10*(9). https://doi.org/10.1371/JOURNAL.PONE.0136891

Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S., Fung, C., Nikolai, L., Lewis, M., Coutouly, M. A., Forsythe, I., Tang, P., Shrivastava, S., Jeroncic, K., Stothard, P., … Querengesser, L. (2007). HMDB: the Human Metabolome Database. *Nucleic Acids Research*, *35*(suppl_1), D521–D526. https://doi.org/10.1093/NAR/GKL923

Worley, B., & Powers, R. (2013). Multivariate Analysis in Metabolomics. *Current Metabolomics*, *1*(1), 92–107. https://doi.org/10.2174/2213235X11301010092

Wragg, W. R., Washbourn, K., Brown, K. N., & Hill, J. (1958). Metamidium: A new trypanocidal drug. *Nature*, *182*(4641), 1005–1006. https://doi.org/10.1038/1821005A0

Xia, J., Psychogios, N., Young, N., & Wishart, D. S. (2009). MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Research*, *37*(suppl_2), W652–W660. https://doi.org/10.1093/NAR/GKP356

Xia, J., & Wishart, D. S. (2016). Using MetaboAnalyst 3.0 for Comprehensive Metabolomics Data Analysis. *Current Protocols in Bioinformatics*, *55*(1), 14.10.1-14.10.91. https://doi.org/10.1002/CPBI.11

Xiong, G., Wu, Z., Yi, J., Fu, L., Yang, Z., Hsieh, C., Yin, M., Zeng, X., Wu, C., Lu, A., Chen, X., Hou, T., & Cao, D. (2021). ADMETlab 2.0: an integrated online platform for accurate and comprehensive predictions of ADMET properties. *Nucleic Acids Research*, *49*(W1), W5–W14. https://doi.org/10.1093/NAR/GKAB255

Yang, M., Derbyshire, M. K., Yamashita, R. A., & Marchler-Bauer, A. (2020). NCBI's Conserved Domain Database and Tools for Protein Domain Analysis. *Current Protocols in Bioinformatics*, *69*(1), e90. https://doi.org/10.1002/CPBI.90

Yang, S., Wenzler, T., Miller, P. N., Wu, H., Boykin, D. W., Brun, R., & Wang, M. Z. (2014). Pharmacokinetic Comparison To Determine the Mechanisms Underlying the Differential Efficacies of Cationic Diamidines against First- and Second-Stage Human African Trypanosomiasis. *Antimicrobial Agents and Chemotherapy*, *58*(7), 4064. https://doi.org/10.1128/AAC.02605-14

Yaro, M., Munyard, K. A., Stear, M. J., & Groth, D. M. (2016). Combatting African Animal Trypanosomiasis (AAT) in livestock: The potential role of trypanotolerance. *Veterinary Parasitology*, *225*, 43–52. https://doi.org/10.1016/J.VETPAR.2016.05.003

APPENDICES

Appendix 1: Multiple protein sequence alignment of glucose transporters
Protein sequence alignment of glucose transporters

Amino acid residues defining the glucose substrate binding site on Homo sapiens solute carrier 2 A1 (SLC2A1) and their corresponding alignments with other glucose transporters. *Bos taurus* GLUT1 (BTGLUT1), *Trypanosoma cruzi* hexose transporter 1 (TcrHT), *Trypanosoma brucei* hexose transporter 1, (THT1) and *Trypanosoma congolense* hexose transporter 1 (TcoHT1). Transmembrane domains of TcoHT1 are underlined and labelled and residues forming the substrate binding site on SLC2A1 are highlighted in yellow.

```
SLC2A1    ------------------------MEPS---SKKLTGRLMLAVGGAVLGSLQFGYNTG 31
BTGLUT1   ------------------------MEPT---SKKLTGRLMLAVGGAVLGSLQFGYNTG 31
TcrHT1    MPS---------KKQTDVSGGDRQPDETPTFCSLENLKVAQV-----QVVGGTLNGFSIG 46
THT1      MTERRDNVSHAPDAIEGPNDGAHAEDTSPGFFSLENLGVAQV-----QVVGGTLNGYVIG 55
TcoHT1    MSDAQEGRERAPDTVMVSPDE-VHEDNAPAFFSAENLGVVQV-----QVIGGTLYGFSIG 54
                                         *    ::* :       1  *:*.   *:  *

SLC2A1    VINAPQ--------------------K-----------VIEEFYNQTWVHRYGE---- 54
BTGLUT1   VINAPQ--------------------K-----------VIEEFYNQTWVQRYGE---- 54
TcrHT1    FVAVYAYFYLMSTDCSMYKKEVACNRVLNAECSWNGTRGECGWNGFTCFWGHGKDKTPCL 106
THT1      YVAVYLLLYLTATECKFTT-EVACGGAKIYGCKWSGTT--------CKFENPKCSE---- 102
TcoHT1    FVAVYLMLYEVSTNCSLAKAEASCVGMKNVMCVWSPDV--------NCTWANACDS---- 102
               :  .                                              ..

SLC2A1    -------------SILPTTLTTLWS-LSVAIFSVGGMIGSFSVGLFVNRFGRRNSMLMMN 100
BTGLUT1   -------------PIPPATLTTLWS-LSVAIFSVGGMIGSFSVGLFVNRFGRRNSMLMMN 100
TcrHT1    DDSRCKWVYSDEECKNPTGYSSSYNGIFAGAMIVGAMIGSIYAGQFAARFGHKVSFLIVG 166
THT1      ------GSDPSDSCKNEVAYTSVYSGIFACAMIVGSMVGSIIAGKCITTFGLKKSFIIVS 156
TcoHT1    ------I-RDPNVCLDKVGYDALHSGIFACSMIVGSMIGSIVTERFITMFGLKKSFLIVA 155
                  .      :    .  2 :  .  :  **.*:**: .      ** :  *::::

SLC2A1    LLAFVSAVLMGFSKLGKSFEMLILGRFIIGVYCGLTTGFVPMYVGEVSPTALRGALGTLH 160
BTGLUT1   LLAFVSAVLMGFSKLGKSFEMLILGRFIIGVYCGLTTGFVPMYVGEVSPTELRGALGTLH 160
TcrHT1    IVGVVSSVMYHVSSATNEFWVLCVGRLLIGVVLGLVNVACPMYVDQNAHPKFLHVDGVLF  226
THT1      ITCTIACVVVQVAIEYNNYYALCTGRVLIGLGVGILCSVCPMYVNENAHPKLCKMDGVLF 216
TcoHT1    IIGVVASAMNHIAVSTDEFWVLCPARLLMGLGLGILCSVCPMYVNENAHSKHRKVDGVMF 215
               :  3 ::..: .:  ..: *  .*.::*: *:     ****.: :      *..:
                                            4

SLC2A1    QLGIVVGILIAQVFGLDSIMGNK---D--LWPLLLSIIFIPALLQ---CIVLPFCPESPR 212
BTGLUT1   QLGIVVGILIAQVFGLDSIMGNQ---E--LWPLLLSVIFIPALLQ---CILLPFCPESPR 212
TcrHT1    QVFTTFGIMFAAAMGLAIGQSVNFDKDTKMDARMQGYCAFSTLLSVLMVALGIFLGESKT 286
THT1      QVFTTLGIMLAAMLGLILDKTGASKEEANMAGRLHVFSAVPLGLSVAMFLVGMFLRESTA 276
TcoHT1    QVFITFGIMLAALLGLALYYTVDYETNVAMVDRLHGFCAVSSILAVVMFIMGIFLRESKT 275
              *:  ..**::*  :**          :  :    :    .   *    6 :  *  **
                     5
```

```
SLC2A1    FLLINRNEENRAKSVLKKLRGTADVTHDLQEMKEESRQMMREKKVTILELFRSPAYRQPI 272
BTGLUT1   FLLINRNEENRAKSVLKKLRGTADVTRDLQEMKEESRQMMREKKVTILELFRSAAYRQPI 272
TcrHT1    KFTS------------GKHEDDGTALD-------------PNE------YSYLQMLGPL 314
THT1      TFSQ------------DDDGKADGGMD------------PNE------YGWGQMLWPL 304
TcoHT1    VVVC------------ENAGKTDGGLD------------PNE------YSWGEMMWPL 303
                .                        .      *              ::        :        *:

SLC2A1    LIAVVLQLSQQLSGINAVFYYSTSIFEKAGVQQPVYATIGSGIVNTAFTVV-SLFVVERA 331
BTGLUT1   LIAVVLQLSQQLSGINAVFYYSTSIFEKAGVQQPVYATIGSGIVNTAFTVV-SLFVVERA 331
TcrHT1    AMGLVTSGTLQLTGINAVMNYAPKIMGNLGMVPLVGNFV----V-MAWNFVTTLVSIPLA 369
THT1      FMGAVTAGTLQLTGINAVMNYAPKITENLGMDPSLGNFL----V-MAWNFVTSLAAIPLA 359
TcoHT1    FMGTMTAGTLQLTGINAVMNYAPKITENLGMDPPLGNFL----V-MMWNFVTALVAIPLA 358
                :.: :  :  **:*****: *: .*   : *:   :    :       *   :..* :* : *
                        7                                              8

SLC2A1    GRRTLH--LIGLAGMAGCAILMTIALALLEQLP---WMSYLSIVAIFGFVAFFEVGPGPI 386
BTGLUT1   GRRTLH--LIGLAGMAGCAVLMTIALALLERLP---WMSYLSIVAIFGFVAFFEVGPGPI 386
TcrHT1    RVLTMRQLFLGASLVASVSCLLLCGVPVYPGVADKNVKNGVAITGIAIFIAAFEIGLGPC 429
THT1      SRFTMRQMFITCSFVASCMCLFLCGIPVFPGVAEEKVKNGVATTGIALFIAAFEFGVGSC 419
TcoHT1    SRFTMRQMFITCSFIASCTCLFLCGIPVFPGVAEESVKNGVATTGIALFIAAFEFGVGSC 418
                 *::   ::   : .:*.    *:   .: :   :         . :: ..*  *:* **.* *
                       9                                          10

SLC2A1    PWFIVAELFSQGPRPAAIAVAGFSNWTSNFIVGMCFQYVEQLC-----------GPYVFI 435
BTGLUT1   PWFIVAELFSQGPRPAAIAVAGFSNWTSNFIVGMCFQYVEQLC-----------GPYVFI 435
TcrHT1    FFVLAQELFPRSFRPRGASFVLLTNFIFNVIINVCYPIATEGISGGPSGNQDKGQAVAFI 489
THT1      FFVLAQDLFPPSFRPKGGSFVVMMQFIFNILINLLYPITTEAISGGPTANQDKGQAVAFI 479
TcoHT1    FFVLAQDLFPPSFRPKGASFVVMMQFVFNILVNLLYPITTEAISGGASGNQDKGQAVSFI 478
                :.:. :** . ** . .:.. : ::  *.::.: :  . :                    **
                                           11

SLC2A1    IFTVLLVLFFIFTYFKVPETKGRTFDEIASGFRQGGASQSDKTPEELFHPLGADSQ---- 491
BTGLUT1   IFTVLLVLFFIFTYFKVPETKGRTFDEIASGFRQGGASQSDKTPEELFHPLGADSQ---- 491
TcrHT1    FFGCIGLVCFVLQVFFLYPWEESTPQN-HGDAN------EESALPERQSPIEVATPGNRQ 542
THT1      LFGLIGLICSVLQFFYLYPYDANQDHE-ND---HGGEPVEQKTYPVEASP------RN-- 527
TcoHT1    LFGLIGLVCFVLQYFYLYPYEAKSSGD-SSEMQNGGSESK-PSEP-ANQP------AN-- 527
                :* : :: :: * :   .      :  .      . :        *
                   12

SLC2A1    V-        492
BTGLUT1   V-        492
TcrHT1    AA        544
THT1      --        527
TcoHT1    --        527
```
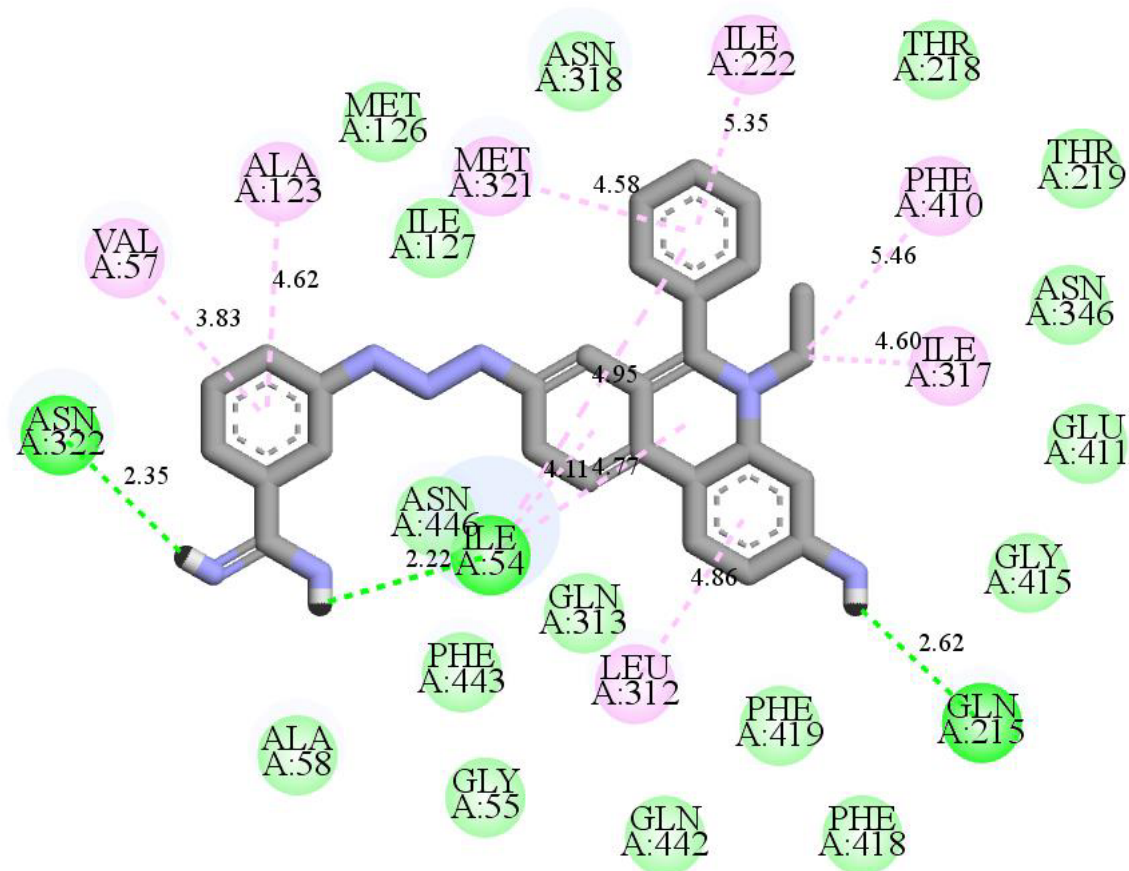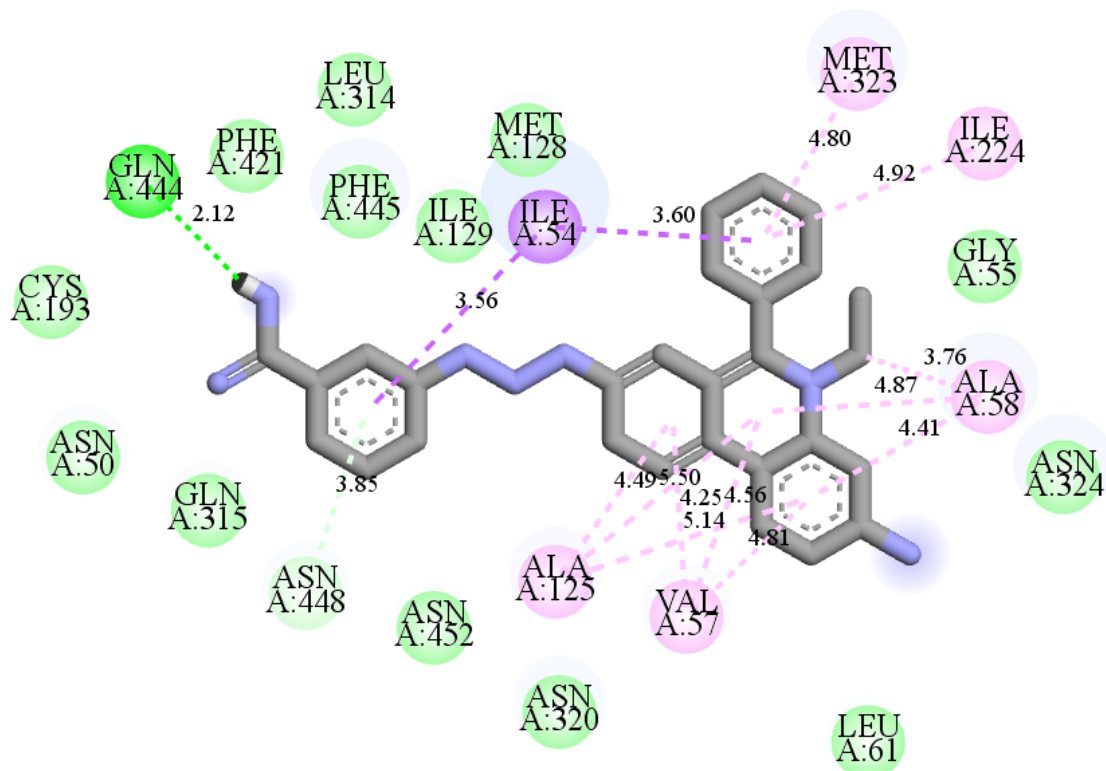
Appendix 2: Molecular docking analysis of isometamidium chloride docked to the *T. vivax* glucose transporter 1 (TvHT1)

Isometamidium chloride interactions with TvHT1

Isometamidium chloride formed the following interactions: hydrogen bonds with Asn 322 (2.35 A˚) and Ile 54 (2.22 A˚), alkyl interactions with Ile 222, Phe 410, Ile 317, Leu 312, Val 57, Ala 123, Met 321 and hydrophobic interactions with Met 126, Ile 127, Asn 318, Thr 218, Thr 219, Asn 346, Glu 411, Gly 415, Phe 418, Phe 419, Gln 442, Gln 313, Phe 443, Gly 55, Ala 58 and Asn 446. The best predicted ISM-protein pose displayed a binding energy of - 10.1 kcal/mol.
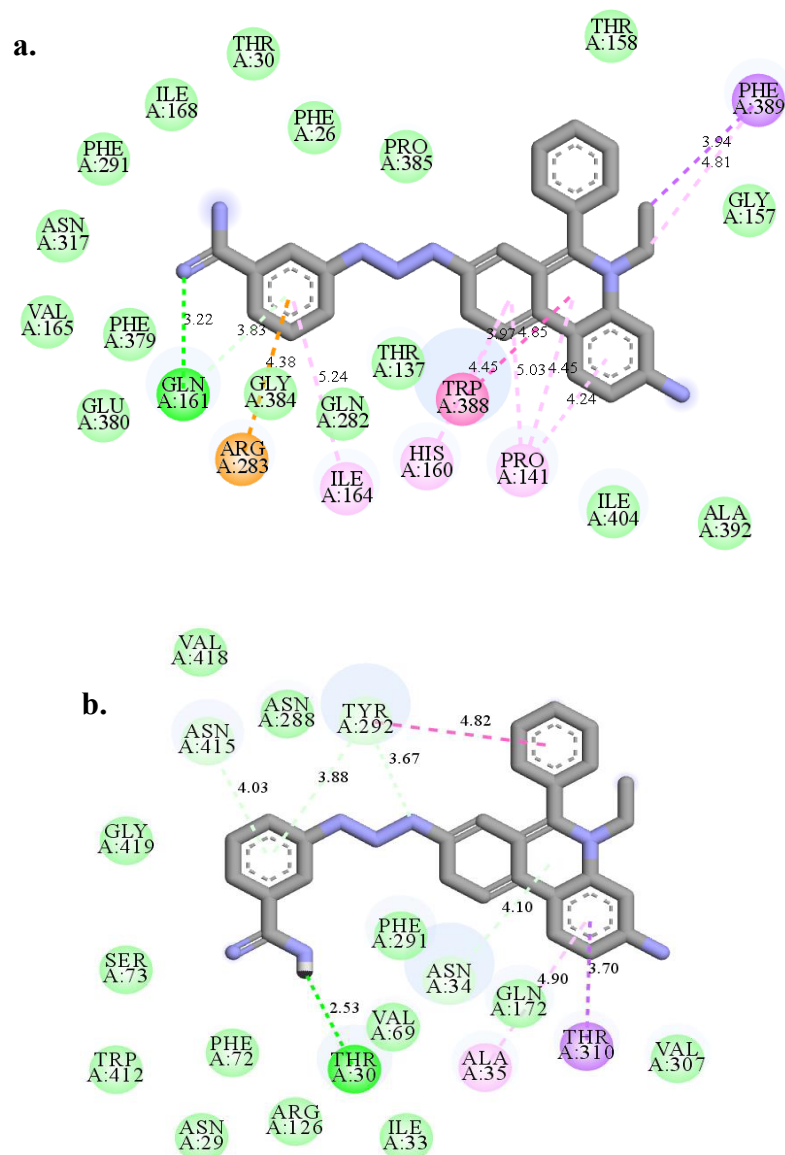
Appendix 3: Molecular docking analysis of isometamidium chloride docked to the *T. brucei* glucose transporter 1 (THT1)

Isometamidium chloride interactions with THT1

Isometamidium chloride formed the following interactions: hydrogen bond with Gln 444 (2.12 A˚), hydrophobic interactions with Phe 412, Leu 314, Phe 445, Ile 129, Met 128, Gly 55, Asn 324, Leu 61, Asn 320, Asn 452, Gln 315, Asn 50 and Cys 193, pi-sigma interactions Ile 54 and alkyl interactions Ala 125, Val 57, Ala 58, Ile 224 and Met 323. The best predicted ISM-protein binding pose displayed a binding energy of -9.9 kcal/mol.

Appendix 4: Molecular docking analysis of isometamidium chloride docked to the *Bos Taurus* and *Capra hircus* glucose transporter 1


Isometamidium chloride interactions with the (a.) *Capra hircus* and (b) *Bos Taurus* glucose transporter 1

(**a**) *Capra hircus* GLUT1 (uniprot entry I0BWL9_CAPHI) interactions with isometamidium chloride. The best ISM-protein predicted pose had a binding affinity of -12.5 kcal/mol while the best predicted glucose-protein pose had a binding affinity of -6.2 kcal/mol (not shown). (**b**) *Bos Taurus* GLUT1 interactions with isometamidium chloride and glucose. The best ISM-protein predicted pose had a binding affinity of -10.2 kcal/mol while the best predicted glucose-protein pose had a binding affinity of -6.3 kcal/mol.

Appendix 5: Molecular docking analysis of the dephospho-CoA bound to the predicted *L. mexicana* dephospho-CoA kinase enzyme (DPCK)
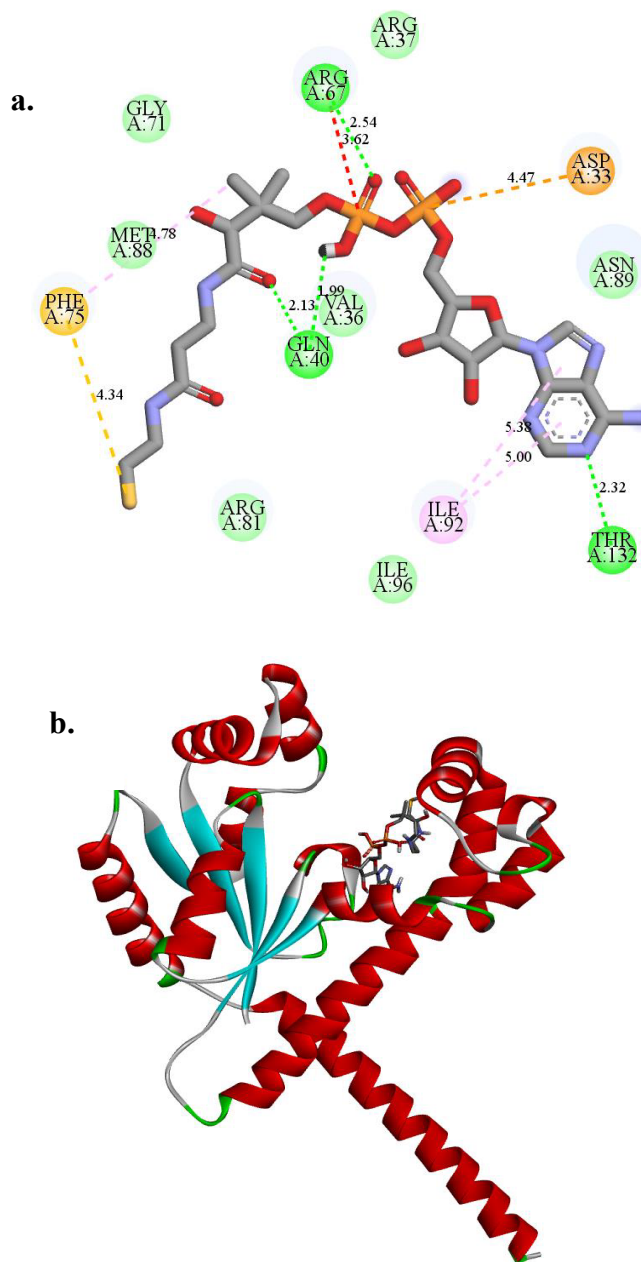
Interaction of dephospho-CoA (dCoA) with DPCK

Dephospho-CoA interacted with Thr 132, Gln 40 and Arg 67 through conventional hydrogen bonding, Gly 71, Met 88, Val 36, Arg 81, Asn 89 and Arg 37 via van der Waals forces, Phe 75 via pi-sulfur interaction, Ile 92 via pi-alkyl interaction, Asp 33 through interactive charge and an additional interaction Arg 37 through unfavorable positive-positive hydrogen bonding (**a**). The best dCoA-DPCK predicted pose displayed a binding energy of -6.2 kcal/mol and was docked to the CoA binding domain (**b**).

Appendix 6: Molecular docking analysis of ATP docked to the predicted *L. mexicana* phosphopantetheine adenylyltransferase enzyme (PPAT)
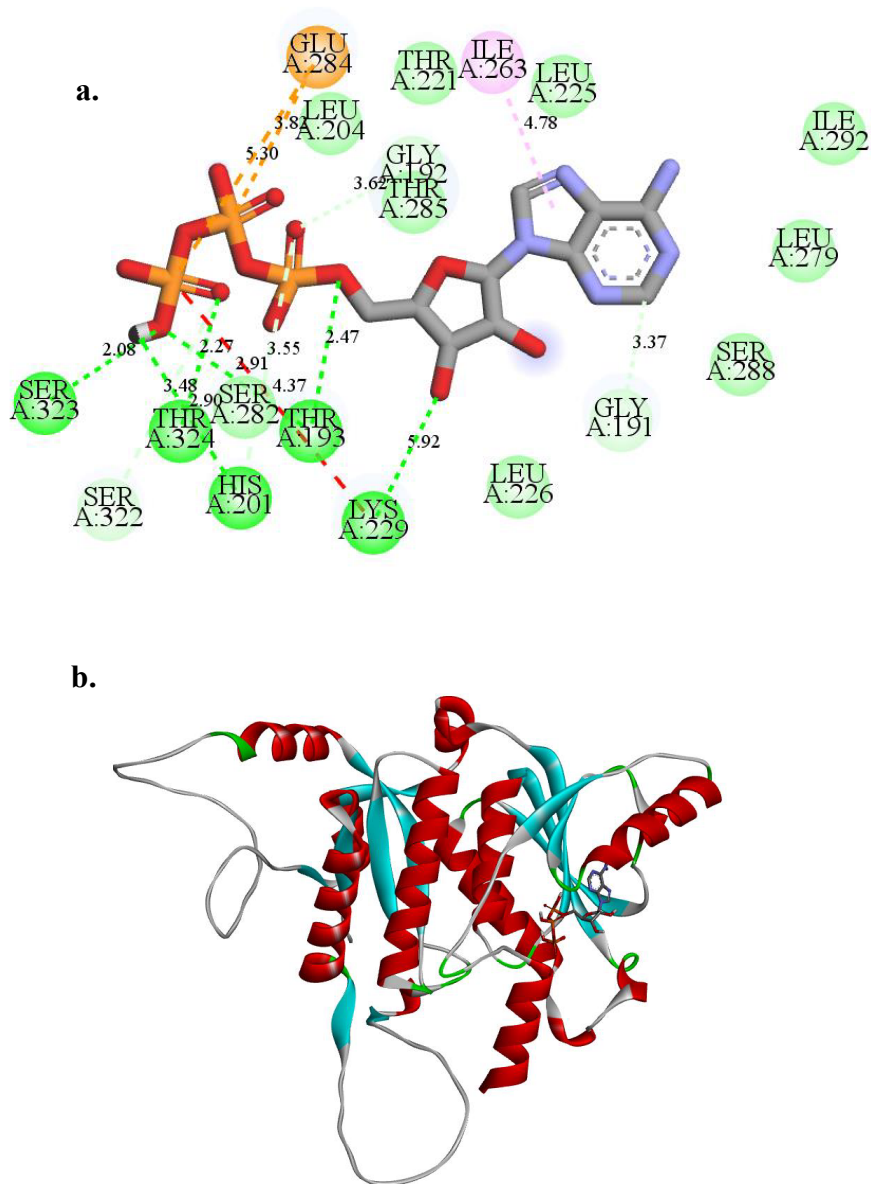

Interaction of ATP with PPAT

ATP interacted with Ser 323, Thr 324, Thr 193, Lys 229, His 201 through conventional hydrogen bonding, Leu 204, Thr 221, Gly 192, Thr 285, Leu 225, Ile 292, Leu 279, Ser 288, Gly 191, Leu 226 and Ser 322 via van der Waals forces, Ile 263 via pi-alkyl interaction, and an additional interaction Glu 284 through attractive charge interaction (**a**). The best ATP-PPAT predicted pose displayed binding energy of -6.4 kcal/mol and was docked to the active site of the enzyme (**b**).

Appendix 7: Multiple protein sequence alignment of *Trypanosoma* and *Leishmania* DPCK sequences

```
L.braziliensis   MILIGLTGGIACGKSAVSRILREEYHIEVIDADLIVRELQAPNAACTRLIAARWPLCVHP  60
L.mexicana       MILIGLTGGIACGKSSVSRILRDEFHIEVIDADLVVRELQTPNSACTRRIAARWPLCVHP  60
L.donovani       MILIGLTGGIACGKSSVSRILRDEFHIEVIDADLVVRELQAPNSACTRRIAARWPLCVHP  60
L.infantum       MILIGLTGGIACGKSSVSRILRDEFHIEVIDADLVVRELQAPNSACTRRIAARWPLCVHP  60
L.major          MILIGLTGGIACGKSSVSRILRDEFHIEVIDADLVVRELQAPNSACTRRIAARWPLCVHP  60
T.vivax          MLLVGLTGGIACGKSTVSTLLQGRHNTIVVDSDRIVRDLQRPCMSCTLKIARRWPNCVDA  60
T.cruzi          MLLIGLTGGIACGKSTVSTMLEKQHHLTVIDADRVVRELQRPSMPCTRKIARRWPGCVNS  60
T.b.brucei       MLLVGLTGGIACGKSTVSLLLKESHHIVVVDSDLVVRELQRPFMPCTRKIARRWPNCVDP  60
T.congolense     MLLVGLTGGIACGKSTVSTILQGRHHITVVDCDKLVRNLQQPFSACARRIARRWPQCVNP  60
                 *:*:************:** :*.  .:  *:*.* :**:** *   *:  ** *** **.

L.braziliensis   ETGELNRAELGKIIFSDAQARRALGKIMNPIIFRVILRRIAAAWWGDLWRSGATSSPAIV 120
L.mexicana       ETGELNRAELGKVVFSDARARRELGKVMNPAIFKAILKRIAAAWWRDLWRSGAVSSPSIV 120
L.donovani       ETGELNRAELGKVVFSDAQARRELGKIMNPAIFKAILKRIAAAWWRDLWRSGAASSPSIV 120
L.infantum       ETGELNRAELGKVVFSDAQARRELGKIMNPAIFKAILKRIAAAWWHDLWRSGAASSPSIV 120
L.major          ETGELNRAELGKIVFSDAQARRALGEIMNPAIFKAILKRIAAAWWHDLWRSGAASSPSIV 120
T.vivax          KTGEINRAALGSVIFSDPAARRELGRIMNTPIFLATMKVLIKLWNRSVWSQAKGESALMV 120
T.cruzi          QTGEIDRAALGEIIFRDPQARRELARIMNFPIFSKVMLLLVRFWWESMKQRMRGEGPLLV 120
T.b.brucei       QSGEVNRAALGSIIFSDPSARRALARIMNFPIFRATMKMVIGLWWQSLRQQLRGQGPLLV 120
T.congolense     LSGEIDRAALGGIIFGDPIARRDLARIMNFPIFCATMKLLLGLWWESLCRQLKGGEPLLV 120
                  :**::** ** ::* *  *** *..:**  **   :  :   ** .:         :*

L.braziliensis   VLDAPTLFETKTFMYFISASVVVSCSEERQIERLRSRNGFSKEEALQRIGSQMALETKRR 180
L.mexicana       VLDAPTLFETKTFTYFISASVVVSCSEQRQIERLRSRDGFSREAALQRIGSQMPLEAKCR 180
L.donovani       VLDAPTLFETKTFMYFVSASVVVSCSEQRQIERLRSRDGFSKKEALQRIGSQMSLEAKRR 180
L.infantum       VLDAPTLFETKTFMYFVSASVVVSCSEQRQIERLRSRDGFSKKEALQRIGSQMSLEAKRR 180
L.major          VLDAPTLFETKTFTYFVSASVVVSCSEQRQIERLRGRDGFSKEEALQRIGSQMSLEAKRR 180
T.vivax          VLDAPLLYESSIYTWFVDCVVVSCTEEKQIARMKARNNLTTEQALQRVRSQMPVSEKCK 180
T.cruzi          VLDAPLLYESNIYTWFIDRVVVVGCKEEEQLARLEKRNGFTREQAMQRVRAQMPIEEKCR 180
T.b.brucei       VLDVPLLYESNIYTWLVDRVVVVSCSEEQQVERMAKRNGLTREQALQRINAQMPISEKCK 180
T.congolense     VLDAPLLYESNIYTWIVDCVMVVACREEQQVERIMKRNGLNREQAVQRVSAQMPISEKCK 180
                 ***.* *:*:.. : ::::.  :**.* *:.*: *:  *:.:. : *:**: :** :. * :

L.braziliensis   LADYIIENDSADDFDQLRGSLRECVAWMSRQSNKRLTCIFV-TVAAAAGVAAVVGYVGY 239
L.mexicana       LADYIIENDCADDLDALRGGVCACVAWMSRQSNKRLTYMFG-TVAVGAVGVAAAVGYACY 239
L.donovani       LADYIIENDCADDLDALRGVVCACVAWMSRQSNKRLTYMFG-TVAAAVGVAAAVGYAGY 239
L.infantum       LADYIIENDCADDLDALRGVVCACVAWMSRQSNKRLTYMFG-TVAAAVGVAAAVGYAGY 239
L.major          LADYIIENDYADDLDALRGVVCACVAWMSRQSNKRLTYMFG-TVATAAVGVAAAVGYVGY 239
T.vivax          LADFVLQNDGS--LDELERLVDKSVNWMRAQRGGRMTKYIALVVL-GTVAVAVSSAYIAR 237
T.cruzi          RADYVIHNSGT--LTELFFSVKDSVEWMRQQSGFKMNTIVFVSAAGGTVCLAAVVHLLCH 238
T.b.brucei       RADRVIHNEES--LSELEHSVADTVAWMQQQSGGRVAFALSGALAVGV-SLGAVVVYCCL 237
T.congolense     RADQVIFNECP--LSELEQLVDDAVLWMRQQSGKQVTRILLATATAGI-GFAAVTAYIVF 237
                  ** :: *.     : *   :   * ** *  .:: .      .   ...

L.braziliensis   RLLLP--  244
L.mexicana       RLLLA--  244
L.donovani       RLLLA--  244
L.infantum       RLLLA--  244
L.major          QLLLA--  244
T.vivax          QLLSSVF  244
T.cruzi          LPW----  241
T.b.brucei       RIVF---  241
T.congolense     RFFV---  241
```

Multiple protein sequence alignment of DPCK sequences.

Conserved amino acids are undermarked with asterisks (*). Conservative amino acid substitutions are marked with a semi-colon (:) while semi-conservative amino acid substitutions are marked with a dot (.). Amino acids that form the CoA binding domain in the *L. mexicana* structure are highlighted in yellow.

Appendix 8: Comparison of the *T. congolense* DPCK and *L. mexicana* DPCK
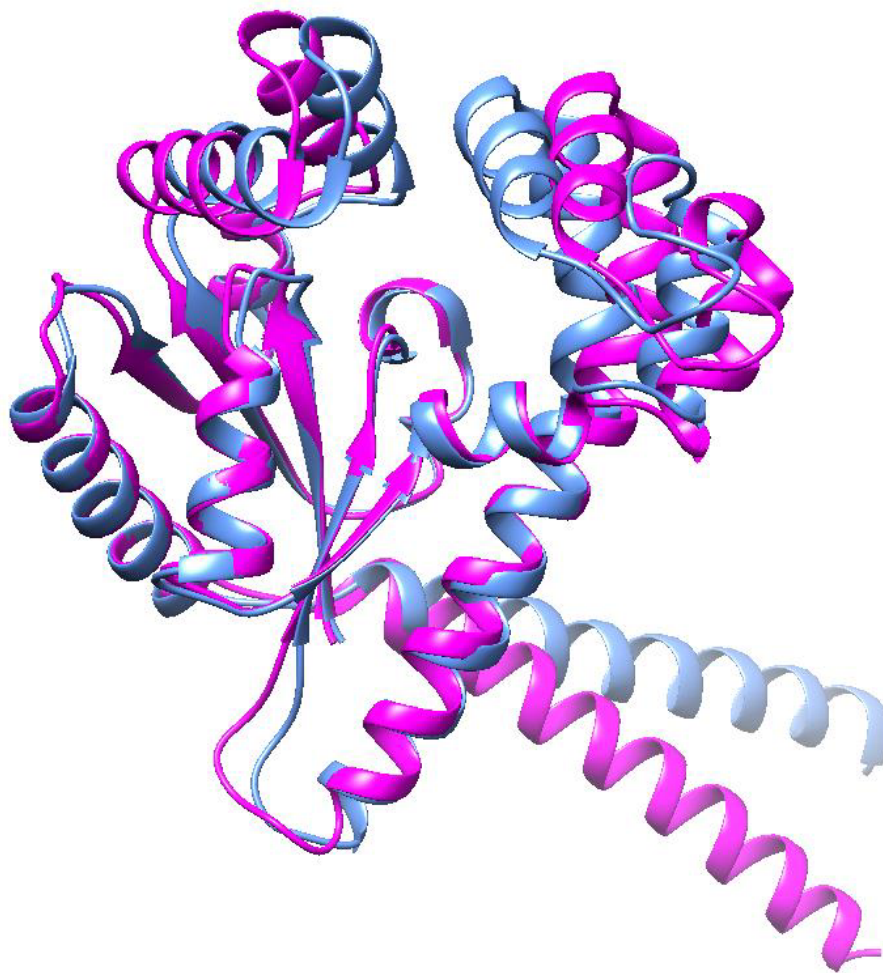
Superposed *L. mexicana* and *T. congolense* DPCKs.

The *L. mexicana* structure is depicted in magenta while the *T. congolense* enzyme is depicted in blue. The structure of *T. congolense* DPCK and *L. mexicana* DPCK were superimposed upon each other using the MatchMaker tool on UCSF chimera (Ballante, 2018). Although both structures were similar; composed of the LID domain, as well as the CoA and nucleotide binding domains, the superimposition revealed a more relaxed and open structure of the *L. mexicana* DPCK compared to the *T. congolense* one. The *T. congolense* DPCK seemed to show a projected LID domain and a somewhat enclosed CoA binding domain.

Appendix 9: Surface analysis of *L. mexicana* DPCK and *T. congolense* DPCK

a.

b.
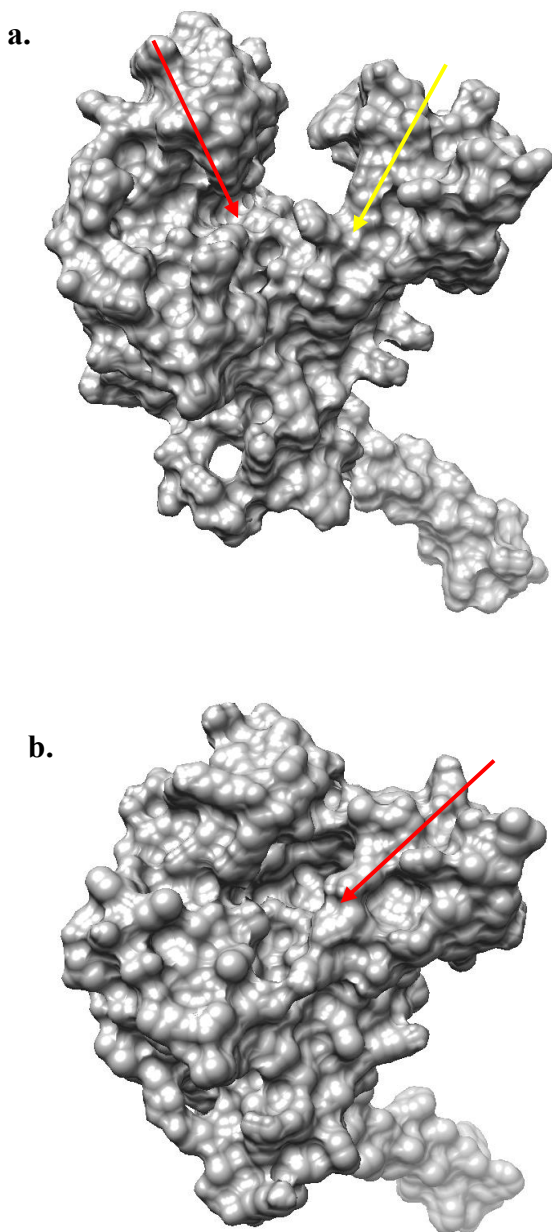
Surface analyses of the (a) *L. mexicana* DPCK and (b) *T. congolense* DPCK.

A surface analysis on chimera showed an open "pore" on the *L. mexicana* structure that could explain why the compounds have access to the CoA binding domain compared to the rather hidden domains in *T. congolense*. On the *L. mexicana* structure the ATP binding site is marked with a red arrow and the CoA binding domain is marked with a yellow arrow while on the *T. congolense* the surface access to the nucleotide and CoA binding domains is marked with a red arrow.