UNIVERSITY OF ZULULAND

Structural Characterization of the Switch domain regions of *Schistosoma mansoni* druggable Adenylate cyclase-stimulating Gα-protein

By

Babatunji Emmanuel OYINLOYE

(201443936)

B.Sc. (Hons) Biochemistry (Babcock), M.Sc. Biochemistry (Ibadan)

Thesis submitted in fulfilment of the requirement for the degree of Doctor of Philosophy (PhD) in Biochemistry

Department of Biochemistry and Microbiology, Faculty of Science and Agriculture,

UNIVERSITY OF ZULULAND

Supervisor: Prof. Abidemi M. Paul KAPPO

March 2017

DECLARATION

I, Babatunji Emmanuel OYINLOYE (Student No: 201443936), declare that apart for the supervisory guidance received "**Structural Characterization of the Switch domain regions of *Schistosoma mansoni* druggable Adenylate cyclase-stimulating Gα-protein**" is the product of my own work and effort. I have acknowledged all sources of information in line with normal academic conventions to the best of my knowledge. I further affirm that the research is original and that the material submitted for examination, is submitted for the degree of Doctor of Philosophy at the University of Zululand and has not previously been submitted for a degree at this or any other higher institution of learning.

Candidate: Babatunji Emmanuel OYINLOYE

Signature:……..

Date:……..

Supervisor: Prof. Abidemi Paul KAPPO

Signature:…….

Date:…….

DEDICATION

This work is dedicated to "**ONISE IYANU" (God of awesome wonders), the Sculptor of my life.** Really, **He** has shown me much more mercy than I deserve. **LORD JESUS CHRIST**, **YOU** are amazing!

ACKNOWLEDGEMENTS

It gives me great pleasure to deeply express my sincere gratitude to my supervisor, Prof. Abidemi P. Kappo, for his excellent guidance, motivation, cooperation, support as well as the constant encouragement I got from him during the entire course of this PhD training. I know I am very fortunate to have an opportunity to be mentored under your tutelage.

My profound and sincere gratitude goes to my parents, Prof. and Dns. G. O. Oyinloye, for all the forms of support I have received from you over the years; especially for allowing me to realize my own potentials by encouraging and funding me to pursue my PhD. Your genuine love and unquantified kindness towards me and my family is highly appreciated. To my amazing wife, Oluwatoyin Oyinloye, your prayers, love, kindness, encouragement and support has strengthened me during this programme. I love you so much. And to my boys, Ayomiposi Enoch and AanuOluwaposimi Emmanuel, for giving me so much joy and happiness, I know I owe you guys a lot.

I need to further thank Dr. (Mrs.) Iyadunni A. Osibote, for her support and kindness, especially for proof reading the thesis in a short period. God bless you abundantly. My heartfelt gratitude goes to Engr. Jolayemi B. Oyinloye and Miss Oyindamola O. Oyinloye, for their much needed support, understanding and encouragement in every possible way. I also wish to express my sincere gratitude to my in-law Elder and Dns. P. O. Akereja for their encouragement and financial support. Dr. (Mrs.) S. O. Nwozo is highly appreciated for her encouragement and mentorship role over the years. Special gratitude goes to Dr. (Mrs.) Christiana Kappo, Daniel and Daniela Kappo for their contributions directly and indirectly to the success of my programme.

The fatherly role of Prof. A.R. Opoku can't be easily forgotten. Your humility, daily smiles, kind words and helping hands as at when needed was a source of encouragement for me during the course of this programme. I deeply appreciate and acknowledge Prof. A.k. Basson, Dr. and Mrs. O. O. Durodolu, Mr. and Mrs. Oki, The Adetiba's, Pastor and Pastor (Mrs.) Foli and Dr. and Mrs. Olowe for their kindness and support. I also wish to acknowledge my spiritual parents, Pastor and Pastor (Mrs.) E. O. Adurokiya for their prayers and support in different areas; especially their financial seeds. I must let you know that the seeds you sowed actually grew, for the jar of flour and jug of oil was not used up, may God bless your home and enlarge your ministry.

I will also like to acknowledge Miss. Priscilla Masamba, my sister and a colleague; you are hard working and energetic, if you continue in this vibe, your limit if any is far away beyond the skies. Our adventure round South Africa is one thing I will not forget in a hurry. Remember Science and Research is not fun when tourism is not incorporated. On this ***Schisto*** project, I also enjoyed the cooperation and benefited from the productivity of my amiable colleague, Mrs. Fatimah Adenowo, more ink to your pen. Alex Adekiya and Taiwo Aruleba, you guys came at the right time. Joining the group was a personal blessing to me.

I will forever be grateful to Dr. and Dr. (Mrs) R. O. Ogunyinka and Dr. and Mrs. F. O. Osunsanmi, for giving me and my family a soft landing in South Africa and for other things too numerous to mention that I have benefited from them. Let me quickly say a very big thank you to Mrs. Ife Ajepe for her invaluable contribution to my career and also for the support I got form her during the course of this study. I also wish to acknowledge all other members of the BSB group: Mrs. Edwina Uzunuigbe, Mrs.

Funmi Akapo, Kho, Makhosi, Nthando, Sbonelo, Philisiwe, Paul and Khuboni your presence and vital help made my research easy and enjoyable.

I also wish to acknowledge the following people for their technical support and assistance and for hosting me in their laboratories during research visits: Prof. A. Shonhai and his research group (especially Dr. Tawanda Zininga and Mr. Stanley Makumire) at the University of Venda; Prof. Ruzvidzo and his research group at North-West University, Mafikeng Campus and Prof. R. Tshentu and Dr. A. S. Ogunlaja at the Nelson Mandela Metropolitan University, Port Elizabeth. I also appreciate the support and kindness of the Pastorate and members of the Redeemed Christian Church of God, Wonderland Parish, Agbowo, Ibadan as well as the Pastorate and members of the Redeemed Christian Church of God, Rock of Salvation Parish, Empangeni, South Africa.

Additionally, I will like to acknowledge Aare Afe Babalola and the management of Afe Babalola University for awarding me a research fellowship for educational advancement towards PhD degree in Biochemistry at the University of Zululand, South Africa. I am also grateful to the National Research Foundation (NRF) South Africa for funding me in 2016 academic session. Lastly, I would like to thank the University of Zululand Research Committee for funding this project.

Above all, I appreciate my LORD and SAVIOUR Jesus Christ, whose strength, grace and mercy made me what I am today. Really, meeting Jesus Christ has made a great difference in my life. Thank you JESUS!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

3D	3-Dimensional
5-HT	5-hydroxytryptamine
AIDS	Acquired Immune Deficiency Syndrome
AMPs	Antimicrobial Peptides
APS	Ammonium Persulphate
ATP	Adenosine Triphosphate
cAMP	Cyclic adenosine Monophosphate
DAG	Diacylglycerol
DALYs	Disability Adjusted Years
DSC	Differential Scanning Calorimetry
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol
ExPASY	Expert Protein Analysis System
FTIR	Fourier Transform Infrared Spectroscopy
GDP	Guanosine-5'-diphosphate

GSP	Guanosine 5'- [gamma-thio] Triphosphate
GNP	Phosphoaminophosphonic Acid-guanylate Ester
GPCRs	G protein-coupled Receptors
GTP	Guanosine Triphosphate
GRAVY	Grand Average of Hydropathicity
HIV	Human Immunodeficiency Virus
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
IARC	International Agency of Research on Cancer
MAP	Multiple Antigen Epitopes
NCBI	National Centre for Biotechnology Information
NMR	Nuclear Magnetic Resonance
NTDs	Neglected Tropical Diseases
NTTs	Neurotransmitter Transporters

PCR	Polymerase Chain Reaction
Phyre2	Protein Homology/analogy Recognition Engine 2
PZQ	Praziquantel
RNA	Ribonucleic Acid
ROS	Reactive Oxygen Species
SPR	Surface Plasmon Resonance
SSA	Sub-Saharan Africa
STDs	Sexually Transmitted Diseases
TGA	Thermogravimetric Analysis
TB	Tuberculosis
UCSF	University of California San Francisco
WHO	World Health Organization

ABSTRACT

Among the neglected tropical diseases (NTDs), schistosomiasis, a common human parasitic disease; continues to rank high as one of the major causes of morbidity and mortality among the labour forces in known endemic regions. Recently, the socio-economic impact of *Schistosoma* infections have been reported to surpass that of malaria in sub-Saharan Africa, as schistosomiasis is known to cause severe damage to various organs thereby impairing childhood development and adult productivity. *Schistosoma mansoni* one of the two main widely distributed causative species parasitizing human, is believed to be responsible for over 90 % of all human schistosomiasis. In the last three decades, praziquantel (PZQ) remains the first-line drug for the treatment of schistosomiasis. Due to drug pressure, some strains of *Schistosoma mansoni* has exhibited worrisome signs of resistance to treatment with PZQ. In search for a novel and/or new anti-schistosomal agent or molecule as an alternative to praziquantel treatment, the two molecular switch domain regions (switch I and switch II) of *S. mansoni* has been identified as a promising druggable target against the parasite. The broad focus of this study was to explore the solubility, stability and druggability of the Gα protein from the switch domain regions of *S. mansoni*. *In-silico* analysis of the switch domain regions using various bioinformatics tools has revealed the existence of two promising functional motifs which are highly conserved in nature. Homology modeling of the switch domain regions and overall stereochemical parameters as evaluated by the Ramachandran plot analysis in RAMPAGE reveals that the modelled structure is stable and of good quality. Docking studies revealed the binding and interaction patterns of the switch domains protein with guanosine-5'-triphosphate, magnesium, calcium and zinc. Recombinant GST-tagged switch domain was heterologously expressed in

Escherichia coli (*E. coli*) JM109 and subsequently purified using glutathione-agarose beads. Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) analysis the recombinant GST-Switch domain protein shows that a high yield soluble fusion protein was obtained, as visible prominent bands (~31.84 kDa) corresponding to the GST-Switch domain protein were observed in the region of 25 – 37 kDa. Preliminary characterization of the GST-Switch domain protein was carried out using fourier transform infrared spectroscopy (FTIR), ultraviolet-visible spectrophotometry, fluorescence spectroscopy, differential scanning calorimeter (DSC), thermogravimetric analysis (TGA), nuclear magnetic resonance spectroscopy (NMR) and isothermal titration calorimetry (ITC). The results obtained from this study, has given us a considerable insight into the nature of the protein in terms of its purity, functional groups, structural and thermal stability as well as its folding state and its binding interaction with two of the hypothesized ligands in literature. In conclusion, the overall results from the computational studies, wet-laboratory experiments and characterization studies are impressive and suggest the possible druggability of the switch domain regions. These findings will enhance our understanding in selective identification of small molecule inhibitors or peptides, which could be developed as novel therapeutic candidates in the control and treatment of human *Schistosoma* infections.

Keywords: Antimicrobial peptides, characterization, docking, drug resistance, human schistosomiasis, neglected tropical diseases, praziquantel, recombinant protein**,** *Schistosoma mansoni,* trematode.

CHAPTER ONE

INTRODUCTION AND LITERATURE REVIEW

1.1 General introduction

Schistosoma mansoni (*S. mansoni*) is an important human parasite, a major causative agent of the debilitating infection schistosomiasis sometimes referred to as bilharzias. In the tropical regions globally, its prevalence is rapidly increasing among the poignant susceptible neglected diseases (Colley *et al.,* 2014). This disease has been estimated to affect about 235 million people while additional 732 million people are projected to be vulnerable to the infection in known endemic regions (Alebie *et al*., 2014; Oyinloye *et al.,* 2014). Obviously the global health impact and economic burden of schistosomiasis is underrated (Engels *et al*., 2002). Schistosomiasis is generally prevalent in developing nations of Asia, South America, Middle East, the Caribbean and Africa (King, 2010; Steinmann *et al*., 2006).

Millions of individuals are known to be infected by the disease predominantly in the rural communities of the tropical regions. This accounts for about 280,000 deaths annually, as a result of stretch in social infrastructures proceeding from population growth, poor sanitation and rising poverty level as well as inadequate facilities and poor health education in these regions (Miranda *et al*., 2013). At present, this disease remains a global public health concern because it affect people of all age with the rates of occurrence higher in boys than girls (school age). Interestingly, the occurrence is higher in female adults than male adults (Zongo *et al*., 2012). The schistosoma lifecycle is complex; it requires series of biochemical transition from one phase to another, interruption of this lifecycle and treatment with the drug

praziquantel has been engaged in the treatment and control of schistosoma infection (Crompton *et al*., 2003).

Infection commences when the free-swimming cercaria enter through the human skin, sheds their tail, transforms into schistosomula and migrates to the hepatic portal system through the lungs (Oyinloye *et al.,* 2014). Routinely, schistosomiasis is diagnosed and confirmed by microscopic examination of urine or stools samples for eggs of this parasite, imaging, polymerase chain reaction (PCR) and serologic tests (Enk *et al*., 2012). Praziquantel (PZQ) as a result of its effectiveness in killing all schistosoma species in a single or few oral doses, remains an important drug in the treatment of all forms of schistosoma infection; the preferred drug due to its excellent tolerability, simple administration, few and transient side effects as well as its high efficacy and competitive cost (Wang *et al*., 2012). However, recent reports have acknowledged that some species of *S. mansoni* from Senegal, Kenya and Egypt have exhibited worrisome signs of resistance to treatment with PZQ due to drug pressure (Doenhoff *et al*., 2002; Ross *et al*., 2002).

Cases of praziquantel resistance together with its inability to prevent schistosomal re-infection as well as its ineffectiveness to annihilate the parasite within two to four weeks after infection, has called for search and development of a novel anti-schistosomal agent as a substitute to treatment with praziquantel (Mbah *et al*., 2012). Mbah and colleagues, in their recent work, elucidated the modulatory role of serotonin (5-hydroxytryptamine: 5-HT) in the motility, muscle contraction, energy metabolism and activation of adenylate cyclase activity in *S. mansoni*. They also describe it as a potential and promising new druggable site in the control and treatment of schistosomal infections (Mbah *et al*., 2012). This is vital, because disruption in any of these mechanisms pose a serious threat to the continuous

survival of the parasite; which may result in the parasite inability to continue with its normal life cycle or its outright death.

Interestingly, six out of the seven different types of 5-hydroxytryptamine (5-HT) receptors belongs to the G protein-coupled receptors (GPCRs) family (Nichols and Nichols, 2008). GPCRs, otherwise known as seven transmembrane receptors are the largest family of signalling receptors, mediating extracellular signals to intracellular signalling cascades (Zamanian *et al*., 2011). The *S. mansoni* genome has been predicted to encode a guanosine triphosphate (GTP) binding protein envisaged to be a member of the G protein α-s subunit, which is accountable for harmonizing adenylate cyclase activity in *S. mansoni.* However, two molecular switches, recognized as switch I and switch II are responsible for the cycling between the GTP-bound (active state) and GDP-bound (inactive state). Identified on these molecular switches are key amino acid residues actually accountable for cycling the protein between GTP-bound (active) and GDP-bound (inactive) states (Mbah *et al*., 2012). Therefore, structural determination and characterization of these molecular switch domains is crucial to the elucidation of the underlying mechanism of activation-deactivation in the trematode worm, which can be exploited in developing novel therapeutic molecules against schistosoma infection.

1.2 Neglected tropical diseases

From time immemorial humanity has been inflicted by a group of diseases categorized as the Neglected Tropical Diseases (NTDs). NTDs are regarded as a group of major chronic infectious diseases, majority of them are parasitic, viral and bacterial infections; predominately occurring in the developing countries of Africa, Asia, and the Americas (Franco-Paredes and Santos-Preciado, 2015; Kappagoda

and Ioannidis, 2014; Lobo *et al*., 2011; Tambo *et al*., 2015). The global burden of NTDs has been reported to be higher than that of malaria, tuberculosis, and other known global diseases such as Human Immunodeficiency Virus/Acquired Immunodeficiency Syndrome (HIV/AIDS). Recently, NTDs have been described has a powerful force in trapping the world bottom billion (the poorest people in the world) in poverty (Alonso and Alvar, 2010; Hotez, 2011; Hotez *et al*., 2009).

With over one billion individuals (world bottom billion) affected with NTDs globally and approximately two billion at risk, it has been estimated that NTDs accounts for about 534,000 deaths and up to 57 million disability-adjusted life years (DALYs) lost annually (Alonso and Alvar, 2010; Kappagoda and Ioannidis, 2014, Martins-Melo *et al*., 2016). Despite the significant disease burden NTDs impose, until recently they have been basically ignored in the global health scheme (Deribe *et al*., 2012). The prevalence of NTDs is known to decrease agricultural productivity and hinder the ability of the communities in many endemic locations to attain sustainable economic development. NTDs have also been reported to have remarkable impacts on the health of the world bottom billion individuals living with NTDs. Added to this, they are believed to reduce children survival and impair childhood growth as well as cognitive development and cause chronic life-long disabilities (Deribe *et al*., 2012; Martins-Melo *et al*., 2016).

Furthermore, they are believed to be responsible for the reduced educational attainment of some individuals in the afflicted populations where they are prevalent as result of disabilities, disfigurement and exacerbating poverty (Alonso and Alvar, 2010; Deribe *et al*., 2012). The World Health Organization has identified 17 major parasitic and other related infections as NTDs; this includes: dengue, rabies, trachoma, buruli ulcer, endemic treponematoses, leprosy, Chagas disease, human

African trypanosomiasis (HAT), leishmaniasis, taeniasis/cysticercosis, dracunculiasis, echinococcosis, food-borne trematodiases, lymphatic filariasis, onchocerciasis, soil-transmitted helminthiases and schistosomiasis (Hotez *et al*., 2016; Mackey *et al*., 2014). It is possible to prevent, control and eliminate these NTDs. The following five strategies: provision of clean water and improved sanitation, intensified disease management, veterinary public health measures for zoonotic neglected diseases, control of disease vectors, and preventive chemotherapy has been recommended by the World Health Organization (WHO) for prevention and controlling as well as possible elimination of these NTDs (Kappagoda and Ioannidis, 2014).

In sub-Saharan Africa (SSA), it has been highlighted that about one-half of the 800 million people who live in this area, live below the World Bank poverty level of US$2 per day (Bethony *et al*., 2011; Hotez and Kamath, 2009). Helminth infections are the most prevalent NTDs in SSA and they are responsible for approximately 85% of the NTDs disease burden in this region. Taken together, in SSA, NTDs disease burden has been estimated to be approximately one-half the disease burden resulting from malaria and one-quarter that of HIV/AIDS. Hookworm infection which is primarily caused by *Necator americanus* and schistosoma infection leading to schistosomiasis (caused predominantly by *Schistosoma haematobium* and *Schistosoma mansoni*) are the most common African helminthiases. It has been documented in literature that about one-third of the world's hookworm infection as well as approximately 90 % of the world's schistosoma infection occur in Sub-Saharan Africa (Bethony *et al*., 2011).

1.3 Historical background of the disease called Schistosomiasis

Among the common neglected tropical diseases (NTDs), in prevalence, human schistosomiasis ranks high (Adenowo *et al*., 2015). Although it is one of the most intensely studied diseases among the NTDs, it continues to remain a life threatening public health challenge in regions where it is endemic (Sady *et al*., 2013). In 1851, a German surgeon working in Cairo, Theodore Maximilian Bilharz (1825 to 1862) was the first to discover and describe the parasite responsible for human schistosomiasis (Aly, 2012; Miranda *et al*., 2013; Tan and Ahana, 2007). He identified the schistosomal worms (male and female) in the bladder and portal system of infected patients during the course of autopsies on patients infected with haematuria in Egypt. Thereafter, the worm (*Distomum* (*Schistosoma*) *haematobium*) which he identified was named by him (Coon, 2005).

The symptom of schistosomiasis was first described in 1847 by Yoshinao Fujii, a Chinese medicine practitioner who lived in Japan (1818 to 1895). He discovered that the villagers in the Katayama District of Japan who worked in the rice fields suddenly developed a rash on their legs. He observed that the workers who developed rash on their legs also had high fever, diarrhea and blood in their stools. Over a period of time, majority of them became emaciated, developed ascites and leg edema and eventually died. Dr. Yoshinao Fujii concluded that this strange disease was associated with something in the water the villagers are exposed to (Sady *et al*., 2013). However, in 1904, a pathologist at Okayama Medical College in Japan described the causative agent of Katayama fever. He called the worm he found in the portal vein of an infected cat, *S. japonicum* to distinguish it from *S. haematobium*. Interestingly, a third species *S. mansoni* (Figure 1.1 and Figure 1.2) was discovered in 1907 in London School of Tropical Medicine by Luigi Westenra Sambon who

named the species in honour of Patrick Manson who was considered as the father of tropical medicine (Sady *et al*., 2013; Wu and Halim, 2000).

Figure 1.1: Female S. mansoni worm ready to lay eggs in canal of male worm. Figure adapted from the website: http://www.genomenewsnetwork.org/articles/10_03/schisto.shtml [Accessed 09 January 2014].

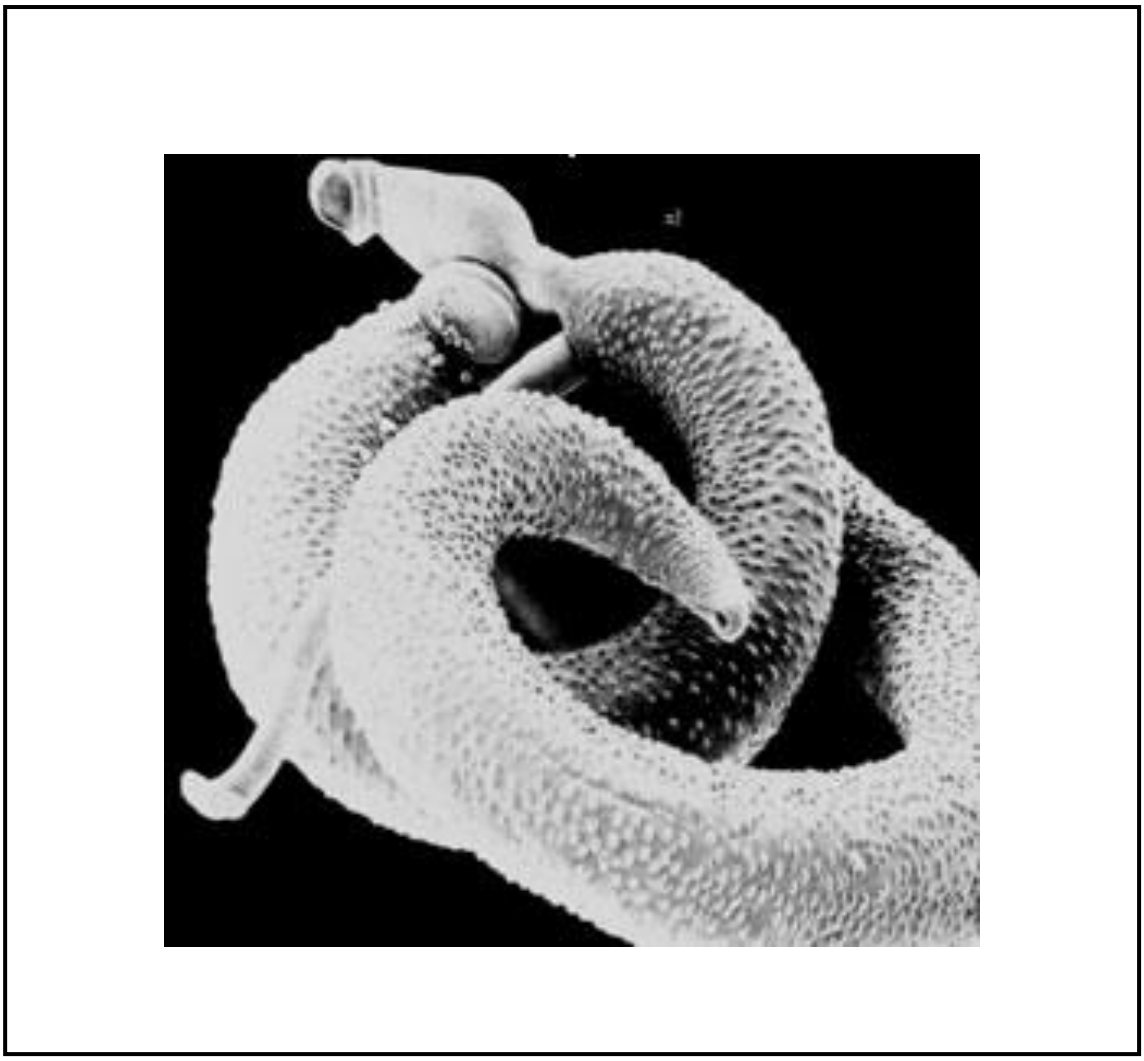


Figure 1.2: Pair of Adult S. mansoni worms. Figure adapted from the website: http://www.genomenewsnetwork.org/articles/10_03/schisto.shtml [Accessed 09 January 2014].

1.3.1 The disease burden and geographical distribution

Although schistosomiasis ranks high among the world's most prevalent infectious diseases, up until now, data on the extent of the burden due to schistosoma infections is scarce, making its effect on the global burden of disease (GBD) controversial (King, 2008; King *et al*., 2005; Mtethiwa *et al*., 2015). However, schistosoma infections and geohelminths have been endorsed by the WHO to account for over 40% of the world tropical disease burden (Adenowo *et al*., 2015). The prevalence of schistosomiasis varies significantly from one geographical region to another as a result of environmental risk factors and lifestyle. In over 75 countries from the Caribbean, South America, Africa and Asia, where this disease is known to be prevalent, it occurs primarily in the underdeveloped areas than the urban centres. Schistosomiasis-endemic nations are located between longitude/latitude 36 ° north and south of the equator (Oyinloye *et al.,* 2014). Added to this, for survival of snails (that serve as intermediate hosts for human schistosomiasis) they have optimal freshwater temperatures (Bica *et al*., 2000; Schutte *et al*., 1994).

Globally, its prevalence has been rising progressively, despite all the efforts to manage and get rid of this disease. Epidemiological studies have revealed that the number of people affected or projected to be infected is still on the increase, in spite of great success that has been achieved in some part of the world concerning control and distribution of this disease. Consequently, human schistosomiasis is regarded as an incapacitating neglected tropical disease (Alebie *et al*., 2014; Engels *et al*., 2002; King, 2010; Oyinloye *et al.,* 2014; Steinmann *et al*., 2006). Due to the fact that it is poorly recognised and diagnosed at early stages, this acute and chronic insidious disease disables men and women during their most productive years (Engels *et al*., 2002). Though, higher rate of prevalence has been noticed and

reported to affect children (less than 14 years of age) in many endemic regions (Alebie *et al*., 2014; Gryseels *et al*., 2006); it is exciting to note that there is an overall decrease in the rate of prevalence among the people in endemic regions, after the age of 20 years, possibly as a result of developing immunity, or perhaps because of behavioural changes associated with reduced exposure to infected water bodies (Conlon, 2005).

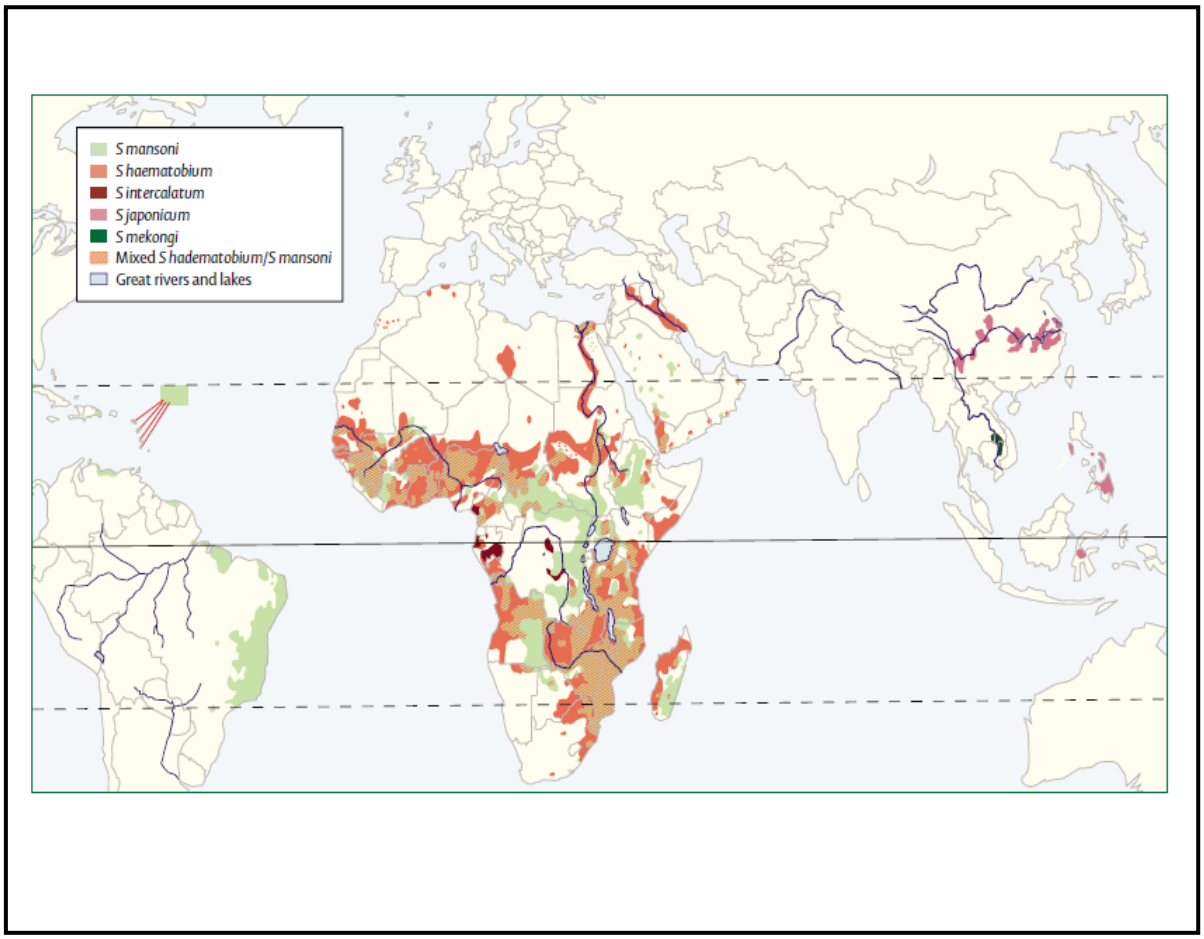


Figure 1.3: Global geographical distribution of schistosomiasis (figure taken from Gryseels *et al*., 2006).

1.4 G-protein coupled receptors (GPCRs)

G-protein coupled receptors (GPCRs) are ubiquitous in nature, they are found in most of the organs of the human body as well as in a wide range of organisms (including mammals, plants, microorganisms, and invertebrates); where they are known to play key roles in biological signalling network that regulates many crucial physiological processes such as, neurotransmission, cellular metabolism, motility, hormone releases, cell growth, immune responses inflammation and blood coagulation just to mention a few (Belmonte and Blaxall, 2012; Covic *et al*., 2002; Karchin *et al*., 2002; Krishnan and Schiöth, 2015; Ramesh and Soliman, 2015). GPCRs are made-up of a large and diverse family of proteins. In the last two decades, notable achievements have been recorded in the study of GPCR; this includes the cloning of the first GPCR genes as well as the successful sequencing of the human genome thereby showing the size of the GPCR family and the number of orphan GPCRs (Kobilka, 2007). As a versatile protein, they are found in the cell membranes; and have been implicated in signal transduction from outside the cell (extracellular signals) to inside the cell (intracellular signals) (Jacoby *et al*., 2006).

The extracellular signals are transmitted to an intracellular molecule known as G-protein (guanine nucleotide-binding protein). About 30 % of all marketed prescription drugs act on GPCRs; added to this, they are the richest historically most successful source of targets for the pharmaceutical industry in recent times (Jacoby *et al*., 2006; Kobilka, 2007; Kroeze *et al*., 2003). The first molecular cloning of GPCRs predicted that the structure of the protein is similar to rhodopsin on the basis of homology. In addition, even though they are remarkably diverse in sequence and function, all GPCRs share a highly conserved topological arrangement predicted to contain

seven membrane spanning helices, an extracellular N-terminus and an intracellular C-terminus. Hence, this gave rise to their other names (seven transmembrane; 7-TM receptors, heptahelical receptors or serpentine receptors (Covic *et al*., 2002; Jacoby *et al*., 2006; Kroeze *et al*., 2003). At the moment, GPCRs are recognized as extremely versatile receptors for extracellular messengers such as biogenic amines, purines and nucleic acid derivatives, lipids, peptides and proteins, odorants, pheromones, tastants, ions like calcium and protons, and even photons in the case of rhodopsin (Jacoby *et al*., 2006).

1.4.1 Structural features of G-protein coupled receptors

Extensive investigation on the protein sequence of G-protein coupled receptors (GPCRs) shows they all have a domain with common protein topology consisting of seven membrane spanning helices bundle, which is believed to accommodate the binding site for low-molecular-weight ligands (Salon *et al*., 2011). On structural basis, GPCRs can be classified as: GPCRs with short N-terminal consisting of about 5 - 80 residues and GPCRs with a long N-terminal made up of about 80 - 600 residues. The long N-terminals are believed to be involved in the ligand recognition; they move into the transmembrane (TM) region to activate the G protein, when they eventually bind to a ligand (Van Neuren *et al*., 1999; Vaidehi *et al*., 2002; Zhao *et al*., 2016). Even though all GPCRs share a similar structural signature, a seven hydrophobic transmembrane region; with an extracellular N-terminal and an intracellular C-terminal domains (Figure 1.4), recent comprehensive investigation of the human genome reveals over 800 distinctive GPCRs, out of which about 460 are predicted to be olfactory receptors (Deupi and Kobilka, 2007; Janero and Thakur, 2016).

It is interesting to note that based on sequence similarity within the transmembrane region, these receptors can be grouped into five main families: the rhodopsin family; having 701 members, the adhesion family; having 24 members, the frizzled/taste family; having 24 members, the glutamate family; having 15 members, and the secretin family; having 15 members (Katritch *et al*., 2013; Kobilka, 2007). The actual function of many of these 800 GPCRs is yet to be known; therefore, these receptors are often referred to as orphan GPCRs. Interestingly, due to the ongoing deorphanization process of non-olfactory GPCRs, the total number of orphan GPCRs continues to decline gradually (Kobilka, 2007). Traditionally, heterotrimeric guanine nucleotide-binding regulatory proteins mediate signal transduction between many membrane bound receptors and intracellular effectors (Figure 1.5). To this end, activated GPCRs act as guanine nucleotide exchange factors (GEFs) for the α-subunits of heterotrimeric G proteins (monomeric G proteins), catalysing the release of GDP and the binding of GTP (Ritter and Hall, 2009; Stanley, and Thomas, 2016; Tall *et al*., 2003).

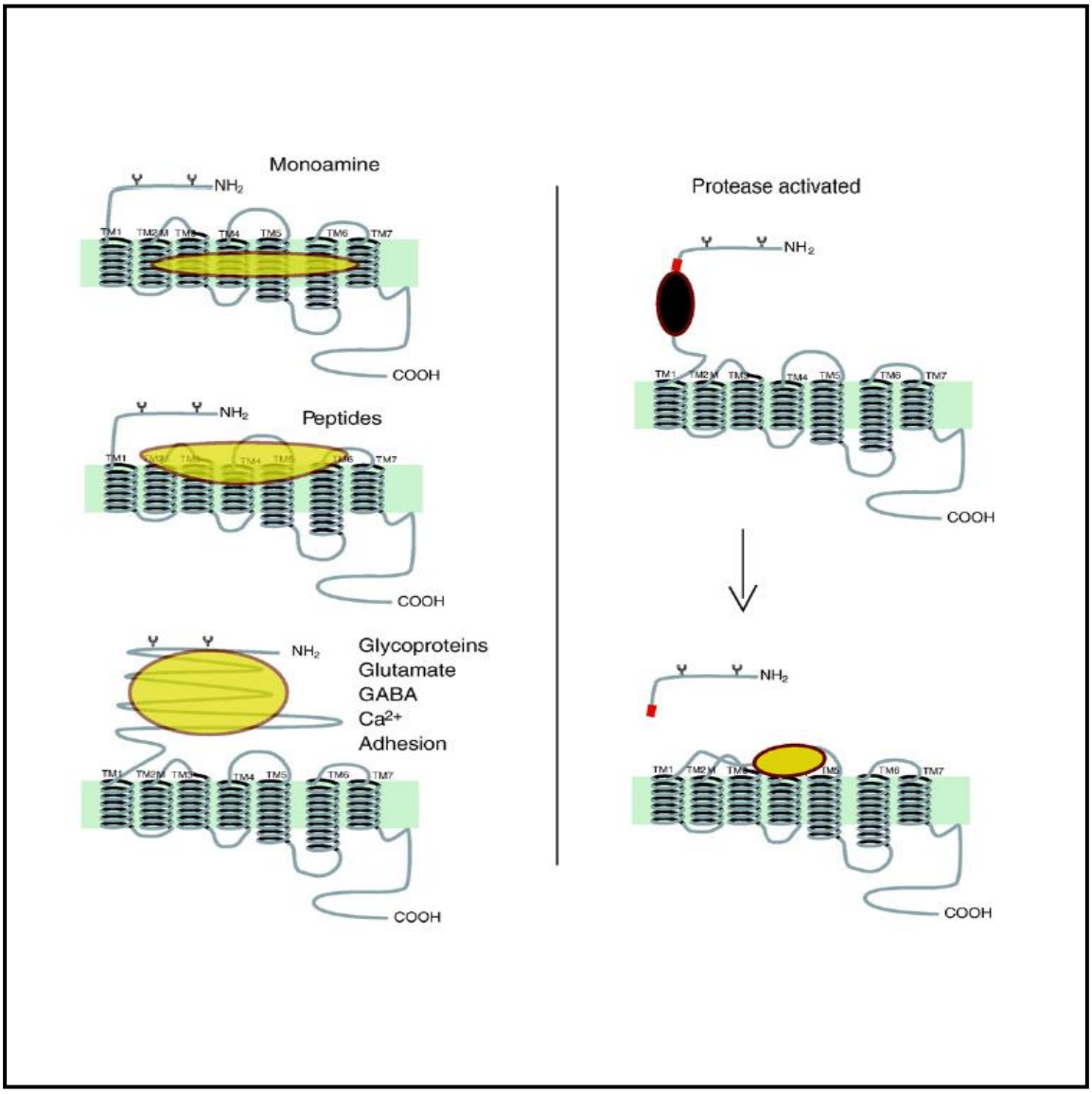


Figure 1.4: Cartoons depicting the secondary structure and the location of agonist binding sites for different GPCRs (figure taken from Kobilka, 2007).

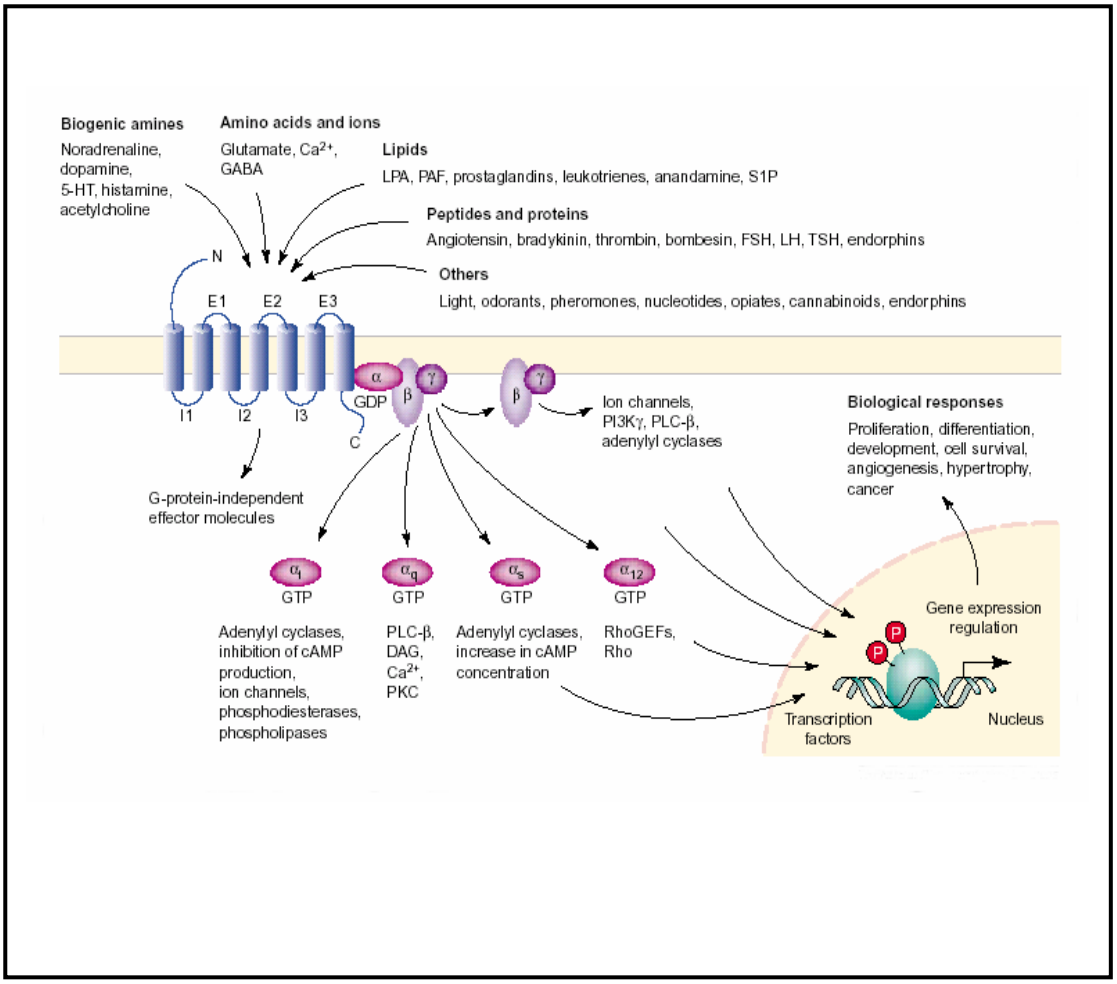


Figure 1.5: A wide variety of ligands use GPCRs to stimulate cytoplasmic and nuclear targets through heterotrimeric G-protein-dependent and -independent pathways. Such signaling pathways regulate key biological functions such as cell proliferation, cell survival and angiogenesis. Figure adapted from the website: http://diposit. ub. edu/ dspace /bits tream/ 24 4 5/36250/2/1.INTRODUCTION.pdf [Accessed 25 February 2016].

1.4.2 Guanine nucleotide-binding proteins (G-proteins)

Guanine nucleotide-binding proteins (G-proteins) also known as GTP-binding proteins or GTPases; are an essential and universal family of intracellular signalling molecules (Stanley, and Thomas, 2016). They are known mediators involved in many cellular processes. G-proteins are anchored on the cytoplasmic cell membrane where they play a role in signal transduction, protein transport, growth regulation, and polypeptide chain elongation just to mention a few. Interestingly, almost all members of this super family of proteins act as a molecular switch (nucleotide-dependent switches that dictate the specificity of their interactions with binding partners), which is activated when GTP is bound and deactivated when GDP is bound (Vogler *et al*., 2008). G-protein regulatory systems are vital components of many intracellular processes, and inaccurate regulation of G-proteins has been implicated in many disease conditions (Hatley *et al*., 2003; Hughes, 1983; Stanley, and Thomas, 2016; Wang *et al*., 2011).

G-proteins are of two major subfamilies incorporating both the small monomeric Ras-like G-proteins and the heterotrimeric G-proteins (Ishii *et al*., 2010). The heterotrimeric G-proteins is made-up of three different subunits (the α-subunit, the β-subunit, and the γ-subunit; Figure 1.6). The α-subunit contains the GTP/GDP binding site. The different subunits in the heterotrimeric G-proteins have important physiological and regulatory roles that are well documented in literature such as regulation of enzyme activity (adenylyl cyclase, phospholipase C, phosphodiesterase, GRK) or modulating channel function (K^+, Ca^{2+}). Additionally, G-proteins bind guanine nucleotides (GDP, GTP) in a strongly conserved nucleotide binding pocket—an ancient mechanism preserved in both eukaryotes and prokaryotes (Ligeti *et al*., 2012; Stanley and Thomas, 2016).

Figure 1.6: The structure of the heterotrimeric G protein, with GDP bound to Gα (cyan) and Gβ (magenta) linking Gα to Gγ (green). Figure adapted from the website: https://www.caymanchem.com/Article/2188 [Accessed 09 January 2014].

1.4.3 Signal transduction through G protein-coupled receptors (GPCRs)

As mentioned earlier, G protein-coupled receptors (GPCRs) are an important class of protein (Stanley, and Thomas, 2016). They are involved in numerous physiological processes and diseases. The pathogenesis of several chronic diseases has been linked to a deficiency in the signal transduction pathways especially the G protein-coupled pathway, which basically comprises of three definite components: the G protein-coupled receptors (GPCRs), G protein, and effectors otherwise known as intracellular second messengers (Spiegel *et al*., 1993). Signal transduction explains the mechanism of transformation of extracellular signals (such as neurotransmitters, hormones and growth factors) to a precise internal cellular response (such as cell division, gene expression and even cell death or apoptosis; Halazy, 2003). GPCRs, also known as seven transmembrane receptors (because they span through the lipid bilayer seven times; use the heterotrimeric G proteins found on the cell surface to deliver intracellular signals) are the largest family of signaling receptors, mediating extracellular signals to intracellular signaling cascades by binding to a ligand or a signaling molecule, which in turn changes the conformation of the target cell (Zamanian *et al*., 2011). Motility in *S. mansoni* within the intermediate and the definitive host is controlled by serotonin activation of GPCRs (Patocka *et al*., 2014).

The G-proteins as described earlier, are of two major subfamilies: the small monomeric Ras-like G proteins and the heterotrimeric G proteins (Grant *et al*., 1997; Ishii *et al*., 2010). The heterotrimeric G proteins are confined to the plasma membrane inner surface made up of Gα, Gβ and Gγ subunits (Figure. 1.6; Tuteja, 2009). They play a pivotal role in cell signaling as molecular switches where they regulate the rate of production, release and degradation of intracellular second messengers. GPCRs receive many extracellular signals and diffuse them to G-

proteins, which in turn diffuse them to the appropriate effector (intracellular second messengers). Ligand binding to a GPCR initiates the activation of the interacting G-protein that results in exchange of GDP for GTP (Ligeti *et al*., 2012; Milligan and Kostenis, 2006; Oldham and Hamm, 2008; Spiegel, 1993).

In the inactive state, the Gα subunit of the heterotrimeric G protein is bound to GDP; conformational changes occurs at the receptor after binding of neurotransmitters or hormones (ligands) to the GPCRs leading to the replacement of GDP with GTP on the Gα subunit causing the inactive heterotrimeric G protein to be released from the receptor as well as a concurrent dissociation of the Gα subunit from the Gβγ subunit (Mbah *et al*., 2012; McCudden *et al*., 2005). The dissociated Gα subunit and Gβγ subunit (Figure 1.7) will in turn activate different effectors such as phosphodiesterases, adenylate cyclase, ion channels and phospholipase C which will in turn regulate intracellular concentrations of secondary messengers, such as diacylglycerol, cAMP, sodium or calcium cations thereby leading to a physiological response. Gα has an inherent GTPase activity which results in rapid hydrolysis of GTP to GDP causing inactivation of the Gα and its subsequent dissociation from adenylyl cyclase resulting in reassociation with the βγ dimer and their binding to the receptor (McCudden *et al*., 2005).

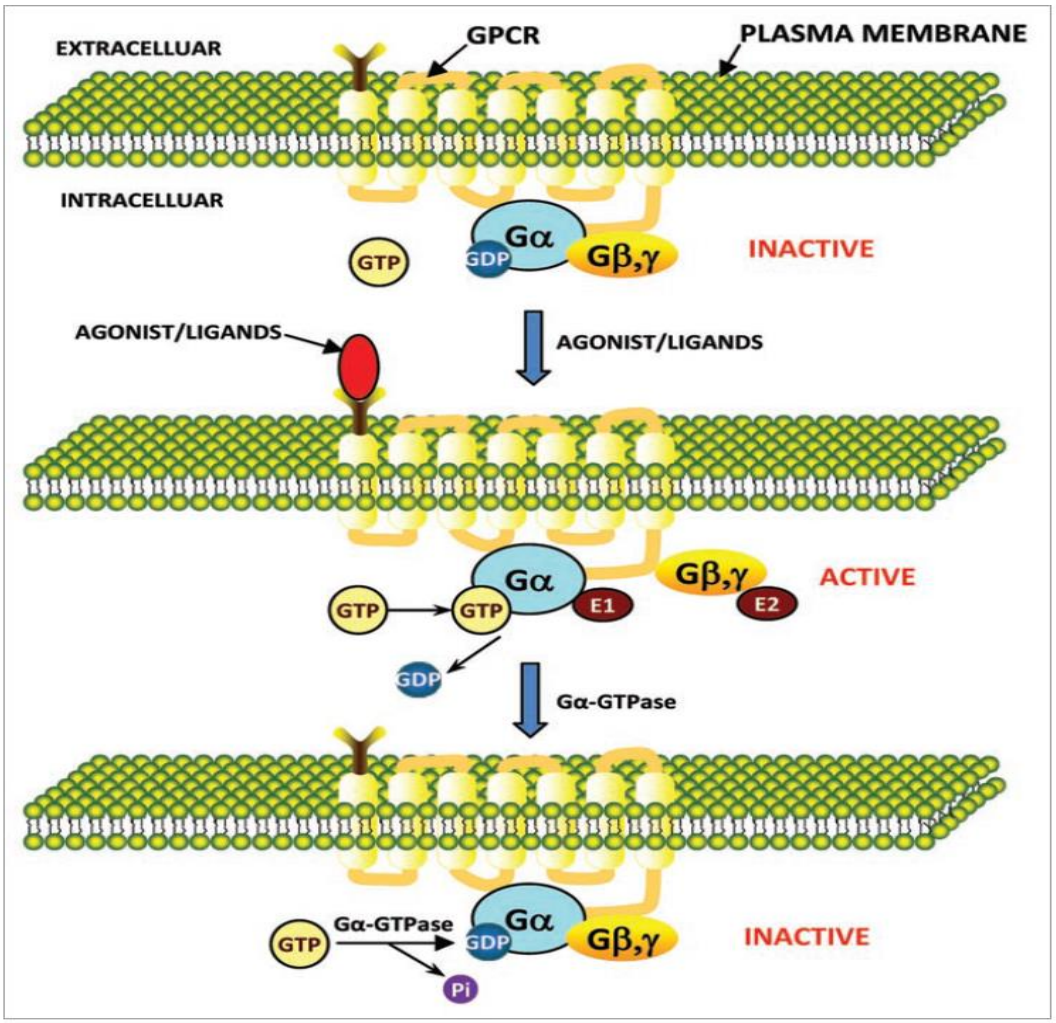


Figure 1.7: Model for Signal transduction through G protein-coupled receptors (figure taken from Tuteja, 2009**).** Gα and Gβγ (heterotrimeric G proteins) subunits in their inactive state are associated with each other, with the Gα subunit bound to GDP. Ligands binding to the receptor lead to conformational changes that result in the exchange of GDP on the Gα subunit with GTP as well as a concurrent dissociation of the Gα from Gβγ. Free Gα and Gβγ will in turn activate different effectors (E1 and E2, respectively) to further transmit the signals and initiate unique intracellular signalling responses. Thereafter, Gα-GTPase hydrolyses causing inactivation of the Gα and its subsequent dissociation from adenylyl cyclase (E1) resulting in reassociation with the βγ dimer and the concommittant binding to the receptor.

1.4.4 Serotonin

Serotonin (5-hydroxytryptamine: 5-HT) is a small ubiquitous neuroactive signalling molecule. It is an important biogenic amine neurotransmitter widely distributed in both vertebrates and invertebrates, which is gotten from the metabolism of amino acids (El-Shehabi *et al*., 2009). In both invertebrates and vertebrates, this neuromodulator and neurotransmitter has been confirmed to perform an extensive role in the management of diverse physiological processes (Boyle and Yoshino, 2005). Serotonin and serotonin receptors have been identified in *S. mansoni*. In the schistosome nervous system, serotonin is one of the most abundant neuroactive substances, that performs very essential modulatory role in the survival mechanism of the parasite; it is implicated in the crucial coordinated activities of the parasite such as motility, energy metabolism (glucose uptake and glycogen utilization) and adenylate cyclase activity (Boyle *et al*., 2000; Patocka *et al*., 2014; Patocka and Ribeiro, 2007).

Linking these roles together (motility and metabolism) in this parasite, it appears that serotonin operate as both a stimulatory neuromodulator and a neurotransmitter. It is also documented that endogenous serotonin plays a crucial role in motility of the helminthic sporocysts and reduced levels of endogenous serotonin subsequently inhibit the release of daughter sporocysts (Boyle *et al*., 2003). The unique tegumental surface of the sporocyst, larvae and adult *S. mansoni* serves as a site of serotonin and serotonin receptor activation, which is important for the parasites survival and existence (Boyle and Yoshino, 2005). Six out of the seven different types of 5-HT receptors belongs to G protein-coupled receptors (GPCRs) family (Nichols and Nichols, 2008).

1.4.5 Second messengers

Intracellular small molecules synthesized within the cell in response to external signals, that coordinate external signals to specific intracellular responses are known as second messengers. Common examples include diacylglycerol (DAG), cyclic adenosine 3′, 5′ -monophosphate (cAMP) and inositol 1, 4, 5-triphosphate (IP3; Bornfeldt, 2006; Leckie *et al*., 1998). Cyclic AMP (cAMP) was the first to be identified among the second messengers. It plays essential roles in regulating cellular function of many hormones and neurotransmitters. The intracellular levels of cAMP are regulated by the balance between the activities of two enzymes: adenylyl cyclase (AC) and cyclic nucleotide phosphodiesterase (Sassone-Corsi, 2012). Adenylyl cyclase, a membrane-associated enzyme is a common target of activated G proteins. When G proteins are activated (GTP-bound alpha subunit) within the cell, they stimulate adenylyl cyclase to synthesize large amount of cAMP from molecules of ATP as shown in Figure 1.8 (Levin and Reed, 1995).

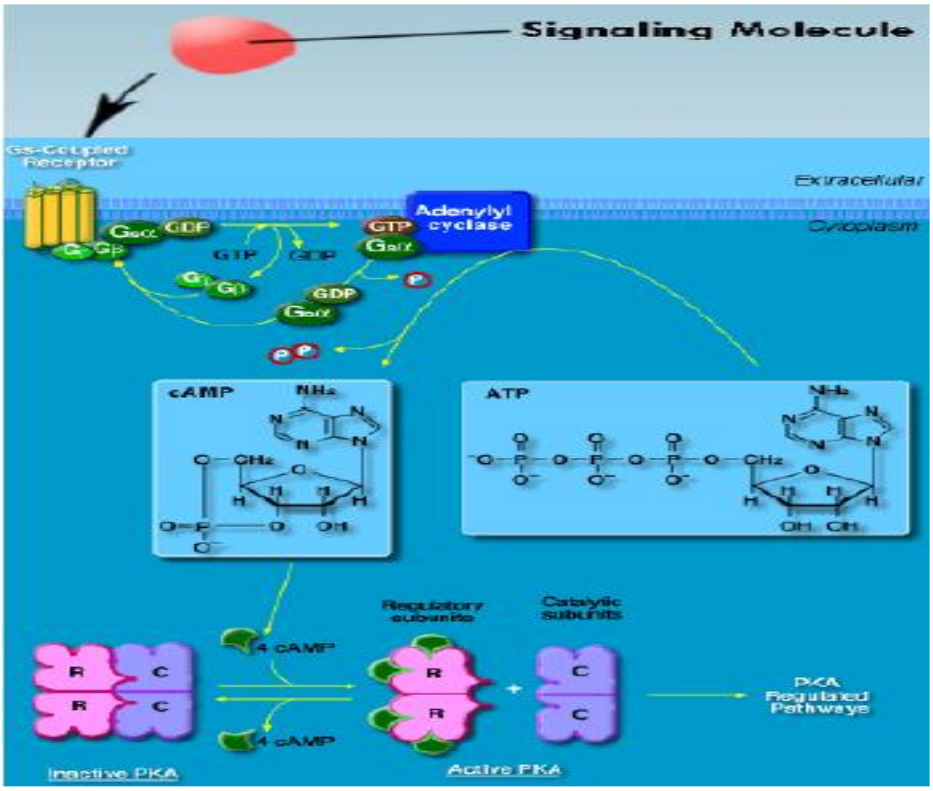


Figure 1.8: Serotonin-activated GPCRs signaling pathway leading to cAMP production (Figure adapted and modified from the website: http://biocarta.com/pathfiles/gspathway.asp [Accessed 30 July 2014]. Upon activation of the Gα subunit of the heterotrimeric G protein by serotonin, cAMP is synthesized from ATP by adenylyl cyclase. Increase in cAMP is translated to cellular responses by activating the cAMP-dependent protein kinase, protein kinase A (PKA).

1.4.6 Adenylate cyclase (AC)

The adenylate cyclases (ACs) are known to be the effector molecule of one of the most widely utilized signal-transduction pathways (Tresguerres *et al*., 2011). Typically, ACs consist of the families of enzymes that catalyse the formation of the ubiquitous second messenger 3′,5′-cyclic adenosine monophosphate (cAMP), which transmits various signals (such as the presence of hormones, nutrient availability and changes in the cellular environment) to effectors such as protein kinases and transcription factors, which in turn generate a cellular response. Up until now, six classes of ACs have been identified, and these are believed to be the product of convergent evolution, since they share no sequence similarities (Hurley, 1999; Linder, 2008). Adenylylate cyclase isoform was first cloned from the brain. The amino acid sequence obtained from the cDNA clone showed several interesting characteristics. The predicted structure of the brain isoform (Figure 1.9) was surprising; hydropathy investigation predicted a molecule consisting of a component of six transmembrane spans linked to a large cytoplasmic domain that was tandemly repeated (Ishikawa and Homcy, 1997).

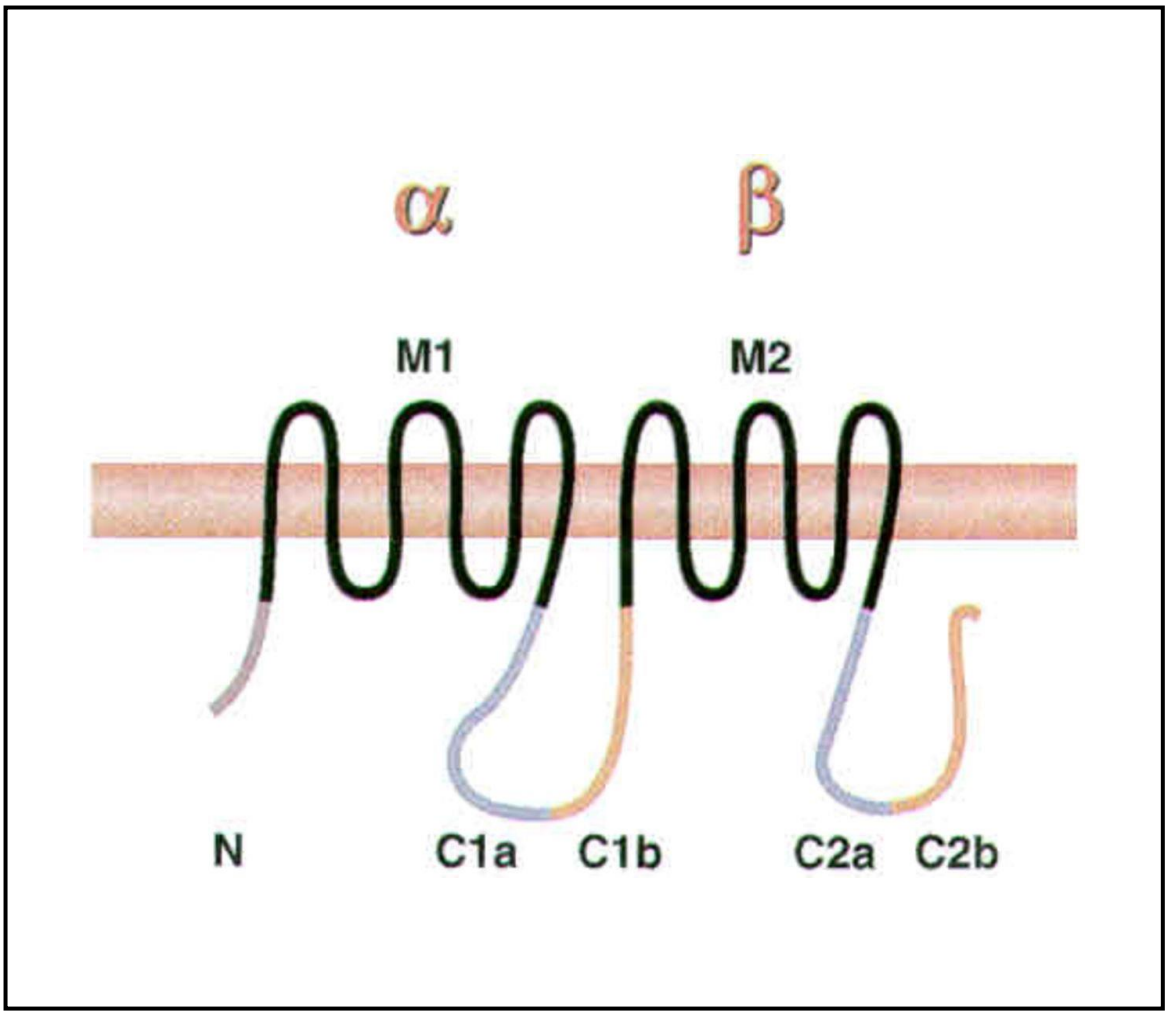


Figure 1.9: Structure of adenylyl cyclase (figure taken from Ishikawa and Homcy, 1997). The entire molecule is subdivided into seven domains: N indicates the N-terminal cytoplasmic domain; M1, the first six transmembrane-spanning domain; C1a, the first cytoplasmic catalytic domain; C1b, the first cytoplasmic linker domain; M2, the second six transmembrane-spanning domain; C2a, the second cytoplasmic catalytic domain; and C2b, the second cytoplasmic noncatalytic domain. Alpha (α) represents the front half of the molecule; beta (β), the back half.

1.4.7 Therapeutic potentials of GPCRs

In recent times, G protein-coupled receptors (GPCRs) are popular among the most heavily studied targets in the pharmaceutical industry for new drug development. In fact, they are documented in literature to account for approximately 20% of all the clinical drug targets (Suzuki and Kaneko-Kawano, 2016). This is as a result of their diversity and cell-specific expression as well as the believed druggability of the GPCR superfamily (Reimann and Gribble, 2016). Added to this, it has become clear that they are important regulators of several cellular functions and pathophysiological processes which includes neurotransmission, metabolism, cell growth, immune or inflammatory responses, olfaction and vision (Hudson *et al*., 2013; Reimann and Gribble, 2016; Suzuki and Kaneko-Kawano, 2016). Actually, many metabolic intermediates, formerly believed to carry out their biological functions through intracellular targets, are actually capable of activating GPCRs. To this end, in spite of all the achievements recorded about GPCRs as drug targets, it should be known that only a minute fraction of the identified GPCRs are presently explored, suggesting that GPCRs remain a largely untapped source with potentials for novel therapeutic development against various human diseases (Hudson *et al*., 2013; Smith, 2012).

1.5 Drug pressure and drug resistance in schistosomiasis

In the last few decades, among the three notable compounds (metrifonate, oxamniquine, and praziquantel) employed in the treatment of human schistosomiasis; praziquantel (PZQ), is gradually becoming the only commercially available anti-schistosomal drug available in the market (Deribew and Petros, 2013). As a matter of fact, PZQ is now widely accepted as a drug of choice preferred above other available anti-schistosomal drugs due to its safety, low cost, single dose

efficacy, and its effectiveness against all five schistosoma species parasitizing human (Almeida *et al*., 2015; Deribew and Petros, 2013; Fallon, 1998; Mendonca *et al*., 2016). At the moment, majority of the research carried out to establish resistance to PZQ treatment are focused on *S. mansoni* (Couto *et al*., 2011; Deribew and Petros, 2013; Mendonca *et al*., 2016).

To this end, due to drug pressure, the prospect of emergence of low cure rate and resistance to PZQ treatment in some *S. mansoni* strains; demonstrated in some endemic regions, particularly in Senegal, Egypt and Brazil is worrisome (Coeli *et al*., 2013; Deribew and Petros, 2013). Impaired cognitive potential in the midst of primary school-age children, hepatosplenomegaly, anemia as well as cancer of the bladder and stunted growth, appears to be some of the devastating effects resulting from schistosoma infection (Deribew and Petros, 2013). Dependence on a single drug (PZQ) for a disease of this magnitude is risky (Greenberg, 2013). Therefore, it is necessary to deal with this issue of drug resistance among schistosomal species by searching for a promising alternative treatment to PZQ; before schistosomes develop a disastrous widespread resistance.

1.6 Mechanism of action of Praziquantel (PZQ)

Praziquantel has been employed in the treatment of all forms of schistosomal infection, yet the actual mechanism of action remains poorly understood (Brindley and Sher, 1990). As shown in Figure 1.11, it is believed that praziquantel elicit its action by causing structural changes in the teguments of adult *S. mansoni*, leading to enhanced permeability to calcium ions associated with a gradually inflow of sodium ions and a diminished inflow of potassium ions. Accumulation of calcium ions in the parasite cytosol results in intense muscular contractions, which ultimately leads to

paralysis of worms. This mechanism suggested that praziquantel might impede inorganic ion transport (see Figure 1.11). Added to this, structural changes in the teguments of the worm, opens up its antigens to the host immune system resulting in the elimination of worms from the host intestine. This can also be linked to the disruption of calcium ion homeostasis (Aragon *et al*., 2009). Documented experimental evidence proposes that neurotransmitter transporters (NTTs) are common target for therapeutic drugs. NTTs have been identified to play a crucial role in the control of neurotransmitter signalling and homeostasis; this concept is perfectly explained by the example of biogenic amine transporters (Ribeiro and Patocka, 2013).

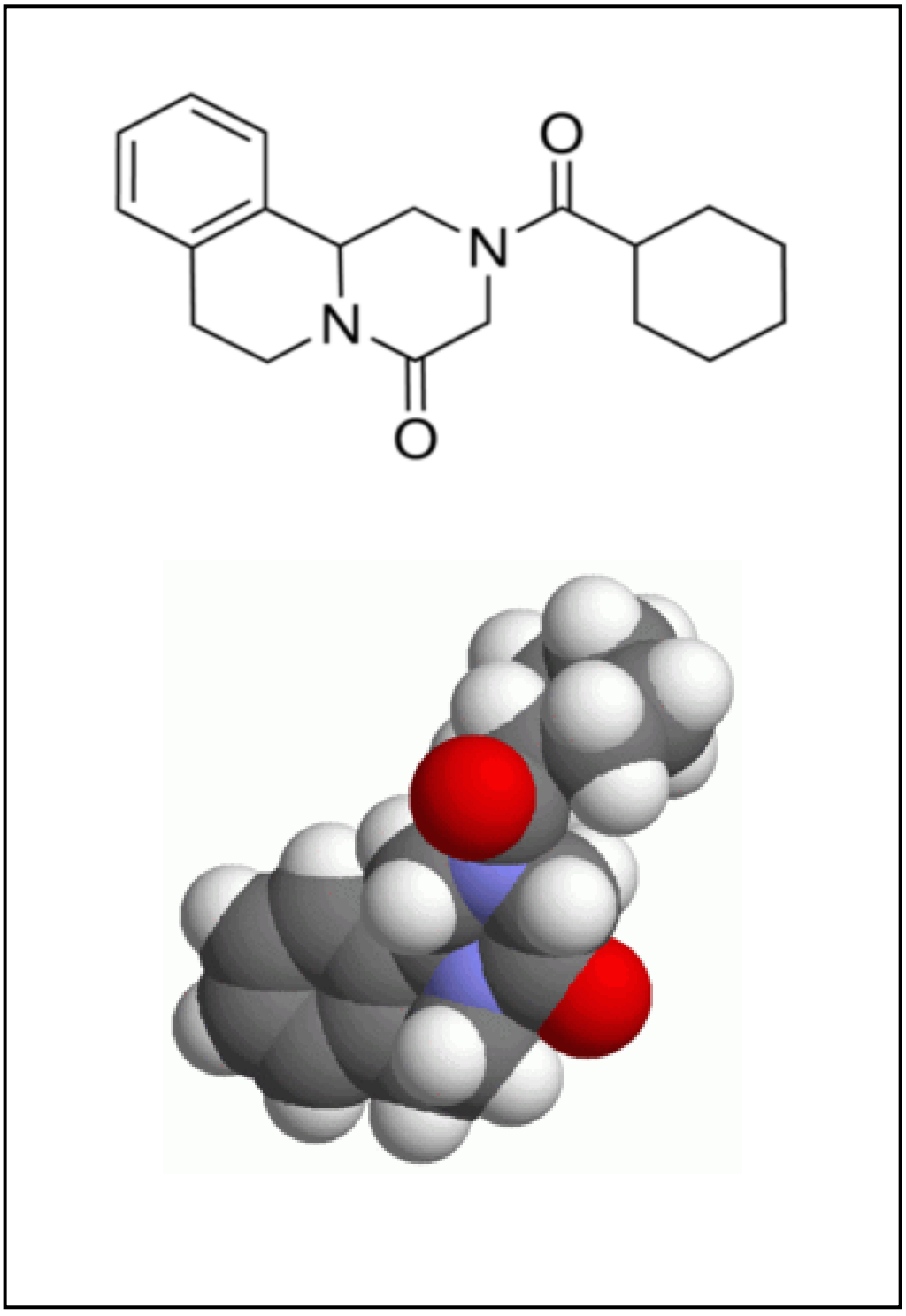


Figure 1.10: Chemical structure of Praziquantel (PZQ). Figure adapted from the website: http://www.sigmaaldrich.com/catalog/product/sial/46648?lang=en&region=ZA, http://www. Wo rld ofchemicals.com/chemicals/chemical-properties/praziquantel.html [Accessed 09 January 2014].

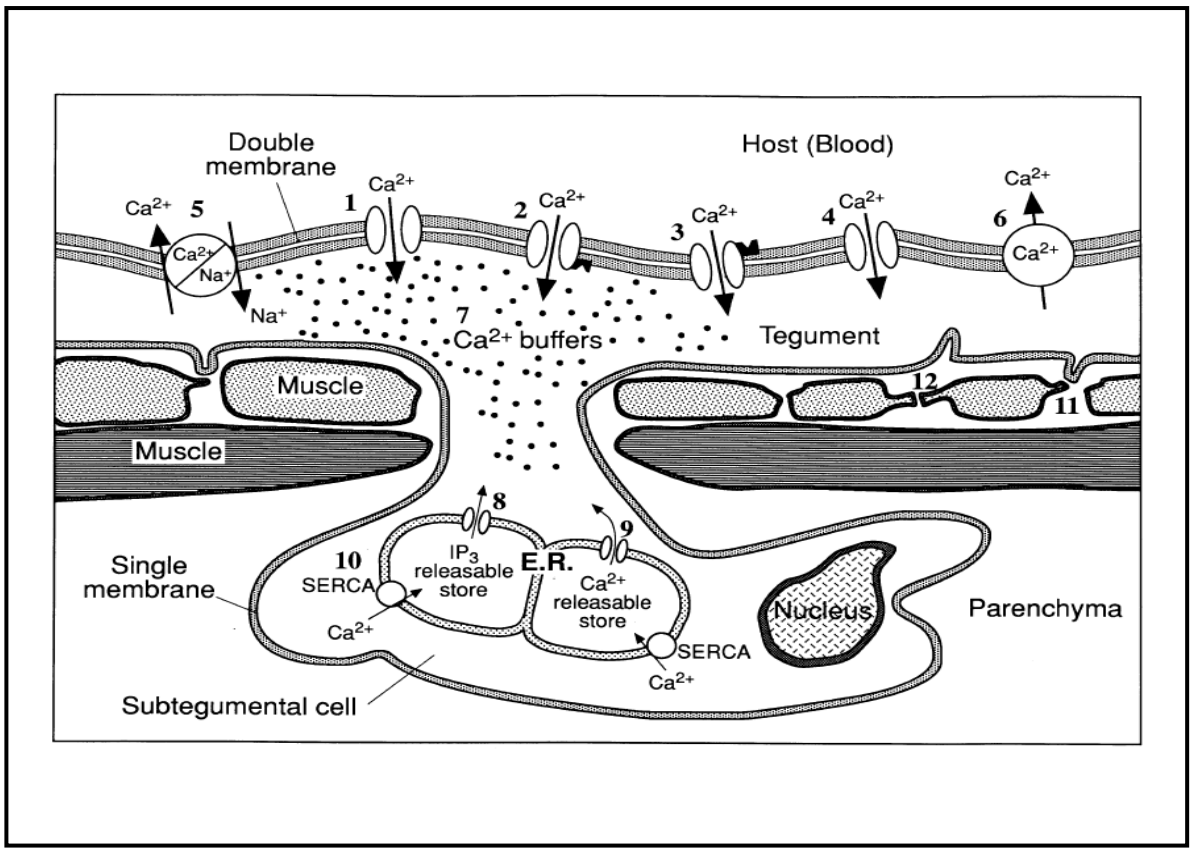


Figure 1.11: Diagram showing the possible sites of action of praziquantel on the body wall of Schistosoma mansoni and mechanisms of Ca2+ transport into and out of the tegument (Figure taken from Martin *et al*., 1997). 1: voltage-activated Ca^{2+} channel. 2: Intracellular messenger activated Ca^{2+} channel. 3: Extracellular receptor operated Ca^{2+} channel. 4: Non-selective cation channel also allowing entry of Ca^{2+}. 5: Na^{+}/Ca^{2+} exchanger. 6: Ca^{2+} ATPase pumping out Ca^{2+}. 7: Intrategumental Ca^{2+} buffers. 8: IP_3 releasable store from the sarcoplasmic reticulum. 9: Ca^{2+}-induced Ca^{2+} release channel (CICR channel). 10: Ca^{2+} ATPase pump: sarcoplasmic endoplasmic reticulum Ca^{2+} (SERCA). 11: Electrical junction between muscle cell and tegument. 12: Electrical junctions between muscle cells. If praziquantel acts like caffeine it would act on site 9.

1.7 Neurotransmitters in *S. mansoni*

Neurotransmitters are chemical messengers that expedite the transmission of nerve impulses from one neuron to another neuron or body cell. They are known to elicit their effect by binding to specific receptors, initiating excitatory and inhibitory biochemical processes. Examples include amino acids, neuropeptides, monoamines otherwise called biogenic amines and other recently classified chemicals such as nitric oxide, and the purines (adenosine and adenosine triphosphate; Ribeiro *et al*., 2005). The nervous system of *S. mansoni* is well developed and is responsible for the major coordinated activities in the parasite such as muscle contraction, which is linked to motility and energy metabolism (El-Shehabi *et al*., 2012).

Among the established classes of biogenic amines neurotransmitter present in *S. mansoni* is serotonin (5-hydroxytryptamine: 5HT) documented to be responsible for majority of the coordinated activities in the parasite, within the intermediate and the definitive host (Mbah *et al.,* 2012; Patocka *et al*., 2014; Patocka and Ribeiro, 2007). Elucidation of the mechanism of action involving this biogenic amine neurotransmitter (5HT) in *S. mansoni* is important in understanding the role of molecular switches, which have been shown to regulate adenylyl cyclase activity in *S. mansoni* (Mbah *et al.,* 2012). Be that as it may, due to the recent reports of increasing widespread resistance of some isolates of *S. mansoni* to praziquantel (PZQ) treatment; new drug targets are urgently needed. The druggable switch domains region of *S. mansoni* is acknowledged to be an attractive target towards chemotherapeutic intervention in the control and treatment of schistosomiasis (Mbah *et al.,* 2012). To this end, antimicrobial peptides (AMPs) are currently gaining prominence as possible source worthy to be explored as we search for a promising

alternative to PZQ treatment (see Chapter two) in the design of molecular templates of new therapeutic drugs in the control and treatment of infectious diseases including schistosomiasis due to their multifunctional properties.

1.8 Problem statement and motivation

Schistosomiasis is second after malaria among the prevalent neglected tropical diseases estimated to affect approximately 235 million people globally with 732 million people estimated to be at risk of infection in known endemic regions; accounting for about 280,000 deaths annually (Hotez and Kamath, 2009). Schistosoma infection is a predisposing factor to cancer and the International Agency for Research on Cancer, in association with World Health Organization; classified *S. haematobium* as a Group 1 carcinogen (Oyinloye *et al.,* 2014). South Africa is a hotspot where schistosomiasis and cancer of the bladder is endemic; KwaZulu-Natal and Mpumalanga remains a focal point of this disease (Appleton and Kvalsvig, 2006; Johnson and Appleton, 2005; Groeneveld *et al*., 1996). *S. mansoni* is known to be responsible for over 90% of all human schistosomiasis. Over the years, praziquantel remained the drug of choice in the treatment of all forms of schistosomal infection due to its high efficacy with only few and transient side effects. Unfortunately, as a result of drug pressure, cases of praziquantel resistance have emerged in some strains of *S. mansoni*. This has necessitated the search for a novel and/or new anti-schistosomal agent or molecule as an alternative to praziquantel treatment.

1.9 The main aim of the study

Structural characterization of the switch domains of the druggable G protein will involve expression and purification of sufficient quantities of the recombinant protein. Hence, the overall aim of this work is to structurally characterize the molecular switch domain regions of *Schistosoma mansoni,* which has been described in literature as a promising druggable target against the parasite. This will enhance our understanding in selective identification of small molecule inhibitors or peptides, which could be developed as novel therapeutic candidates in the control and treatment of human schistosoma infections.

1.10 Objectives

The specific objectives of this research work are as follows:

1. To perform extensive *in-silico* analysis of the switch domains of *S. mansoni* Gα protein.
2. To express and purify the switch domains of *S. mansoni* Gα protein in high quanties and in a soluble active form in *Escherichia coli* for structural studies.
3. To structurally characterize the expressed Gα protein from the switch domains of *S. mansoni*.
4. To evaluate the thermal stability of the expressed Gα protein from the switch domains of *S. mansoni*.
5. To assess the ligand binding interaction of the expressed Gα protein from the switch domains of *S. mansoni*.

REFERENCES

Adenowo A.F., Oyinloye B.E., Ogunyinka B.I. and Kappo A.P. (2015). Impact of human schistosomiasis in sub-Saharan Africa. *Brazilian Journal of Infectious Diseases.* **19**(2): 196-205.

Alebie G., Erko B., Aemero M. and Petros B. (2014). Epidemiological study on Schistosoma mansoni infection in Sanja area, Amhara region, Ethiopia. *Parasites Vectors*. **7**:15.

Almeida G.T., Lage R.C., Anderson L., Venancio T.M., Nakaya H.I., Miyasato P.A., Rofatto H.K., Zerlotini A., Nakano E., Oliveira G. and Verjovski-Almeida S. (2015). Synergy of omeprazole and praziquantel in vitro treatment against Schistosoma mansoni adult worms. *PLoS Neglected Tropical Disease*. **9**(9): e0004086.

Alonso L.M. and Alvar J. (2010). Stigmatizing neglected tropical diseases: a systematic review. *Social Medicine*. **5**(4): 218-227.

Aly H.F. (2012). Challenges in developing vaccines in parasitic diseases with special references to schistosomiasis. *International Journal of Academic Research*. **4**(4): 130-43.

Appleton C.C., and Kvalsvig J.D. (2006). A school-based helminth control programme successfully implemented in KwaZulu-Natal. *Southern African Journal of Epidemiology and Infection* **21**(2): 55-67.

Aragon A.D., Imani R.A., Blackburn V.R., Cupit P.M., Melman S.D., Goronga T., Webb T., Loker E.S. and Cunningham C. (2009). Towards an understanding of the mechanism of action of praziquantel. *Molecular and Biochemical Parasitology*. **164**(1): 57-65.

Belmonte S.L. and Blaxall B.C. (2012). Conducting the G-protein coupled receptor (GPCR) signaling symphony in cardiovascular diseases: New therapeutic approaches. *Drug Discovery Today: Disease Models*. **9**(3): e85-e90.

Bethony J.M., Cole R.N., Guo X., Kamhawi S., Lightowlers M.W., Loukas A., Petri W., Reed S., Valenzuela J.G. and Hotez, P. J. (2011). Vaccines to combat the neglected tropical diseases. *Immunological reviews*. **239**(1): 237-270.

Bornfeldt K.E. (2006). A single second messenger: several possible cellular responses depending on distinct subcellular pools. *Circulation Research* **99**: 790-792.

Boyle J.P., and Yoshino T.P. (2005). Serotonin-induced muscular activity in Schistosoma mansoni larval stages: importance of 5-HT transport and role in daughter sporocyst production. *Journal of Parasitology* **91**(3): 542-550.

Boyle J.P., Hillyer, J.F. and Yoshino, T.P. (2003). Pharmacological and autoradiographical characterization of serotonin transporter-like activity in sporocysts of the human blood fluke, Schistosoma mansoni. *Journal of Comparative Physiology A* **189**(8): 631-641.

Boyle J.P., Zaide J.V., and Yoshino T.P. (2000). Schistosoma mansoni: Effects of serotonin and serotonin receptor antagonists on motility and length of primary sporocysts in vitro. *Experimental Parasitology* **94:** 217–226.

Brindley P.J. and Sher A. (1990). Immunological involvement in the efficacy of praziquantel. *Experimental parasitology*. **71**(2): 245-248.

Coeli R., Baba E.H., Araujo N., Coelho P.M. and Oliveira G. (2013). Praziquantel treatment decreases Schistosoma mansoni genetic diversity in experimental infections. *PLoS Neglected Tropical Disease*. **7**(12): e2596.

Colley D.G., Bustinduy A.L., Secor W.E. and King C.H. (2014). Human schistosomiasis. *Lancet* **383**(9936): 2253-2264.

Coon D.R. (2005). Schistosomiasis: Overview of the history, biology, clinicopathology, and laboratory diagnosis. *Clinical Microbiology Newsletter*. **27**(21): 163-168.

Couto F.F., Coelho, P.M.Z., Araújo N., Kusel J.R., Katz N., Jannotti-Passos L.K. and Mattos A.C.A. (2011). Schistosoma mansoni: a method for inducing resistance to praziquantel using infected Biomphalaria glabrata snails. *Memorias do Instituto Oswaldo Cruz* **106**(2): 153-157.

Covic L., Gresser A. L., Talavera J., Swift S. and Kuliopulos A. (2002). Activation and inhibition of G protein-coupled receptors by cell-penetrating membrane-tethered peptides. *Proceedings of the National Academy of Sciences*. **99**(2): 643-648.

Crompton D.W.T., Engels D., Montresor A., Neira M.P. and Savioli L. (2003). Action starts now to control disease due to schistosomiasis and soil-transmitted helminthiasis. *Acta Tropica*. **86**: 121-24.

Deribe K., Meribo K., Gebre T., Hailu A., Ali A., Aseffa A., and Davey G. (2012). The burden of neglected tropical diseases in Ethiopia, and opportunities for integrated control and elimination. *Parasites & vectors* **5**(1), 240.

Deribew K., and Petros B. (2013). Efficacy of praziquantel for the treatment of schistosomiasis in Ethiopia. *International Journal of Medicine and Medical Sciences* **5**(3), 131-139.

Deupi X. and Kobilka, B. (2007). Activation of G protein–coupled receptors. *Advances in Protein Chemistry* **74**: 137-166.

Doenhoff M.J., Kusel J.R., Coles G.C., and Cioli D. (2002) Resistance of Schistosoma mansoni to praziquantel: is there a problem? *Transactions of the Royal Society of Tropical Medicine and Hygiene* **96**: 465-69.

El-Shehabi F., Taman A., Moali L.S, El-Sakkary N., and Ribeiro P. (2012). A novel G protein-coupled receptor of Schistosoma mansoni (SmGPR-3) is activated by dopamine and is widely expressed in the nervous system. *PLoS neglected tropical diseases* **6**(2): e1523.

Engels D., Chitsulo L., Montresor A. and Savioli L. (2002). The global epidemiological situation of schistosomiasis and new approaches to control and research. *Acta Tropica*. **82**: 139-46.

Enk M.J., Oliveira-Silva G. and Rodrigues N.B. (2012). Diagnostic Accuracy and Applicability of a PCR System for the Detection of Schistosoma mansoni DNA in Human Urine Samples from an Endemic Area. *PLoS One*. **7**(6): e38947.

Fallon P.G. (1998). Schistosome resistance to praziquantel. *Drug Resistance Updates*. **1**(4): 236-241.

Franco-Paredes C. and Santos-Preciado J.I. (2015). The Neglected Tropical Diseases in Latin America and the Caribbean: Burden of Disease and Approaches for Elimination and Control. In *Neglected Tropical Diseases-Latin America and the Caribbean* (pp. 1-10). Springer Vienna.

Grant K.R., Harnett W., Milligan G., and Harnett M.M. (1997). Differential G protein expression during B- and T-cell development. *Immunology* **90**: 564–571.

Greenberg R.M. (2013). New approaches for understanding mechanisms of drug resistance in schistosomes. *Parasitology*. **140**(12): 1534-1546.

Groeneveld A.E., Marszalek W.W., and Heyns C.F. (1996). Bladder cancer in various population groups in the greater Durban area of KwaZulu-Natal, South Africa. *British Journal of Urology* **78**(2): 205-208.

Gryseels B., Polman K., Clerinx J. and Kestens L. (2006). Human schistosomiasis. *The Lancet*. **368**(9541): 1106-1118.

Halazy S. (2003). Signal transduction: an exciting field of investigation for small molecule drug discovery. *Molecules* **8:** 349–358.

Hatley M.E., Lockless S.W., Gibson S.K., Gilman A.G. and Ranganathan R. (2003). Allosteric determinants in guanine nucleotide-binding proteins. *Proceedings of the National Academy of Sciences*. **100**(24): 14445-14450.

Hotez P. (2011). Enlarging the “audacious goal”: elimination of the world's high prevalence neglected tropical diseases. *Vaccine*. **29**: D104-D110.

Hotez P.J. and Kamath A. (2009). Neglected tropical diseases in sub-saharan Africa: review of their prevalence, distribution, and disease burden. *PLOS Neglected Tropical Diseases*. **3**(8): e412.

Hotez P.J. Pecoul B. Rijal S., Boehme C., Aksoy S., Malecela M., Tapia-Conyer R. and Reeder J.C. (2016). Eliminating the neglected tropical diseases: translational science and new technologies. *PLOS Neglected Tropical Diseases*. **10**(3): e0003895.

Hotez P.J., Fenwick A., Savioli L. and Molyneux, D.H. (2009). Rescuing the bottom billion through control of neglected tropical diseases. *Lancet*. **373**(9674): 1570-1575.

Hudson D.B., Ulven T. and Milligan G. (2013). The therapeutic potential of allosteric ligands for free fatty acid sensitive GPCRs. *Current Topics in Medicinal Chemistry*. **13**(1): 14-25.

Hughes S.M. (1983). Are guanine nucleotide binding proteins a distinct class of regulatory proteins?. *FEBS Letters*. **164**(1): 1-8.

Hurley J.H. (1999). Structure, mechanism, and regulation of mammalian adenylyl cyclase. *Journal of Biological Chemistry*. **274**(12): 7599-7602.

Ishii J., Fukuda, N., Tanaka, T., Ogino, C. and Kondo, A. (2010). Protein–protein interactions and selection: yeast-based approaches that exploit guanine nucleotide-binding protein signaling. *FEBS Journal*, **277**(9): 1982-1995.

Ishikawa Y., and Homcy C.J. (1997). The adenylyl cyclases as integrators of transmembrane signal transduction. *Circulation Research*. **80**(3): 297-304.

Jacoby E., Bouhelal R., Gerspacher M., and Seuwen, K. (2006). The 7 TM G-protein-coupled receptor target family. *ChemMedChem*. **1**(8): 760-782.

Janero D.R. and Thakur G.A. (2016). Leveraging allostery to improve G protein-coupled receptor (GPCR)-directed therapeutics: cannabinoid receptor 1 as discovery target. *Expert Opinion on Drug Discovery* **11**(12): 1223-1237.

Johnson C.L., Appleton C.C. (2005). Urban schistosomiasis transmission in Pietermaritzburg, South Africa. *South Africa Journal of Epidemiology and Infection* **20**(3): 103-107.

Kappagoda S., and Ioannidis J. (2014). Prevention and control of neglected tropical diseases: overview of randomized trials, systematic reviews and meta-analyses. *Bulletin of the World Health Organization*. **92**(5): 356-366C.

Karchin R., Karplus K., and Haussler D. (2002). Classifying G-protein coupled receptors with support vector machines. *Bioinformatics*. **18**(1): 147-159.

Katritch V., Cherezov V., and Stevens R.C. (2013). Structure-function of the G protein–coupled receptor superfamily. *Annual Review of Pharmacology and Toxicology*. **53**: 531-556.

King C.H. (2010). Parasites and poverty: the case of schistosomiasis. *Acta Tropica*. **113**: 95-104.

King C.H. and Bertino A.M. (2008). Asymmetries of poverty: why global burden of disease valuations underestimate the burden of neglected tropical diseases. *PLOS Neglected Tropical Diseases* **2**(3): e209.

King C.H., Dickman K. and Tisch D.J. (2005). Reassessment of the cost of chronic helmintic infection: a meta-analysis of disability-related outcomes in endemic schistosomiasis. *Lancet*. **365**(9470): 1561-1569.

Kobilka B.K. (2007). G protein coupled receptor structure and activation. *Biochimica et Biophysica Acta (BBA)-Biomembranes*. **1768**(4): 794-807.

Krishnan A. and Schiöth H. B. (2015). The role of G protein-coupled receptors in the early evolution of neurotransmission and the nervous system. *Journal of Experimental Biology*. **218**(4): 562-571.

Kroeze W.K., Sheffler D.J. and Roth B.L. (2003). G-protein-coupled receptors at a glance. *Journal of Cell Science*. **116**(24): 4867-4869.

Lambright D.G., Sondek J., Bohm A. and Skiba N.P. (1996). The 2.0 angstrom crystal structure of a heterotrimeric G protein. *Nature*. **379**(6563): 311- 319.

Leckie C.P., McAinsh M.R., Montgomery L., Priestley A.J, Staxen I., Webb A.A.R., and Hetherington A.M. (1998). Second messengers in guard cells. *Journal of Experimental Botany* **49**: 339-349.

Levin L.R., and Reed R.R. (1995). Identification of functional domains of adenylyl cyclase using in vivo chimeras. *Journal of Biological Chemistry* **270**(13): 7573-7579.

Ligeti E., Welti S. and Scheffzek K. (2012). Inhibition and termination of physiological responses by GTPase activating proteins. *Physiological Reviews*. **92**(1): 237-272.

Linder J.U. (2008). Structure–function relationships in Escherichia coli adenylate cyclase. *Biochemical Journal*. **415**(3): 449-454.

Lobo D.A., Velayudhan R., Chatterjee P., Kohli H. and Hotez P. J. (2011). The neglected tropical diseases of India and South Asia: review of their prevalence,

distribution, and control or elimination. *PLOS Neglected Tropical Diseases*. **5**(10): e1222.

Mackey T.K., Liang B.A., Cuomo R., Hafen R., Brouwer K.C. and Lee D.E. (2014). Emerging and re-emerging neglected tropical diseases: a review of key characteristics, risk factors, and the policy and innovation environment. *Clinical microbiology reviews*. **27**(4): 949-979.

Martin R.J., Robertson A.P. and Bjorn H. (1997). Target sites of anthelmintics. *Parasitology*. **114**(7): 111-124.

Martins-Melo F.R., Ramos Jr A.N., Alencar C.H. and Heukelbach J. (2016). Mortality from neglected tropical diseases in Brazil, 2000–2011. *Bulletin of the World Health Organization*. **94**(2): 103-110.

Mbah A.N., Kamga H.L., Awofolu O.R. and Isokpehi R. D. (2012). Drug target exploitable structural features of adenylyl cyclase activity in Schistosoma mansoni. *Drug Target Insights*. **6**: 41-58.

McCudden C.R., Hains M.D., Kimple R.J., Siderovski D.P., and Willard F.S. (2005). Gprotein signaling: Back to the future. *Cellular and Molecular Life Sciences* **62**: 551–577.

Mendonca A.M., Feitosa A.P., Veras D.L., Matos-Rocha T.J., Cavalcanti M.G., Barbosa C.C., Brayner F.A. and Alves L.C. (2016). The susceptibility of recent isolates of Schistosoma mansoni to praziquantel. *Revista do Instituto de Medicina Tropical de São Paulo*. **58**: 7.

Milligan G., and Kostenis E. (2006). Heterotrimeric G-proteins: a short history. *British Journal of Pharmacology* **147** (Suppl. 1): S46-S55.

Miranda M.A., Kuehn C.C., Cardoso J.F.R., Oliveira L.G.R., Magalhaes L.G., Tiossi R.F.J., Rodrigues V., Zucolloto S., Prado J.C., McChesney J.D. and Bastos J.K. (2013). Immunomodulatory effect of the alkaloidic extract of Solanum lycocarpum

fruits in mice infected with Schistosoma mansoni. *Experimental Parasitology*. **133**: 396-402.

Mtethiwa A.H., Nkwengulila G., Bakuza J., Sikawa D. and Kazembe A. (2015). Extent of morbidity associated with schistosomiasis infection in Malawi: a review paper. *Infectious diseases of poverty*. **4**(1): 25.

Nichols D.E. and Nichols C.D. (2008). Serotonin receptors. *Chemical reviews*. **108**(5): 1614-1641.

Oldham W.M., and Hamm H.E. (2008). Heterotrimeric G protein activation by G-protein-coupled receptors. *Nature Reviews Molecular Cell Biology* **9**(1): 60-71.

Oyinloye B., Adenowo F., Gxaba N. and Kappo A. (2014). The Promise of Antimicrobial Peptides for Treatment of Human Schistosomiasis. *Current Drug Targets*. **15**(9): 852-859.

Patocka N. and Ribeiro P. (2007). Characterization of a serotonin transporter in the parasitic flatworm, Schistosoma mansoni: cloning, expression and functional analysis. *Molecular and Biochemical Parasitology*. **154**: 125-133.

Patocka N., Sharma N., Rashid M. and Ribeiro P. (2014). Serotonin Signaling in Schistosoma mansoni: A Serotonin-Activated G Protein-Coupled Receptor Controls Parasite Movement. *PLoS Pathogens*. **10**(1): e1003878.

Ramesh M. and E Soliman M. (2015). G-protein coupled receptors (GPCRs): a comprehensive computational perspective. *Combinatorial Chemistry and High Throughput Screening*. **18**(4): 346-364.

Reimann F. and Gribble F.M. (2016). G protein-coupled receptors as new therapeutic targets for type 2 diabetes. *Diabetologia*. **59**(2): 229-233.

Ribeiro P. and Patocka N. (2013). Neurotransmitter transporters in schistosomes: structure, function and prospects for drug discovery. *Parasitology international*. **62**(6): 629-638.

Ribeiro P., El-Shehabi F. and Patocka N. (2005). Classical transmitters and their receptors in flatworms. *Parasitology*. **131**: S19-S40.

Ritter S.L. and Hall R.A. (2009). Fine-tuning of GPCR activity by receptor-interacting proteins. *Nature Reviews Molecular Cell Biology*. **10**(12): 819-830.

Ross A.G., Bartley P.B., Sleigh A.C., Olds G.R., Li Y., Williams G.M., and McManus D.P. (2002). Schistosomiasis. *New England Journal of Medicine***. 346**: 1212-20.

Sady H., Al-Mekhlafi H.M., Mahdy M.A., Lim Y.A., Mahmud R. and Surin J. (2013). Prevalence and associated factors of schistosomiasis among children in Yemen: implications for an effective control programme. *PLOS Neglected Tropical Diseases*. **7**(8): e2377.

Salon J.A., Lodowski, D.T. and Palczewski, K. (2011). The significance of G protein-coupled receptor crystallography for drug discovery. *Pharmacological Reviews* **63**(4): 901-937.

Sassone-Corsi P. (2012). The cyclic AMP pathway. *Cold Spring Harbor perspectives in biology* **4**(12): a011148.

Schistosomiasis: Sequencing the Parasites' Genes. http://www.genomenewsnetwork.org/articles/10_03/schisto.shtml [Accessed 09 January 2014].

Smith N.J. (2012). Low affinity GPCRs for metabolic intermediates: challenges for pharmacologists. *Frontiers in Endocrinology*. **3**:1.

Stanley R.J. and Thomas G.M. (2016). Activation of G proteins by guanine nucleotide exchange factors relies on GTPase activity. *PloS One*. **11**(3): e0151861.

Steinmann P., Keiser J., Bos R., Tanner M. and Utzinger J. (2006). Schistosomiasis and water resources development: systematic review, meta-analysis, and estimates of people at risk. *Lancet Infectious Disease*. **6**: 411-425.

Suzuki K. and Kaneko-Kawano T. (2016). Biological roles and therapeutic potential of G protein-coupled receptors for free fatty acids and metabolic intermediates. *Journal of Physical Fitness and Sports Medicine*. **5**(3): 213-227.

Tall G.G., Krumins A.M. and Gilman A.G. (2003). Mammalian Ric-8A (synembryn) is a heterotrimeric Gα protein guanine nucleotide exchange factor. *Journal of Biological Chemistry*. **278**(10): 8356-8362.

Tambo E., Khater E. I., Chen J. H., Bergquist R. and Zhou X. N. (2015). Nobel prize or the artemisinin and ivermectin discoveries: a great boost towards elimination of the global infectious diseases of poverty. *Infectious Diseases of Poverty*. **4**(1): 58.

Tan S.Y. and Ahana A. (2007). Theodor Bilharz (1825-1862): discoverer of schistosomiasis. *Singapore Medical Journal*. **48**(3): 184–185.

Tresguerres M., Levin L.R. and Buck J. (2011). Intracellular cAMP signaling by soluble adenylyl cyclase. *Kidney International*. **79**(12): 1277-1288.

Tuteja N. (2009). Signaling through G protein coupled receptors. Plant Signaling and Behaviour **4**(10): 942-947.

Vaidehi N., Floriano W.B., Trabanino R., Hall S.E., Freddolino P., Choi E.J., Zamanakos G. and Goddard W.A. (2002). Prediction of structure and function of G protein-coupled receptors. *Proceedings of the National Academy of Sciences*. **99**(20): 12622-12627.

Van Neuren A.S., Müller G., Klebe G. and Moroder L. (1999). Molecular modelling studies on G protein-coupled receptors: from sequence to structure?. *Journal of Receptors and Signal Transduction*. **19**(1- 4): 341-353.

Vogler O., Barcelo J.M., Ribas, C. and Escriba, P.V. (2008). Membrane interactions of G proteins and other related proteins. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1778**(7): 1640-1652.

Wang D.Z., Dong H.P., Li C., Xie Z.X., Lin L. and Hong H.S. (2011). Identification and characterization of cell wall proteins of a toxic dinoflagellate Alexandrium catenella using 2-D DIGE and MALDI TOF-TOF mass spectrometry. *Evidence-Based Complementary and Alternative Medicine*. **2011**: 984080.

Wang W., Wang L. and Liang Y.S. (2012). Susceptibility or resistance of praziquantel in human schistosomiasis: a review. *Parasitology Research*. **111**: 1871-877.

Wu G.Y. and Halim M.H. (2000). Schistosomiasis: progress and problems. *World journal of gastroenterology*. **6**(1): 12.

Zamanian M., Kimber M.J., McVeigh P., Carlson S.A., Maule A.G. and Day T.A. (2011). The repertoire of G protein-coupled receptors in the human parasite Schistosoma mansoni and the model organism Schmidtea mediterranea. *BMC Genomics*. **12**: 596.

Zhao J., Deng Y., Jiang Z. and Qing, H. (2016). G Protein-Coupled Receptors (GPCRs) in Alzheimer's Disease: A Focus on BACE1 Related GPCRs. *Frontiers in Aging Neuroscience* **8**: 58.

Zongo D., Kabre B.G., Dayeri D., Savadogo B. and Poda J.N. (2012). Comparative study of schistosomiasis transmission (urinary and intestinal forms) at 10 sites in Burkina Faso (in sub-Saharan Africa). *Medecine et Sante Tropicales*. **22**(3): 323-329.

CHAPTER TWO

THE PROMISE OF ANTIMICROBIAL PEPTIDES FOR TREATMENT OF HUMAN SCHISTOSOMIASIS

A review paper published in *Current Drug Targets,* 2014*, 15,* 852-859

Babatunji Oyinloye[*], Fatimah Adenowo, Nomagugu Gxaba and Abidemi Kappo[#]

2.1 Abstract

Schistosomiasis sometimes referred to as bilharzia ranks high among the common neglected human tropical diseases. Parasitic trematode flatworm belonging to the genus Schistosoma is responsible for this acute and chronic disease. Its prevalence is rapidly increasing in the tropical regions worldwide; however, its economic and global health impact is grossly underestimated. There are five recognized species of schistosome parasitizing humans but the common species causing the disease are *S. haematobium, S. japonicum* and *S. mansoni*. Over the years, praziquantel, due to its advantage over other drugs employed in the treatment of schistosomiasis especially because of its effectiveness against all schistosoma species has remained the drug of choice. Unfortunately, due to drug pressure, some reports of praziquantel resistance in the treatment of some strains of *S. haematobium* and *S. mansoni* have been documented in literature. This has necessitated the search for novel

antischistosomal agents as alternatives to praziquantel treatment. Currently, antimicrobial peptides are gaining prominence as possible sources of novel drugs in the control and treatment of schistosomiasis. A major driving force for evaluating antimicrobial peptides is their general diverse therapeutic applications, basically due to their size and properties as well as their broad spectrum of activities. Therefore, antimicrobial peptides are attractive candidates with promising results that may overcome drug resistance problems in search for novel therapeutic agents for the control and treatment of schistosomiasis.

Keywords: Antimicrobial peptides, cercaria, eggs, neglected tropical diseases, praziquantel, schistosomiasis, trematode.

2.2 Introduction

Schistosomiasis is chief among the poignant vulnerable neglected diseases that is rapidly increasing in prevalence in the tropical regions worldwide, as a consequence of stretch in social infrastructures arising from population growth, increasing poverty as well as poor health facilities and poor sanitation due to nonchalant attitude of the people in tropical regions. The economic and global health impact of this disease is grossly underestimated [1]. Data available in the literature revealed that in known endemic regions, on the average about 235 million people are suffering from infections with schistosomiasis worldwide while there is possibility for about 732 million people to also get infected [2]. The disease is commonly found in developing countries in Asia, South America, Middle East, the Caribbean and Africa [3, 4].

Schistosomiasis is second after malaria among the prevalent neglected tropical diseases. Schistosoma infection is a predisposing factor to cancer and International Agency for Research on Cancer, in association with World Health Organization; classified *S. haematobium* as Group 1 carcinogen. Over the years, praziquantel, due to simple administration, high efficacy, excellent tolerability, few and transient side effects as well as its competitive cost, remained the preferred drug administered to patients in the hospitals during schistosomiasis treatment, basically because it is effective in combating all schistosoma species in a single or a few oral doses [5].

Schistosomiasis, sometimes called bilharzia, is an acute and chronic disease associated with infection by parasitic (trematode) flatworms belonging to the genus Schistosoma; it continues to affect millions of people, especially in the rural areas of the tropical regions and it accounts for more than 280,000 deaths annually. Theodore Bilharz, a German surgeon discovered the parasite responsible for

schistosomiasis in Egypt in 1851 [8, 9]. Five species of schistosome have been recognized to affect humans and these are: *S. japonicum, S. intercalatum, S. haematobium, S. mekongi* and *S. mansoni*. The common species associated with the disease are *S. japonicum*, *S. haematobium* and *S. mansoni* [10, 11]. Despite intense progress in research over the past decade, this disease still remains a public health concern. Recently, Zongo and co-workers reported that the prevalence rates are higher in school-aged boys than girls, but among adults the prevalence is higher in females than males [12].

Diagnosis of schistosomiasis is routinely confirmed by microscopic examination of stools or urine samples for parasite eggs, serologic tests, PCR and imaging [13]. Interruption of the parasites lifecycle such as application of molluscicides, biological control (introduction of competitor snail species), alteration of the aquatic environment and treatment with the drug praziquantel have been employed in the control and treatment of schistosomiasis [14].

Unfortunately, due to drug pressure, in recent reports authors documented that some strains of *S. haematobium* and *S. mansoni* in Egypt, Kenya and Senegal have shown worrisome signs of resistance to treatment with praziquantel [6, 7]. This has necessitated the search for a novel anti-schistosomal agent as an alternative to praziquantel treatment. In many communities worldwide, several extracts with bioactive constituents originating from living organisms have been engaged in the treatment of parasitic diseases, including schistosomiasis. Currently, antimicrobial peptides are gaining prominence as possible sources in the design of molecular templates of novel drugs in the control and treatment of schistosomiasis. Antimicrobial peptides have been extracted from different plant and animal sources.

Of great interest are those that could be isolated from different vertebrate species, including amphibians, fish and mammals.

Antimicrobial peptides have been reported to play a crucial role in host defence in the presence of an adaptive immune response [15]. A major driving force for evaluating antimicrobial peptides is their general diverse therapeutic applications, basically due to their size and properties as well as their broad spectrum of activities [16]. Therefore, antimicrobial peptides are attractive candidates with promising results that may overcome drug resistance problems in search for novel therapeutic agents for the treatment of schistosomiasis, due to their size and properties.

2.3 Epidemiology

Schistosomiasis remains a global public health challenge and a major tropical neglected disease. The occurrence of schistosomiasis differs considerably from one geographical location to the other as a result of lifestyle and environmental risk factors. Schistosomiasis is known to be prevalent in over 75 countries found in the regions of Asia, South America, the Caribbean and Africa, where it occurs predominantly in the rural areas than the urban centres. The most interesting thing about these countries where this disease is rampant is that; they are located between 36° north and south of the equator. Added to this, they have optimal freshwater temperatures for survival of the snails that serve as intermediate hosts for human Schistosomiasis [17, 18].

Despite all efforts to control and eradicate this disease, its prevalence has been rising progressively all over the world. Literature search has shown that even though great success has been recorded in some part of the world concerning control and distribution of schistosomiasis, the number of people infected or projected to be

affected is still on the increase [1-4]. Consequently, human schistosomiasis is regarded one of the most prevalent and incapacitating neglected diseases in tropical and subtropical regions of the world [4].

Schistosomiasis is an acute and chronic insidious disease because it disables men and women during their most productive years due to the fact that it is poorly recognised and diagnosed at early stages [1]. Although, higher rate of occurrence has been observed and documented to affect children less than 14 years of age in many endemic areas [2, 8]; it is interesting to note that after the age of 20 years, there is a general reduction in the rate of occurrence among the subjects in endemic areas, possibly as a result of increasing immunity, or perhaps because of behavioural changes leading to reduced exposure to fresh water [19].

2.4 Classification of causative agents

Parasitic flatworm of the genus Schistosoma transmits schistosomiasis. The parasite Schistosoma has been reported to display some forms of disparities among species, strains and genera, such as, the levels of pathogenicity, infectivity and immunogenicity. These factors are said to differ among parasite populations [20-22]. Three out of the five recognized species of schistosome parasitizing humans are of major pathological importance and have been reported in literature [7, 8, 23]. *S. mansoni* and *S. haematobium* are of major importance in human schistosomiasis because they account for about 95 % of all human schistosoma infection [2, 24].

S. mansoni is associated (causative agent) with intestinal schistosomiasis, it is predominantly found in Africa, foci in Middle East, as well as some part of Caribbean Islands and South America. It can infect all age groups, but the prevalence is said to be high in children less than 14 years of age [8]. Adult *S. mansoni* worms (male and

female) reside in the mesenteric veins of the infected host, the male surrounding the female in a gynaecophoric canal. Approximately, the adult female worms produce 300 eggs per pair daily [25, 26]. De Man and co-workers reported that *S. mansoni* infection might result in severe gastrointestinal disturbances, which are characterised by abdominal pain, hyperactivity of intestinal muscle, vomiting, nausea and diarrhoea [27].

S. haematobium is known to be responsible for urinary schistosomiasis. It is most frequent in Africa with smaller foci in the Middle East, Turkey, and India. It lives in the blood vessels of the bladder and genitalia of the host, it is the most widely spread among the schistosoma infection that affect humans and it is distinguished by the presence of blood in urine [28]. *S. haematobium* commonly develop rap-idly in the venous plexus of the urinary bladder of humans. In the process, the eggs deposited by adult worms eventually block the venous plexus, obstructing free blood flow. Eventually, the veins become weak and burst, allowing blood flow and eggs to be deposited in the bladder. This results in chronic granulomatous inflammation in the mucosal and submucosal of the bladder [29, 30].

S. japonicum is common in Asia, mainly in China, the Philippines, Thailand and Indonesia. It is responsible for hepatic or intestinal schistosomiasis. Of the three schistosome species parasitizing humans, *S. japonicum*, is recognised to cause the most serious disease and it is reported to be the most difficult to control [31]. This may be due to the fact that *S. japonicum* female worms have a much higher egg output and that the eggs are laid in large aggregates that induce intensive tissue reactions in host organs as well as its zoonotic nature. In addition, the life span of the adult worm is probably the longest [32-34]. Acute infection may result in fever, weakness, diarrhoea, abdominal pain and hepatomegaly. Chronic disease involves

granuloma formation, tissue inflammation, liver lesions and fibrosis, which may persist even after infection has been cleared [12, 35].

2.5 Life cycle of schistosome

Schistosomiasis is characterized by a complex life cycle alternating between two important phases: a sexual phase in humans (definitive host) and an asexual phase in snail (intermediate host). The life cycles of the five recognized species of schistosome known to parasitize humans are similar. The life cycle begins when the eggs of the parasites are released from an infected individual in urine or feces; the eggs find their way into fresh water, on contact with fresh water they hatch and become free-swimming miracidia. The miracidia will in turn be transformed into sporocyst when they find their way into the snail's tissue. The sporocyst will produce new parasites, known as cercariae (larvae) capable of infecting mammals [36, 37].

While inside the water or shortly after leaving; infection occurs as the cercaria penetrates human skin and loses its tail in the process, and transforms into schistosomulum which enters the bloodstream and migrates through various tissues to the site of maturation where it develops into adult male or female schistosomes. Depending on the schistosoma species, adult worms mature in pairs in the veins surrounding the bladder, intestines or liver. *S. japonicum* and *S. mansoni* mature in veins around the liver and intestines while *S. haematobium* matures in veins around the bladder. Female worms (*S. japonicum* and *S. mansoni*) produce eggs, which are transported from the veins to the lumen of the liver or intestine whereas in *S. haematobium* the eggs produced are transported to the bladder. Majority of the eggs are eliminated from the body in urine or feces, the eggs then hatch in fresh water, liberating miracidia to restart the life cycle [38, 39].

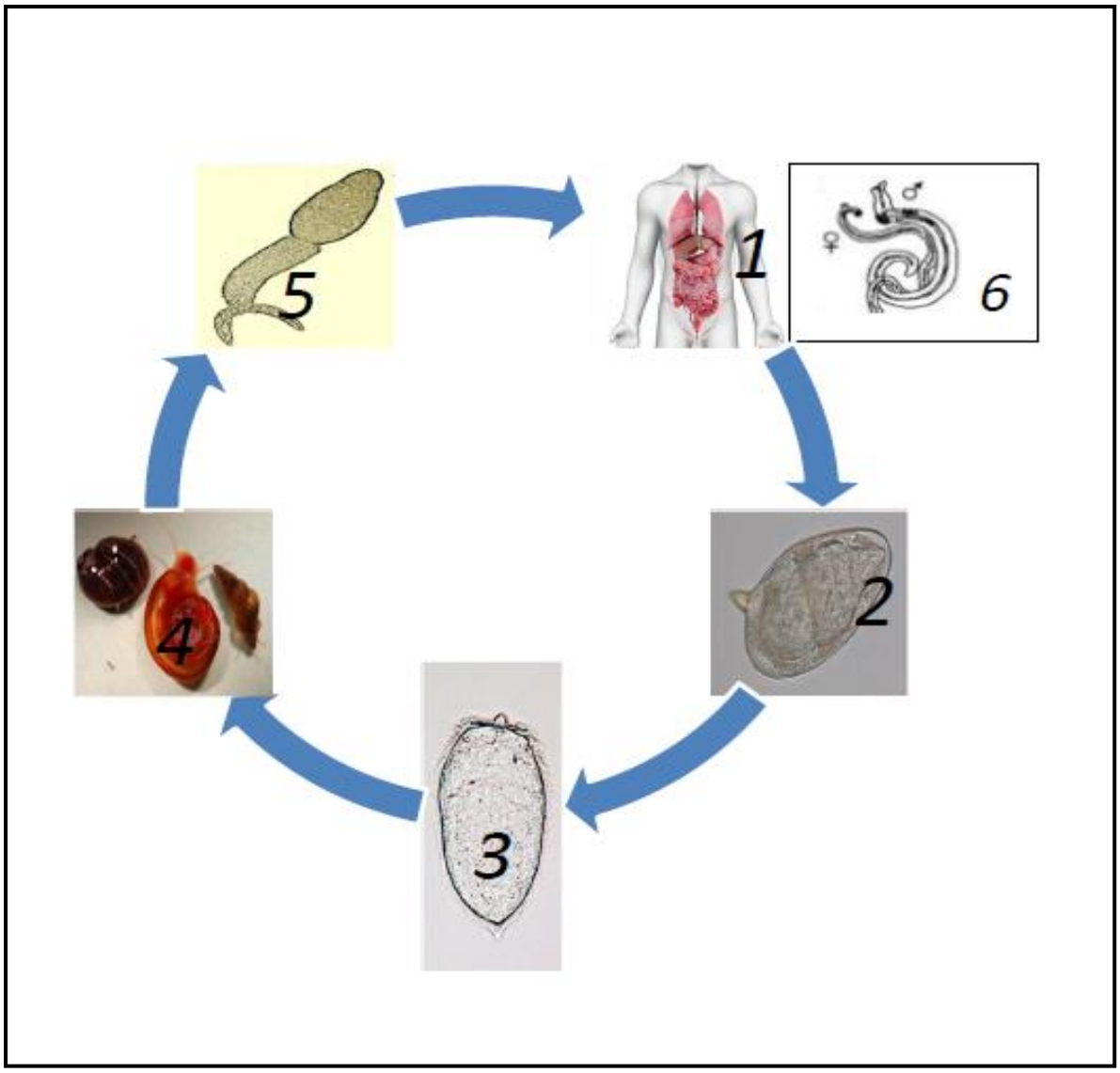


Figure 2.1: Schistosome life cycle1. Definitive host, 2. Schistosome eggs released in urine or feces of definitive host, 3. Free-swimming miracidia from eggs, 4. Intermediate host, 5. Cercaria (penetrates skin, sheds its tail and transforms into schistosomulum), 6. Paired adult worm (schistosomulum migrates to the hepatic portal system; adult worms mature in pairs in the veins surrounding the bladder, intestines or liver. They produced eggs, which majorities are eliminated in urine or feces of definitive host to the environment. The eggs hatch, liberating miracidia to restart the life cycle

2.6 Pathogenesis and mechanism of action in schistosome infection

It is well documented in literature that various organs within the human body can be affected as a result of schistosome infection [40]. The organs affected depend largely on the schistosomal species responsible for the infection and this also dictates the clinical presentation of the disease. Progression of schistosomiasis is a multifactorial and multistep process that involves many environmental risk factors. Pathogenesis and clinical manifestation of the disease involves three stages; these are: the early stage which begins with the initial penetration of human skin by cercaria and its transformation to schistosomula, resulting in hypersensitivity reaction against the migrating schistosomula and skin rash with pruritus, lasting for 1-2 days [41].

The next stage is the acute (intermediate) stage where the schistosomula moves through the bloodstream and migrates through various tissues to the appropriate site, depending on the Schistosoma species. The syndrome is often referred to as Katayama fever, with rash, cough, abdominal pain, fever, malaise, diarrhoea, nausea, lymphadenopathy, and eosinophilia as noticeable symptoms four to 8 weeks after exposure. In acute infections, mucoid bloody diarrhoea together with tender hepatomegaly may occur due to heavy exposure to *S. mansoni* or *S. japonicum*. [41].

The third stage is known as the chronic stage commonly referred to as chronic schistosomiasis. Secretion of proteolytic enzymes, which induce chronic eosinophilic inflammatory and granulomatous reactions, occurs as a result of retained or entrapped eggs in the various infected organs and this is responsible for the disease

syndrome [41]. The worm burden is associated with the severity of symptoms displayed in chronic schistosomiasis. At this stage of infection, obstruction to blood flow is common, even though persons with low or moderate worm burdens may be without obvious symptoms [41]. Various organs within the human body can be affected by chronic infection including the reproductive organs, brain, gut and the lungs giving rise to the development of urinary/genital schistosomiasis, neuro-schistosomiasis, pulmonary schistosomiasis and hepatointestinal schistosomiasis respectively. Recently, chronic urinary schistosomiasis has been associated with squamous bladder cancer [42, 43]. Both immune and non-immune mechanisms may contribute to the pathogenesis of schistosomiasis [44]. Ability to infect, invade and evade the host's immune mechanism is inherent in the entire *Schistosoma* species parasitizing human [40]. These *Schistosoma* species appears to have developed several mechanisms by which the host's immune system is down-regulated in order to enhance their own survival while they migrate to their various sites of maturation [45].

2.7 Role of schistosome in cancer of the bladder

Incidence of cancer of the bladder is four times higher in men than in women, and this is the commonest malignancy of the urinary tract [46]. Schistosoma infection has been linked with various types of malignancy such as carcinoma of the bladder, liver, uterus and the intestine [47]. Bladder cancer is frequently linked to urinary schistosomiasis caused as a result of infection by *S. haematobium* [48, 49]. In fact, bladder cancer is the most common cancer in men and the second in women in regions where *S. haematobium* is endemic, accounting for as much as 30 % of all documented cancer incidence [50]. International Agency for Research on Cancer, in

association with World Health Organization; classified *S. haematobium* as Group 1 carcinogen [51, 52]. However, underlying molecular and cellular mechanisms associating *S. haematobium* infection with cancer development is still poorly elucidated [53].

Nevertheless, in relation to schistosoma infection, observed chronic inflammation and irritation in urinary bladder was reported recently. This was believed to have played a prominent role in facilitating cancer initiation at the site of inflammation [54, 55]. Macrophages and neutrophils (inflammatory cells) are important sources of endogenous free radicals, these radicals are associated with the formation of carcinogenic N-nitrosamines [56], which may elicit deleterious effects, such as mutations [57], sister chromatid exchanges [58], and breaks of the DNA strand [59]. Furthermore, inflammatory cells have also been implicated in the activation of pro-carcinogens, such as aromatic amines and polycyclic aromatic hydrocarbons, to their final carcinogenic metabolites with carcinogenic effect [60].

In another report, it was observed that the levels of N-nitrosamines in urine and nitrate-reducing bacteria were higher in patients with schistosomal infections when com-pared with healthy individual [50, 61-63]. This observation gives credibility to the notion that nitrosamines play a crucial role in bladder cancers associated with schistosomal infections [64-67]. Therefore, it is believed that inflammatory cells mediated by schistosomal infections in the urinary bladder of patients suffering from schistosomiasis may enhance the potential of cancer formation of aromatic amines [50].

2.8 Antischistosomal drugs and challenges

In the last decade, notable progress has been made in the reduction of morbidity and mortality due to schistosoma infection, as a result of chemotherapy employed globally as the major approach in the control and treatment of human schisosomiasis. At the moment, only three effective antischistosomal drugs are used: metrifonate, oxamniquine and praziquantel [68, 69]. Metrifonate is only efficient in the treatment of *S. haematobium* infection and inefficient in the treatment of other human schistosomal infection. Oxamniquine is only efficient in the treatment of S. *mansoni* infection, it is known to be more active against male than female worms, and has little effect on juvenile worms [70]. Praziquantel (PZQ) is efficient when administered in the treatment of all forms of schistosomal infection. It has become the most acceptable, effective, common and widely used drug administered in the treatment of all forms of schistosomal infection; it has been employed in the treatment of schistosomiasis for over 20 years [18, 71, 72], but its effectiveness is limited by its known side effects [68] as well as its inadequacy to destroy parasitic infection within 2 to 4 weeks post-infection [38]. Common side effects experienced during treatment with PZQ, which necessitated the quest for alternative drug in treating this infection include but not limited to dizziness, mucoid diarrhoea, abdominal pain, vomiting, weakness, and headache, while uncommon side effects such as joint pains and swellings, myalgia and peritibial oedema have also been experienced during treatment [88].

Recent documented accounts on the incidence of resistance to chemotherapy, especially PZQ have caused much more concern [13, 18], however, some researchers does not believe this resistance exist [73]. Resistance varies according

to the schistosomal species, sex, stage of development, and geographic origin of the parasite as well as host factors, especially immunologic compromise [74, 75]. Due to drug pressure, resistance to PZQ and oxamniquine have been reported to exist in some isolates of *S. mansoni* in Brazil, Puerto Rico, Kenya, Senegal and Egypt [18, 76].

2.9 Antimicrobial peptides (amps) and their classification

As a result of drug resistance associated with drug pres-sure globally, antimicrobial peptides (AMPs) are currently considered as a vital source in new therapeutic design for infectious diseases due to their multifunctional properties [77]. AMPs can act as drug delivery vector, signalling molecule, contraceptive agent, immunomodulatory agent, mitogenic and antitumour agent [78]. AMPs are evolutionarily conserved molecules involved in the defence mechanisms of a wide range of organisms [79]. Usually, they are proteins with small molecular weight, widely distributed in nature, produced in bacteria, insects, plants and vertebrates [79, 80].

These peptides have been said to possess the ability to protect against (or attack) bacteria, viruses, fungi, parasites and tumour cells [81]. They are usually positively (cationic) charged with a net charge varying from +2 to +9 at neutral pH and have both a hydrophobic and hydrophilic side that enables the molecule to be soluble in aqueous environments yet also enable them interact with and disrupt lipid-rich membranes [80, 82].

Different AMPs possess amino acid sequences that are highly heterogeneous with their secondary structures having great variation. Majority are very short in length, containing 5 to 40 amino acid residues, while a few of them contain more than 40 residues [81]. Novel biological potentials of AMPs have been documented recently,

such as neutralization, endotoxin chemotactic and immune-modulating activities as well as induction of angiogenesis and wound repair [79]. Basically, classification of AMPs is difficult as a result of their diversity. Based on their amino acid composition, size and conformational structures, AMPs can be can be classified into four major classes: peptides with α-helix structures, β-sheet, extended and loop structures [79, 82].

Table 2.1: Some examples of Antimicrobial peptides and their composition

Antimicrobial Peptides	Composition
Cecropins	Contains 31–39 amino acids with an amphipathic, basic N-terminal domain and a hydrophobic C-terminal domain.
Mellitin	Contains 26 amino acid residue peptide with distinct hydrophilic and hydrophobic domains
Maximins 1, 2, 3, 4 and 5	Contains 27 amino acid residues
Abaecin	Contains 34 amino acids and contains almost 30% proline making it the largest proline-rich antimicrobial peptide character-ized, with broad spectrum of activity.
Magainin	Contains 23 amino acids residues
Hymenoptaecin	A glycine-rich antimicrobial peptide, containing 93 amino acids, with 2-pyrrolidone-5-carboxylic acid at the N-terminus.
Protegrins	Contains 16-18 amino-acid residues with four invariant cysteine residues, which form two disulfide bonds
Pleurocidin	Contains 12 amino acid residues
Indolicidin	Composed of 13 amino acid residues containing 5 tryptophan and 3 proline residues
Bactenecin	Composed of 12 amino acid residues, including 4 arginines, 2 cysteines and 6 other hydrophobic residues

2.10 Antimicrobial peptides (AMPs): mechanism of action

Basically, AMPs employ varying mode of action and molecular target mechanism in eliciting their protective effects and this is associated with their structural and physicochemical properties (Figure 2.2). The main hypothesis for their mechanism of action is their ability to disrupt and permeate the cell membrane through interaction with lipid molecules on the cell membrane surface, which may occur *via* several mechanisms including; barrel-stave model, carpet model, toroidal-pore wormhole model and detergent-type membrane lytic mechanism (Hoskin and Ramamoorthy, 2008; Marsh *et al*., 2009). This is achieved because of electrostatic attraction that occurs between the cationic peptides and the negatively charged molecules such as anionic phospholipids, lipopolysaccharides (LPS) (Gram-negative) and teichoic acid (Gram-positive), which are located asymmetrically in the cell membrane (Pelegrini *et al*., 2011).

Apart from membrane disruption, other mechanisms have been described by which AMPs elicit their protective effects. These mechanisms involve an alternative pathway; they can either up regulate the host immune response (Torrent *et al*., 2012), or inhibit and disturb the intracellular targets. That is, they have the ability to bind DNA, RNA and proteins, thereby, inhibiting nucleic-acid synthesis, cell wall synthesis, and protein synthesis as well as inhibit enzymatic activity (Li *et al*., 2012). Considering the mode of action (cell targets), AMPs can be classified into two major groups. The first group are said to be highly active against bacterial and cancerous cells but not affecting the normal healthy human cells while the second group are active against bacterial and cancerous cells as well as normal healthy human cells (Hoskin and Ramamoorthy, 2008).

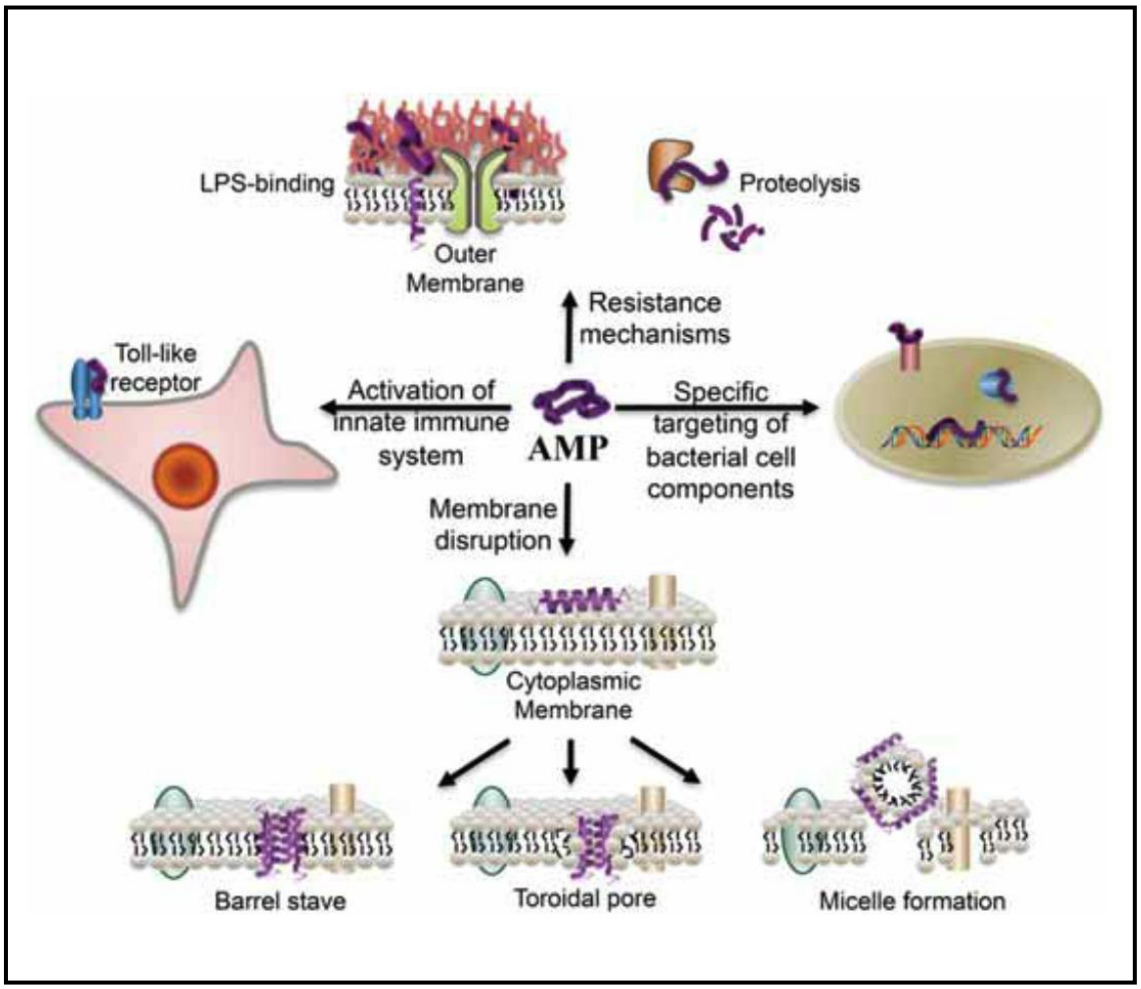


Figure 2.2: Overview on the mode of action of Antimicrobial Peptides (figure taken from Marsh et al., 2009).

2.11 Relationship between the mechanisms of action of schistosomes and that of AMPs

Several articles have reported that *Schistosoma* species have developed several mechanisms by which they parasitize their host. One of such mechanisms is their ability to mimic and manipulate their host immune system. Consequently, developing favourable environments which enhance their own survival; and co-habitation between both parties in an adaptive complex mechanism (Nayak and Kishore, 2013; Silva *et al*., 2014). Immunosuppression as a result of co-habitation between both parties may in turn lead to serious complications in schistosomiasis as well as also make their host vulnerable to other infections.

Immunostimulatory potentials of some evolutionarily conserved molecules have been documented recently. These evolutionarily conserved molecules are structurally and functionally diverse in nature. They are component of innate immune system responsible for up regulating the host immune system in attempt to inactivate or destroy certain parasitic infection or invaders (Guaní-Guerra *et al*., 2010); a good example is the AMPs. AMPs are involved in vital biological processes, such as endotoxin neutralization, chemotactic and immunomodulating activities as well as induction of angiogenesis and wound repair (Guaní-Guerra *et al*., 2010).

Interestingly, comparing the mechanism of action of Schistosomes and that of AMPs, the relationship that exist are as follow: Schistosomes are immunosuppressors while AMPs can act as immunostimulators (Hoskin and Ramamoorthy, 2008; Nayak and Kishore, 2013). Schistosomes are capable of generating reactive oxygen species (ROS) due to complications arising from their infection while AMPs has the potential

to delay the formation of ROS or scavenge them in the system (Pushpanathan *et al*., 2013; Shacter *et al*., 1988). Complications in chronic schistosoma infection may result in abnormal cell proliferation and eventually degenerate into cancer in affected organs while AMPs has selective cytotoxic potential to destroy the cancerous cells leaving the normal body cells (Li *et al*., 2013; Mostafa *et al*., 1999; Shacter *et al*., 1988).

2.12 What level of schistosomal infection do we think amp can be functional against?

In my own opinion, based on recent documented reports on the mechanisms of action of schistosomes and that of AMPs, I propose that AMPs will be an excellent candidate with the ability to control and treat schistosomal infections in the following ways. Firstly, by up regulating the definitive host immune system through immune complex formation within the definitive host. Thereby, attenuating or preventing the penetration of cecaria and migration of schistosomula in the erythrocyte to the site of maturation.

Secondly, AMPs can target pathogenesis and progression of schistosomiasis by selective cytotoxic elimination of schistosomula, preventing it from developing into adult worm thereby causing a great reduction in egg deposition in different organs while preventing complications that may arise as a result of the disease. Thirdly, AMPs can target schistosomiasis by selectively disrupting the cell membrane through various mechanisms, which may result in induction of apoptosis in affected cells. Lastly, AMPs can act as ROS scavenger, delaying ROS production as well as scavenging ROS generated in chronic schistosomiasis; thereby eliciting their anti-tumour effect.

2.13 Conclusion

Taking all of the above into account, AMPs are excellent candidates with promising results that may overcome drug resistance problems in the search for novel therapeutic agents for the control and treatment of schistosomiasis, basically due to their size and properties as well as their broad spectrum of activities.

REFERENCES

Alebie G, Erko B, Aemero M, Petros B. Epidemiological study on Schistosoma mansoni infection in Sanja area, Amhara region, Ethiopia. Parasit Vectors 2014; 7: 15.

Ali SA. Natural products as therapeutic agents for schistosomiasis. Res J Med Plant 2011; 5: 1-20.

Aly HF. Challenges in developing vaccines in parasitic diseases with special references to schistosomiasis. Int J Acad Res 2012; 4(4): 130-43.

Badawi AF, Mostafa MH. Possible mechanisms of alteration in the capacities of carcinogen metabolizing enzymes during schistosomi-asis and their role in bladder cancer induction. J Int Med Res 1993; 21: 281-305.

Barsoum RS, Esmat G, El-Baz T. Human Schistosomiasis*:* Clinical Perspective*:* Review. J Adv Res 2013; 4: 433-44.

Bennett JL, Day T, Feng-Tao L, Ismail M, Farghaly A. The devel-opment of resistance to anthelmintics: A perspective with an em-phasis on the antischistosomal drug praziquantel. Exp Parasitol 1997; 87 (3): 260-7.

Bica I, Hamer DH, Stadecker, MJ. Hepatic schistosomiasis. Infect Dis Clin North Am 2000; 14: 583-604.

Boros DL. Immunopathology of Schistosoma mansoni infection. Clin. Microbiol Rev 1989; 2(3): 250-69.

Botelho MC, Soares R, Vale N, *et al.* Schistosoma haematobium: identification of new estrogenic molecules with estradiol antagonis-tic activity and ability to inactivate estrogen receptor in mammalian cells. Exp. Parasitol 2010; 126; 526-35.

Botros S, Sayed H, El-Dusoki H, *et al.* Efficacy of mirazid in com-parison with praziquantel in Egyptian Schistosoma mansoni-infected school children and households. Am J Trop Med Hyg 2005; 72(2): 119-23.

Brindley PJ. Relationships between chemotherapy and immunity in schistosomiasis. Adv Parasitol 1994; 34: l34-63.

Cheever AW, Kuntz RE, Moore JA, Bryan GT, Brown RR. Animal model of human disease: carcinoma of the urinary bladder in *Schis-tosoma haematobium* infection. Am J Pathol 1976; 84: 673-6.

Cheever AW. Schistosomiasis and neoplasia. J Natl Cancer Inst 1978; 61: 13-18.

Chen MG, Mott KE. Progress in assessment of morbidity due to Schistosoma japonicum infection: a review of recent literature. Trop Dis Bull 1988; 85(6): R1-R45.

Chen MG. In: Human Schistosomiasis; Jordan, Webbe, Sturrock, Eds. Wallingford: C.A.B International; pp. 237-70.

Conlon CP. Schistosomiasis. Medicine 2005; 33(8): 64-67.

Crompton DWT, Engels D, Montresor A, Neira MP, Savioli L. Action starts now to control disease due to schistosomiasis and soil-transmitted helminthiasis. Acta Tropica 2003; 86: 121-4.

De Man JG, Chatterjee S, De Winter BY, *et al.* Effect of soma-tostatin on gastrointestinal contractility in Schistosoma mansoni in-fected mice. Int J Parasitol 2002; 32: 1309-320.

Doenhoff MJ, Kusel JR, Coles GC, Cioli D. Resistance of Schisto-soma mansoni to praziquantel: is there a problem? Trans R Soc Trop Med Hyg 2002; 96: 465-9.

Ekpo UF, Laja-Deile A, Oluwole AS, Sam-Wobo SO, Mafiana CF. Urinary schistosomiasis among preschool children in a rural com-munity near Abeokuta, Nigeria. Parasit Vectors 2010; 3: 58.

Elder K, Baker D, Ribes J. Infections, fertility and assisted Repro-duction, 1st ed. Cambridge: Cambridge University Press 2005.

Engels D, Chitsulo L, Montresor A, Savioli L. The global epidemi-ological situation of schistosomiasis and new approaches to control and research. Acta Tropica 2002; 82: 139-46.

Enk MJ, Oliveira e Silva G, Rodrigues NB. Diagnostic Accuracy and Applicability of a PCR System for the Detection of Schisto-soma mansoni DNA in Human Urine Samples from an Endemic Area. PLoS ONE 2012; 7(6): e38947.

Fenwick A, Keiser J, Utzinger J. Epidemiology, burden and control of schistosomiasis with particular consideration to past and current treatment trends. Drugs Future 2006; 31: 413-25.

Ferrari MLA, Coelho PMZ, Antunes CMF, Tavares CAP, Cunha AS. Efficacy of oxamniquine and praziquantel in the treatment of *Schistosoma mansoni* infection: a controlled trial. Bull. WHO 2003; 81: 190-6.

Fulford AJ, Ouma JH, Kimani G, Dunne DW. Puberty and age-related changes in susceptibility to schistosome infection. Parasitol Today 1998; 14(1): 23-6.

Gentile R, Oliveira G. Brazilian studies on the genetics of Schisto-soma mansoni. Acta Trop 2008; 108; 175-78.

Giuliani A, Pirri G, Nicoletto SF. Antimicrobial peptides: an over-view of a promising class of therapeutics. Cen Eur J Biol 2007; 2(1): 1-33.

Gray DJ, Ross AG, Li YS, McManus DP. Diagnosis and manage-ment of schistosomiasis. Br Med J 2011; 342: d2651.

Gryseels BK, Polman JC, Kestens L. Human schistosomiasis. Lancet 2006; 368: 1106-18.

Guaní-Guerra E, Santos-Mendoza T, Lugo-Reyes SO, Terán LM. Antimicrobial peptides: general overview and clinical implications in human health and disease. Clin Immunol 2010; 135(1): 1-11.

Hicks RM, Ismail MM, Walters CL, *et al.* Association of bacteri-uria and urinary nitrosamine formation with *Schistosoma haemato-bium* infection in the Qalyub area of Egypt. Trans R Soc Trop Med Hyg 1982; 76: 519-27.

Hoskin DW, Ramamoorthy A. Studies on anticancer activities of antimicrobial peptides. Biochimica et Biophysica Acta 2008; 1778(2): 357-75.

Ingram K, Ellis W, Keiser K. Antischistosomal activities of Meflo-quine-related arylmethanols. Antimicrob. Agents Chemother 2012; 56(6): 3207.

Izadpanah A, Gallo RL. Antimicrobial peptides. J Am Acad Der-matol 2005; 52(3): 381-92.

Jenkins SJ, Hewitson JP, Jenkins GR, Mountford AP. Modulation of the host's immune response by schistosome larvae. Parasite im-munol 2005; 27(10_11): 385-93.

Jenssen H, Hamill P, Hancock RE. Peptide Antimicrobial Agents. Clin Microbiol Rev 2006; 19(3): 491-511.

Johansson SL, Cohen SM. Epidemiology and etiology of bladder cancer. Semin Surg Oncol 1997; 13: 291-98.

King CH, Mahmoud AA. Drugs five years later: praziquantel. Ann Intern Med 1989; 110: 290-6.

King CH. Parasites and poverty: the case of schistosomiasis. Acta Trop 2010; 113: 95-104.

Lehman JS, Farid Z, Smith JH, Bassily S, el-Masry NA. Urinary schistosomiasis in Egypt: Clinical, radiological, bacteriological and parasitological correlations. Trans R Soc Trop Med Hyg ***1973***; 67(3): 384-99.

Li Y, Qu H, Xiang Q, Su Z, Huang Y. Overview on the recent study of antimicrobial peptides: origins, functions, relative mecha-nisms and application. Peptides 2012; 37(2): 207-15.

Li YS, Li Y, Yu DB, *et al.* A multivariate analysis of the relation-ship between the workability and S. japonicum infection in China. Rev Inst Med Trop Sao Paulo 1993; 35: 347-53.

Li YS, Sleigh AC, Ross AGP, *et al.* Epidemiology of *Schistosoma japonicum* in China: morbidity and strategies for control in the Dongting Lake region. Int J Parasitol 2000; 30: 273-81.

Madbouly KM, Senagore AJ, Mukerjee A, *et al.* Colorectal cancer in a population with endemic Schistosoma mansoni: is this an at-risk population? Int J Colorectal Dis 2007; 22: 175-81.

Marletta MA. Mammalian synthesis of nitrite, nitrate, nitric oxide, and N-nitrosating agents. Chem Res Toxicol 1988; 1: 249-57.

Marsh EN, Buer BC, Ramamoorthy A. Fluorine - a new element in the design of membrane-active peptides. Mol Biosyst 2009; 5: 1143-47

Mazigo HD, Nuwaha F, Wilson S, *et al.* Epidemiology and interactions of Human Immunodeficiency Virus - 1 and Schistosoma mansoni in sub-Saharan Africa. Infect Dis Poverty 2013; 2(1): 2.

Mbah AN, Mahmud O, Awofolu OR, Isokpehi RD. Inferences on the biochemical and environmental regulation of universal stress proteins from Schistosomiasis parasites, Adv Appl Bioinform Chem 2013; 6: 15-27.

McManus DP, Loukas A. Current status of vaccines for schis-tosomiasis. Clin Microbiol Rev 2008; 21: 225-42.

Mekonnen A, Legesse M, Belay M, Tadesse K, Torben W, Tekle-mariam Z, Erko B. Efficacy of Praziquantel against *Schistosoma haematobium* in Dulshatalo village, western Ethiopia. BMC Res Notes 2013; 6: 392.

Melman SD, Steinauer ML, Cunningham C, *et al.* Reduced suscep-tibility to praziquantel among naturally occurring Kenyan isolates of *Schistosoma mansoni*. PLoS Negl Trop Dis 2009; 3(8): e504.

Michaud DS, Platz EA, Giovannucci E. Gonorrhoea and male blad-der cancer in a prospective study. Br J Cancer 2007; 96: 169-71.

Miranda MA, Kuehn CC, Cardoso JFR, *et al.* Immunomodulatory effect of the alkaloidic extract of Solanum lycocarpum fruits in mice infected with Schistosoma mansoni. Exp Parasitol 2013; 133: 396-402.

Mostafa MH, Helmi S, Badawi AF. Nitrate, nitrite and volatile N-nitroso compounds in the urine of *Schistosoma haematobium* and *Schistosoma mansoni* infected patients. Carcinogenesis 1994; 15: 619-25.

Mostafa MH, Sheweita SA, O'Connor PJ. Relationship between schistosomiasis and bladder cancer. Clin Microbiol Rev 1999; 12: 97-111.

Nayak A, Kishore U. In: Microbial Pathogenesis: Infection and Immunity; Kishore and Nayak, Eds. New York: Landes Bioscience Springer 2013; Vol. 798: pp 255- 75.

O'Brien PJ. Radical formation during the peroxidase catalyzed metabolism of carcinogens and xenobiotics: the reactivity of these radicals with GSH, DNA, and unsaturated lipid. Free Radic Biol Med 1988; 4: 169-83.

Pelegrini PB, Del Sarto RP, Silva ON, Franco OL, Grossi-De-Sa MF. Antibacterial Peptides from Plants: What they are and how they probably work. Biochem Res Int 2011; 2011: 250349.

Ploeg M, Aben KKH, Kiemeney LA. The present and future bur-den of urinary bladder cancer in the world. World J Urol 2009; 27: 289-93.

Pushpanathan M, Gunasekaran P, Rajendhran J. Antimicrobial Peptides: Versatile Biological Properties. Int J Pept 2013: 1-15.

Rambau PF, Chalya PL, Jackson K. Schistosomiasis and urinary bladder cancer in north west Tanzania; a retrospective review of 185 patients. Infect Agent Cancer 2013; 8(1): 19.

Rollinson D, Kaukas A, Johnston DA, Simpson AJ, Tanaka M. Some molecular insights into schistosome evolution. Int J Parasitol 1997; 27: 11-28.

Rosin MP, Anwar WA, Ward AJ. Inflammation, chromosomal instability, and cancer: the schistosomiasis model. Cancer Res 1994a; 54: 1929s-33s.

Rosin MP, El-Din SS, Ward AJ, Anwar WA. Involvement of in-flammatory reactions and elevated cell proliferation in the devel-opment of bladder cancer in schistosomiasis patients. Mutat Res 1994b; 305: 283-92.

Ross APG, Bartley PB, Sleigh AC, *et al.* Schistosomiasis. N Engl J Med 2002; 346: 1212-20.

Samaras V, Rafailidis PI, Mourtzoukou EG, Peppas G, Falagas ME. Chronic bacterial and parasitic infections and cancer: a re-view. J Infect Dev Ctries 2010; 4: 267-81.

Schutte CHJ, Pienaar R, Becker P, *et al.* Observations on the tech-niques used in the qualitative and quantitative diagnosis of schis-tosomiasis. Ann Trop Med Parasitol 1994; 88 (3): 305-16.

Seo MD, Won HS, Kim JH, Mishig-Ochir T, Lee BJ. Antimicro-bial peptides for therapeutic applications: a review. Molecules 2012; 17(10): 12276-86.

Shacter E, Beecham EJ, Covey JM, Kohn KW, Potter M. Activated neutrophils induce prolonged DNA damage in neighboring cells. Carcinogenesis 1988; 9: 2297-304.

Shaker OG, Hammam OA, El Leithy TR, *et al.* Molecular markers and bladder carcinoma: Schistosomal and non-schistosomal. Clin Biochem 2011; 44: 237-44.

Shekhar KC. Schistosomiasis drug therapy and treatment consid-erations. Drugs 1991; 42(3): 379-405.

Silva JP, Furtado AP, Santos JN. Proteomic profile of Ortleppas-caris sp.: A helminth parasite of Rhinella marina in the Amazonian region. Int J Parasitol Parasites Wildl 2014; 3(2): 67-74.

Sire C, Durand P, Pointier JP, Theron A. Genetic diversity and recruitment pattern of *Schistosoma mansoni* in a *Biomphalaria glabrata* snail population: a field study using random-amplified polymorphic DNA markers. J Parasitol 1999; 85: 436-41.

Steinmann P, Keiser J, Bos R, Tanner M, Utzinger J. Schistosomia-sis and water resources development: systematic review, meta-analysis, and estimates of people at risk. Lancet Infect Dis 2006; 6: 411-25.

Tadesse D, Tsehaye A, Mahmud A. Efficacy of Praziquantel in treating *S.mansoni* infected school children in Tumuga and Waja, North Ethiopia. Momona Ethiopian Journal of Science 2010; 2: 3-11.

Torrent M, Pulido D, Rivas L, Andreu D. Antimicrobial peptide action on parasites. Curr Drug Targets 2012; 13(9): 1138-47.

Tricker AR, Mostafa MH, Spiegelhalder B, Preussmann R. Urinary excretion of nitrate, nitrite and N-nitroso compounds in Schis-tosomiasis and bilharzia bladder cancer patients. Carcinogenesis 1989; 10: 547-52.

Ugochukwu DO, Onwuliri CO, Osuala FOU, *et al.* Endemicity of schistosomiasis in some parts of Anambra State, Nigeria. J Med Lab Diagn 2013; 4(5): 54-61.

Ukoroije BR, Abowei JFN. Some Occupational Diseases in Culture Fisheries Management and Practices Part Two: Schistosomiasis and Filariasis. Int J Fish Aquat Sci 2012; 1(1): 64-71.

Vennervald BJ, Polman K. Helminths and malignancy. Parasite Immunol 2009; 31: 686-96.

Wang W, Wang L, Liang YS. Susceptibility or resistance of praziquantel in human schistosomiasis: a review. Parasitol Res 2012; 111: 1871-7.

Warren KS. The immunopathology of schistosomiasis: a multidis-ciplinary approach. Trans R Soc Trop Med Hyg 1972; 66: 417-34.

Weinstock JV. The pathogenesis of granulomatous inflammation and organ injury in schistosomiasis: interactions between the schis-tosome ova and the host. Immunol Invest 1992; 21: 455-75.

Weitberg AB. Effect of combinations of antioxidants on phagocyte-induced sister-chromatid exchanges. Mutat Res 1989; 224: 1-4.

Weitzman SA, Stossel TP. Mutation caused by human phagocytes. Science 1981; 212: 546-7.

Zasloff M. Antimicrobial peptides of multicellular organisms. Nature 2002; 415: 389-95.

Zongo D, Kabre BG, Dayeri D, Savadogo B, Poda JN. Compara-tive study of schistosomiasis transmission (urinary and intestinal forms) at 10 sites in Burkina Faso (in sub-Saharan Africa). Med Sante Trop 2012; 22(3): 323-9.

CHAPTER THREE

IDENTIFICATION OF THE STRUCTURAL FEATURES OF THE SWITCH DOMAIN REGIONS FROM *SCHISTOSOMA MANSONI* USING BIOINFORMATICS ANALYSIS

3.1 Introduction

Elucidating protein structure and function in biomedical research has become highly automated with the advent of bioinformatics. Interestingly, bioinformatics tools have turn out to be an essential component as well as an integral part in many areas of structural biology. Basically, structural biology and bioinformatics are dynamic area of scientific research attracting high level of attention in recent years (Elumalai and Eswaraiah, 2013; Kanehisa and Bork, 2003). The impact of bioinformatics in the last two decades and its current contribution to our understanding of the molecular events (mechanisms) of various diseases has positioned it at the forefront in science and technology (Chen *et al*., 2014; Scholz *et al*., 2012). In order to develope novel drug candidates in the treatment and management of diseases, identification and characterization of the structure and function as well as the active site of the proteins involved in the disease is important in order to design specific and selective small-molecule inhibitors, ligands or peptides that are capable of modulating the protein activity (Grant, 2011).

Consequently focusing on elucidating the exact biological and biochemical functions of proteins and subsequently analyzing their interactome patterns, can be complicated and time consuming using the conventional methods (in vitro or in vivo)

available. Until recently, majority of the newly sequenced proteins have unknown structures and functions. Ironically, the lack of information on a protein structure and function restricts its further utility. Therefore, *in-silico* methods for predicting protein interaction patterns in various cellular environments proffer an alternative practical solution (Nadzirin and Firdaus-Raih, 2012; Roche *et al*., 2015). The various types of interactions that exist between proteins are important in the discovery of a variety of biological functions and pathways within the living system (Chang *et al*., 2010). Traditionally, proteins carry out their functions by interacting with other proteins. Protein-protein interaction play a crucial role in a variety of cellular biochemical processes like DNA transcription and replication, cellular defences, metabolic cycles, enzymatic catalysis and signalling cascades just to mention a few. Therefore, correct identification and characterization of such proteins as well as detailed investigation of protein network interaction partners can greatly increase our understanding toward the elucidation of protein functions at the molecular level within the cell (Ben-Hur and Noble, 2005; Hue *et al*., 2010; You *et al*., 2014).

Schistosoma mansoni is one of the causative agents accountable for one of the important neglected tropical disease known as schistosomiasis. The disease is a key public health challenge in many developing countries (Adenowo *et al*., 2015; Oliveira *et al*., 2009). In order to reinforce conventional methods and make available new information to improve success towards schistosomiasis control, the scientific community joined efforts to assess the Schistosoma genomic information which started in 1994, as an initiative of WHO/TDR (Nahum *et al*., 2012). The general aim of the project was to create the genomic database and make it accessible for the development of novel control measures such as vaccines and drugs. The development of vaccines and new drugs still remains a huge challenge even though

many researches are ongoing and different research groups are taking advantage of the database to develop the much-needed control intervention (Mourão *et al*., 2012).

Schistosoma mansoni nuclear genome sequence was determined by whole genome shotgun sequencing and assembled into 5,745 scaffolds greater than 2 kilobases (kb), totalling 363 megabases (Mb). Though 40 % of the genome was repetitive, 50 % was assembled into scaffolds of at least 824.5 kb. Added to this, 43 % of the genome assembly (spread over 153 scaffolds) was definitely assigned to chromosomes (seven autosomal, plus ZW sex determination pairs) using fluorescence in situ hybridization (Berriman *et al*., 2009). The main aim of this chapter is to use *in-silico* tools to elucidate structural and functional motifs of the switch domains of *Schistosoma mansoni* Smp_059340.1 as well as to identify its potential interactive partner.

The specific objectives of this study are as follows:

i. To map out the phylogeny of the switch domains of *S. mansoni* Smp_059340.1 and putative homologues;
ii. To predict the secondary structure of the switch domains of *S. mansoni* Smp_059340.1;
iii. To generate the three dimensional homology model of the switch domains of *S. mansoni* Smp_059340.1 and to validate the model;
iv. To identify domains and residues that could be involved in the interaction of the switch domains of *S. mansoni* Smp_059340.1 with guanosine triphosphate (GTP) and divalent cations (Mg^{2+}, Ca^{2+} and Zn^{2+}).

3.2 Materials and Methods

3.2.1 Sequence retrieval

The protein sequence (in the FASTA format) of *S. mansoni* Smp_059340.1 with accession number XP_018644269.1 (379 amino acid long) was retrieved from NCBI (National Centre for Biotechnology Information). Protein and gene sequence similarity was analyzed by BLAST tools (Altschul *et al*., 1997) form searching the NCBI non-redundant sequence database (https://blast.ncbi.nlm.nih.gov/). The *S. mansoni* Smp_059340.1 sequence obtained was aligned manually using BIOEDIT (Hall, 1999) and the molecular phylogenetic tree for *S. mansoni* Smp_059340.1 was built using the ClustalW (Thompson *et al*., 1994).

3.2.2 Analysis of probable functional motifs

MOTIF Finder (http://www.genome.jp/tools/motif/) an online search engine was used to determine the putative motifs present in the switch domains of *S. mansoni* Smp_059340.1 protein structure by analysing the protein sequence. The software predicts the probable motifs within the protein with a statistically significant cut off score (E value).

3.2.3 Evolutionary conservation analysis of the switch domains of S. mansoni Smp_059340.1

Evolutionary conservation analysis of the switch domains of *S. mansoni* Smp_059340.1 was performed using the ConSurf prediction software (http://consurf.tau.ac.il/) (Ashkenazy *et al*., 2010).

3.2.4 Investigation of the secondary structure of the switch domains of S. mansoni Smp_059340.1

A prediction of the overall secondary structure (α-helices, β-sheets and random coils) of the switch domains of *S. mansoni* Smp_059340.1 was done using the Jpred 4 secondary structure prediction server. JPred4 (http://www. compbio. dundee.ac. uk/jpred4) is the latest version of the popular JPred protein secondary structure prediction server which provides predictions by the JNet algorithm, one of the most accurate methods for secondary structure prediction (Drozdetskiy *et al*., 2015).

3.2.5 Homology modeling of the switch domains of S. mansoni Smp_059340.1

Homology modeling of the switch domains of *S. mansoni* Smp_059340.1 was carried out to predict its three dimensional (3D) structure as there was no crystal structure of the switch domains of *S. mansoni* Smp_059340.1 accessible. Therefore, the available protein sequence was used for homology based modeling. The 3D structure of the switch domains of *S. mansoni* Smp_059340.1 protein was predicted using web based server RaptorX (http://raptorx.uchicago.edu/StructurePrediction/predict/) (Källberg *et al*., 2014) and Phyre2 (http://www.sbg.bio.ic.ac.uk/phyre2) (Kelley *et al*., 2015).

3.2.6 Validation of the 3D model of the switch domains of S. mansoni Smp_059340.1

The accuracy and protein structure stereochemistry feature of the predicted models of the switch domains of *S. mansoni* Smp_059340.1 were validated with RAMPAGE

(http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) which generates Ramachandran plots (Lovell *et al*., 2003).

3.2.7 Energy minimization of modelled structures of the switch domains of S. mansoni Smp_059340.1

Structural refinement and energy minimization of the predicted model of the switch domains of *S. mansoni* Smp_059340.1 was performed using YASARA energy minimization server (http://www.yasara.org/minimizationserver.htm) and GalaxyRefine web server for protein structure refinement (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE). All the modelled structures were visualised and further analysed by using University of California San Francisco (UCSF) Chimera (Pettersen *et al*., 2004).

3.2.8 Network interaction analysis of the switch domains of S. mansoni Smp_059340.1

A protein-protein interaction partner of the switch domains of *S. mansoni* Smp_059340.1 was identified using STRING (http://string-db.org/), a biological database which is used to construct protein-protein interaction network for different known and predicted protein interactions (Franceschini *et al*., 2013; Snel *et al*., 2000).

3.2.9 Active site analysis and residues recognition

In order to carry out ligand binding studies by docking, prediction of ligand binding site of the switch domains of *S. mansoni* Smp_059340.1 was evaluated by submitting the modelled structure to COACH protein-ligand binding prediction server, a meta-server approach to protein-ligand binding site prediction

(http://zhanglab.ccmb.med.umich.edu/COACH/) and GalaxySite web server (http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=SITE) (Heo *et al*., 2014). COACH will generate complementary ligand binding site predictions using two comparative methods, TM-SITE and S-SITE, which recognize ligand-binding templates from the BioLiP protein function database by binding-specific substructure and sequence profile comparisons (Yang *et al*., 2013).

3.2.10 *In-silico* molecular docking

The docking study of the switch domains of *S. mansoni* Smp_059340.1 was done using the PatchDock automatic server for molecular docking. The clustering RMSD was 4.0 A˚ and the complex type was set to default. (Duhovny *et al*., 2002; Schneidman-Duhovny *et al*., 2005).

3.3 Results

3.3.1 Analysis of sequence similarity between the S. mansoni Smp_059340.1 and its homologues

Multiple sequence alignment was performed on the full length sequence of *S. mansoni* Smp_059340.1 protein and its homologues (Figure 3.1) as well as on the protein sequence corresponding to the switch domains of the *S. mansoni* Smp_059340.1 and its homologues (Figure 3.2). Figure 3.3 presents the phylogenetic analysis of the switch domains of *S. mansoni* Smp_059340.1, based on its amino acid sequences. The result shows the presence of the conserved sequence regions and their relationship. Table A5 and Table A6 (Appendix section) shows sequence identities and similarities for the multiple sequence alignment.

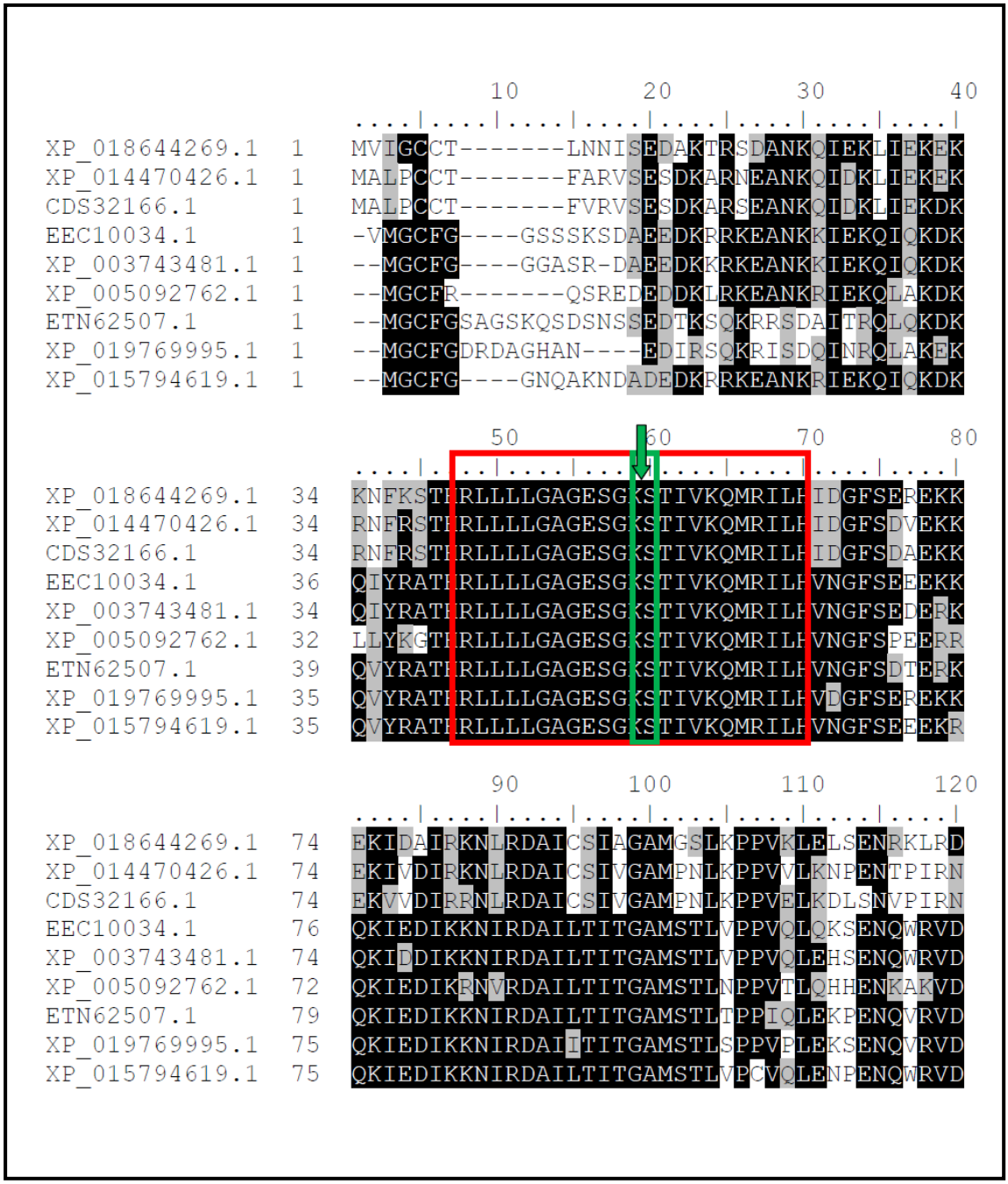


Figure 3.1: Multiple sequence alignment of S. mansoni Smp_059340.1.

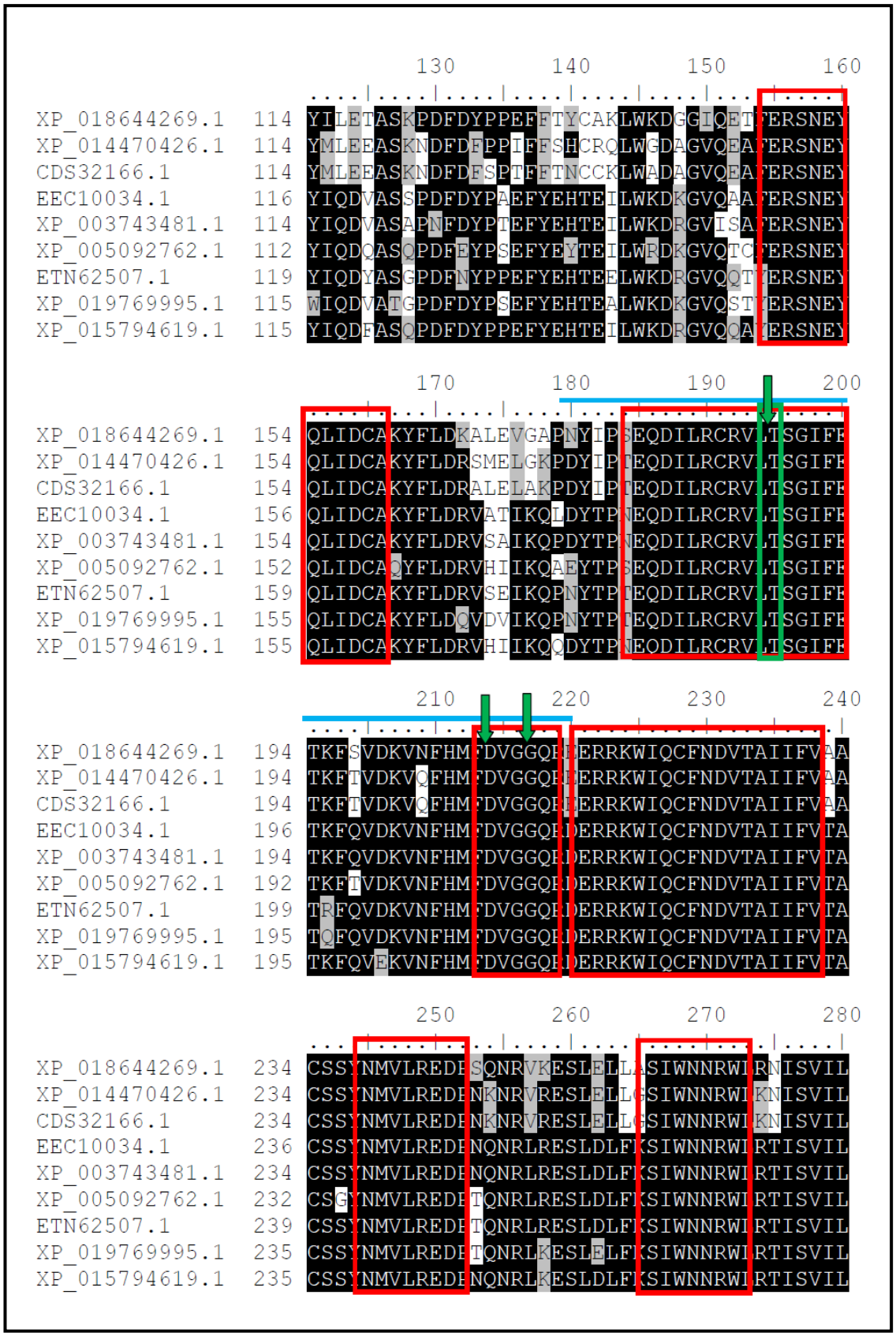


Figure 3.1: (Continued): Multiple sequence alignment of S. mansoni Smp_059340.1.

Figure 3.1 (Continued): Multiple sequence alignment of S. mansoni Smp_059340.1. Table A1 under appendix shows the names and accession numbers of the organisms retrieved and used in the construction of multiple sequence alignment. Multiple sequence alignment was done using BIOEDIT.

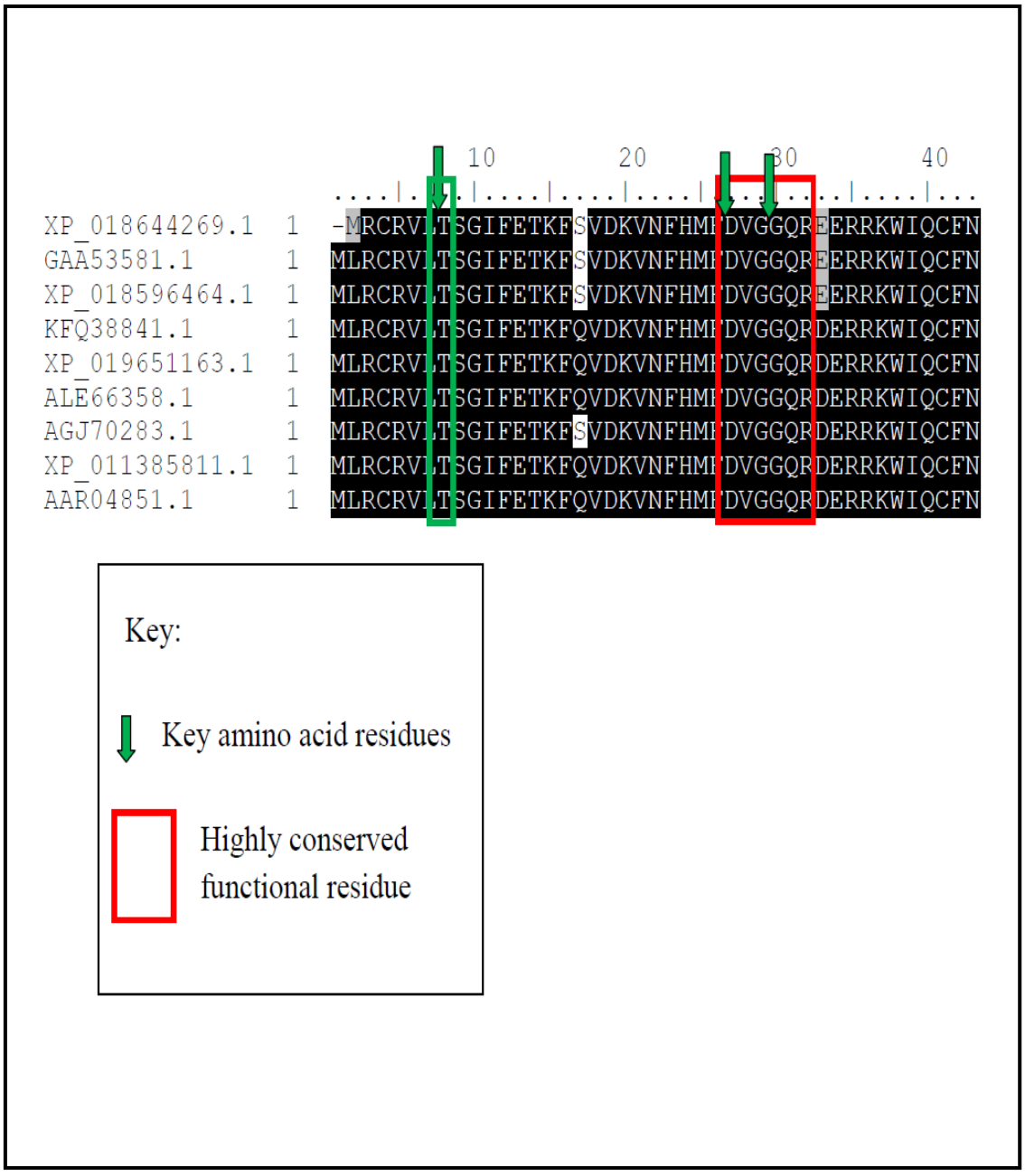


Figure 3.2: Multiple sequence alignment of the switch domains of S. mansoni Smp_059340.1. Table A2 under appendix shows the names and accession numbers of the organisms retrieved and used in the construction of multiple sequence alignment. Multiple sequence alignment was done using BIOEDIT.

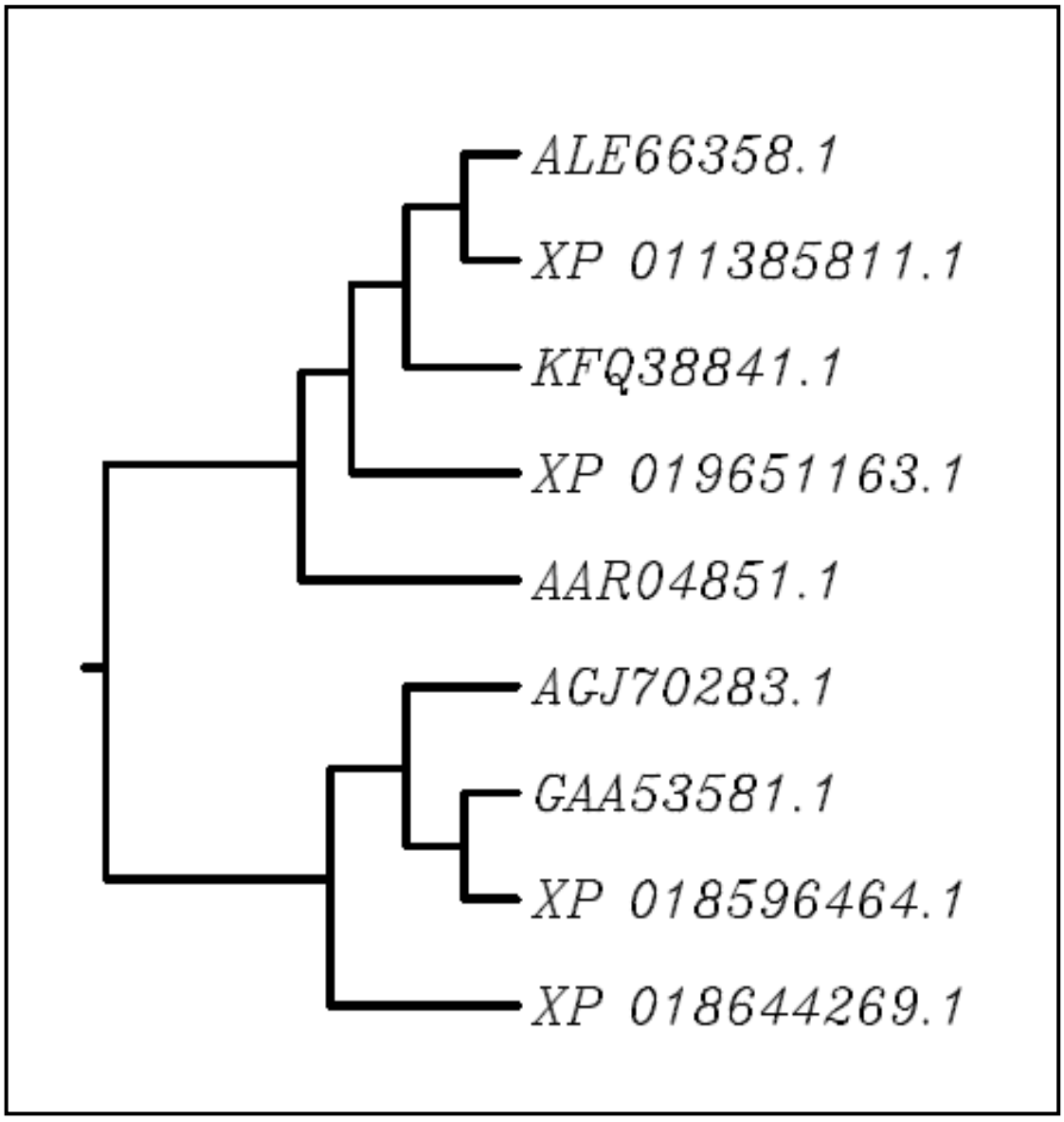


Figure 3.3: Phylogenetic analysis of the switch domains of S. mansoni Smp_059340.1 based on amino acid sequences. Table A2 under appendix contains the names and accession numbers of the organisms retrieved and used in the construction of the phylogenetic tree of the switch domains of *S. mansoni* Smp_059340.1. Phylogenetic analysis was done using ClaustalW.

3.3.2 Analysis of probable functional motifs

The probable motif investigation of the amino acid sequences of *S. mansoni* Smp_059340.1 (XP_018644269.1) proteins corresponding to the switch domains of *S. mansoni* Smp_059340.1 (183 – 223 aa) performed on the MOTIF finder server revealed that the protein consists of two promising functional motifs. The first promising motif is the G-alpha motif located between position 2 to 41 amino acid stretch of the protein while the second promising motif is the ADP-ribosylation factor family (Arf) located between position 8 to 41 amino acid stretch of the protein (Table 3.1 and Figure 3.4).

Table 3.1: The conserved motifs present in the switch domains of S. mansoni Smp_059340.1 predicted by MOTIF finder.

S/N	Pfam ID	Position	i-E value	Description
1	G-alpha	2 – 41	3.5e-14	G-protein alpha subunit
2	Arf	8 – 41	4.2e-05	ADP-ribosylation factor family

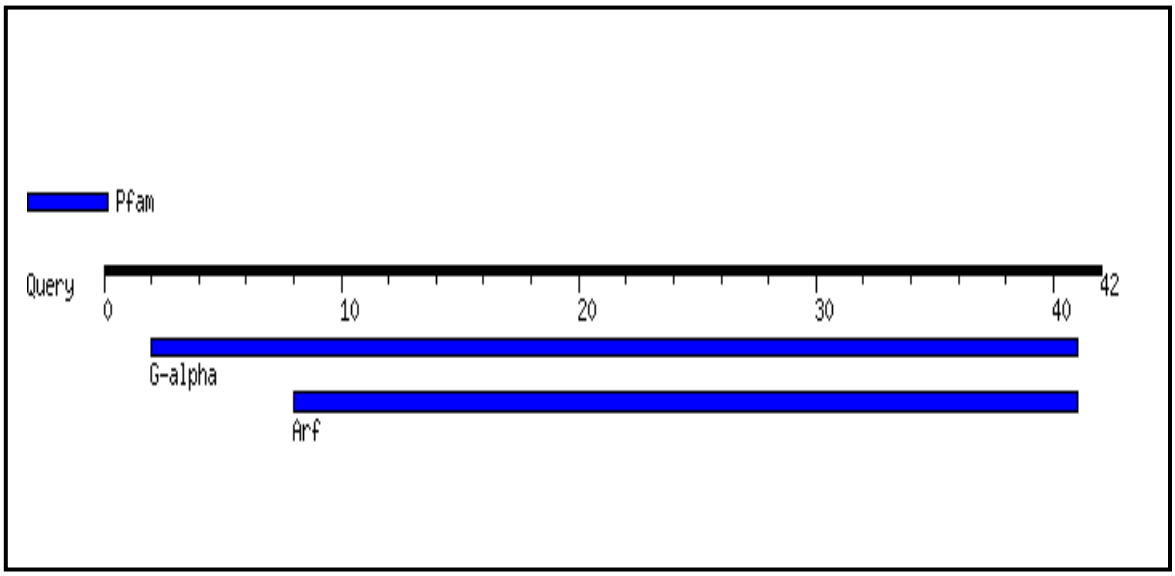


Figure 3.4: The predicted promising functional motifs generated by MOTIF finder.

3.3.3 Evolutionary conservation analysis of the switch domains of S. mansoni Smp_059340.1

Figure 3.5 presents the results obtained from the ConSurf prediction software. The result depicts a structural illustration of the protein which contains a colorimetric conservation score. For any protein the information concerning evolutionary feature is extremely significant for predicting highly conserved residues, variable residues etc. (Goswami, 2017). The ConSurf prediction software results showed that the functional regions of the switch domains of *S. mansoni* Smp_059340.1 are highly conserved.

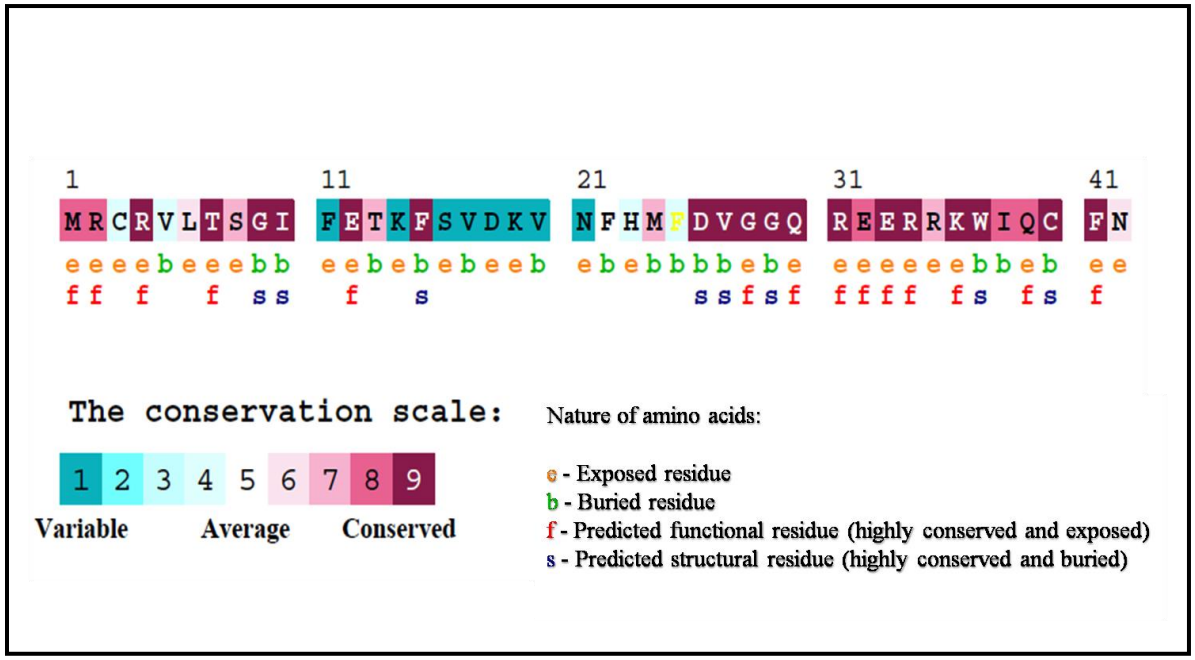


Figure 3.5: Evolutionary conservation analysis of the switch domains of S. mansoni Smp_059340.1 by ConSurf. The colour coding bar represents the colouring pattern of conservation score.

3.3.4 Secondary structure prediction

The secondary structure analysis conducted on JPred 4 server revealed that the switch domains of *S. mansoni* Smp_059340.1 contain 1 α-helices, 2 β-sheets and 4 random coils (Figure 3.6).

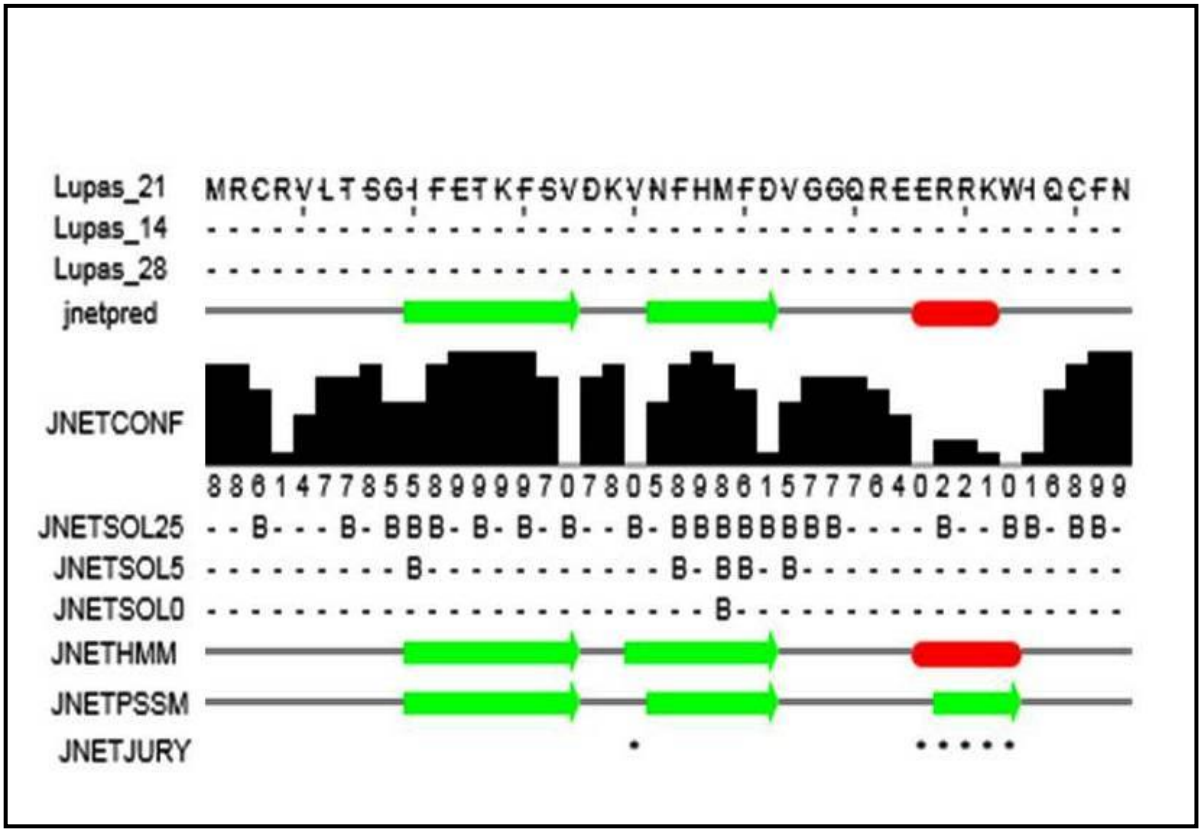


Figure 3.6: Predicted Secondary structure of the switch domains of S. mansoni Smp_059340.1 generated by JPred 4.

3.3.5 Homology modeling of the switch domains of S. mansoni Smp_059340.1

Homology modeling of the switch domains of *S. mansoni* Smp_059340.1 was carried out using Raptor X and Phyre 2 web base server. The 3D structure of the protein obtained from Raptor X and Phyre 2 are presented in Figure 3.7 and Figure 3.8 respectively. The availability of a structural model of a protein is important in order to identify its interactive partners as well as to understand its biological functions at the molecular and cellular levels. The structural quality of the models was evaluated by Ramachandran plot analysis in RAMPAGE.

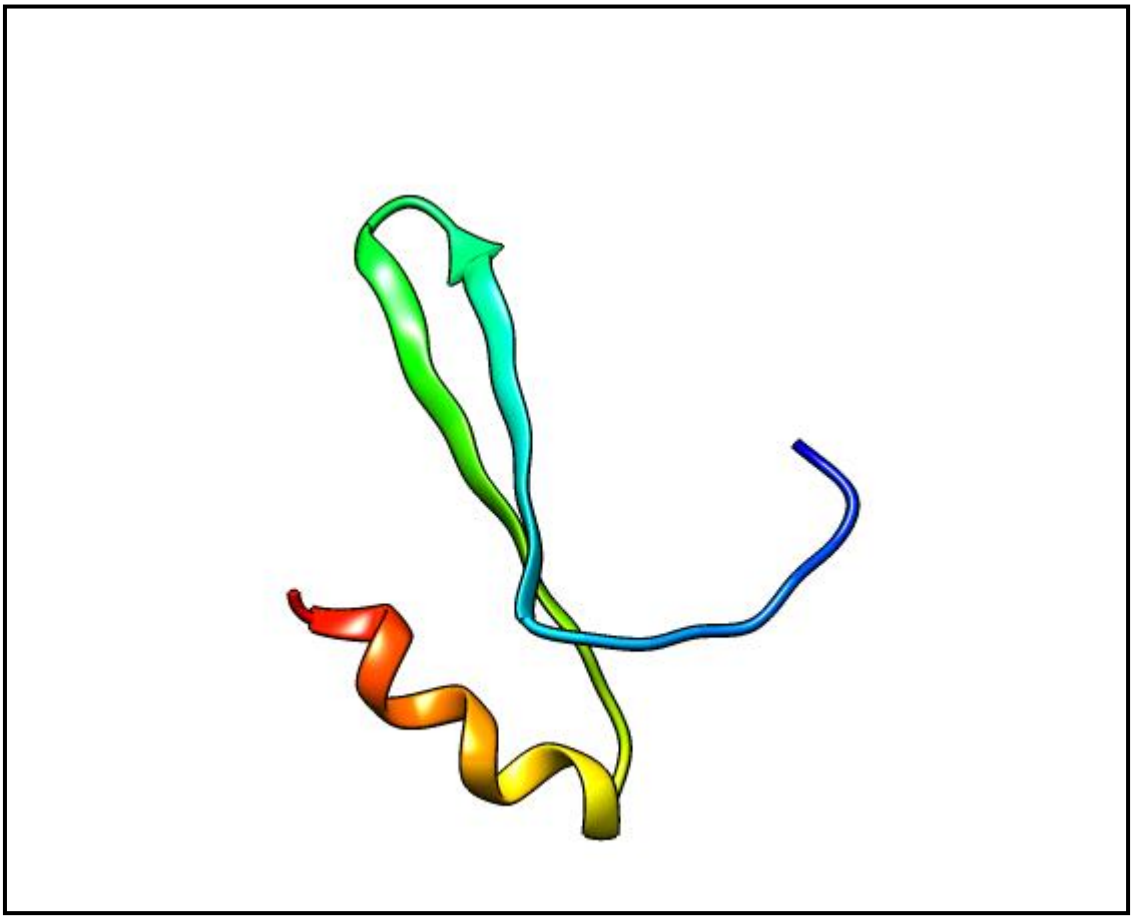


Figure 3.7: Cartoon representation of the homology model of the switch domains of S. mansoni Smp_059340.1 generated by Raptor X.

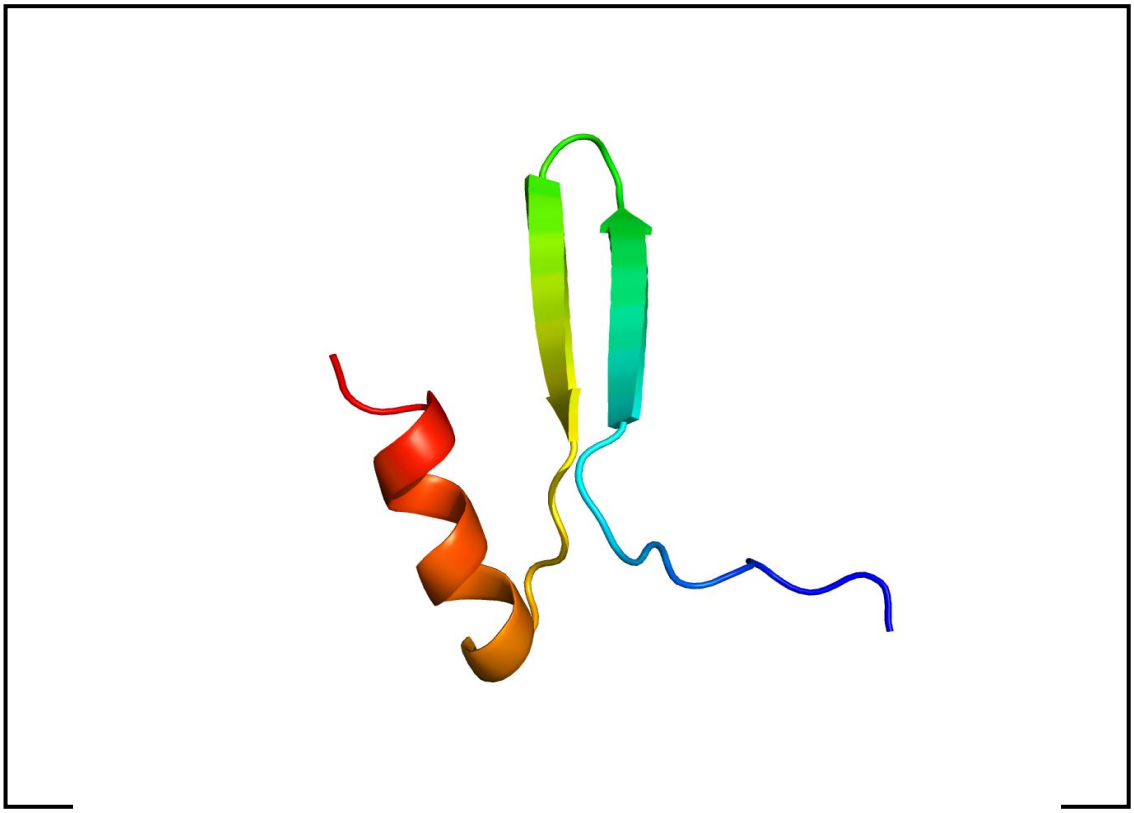


Figure 3.8: Cartoon representation of the homology model of the switch domains of S. mansoni Smp_059340.1 generated by Phyre 2.

3.3.6 Structure validation of the 3D model of the switch domains of *S. mansoni* Smp_059340.1

The overall stereochemical parameters as evaluated by the Ramachandran plot analysis in RAMPAGE showed that Raptor X model has 95.0 % of the residues in the favoured region, 5.0 % of the residues are in allowed region and none of the residues (0.0 %) in the outlier region (Figure 3.9 A and Figure 3.9 B). On the other hand, Phyre 2 model has 97.5 % of the residues in the favoured region, 2.5 % residues in the allowed region and none of the residues (0.0 %) in the outlier region (Figure 3.10 A and Figure 3.10 B). The results obtained from the Ramachandran plot reveals that both models are stable and of good quality, but Phyre 2 model was identified to be slightly superior in nature. All further work was carried out using the Phyre 2 structure of the switch domains of *S. mansoni* Smp_059340.1.

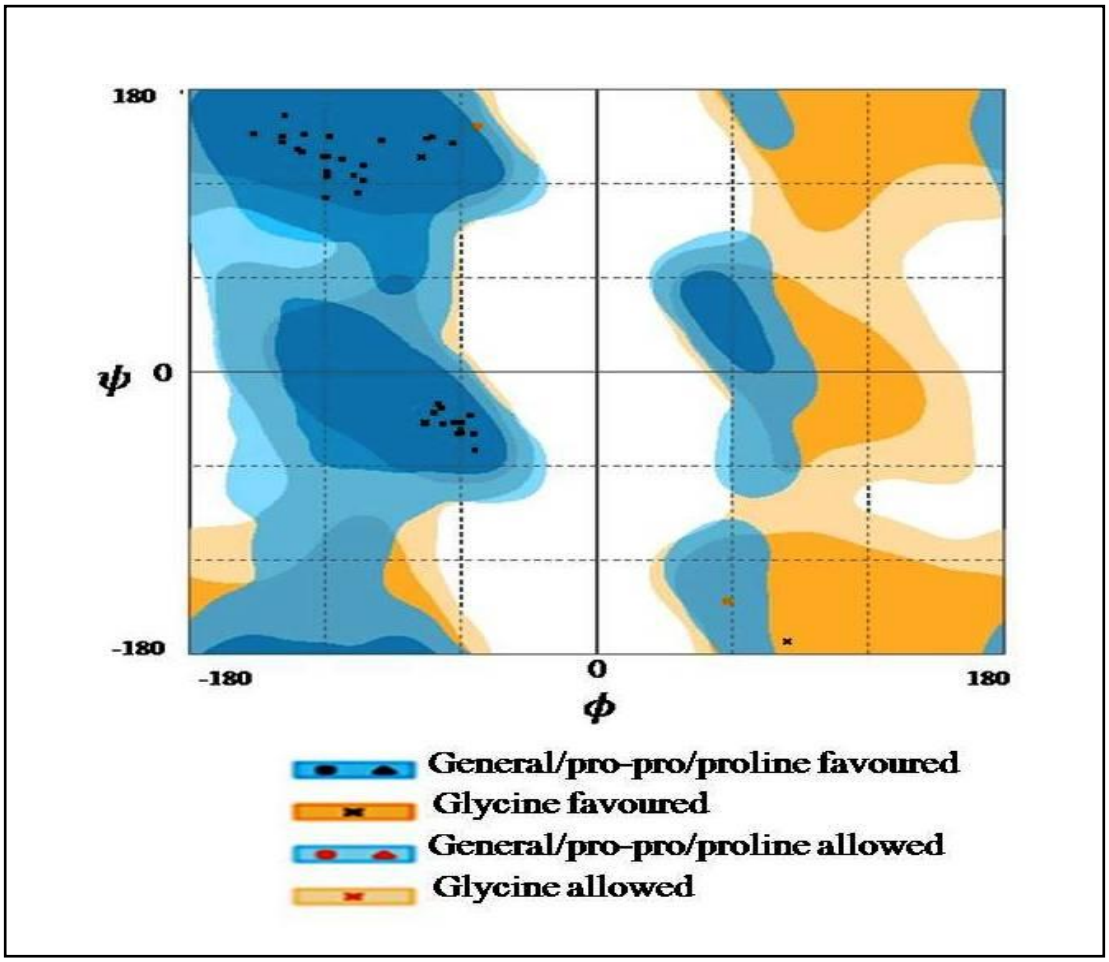


Figure 3.9A: Ramachandran plot of the predicted model of the switch domains of S. mansoni Smp_059340.1 generated by Raptor X.

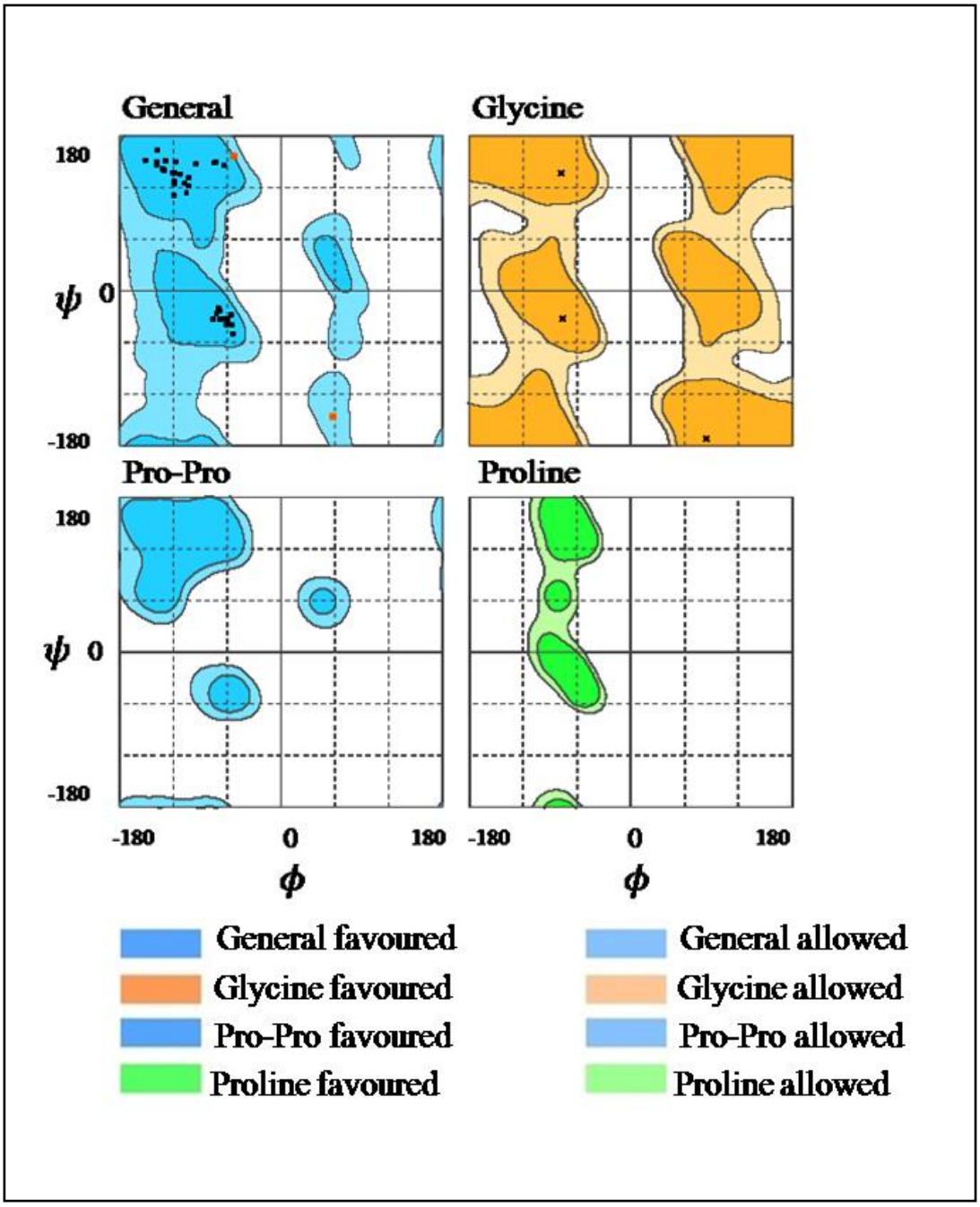


Figure 3.9B: Favoured and allowed amino acid residues in Ramachandran plot of the switch domains of S. mansoni Smp_059340.1 generated by Raptor X in RAMPAGE.

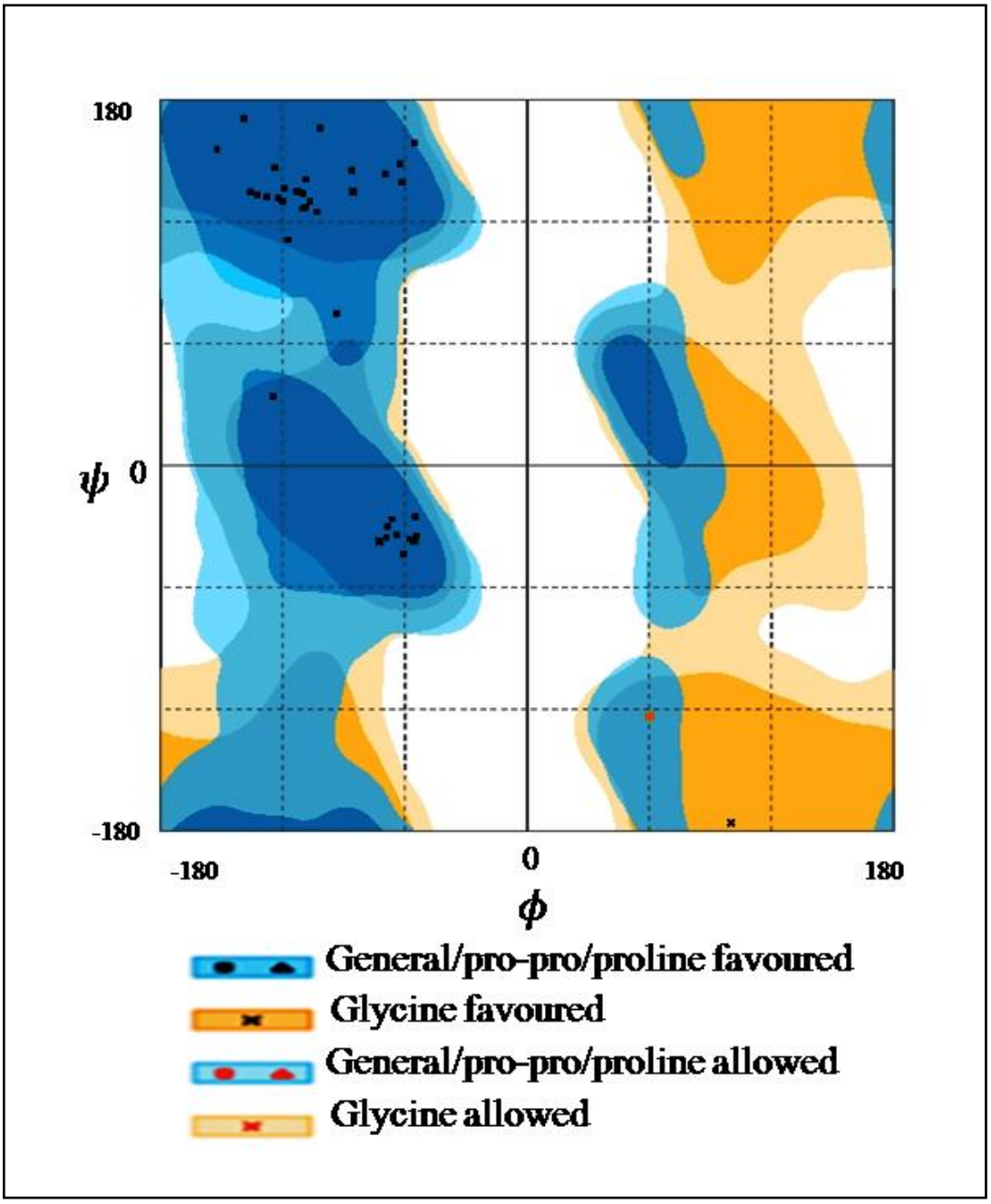


Figure 3.10A: Ramachandran plot of the predicted model of the switch domains of S. mansoni Smp_059340.1 generated by Phyre 2.

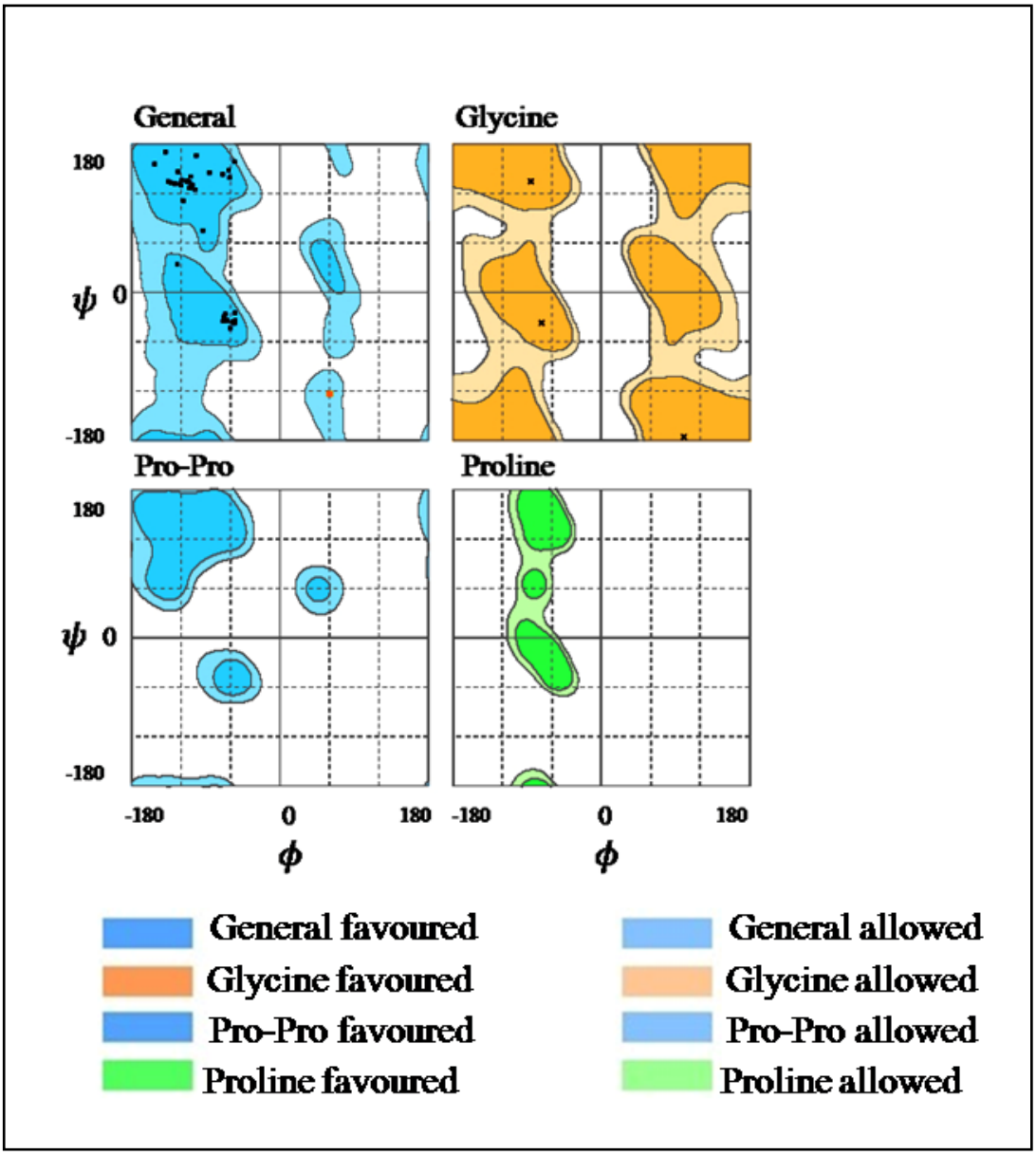


Figure 3.10B: Favoured and allowed amino acid residues in Ramachandran plot of the switch domains of S. mansoni Smp_059340.1 generated by Phyre 2 in RAMPAGE.

3.3.7 Energy minimization of modelled structures of the switch domains of S. mansoni Smp_059340.1

Figure 3.11 and Figure 3.12 shows the energy minimized model of the switch domains of *S. mansoni* Smp_059340.1. The energy minimization with YASARA energy minimization server and GalaxyRefine server for protein structure refinement eventually results into a thermodynamically stable model.

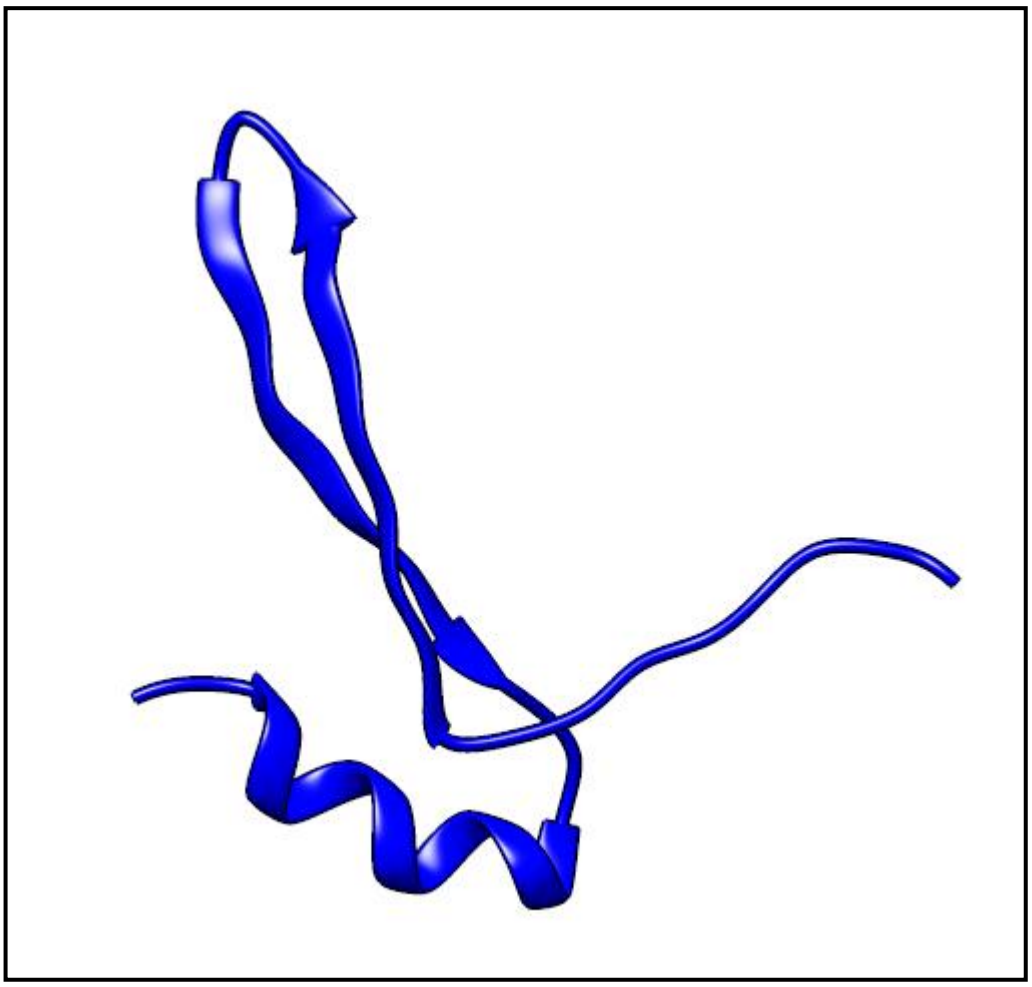


Figure 3.11: Energy minimized model of the switch domains of S. mansoni Smp_059340.1 by YASARA energy minimization server.

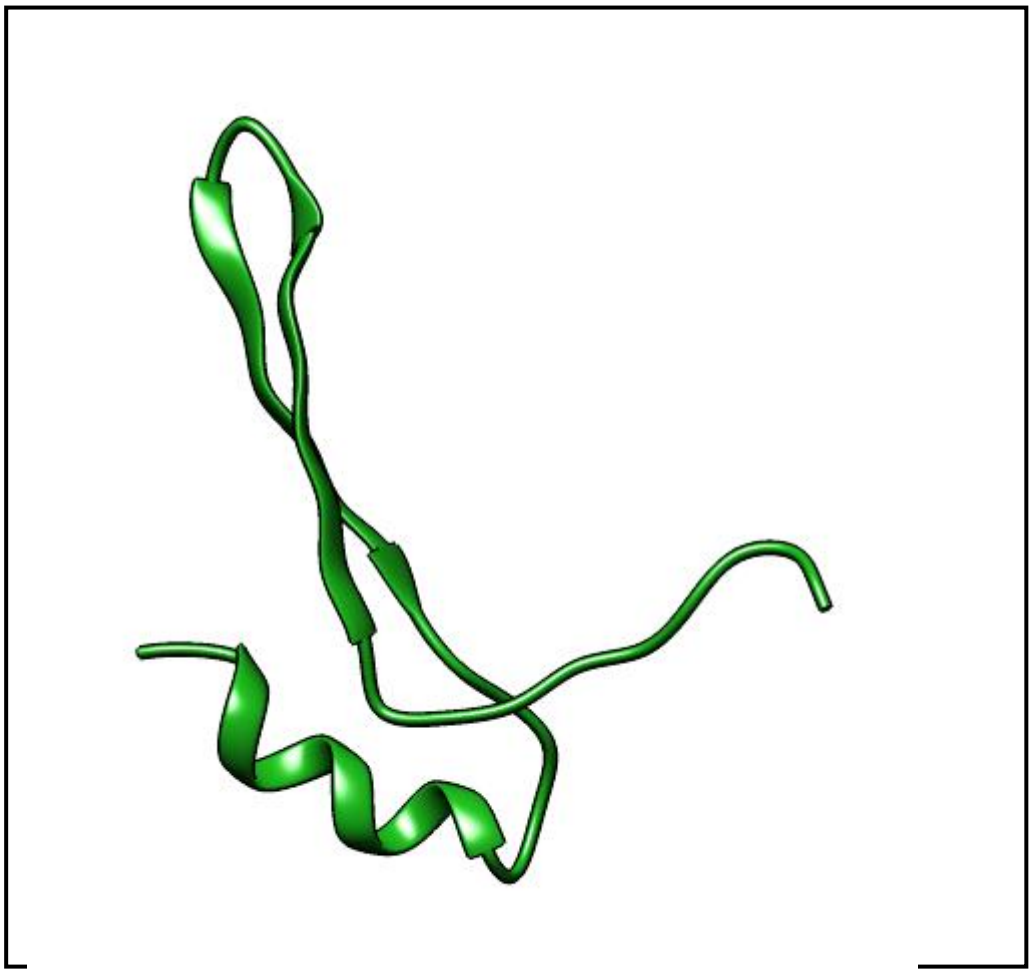


Figure 3.12: Energy minimized model of the switch domains of S. mansoni Smp_059340.1 by GalaxyRefine server for protein structure refinement.

3.3.8 Network interaction analysis of the switch domains of S. mansoni Smp_059340.1

The protein network interaction analysis determined by STRING (Figure 3.13) reveals that *S. mansoni* Smp_059340.1 interacts with ten other proteins; among them were Putative camp-dependent protein kinase type II regulatory subunit (Smp_131050)**,** putative camp-dependent protein kinase regulatory chain (Smp_019280), putative camp-dependent protein kinase type II-alpha regulatory subunit, (Smp_079010), putative camp-dependent protein kinase type I-beta regulatory subunit (Smp_030400), camp-dependent protein kinase type I-beta regulatory subunit (Smp_022100), putative camp-dependent rap1 guanine-nucleotide exchange factor (Smp_176220).

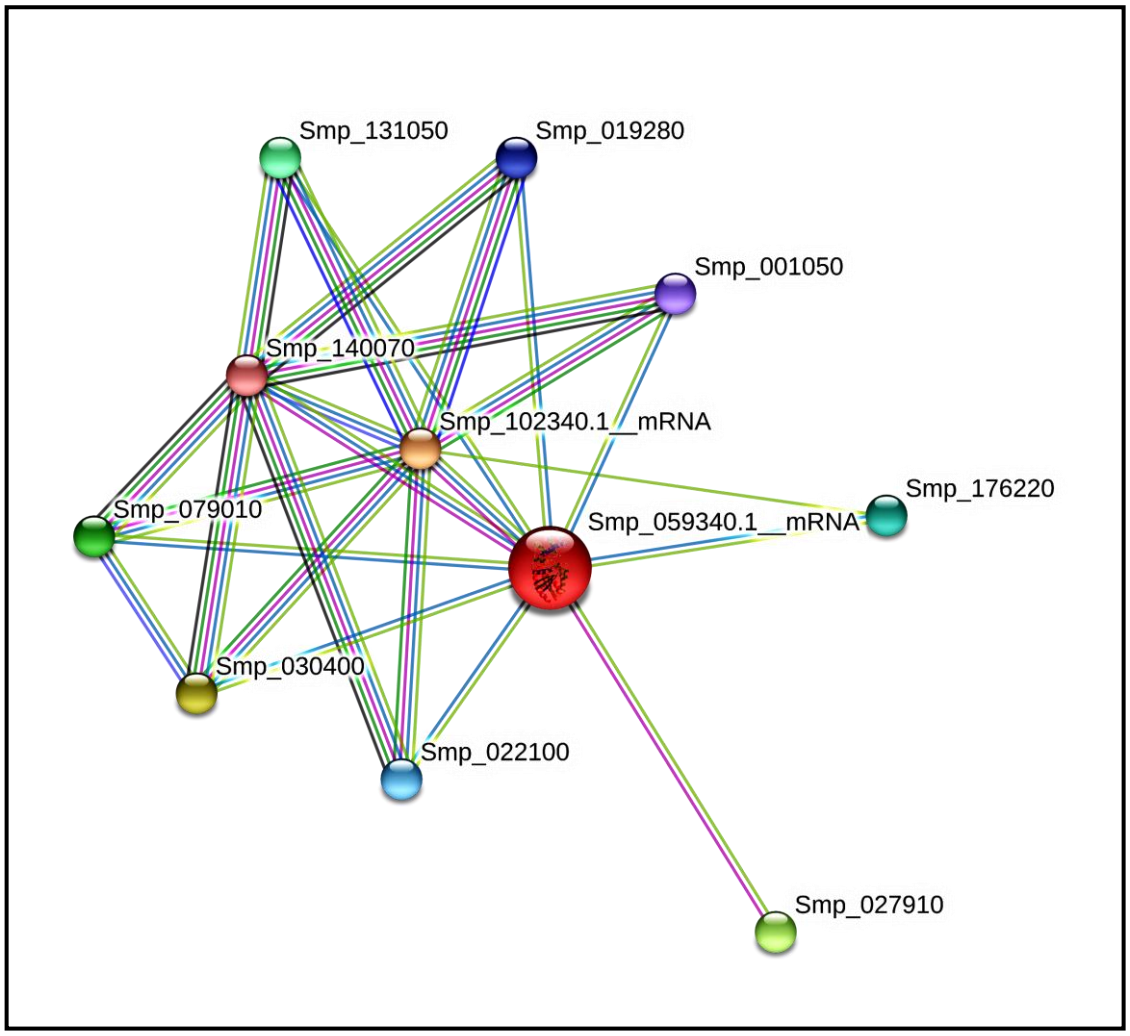


Figure 3.13: Protein-protein interaction network of S. mansoni Smp_059340.1 generated by STRING.

3.3.9 Active site analysis and residues recognition

The ligand binding sites of the switch domains of *S. mansoni* Smp_059340.1 was predicted using the COACH protein-ligand binding prediction server and the GalaxySite web server. As predicted by the COACH server, the residues involved in forming the binding pockets for ligands are presented in Table 3.2. Added to this, the residues involved in forming the binding pockets; which may play important roles in protein-ligand interaction and their hydrophobic representation (viewed in Chimera), as predicted by GalaxySite web server are presented in Figure 3.14 A, B, C, D, E, F, G and H respectively.

Table 3.2: Predicted ligand binding sites in the switch domains of S. mansoni Smp_059340.1 using the homology model generated by Phyre2

Name of Server	Name of ligand	Residue Number	C-Score
COACH Server	Guanosine 5'- [gamma-thio] triphosphate (GSP)	1, 2, 4, 6, 7, 28, 29, 30	0.35
	Guanosine-5'-diphosphate (GDP)	1, 2, 3, 4	0.09
	Magnesium (Mg^{2+})	7, 26, 30, 33	0.04
	Calcium (Ca^{2+})	12, 23	0.02
	Peptide	27, 28, 31, 32, 33, 34, 37, 38, 41	0.03
TM-Site Results	Magnesium (Mg^{2+})	7, 30, 31, 33, 34	0.23
	Calcium (Ca^{2+})	12, 23, 31, 34	0.23
	Manganese (Mn^{2+})	12, 23	0.22
	Zinc (Zn^{2+})	12, 23	0.22
	Peptide	10, 25, 39, 42	0.16
S-Site Results	Guanosine 5'- [gamma-thio] triphosphate (GSP)	1, 2, 4, 5, 6, 7, 27, 28, 29, 30	0.42
	Guanosine-5'-diphosphate (GDP)	1, 2, 3, 4	0.19
	Magnesium (Mg^{2+})	1, 2, 4, 5, 6, 7, 10, 11, 12, 13, 23, 25, 26, 27, 28, 29, 30	0.28
	Peptide	1, 2, 4, 5, 6, 7, 27, 28, 29, 30	0.42

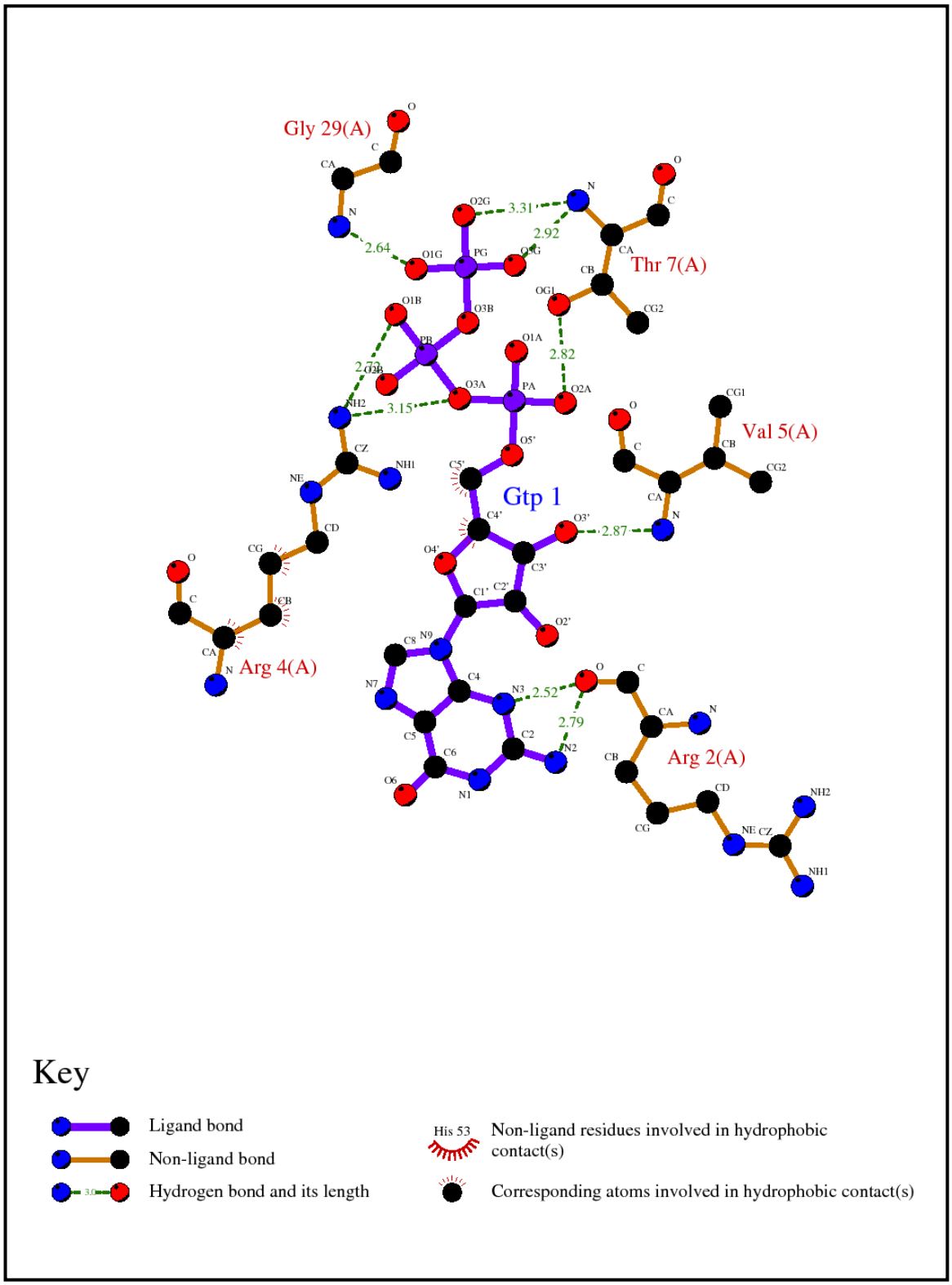


Figure 3.14A: GalaxySite web server prediction of interaction analysis between ligand (Guanosine-5'-triphosphate; GTP) predicted to bind the switch domains of S. mansoni Smp_059340.1 and ligand binding residues involved.

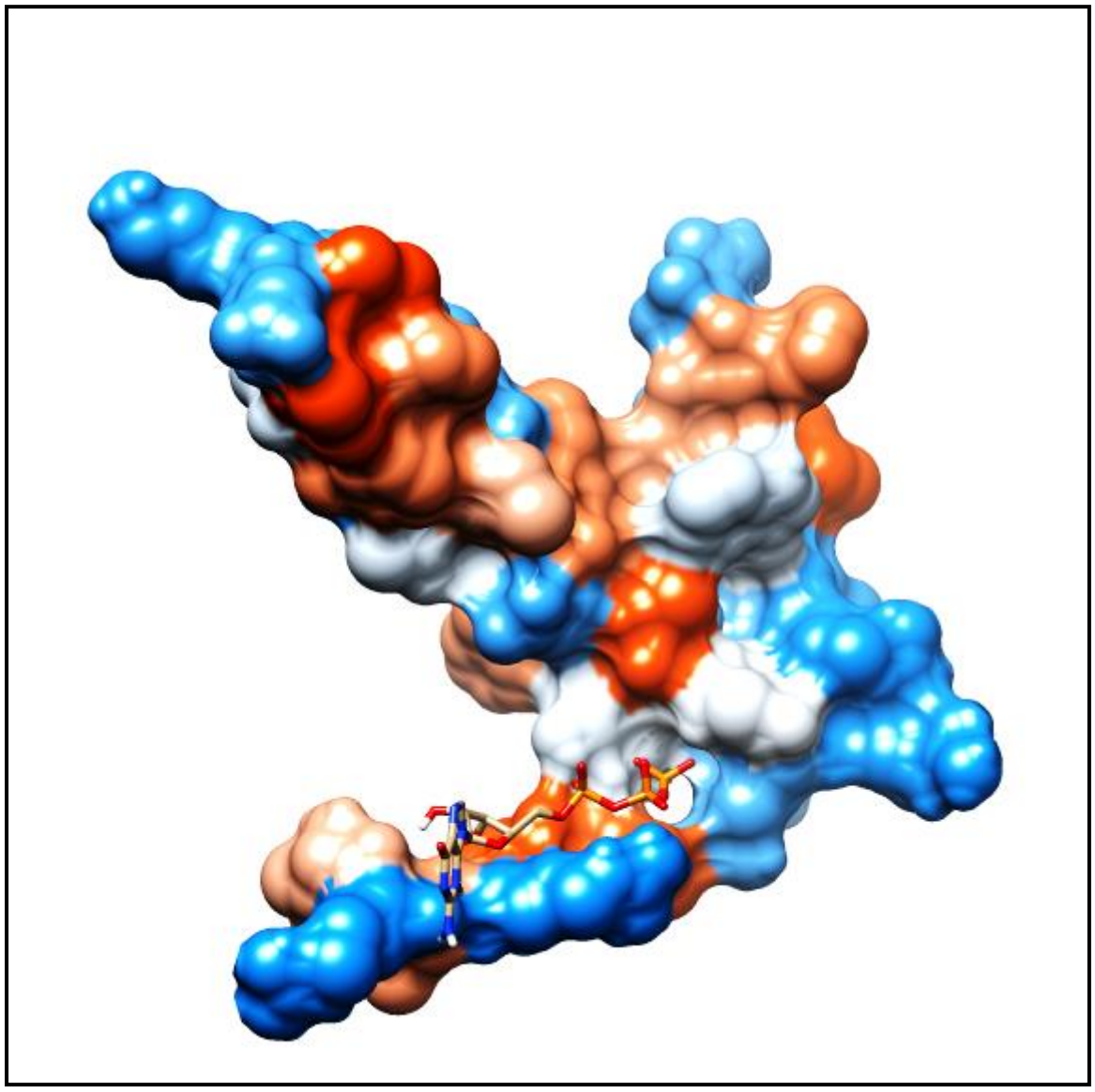


Figure 3.14B: Hydrophobic representation of GalaxySite web server prediction of interaction analysis between ligand (Guanosine-5'-triphosphate; GTP) predicted to bind the switch domains of S. mansoni Smp_059340.1 and ligand binding residues involved.

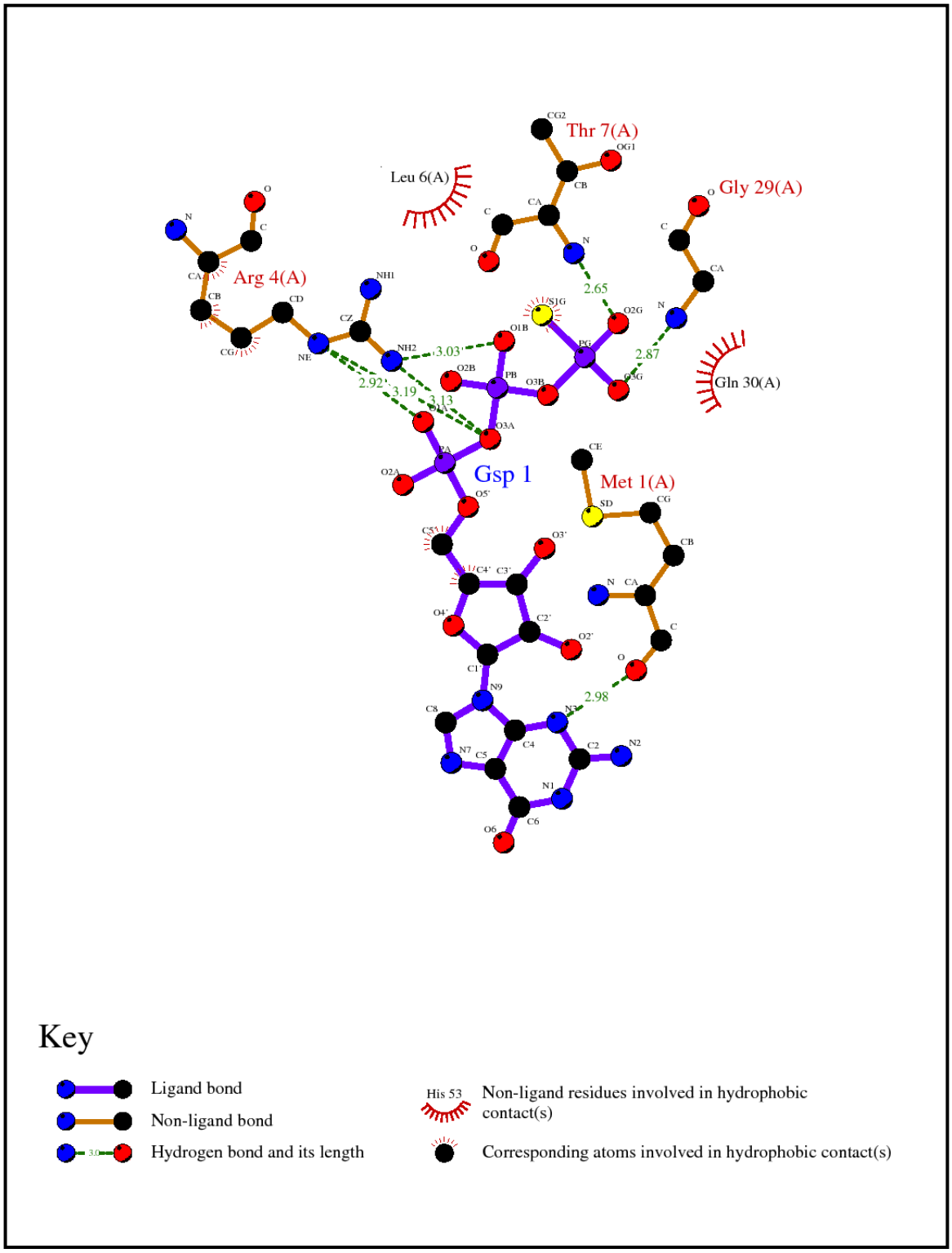


Figure 3.14C: GalaxySite web server prediction of interaction analysis between ligand (5'-guanosine-diphosphate-monothiophosphate; GSP) predicted to bind the switch domains of S. mansoni Smp_059340.1 and ligand binding residues involved.

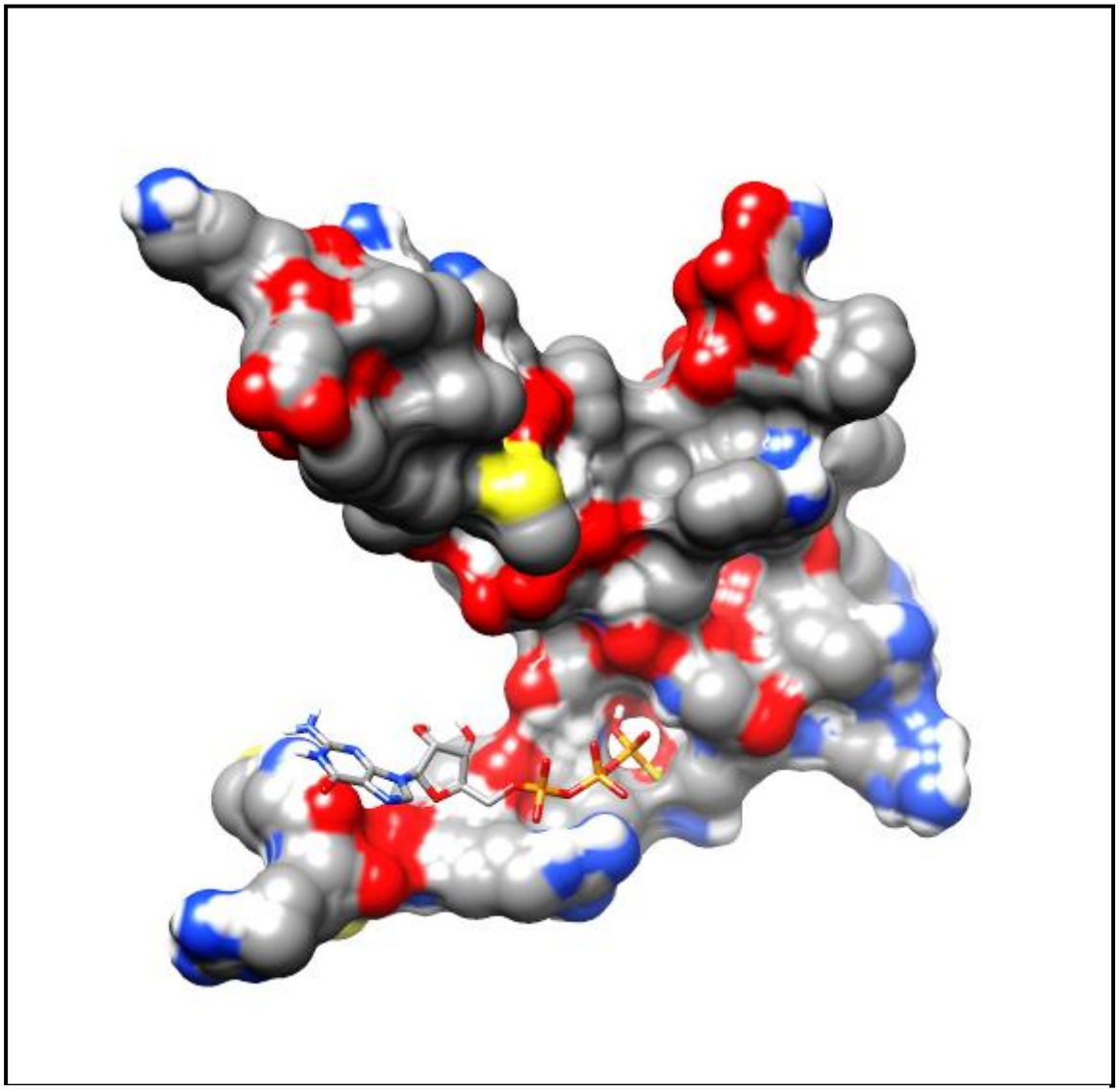


Figure 3.14D: Hydrophobic representation of GalaxySite web server prediction of interaction analysis between ligand (5'-guanosine-diphosphate-monothiophosphate; GSP) predicted to bind the switch domains of S. mansoni Smp_059340.1 and ligand binding residues involved

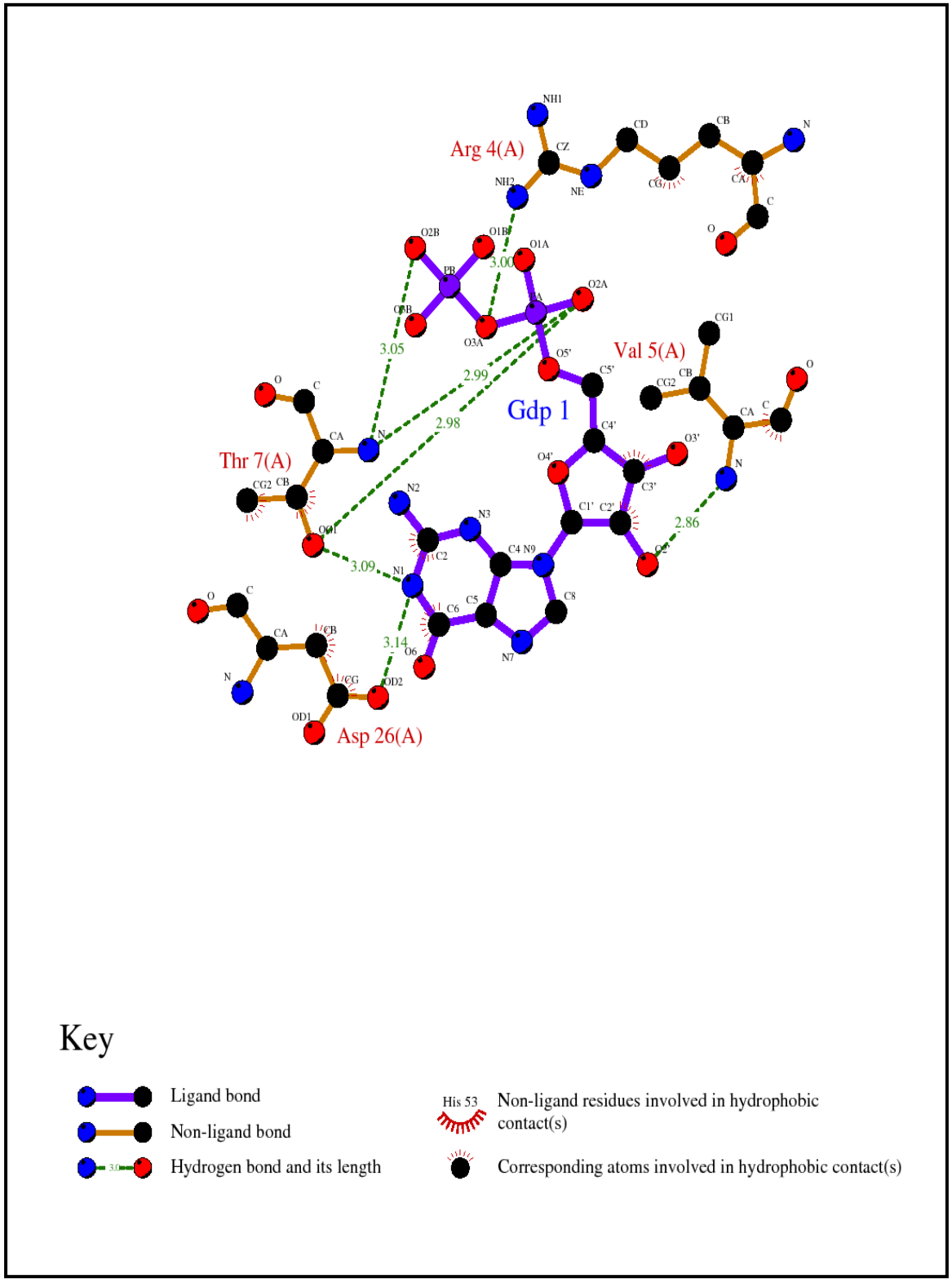


Figure 3.14E: GalaxySite web server prediction of interaction analysis between ligand (Guanosine-5'-diphosphate; GDP) predicted to bind the switch domains of S. mansoni Smp_059340.1 and ligand binding residues involved.

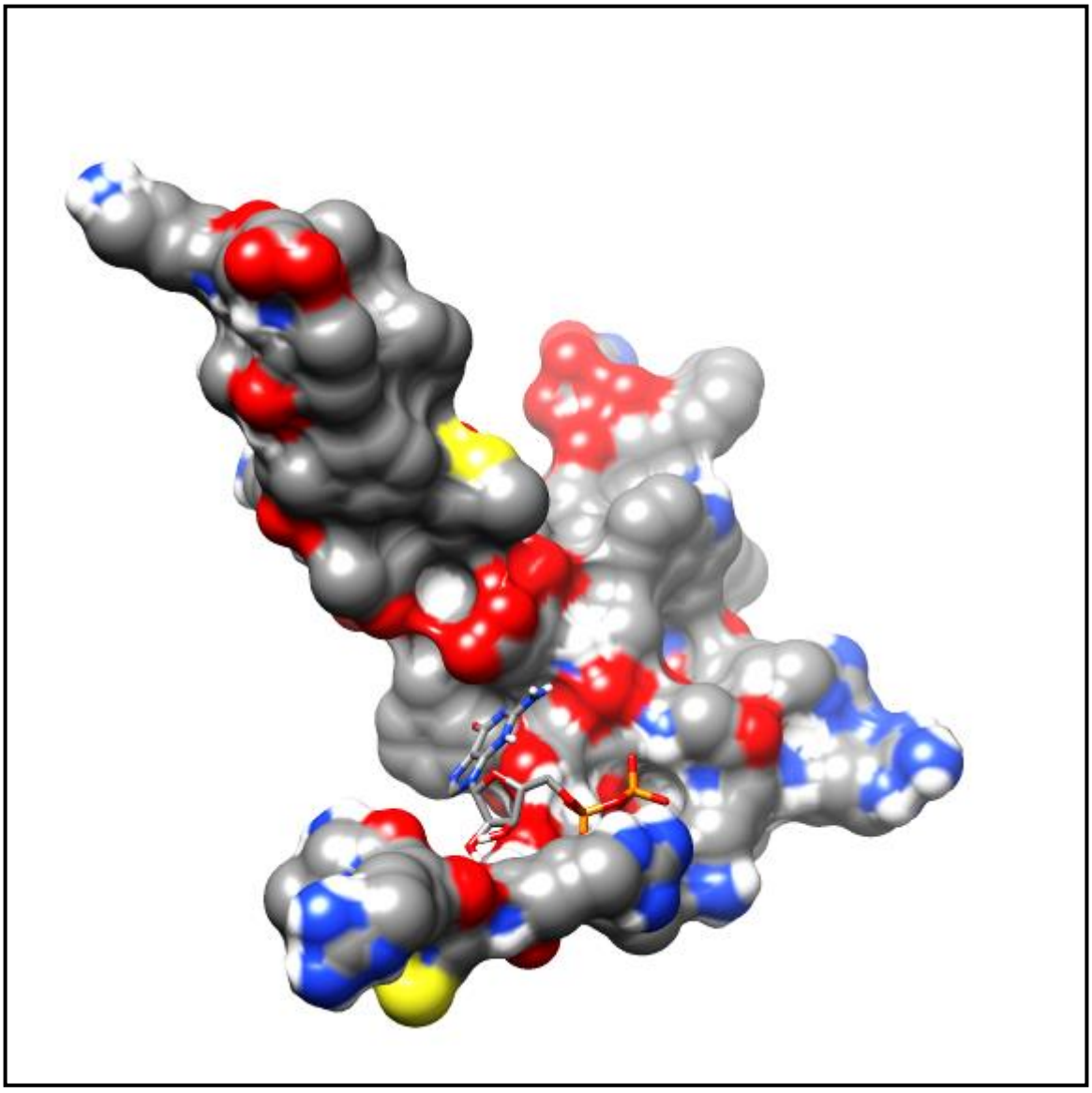


Figure 3.14F: Hydrophobic representation of GalaxySite web server prediction of interaction analysis between ligand (Guanosine-5'-diphosphate; GDP) predicted to bind the switch domains of S. mansoni Smp_059340.1 and ligand binding residues involved.

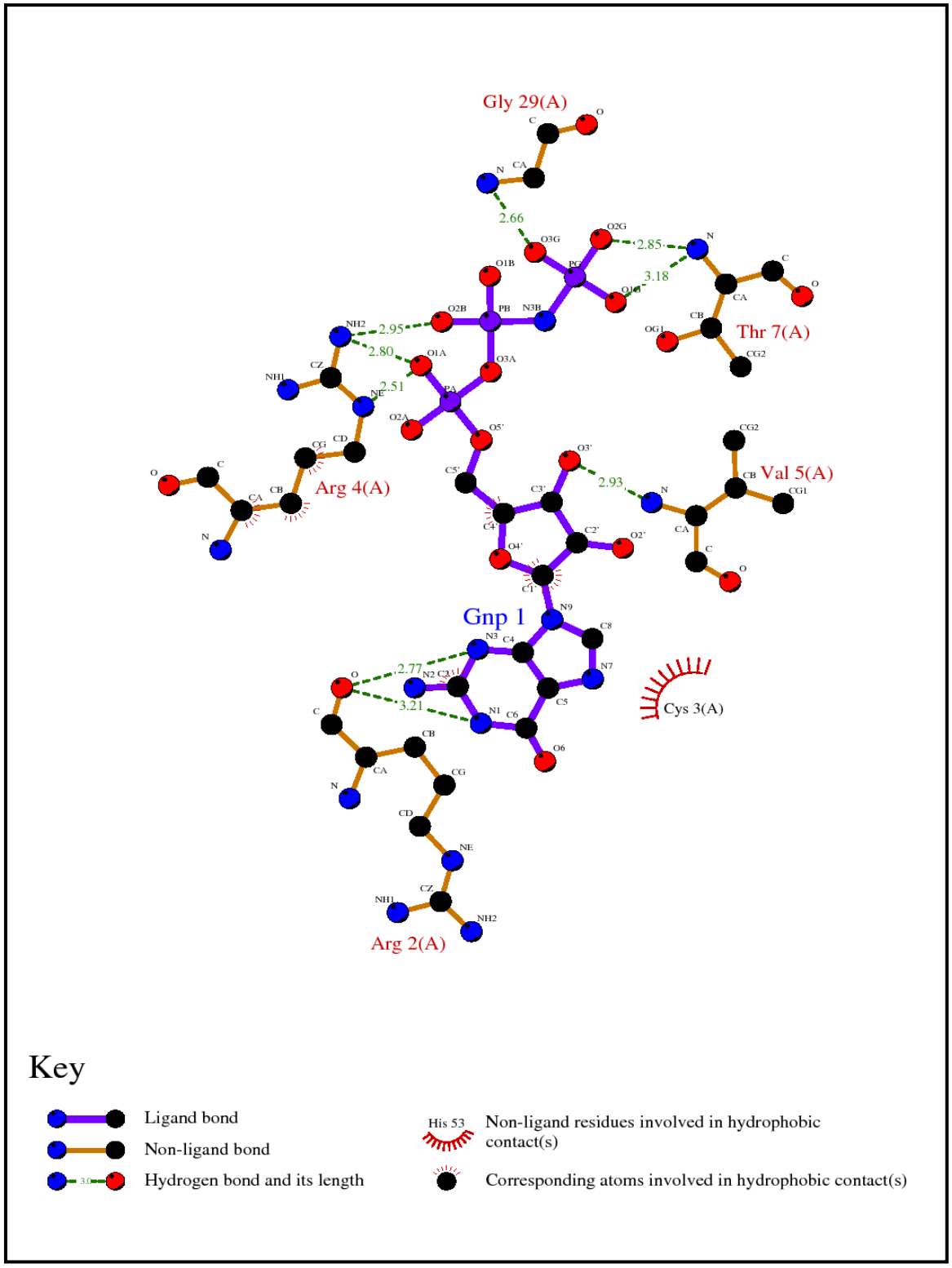


Figure 3.14G: GalaxySite web server prediction of interaction analysis between ligand (Phosphoaminophosphonic acid-guanylate ester; GNP) predicted to bind the switch domains of S. mansoni Smp_059340.1 and ligand binding residues involved.

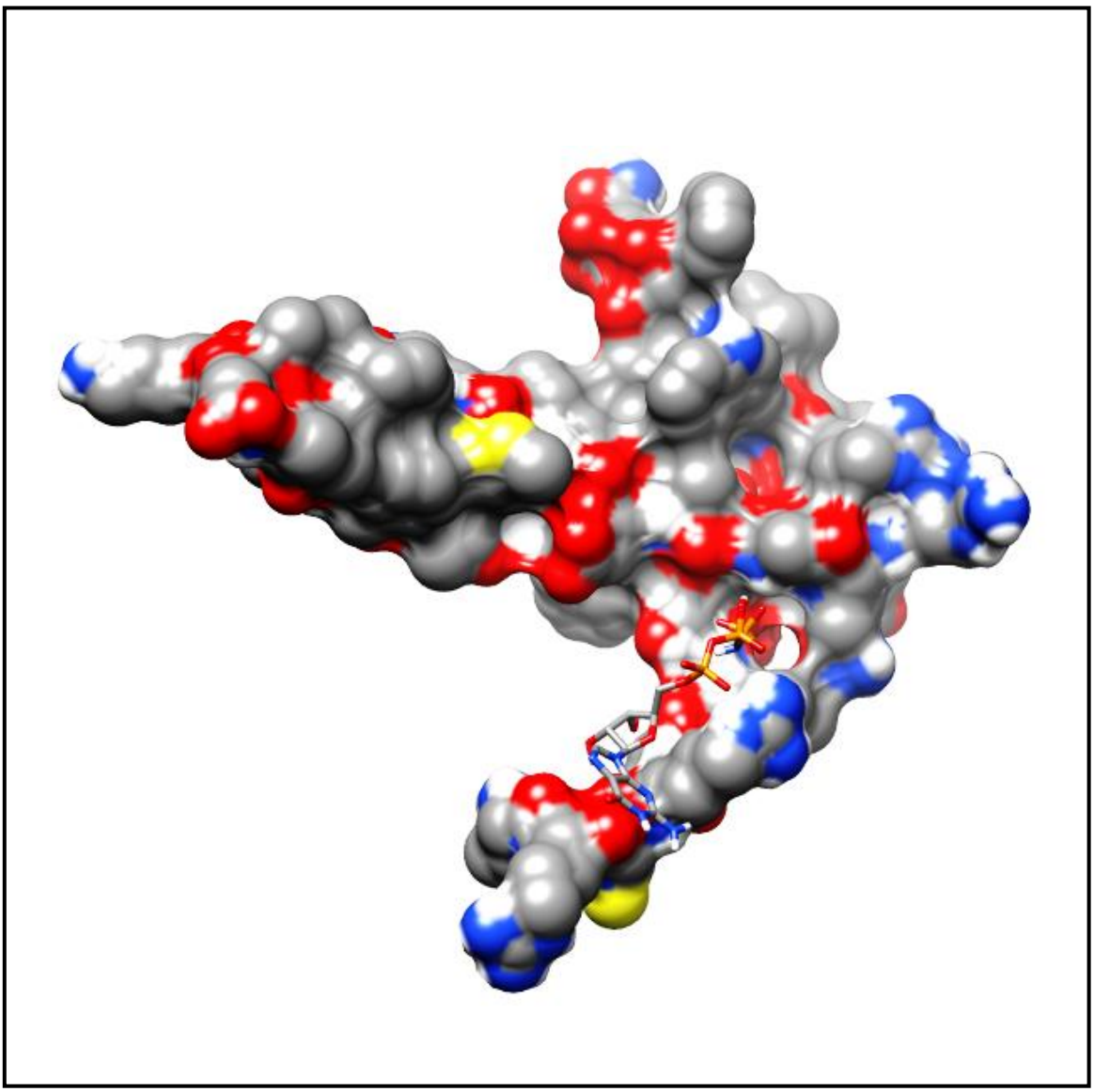


Figure 3.14H: Hydrophobic representation of GalaxySite web server prediction of interaction analysis between ligand (Phosphoaminophosphonic acid-guanylate ester; GNP) predicted to bind the switch domains of S. mansoni Smp_059340.1 and ligand binding residues invo

3.3.1 Molecular docking

The results from docking studies (Figure 3.15 A, B, C and D) using the PatchDock automatic server for molecular docking revealed the binding and interaction patterns of Guanosine-5'-triphosphate (GTP), Magnesium (Mg^{2+}), Calcium (Ca^{2+}) and Zinc (Zn^{2+}) with the switch domains of *S. mansoni* Smp_059340.1. Table 3.3 presents the scores and ACE of the docked complexes. The switch domains of *S. mansoni* Smp_059340.1 showed significant interactions with ligands, notably with GTP and Ca^{2+} with a score of 3266 and 778 and an ACE of -22.57 and -5.93 respectively.

Table 3.3: Scores and atomic contact energy (ACE) of the docked complexes

		Switch domains of *S. mansoni* Smp_059340.1	
	Ligand	**Score**	**ACE Kcal/mo**
1	**Guanosine-5'-triphosphate (GTP)**	3266	-22.57
2	**Magnesium (Mg^{2+})**	342	0.00
3	**Calcium (Ca^{2+})**	778	-5.93
4	**Zinc (Zn^{2+})**	342	0.00

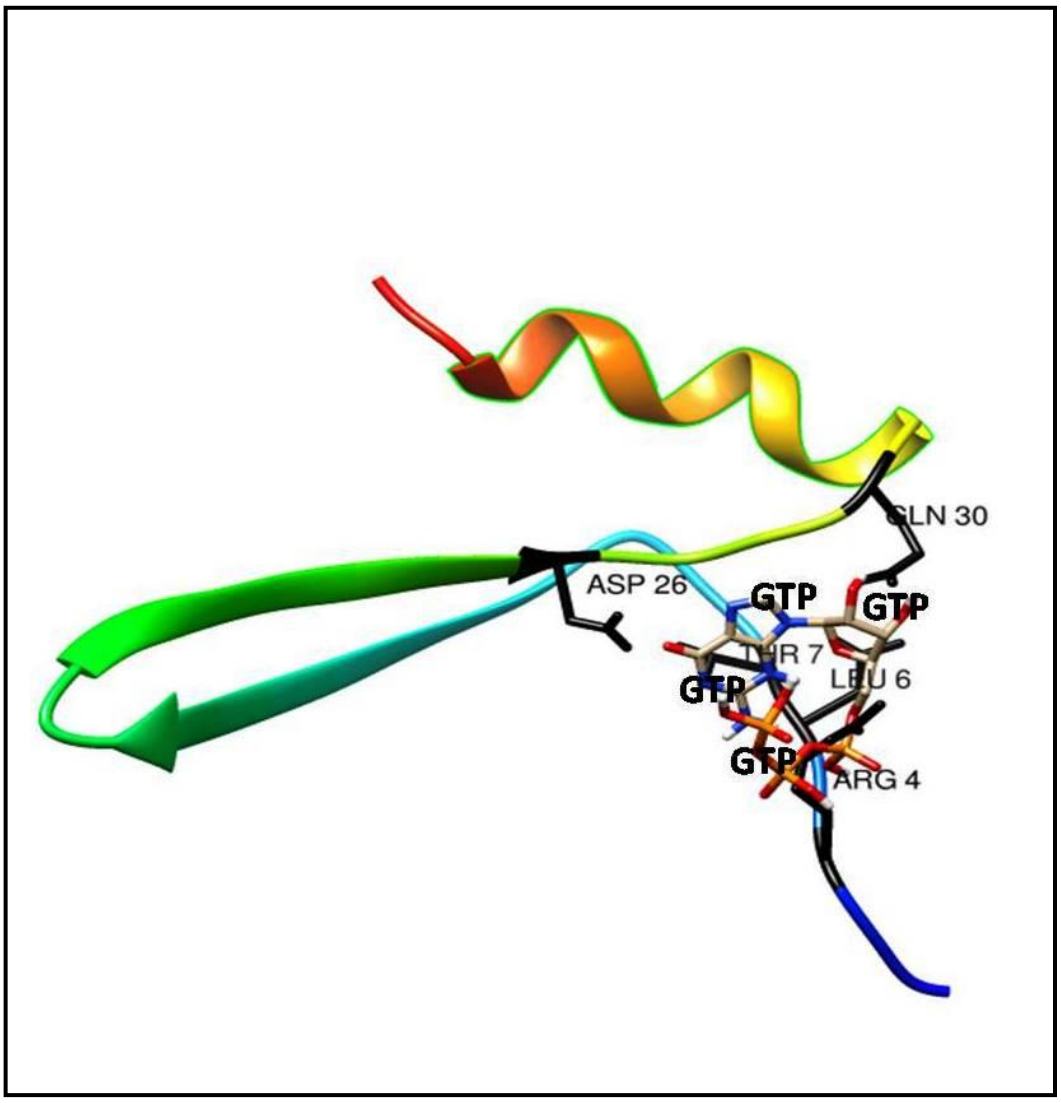


Figure 3.15A: Detailed molecular docking interaction of the switch domains of S. mansoni Smp_059340.1 and Guanosine-5'-triphosphate (GTP).

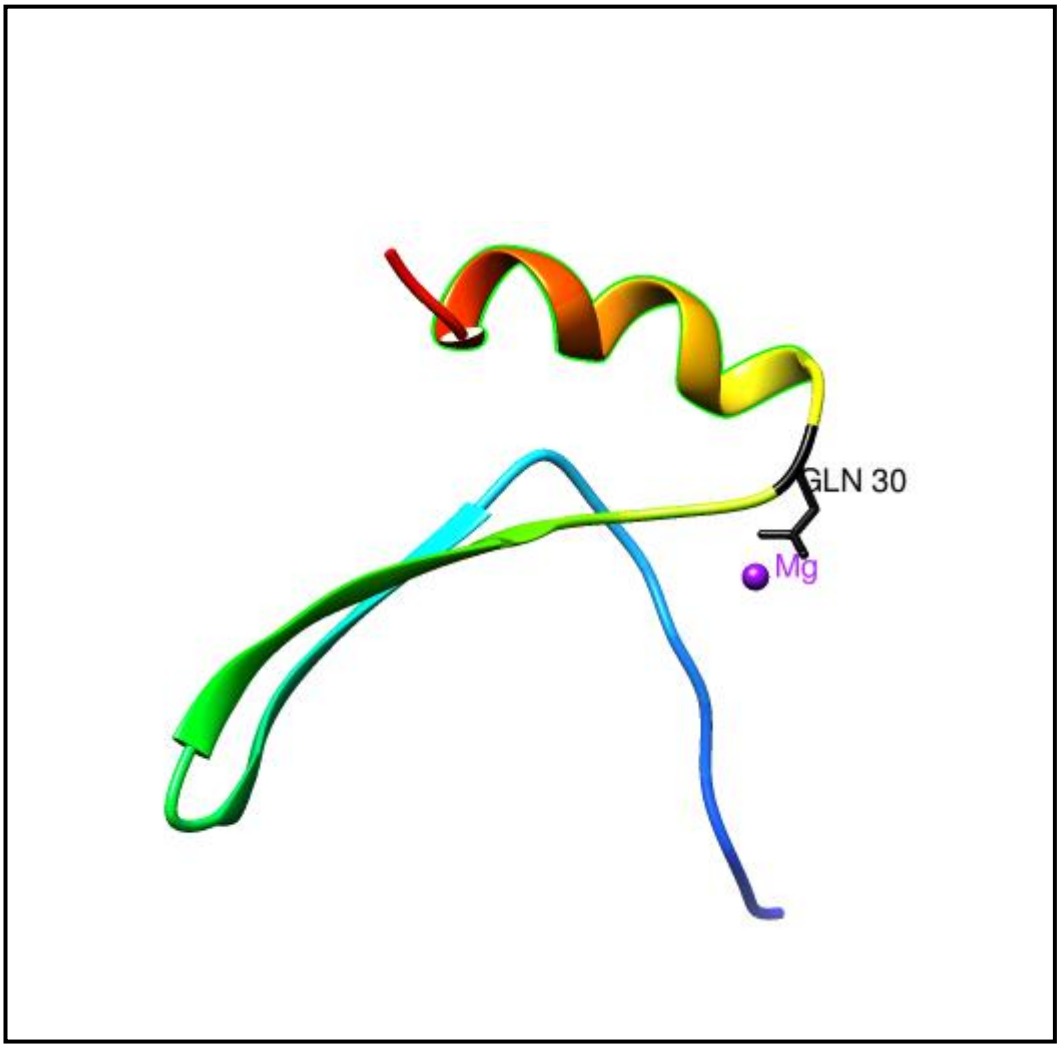


Figure 3.15B: Detailed molecular docking interaction of the switch domains of S. mansoni Smp_059340.1 and Magnesium (Mg2+).

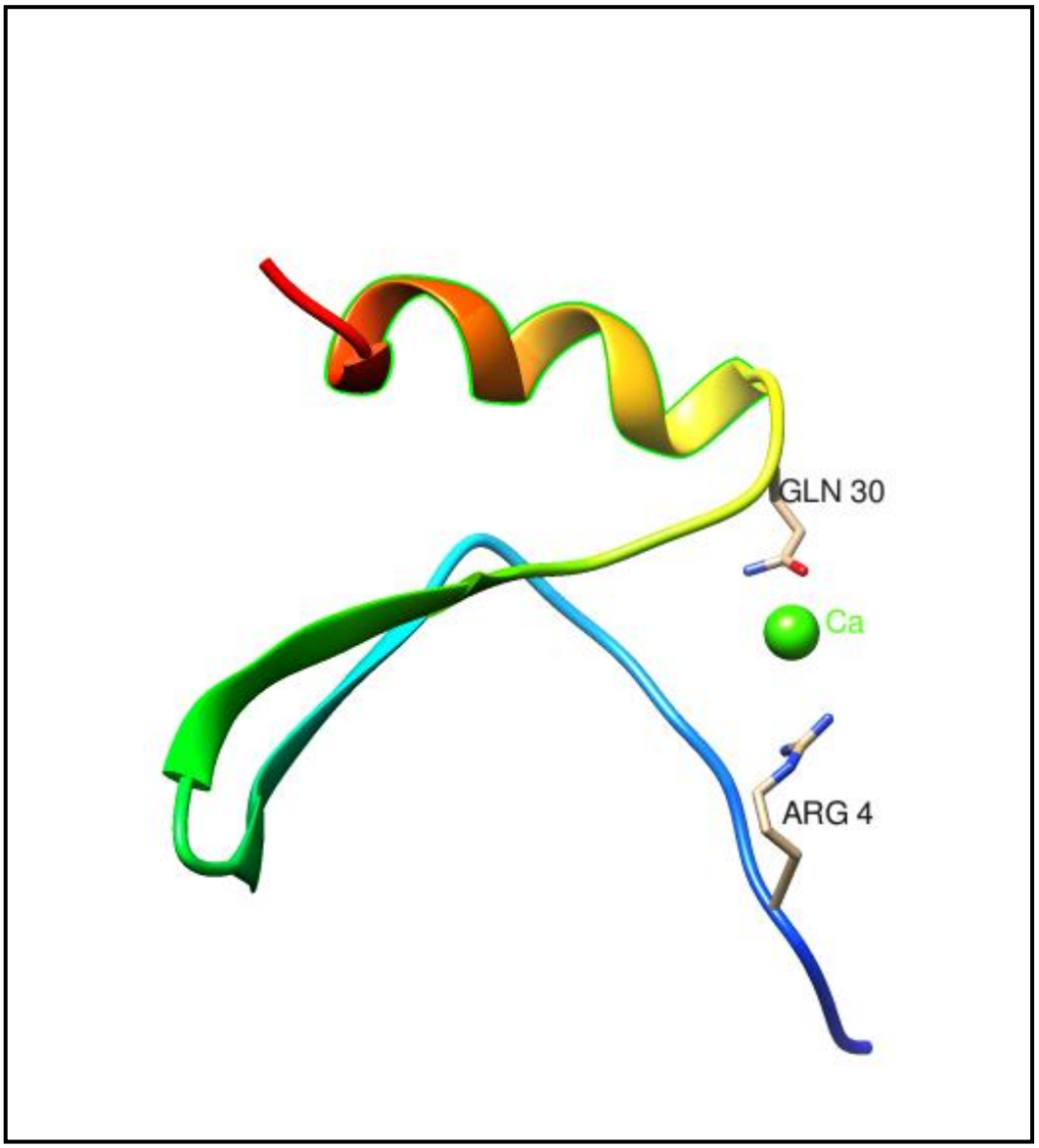


Figure 3.15C: Detailed molecular docking interaction of the switch domains of S. mansoni Smp_059340.1 and Calcium (Ca2+).

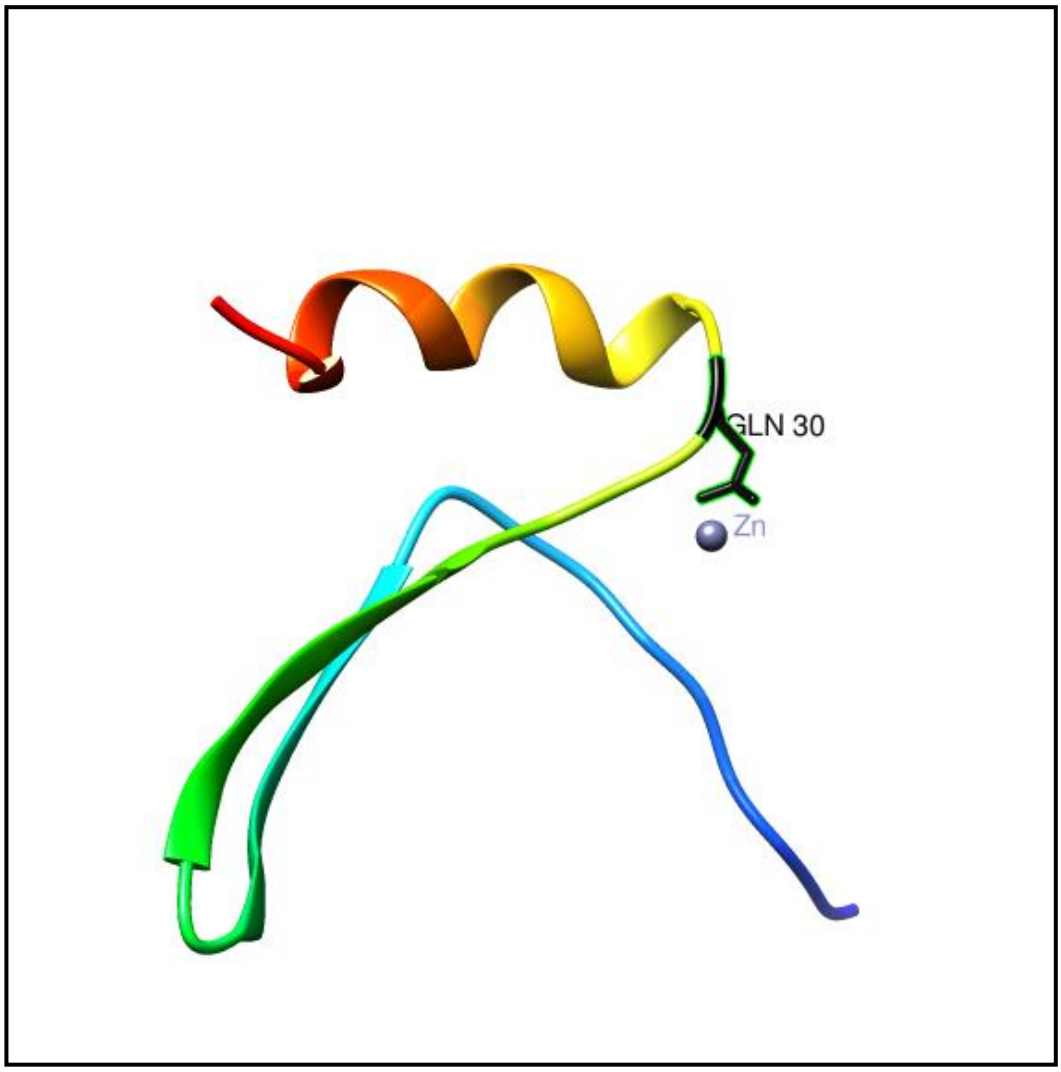


Figure 3.15D: Detailed molecular docking interaction of the switch domains of S. mansoni Smp_059340.1 and Zinc (Zn2+).

3.4 Discussion

The recent advances in *in-silico* methods have resulted in a paradigm shift in biomedical research. This paradigm shift have been identified to play a crucial role in the identification of new lead molecule-inhibitors, ligands and peptides not only in drug and vaccine discovery but also in drug and vaccine development (Gill *et al*., 2016; Shekhar, 2008; Wadood *et al*., 2013). *In-silico* methods are basically used to complement (alongside) the data generated by *in vitro* experimental studies. Results from both studies (*in-silico* and *in vitro*) are used to create a model and to test the accuracy of the model created (Ekins *et al*., 2007). Time and again it has been established that bioinformatics analysis greatly assist in predicting the function of proteins as well as predicting the protein structure through homology modeling, determination of the coding regions of nucleic acid sequences, identifying promising and suitable drug compounds through data mining and molecular docking as well as rapid analysis and interpretation of data generated, thereby, ultimately reducing the cost and time involved in drug discovery (Gill *et al*., 2016; Schwede, 2013).

In this study, it was important to employ various computational methods and tools to analyze the protein sequence in order to gain insight into the basic information of the protein stability and function. Multiple sequence analysis of the *S. mansoni* Smp_059340.1 protein shows that it belongs to the families of proteins with highly conserved functional or structural motifs with significant biological attributes (Doğan and Karaçalı, 2013). Multiple sequence alignments are important and extensively used computational procedure for biological sequence analysis in molecular biology, computational biology, and bioinformatics. Traditionally, multiple sequence alignments are used primarily to perform phylogenetic tree reconstruction, predict

protein secondary and tertiary structure as well as protein function prediction analysis. Multiple sequence alignments describe the relationships and homology between different sequences and also provide useful information for its downstream application, which can be used to further identify new members of the protein families (Blackburne and Whelan, 2012; Daugelaite *et al*., 2013; Kemena and Notredame, 2009).

Interestingly, analysis of probable motif using MOTIF finder server reveals the presence of two promising motifs: the G-protein alpha subunit and the ADP-ribosylation factor family. The G-protein alpha (Gα) subunit is known and considered as a bimodal, GTP-hydrolyzing switch, controlling the duration of signal transduction (Siderovski and Willard, 2005). Generally, G-proteins (guanine nucleotide-binding proteins) belongs to a family of proteins that act as molecular switches and timers that cycle between inactive guanosine diphosphate (GDP) bound and active guanosine triphosphate (GTP) bound states (Vetter and Wittinghofer, 2001). In relation to the above findings, evolutionary conservation analysis of the switch domains of *S. mansoni* Smp_059340.1 revealed a universally conserved residue (DVGGQ) identified to be a GTP binding site. Evolutionary conservation analysis of proteins have been commonly used to characterize proteins in order to identify important residues with functional and structural features, to identify protein–protein interactive patterns as well as to predict possible interactions between a protein and its ligand (Goswami, 2017; Choi *et al*., 2010).

In order to gain insight into protein structure and function; prediction of the protein SS is very important. Protein secondary structure (SS) refers to the local conformation of the polypeptide backbone of proteins and this is crucial in determination of its function and many of its downstream applications like drug and

vaccine design. In recent times, protein SS prediction has been extensively studied using various methods and tools. Theoretically, it has been suggested that proteins exist in two regular protein SS states: alpha-helix and beta-strand and one irregular SS type known as the coil region (Wang *et al*., 2016). Results from this study revealed that the switch domains of *S. mansoni* Smp_059340.1 contain 1 α-helices, 2 β-sheets and 4 random coils. On the whole, protein secondary structure can be considered the bridge that connects the primary sequence and tertiary structure and therefore, is used by many structure and functional analysis tools. Therefore, accurate prediction of protein structure and function relies, in part, on the accuracy of secondary structure prediction (Wang *et al*., 2016).

Protein structure determination using experimental methods such as X-ray crystallography or Nuclear magnetic resonance (NMR) spectroscopy is time consuming and not always successful with all proteins, due to the fact that many proteins are basically too large for NMR analysis and cannot be crystallized for X-ray diffraction. Therefore, protein modeling using any of the available computational methods is the only way to obtain structural information to complement the results obtained from experimental methods (Krieger *et al*., 2003; Vyas *et al*., 2012). In this study, the 3D structure of the protein obtained from Raptor X and Phyre 2 server showed that both models are stable and of good quality. However, the overall evaluation of the stereochemical parameters using Ramachandran plot analysis in RAMPAGE showed that Phyre 2 model was slightly superior in nature, because the model has 97.5 % of its residues in the favoured region while 2.5 % of its residues are located in the allowed region as compared with Raptor X model which has 95.0 % of its residues in the favoured region and 5.0 % of its residues located in the

allowed region. The energy minimization of the modelled protein structure eventually results in a thermodynamically stable model, confirming the accuracy of the model.

Proteins interact with a wide variety of biomolecules within a cell. Understanding their interaction network reveals important biological information about their signaling cascades as well as their predicted functions in various pathways in biological processes (Mier *et al*., 2017; Turanalp and Can, 2008). The protein network interaction analysis determined by STRING shows that *S. mansoni* Smp_059340.1 interacts with ten other proteins. Additionally, the COACH protein-ligand binding prediction server and the GalaxySite web server predicted the residues and the ligand binding sites of the switch domains of *S. mansoni* Smp_059340.1; that may be involved in forming the binding pockets for ligand interaction. Active site determination through COACH and GalaxySite web server suggests that the modelled structure can be utilized as a potential drug target. Protein binding sites are the places where molecular interactions take place. Therefore, the analysis of protein binding sites is of crucial importance to understand the biological processes proteins are involved in (Nisius *et al*., 2012).

The recent reports on the emergence of the prospect of praziquantel (PZQ) resistance in some *Schistosoma* species; calls for an urgent need to screen and develop novel biologically active compounds that can be developed as new drugs to combat schistosomiasis (Qi and Cui, 2013). Molecular docking has become an increasingly important tool for drug discovery. The docking procedure involves two major steps: prediction of the ligand conformation as well as its position and orientation within these sites (often referred to as pose) and assessment of the binding affinity (Meng *et al*., 2011). Molecular docking studies using the PatchDock automatic server for molecular docking revealed the binding and interaction patterns

of Guanosine-5'-triphosphate (GTP: at residue ARG 4, LEU 6, THR 7, ASP 26, GLN 30), Magnesium (Mg^{2+}: at residue GLN 30), Calcium (Ca^{2+} : at residue ARG 4, GLN 30) and Zinc (Zn^{2+} at residue GLN 30) with the switch domains of *S. mansoni* Smp_059340.1. These interactions are believed to be of significant importance because the residues involved in the interaction are known to be evolutionarily conserved and functionally important residues. Hence, it provides the opportunity to further understand and explain the structural features of the switch domains of *S. mansoni* Smp_059340.1.

In conclusion, based on bioinformatics analysis of the switch domains of *S. mansoni* Smp_059340.1, the evolutionary conserved regions have been identified along with the basic structural features having a GTP binding site (DVGGQ). Homology models generated to gain insight into the three dimensional space demonstrates a stable model that can play a fundamental role in further elucidation of the protein structure and function. Added to this, the detection of active site and residues recognition in the switch domains of *S. mansoni* Smp_059340.1 is impressive, as it implies its druggability and provides a starting point for rational drug design of new antischistosomal agents with a novel mode of action.

REFERENCES

Adenowo A.F., Oyinloye B.E., Ogunyinka B.I. and Kappo A.P. (2015). Impact of human schistosomiasis in sub-Saharan Africa. *Brazilian Journal of Infectious Diseases*. **19**(2): 196-205.

Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W. and Lipman D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*. **25**: 3389-3402.

Ashkenazy H., Erez E., Martz E., Pupko T. and Ben-Tal N. (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Research*. **38**: W529–W533.

Ben-Hur A. and Noble W.S. (2005). Kernel methods for predicting protein–protein interactions. *Bioinformatics*. **21**(1): i38-i46.

Berriman M., Haas B.J., LoVerde P.T., Wilson R.A., Dillon G.P., Cerqueira G.C., Mashiyama S.T., Al-Lazikani B., Andrade L.F., Ashton P.D., Aslett M.A., Bartholomeu D.C., Blandin G., Caffrey C.R., Coghlan A., Coulson R., Day T.A., Delcher A., DeMarco R., Djikeng A., Eyre T., Gamble J.A., Ghedin E., Gu Y., Hertz-Fowler C., Hirai H., Hirai Y., Houston R., Ivens A., Johnston D.A., Lacerda D., Macedo C.D., McVeigh P., Ning Z., Oliveira G., Overington J.P., Parkhill J., Pertea M., Pierce R.J., Protasio A.V., Quail M.A., Rajandream M.A., Rogers J., Sajid M., Salzberg S.L., Stanke M., Tivey A.R., White O., Williams D.L., Wortman J., Wu W., Zamanian M., Zerlotini A., Fraser-Liggett C.M., Barrell B.G. and El-Sayed N.M. (2009). The genome of the blood fluke Schistosoma mansoni. *Nature*. **460**: 352-328.

Blackburne B.P. and Whelan S. (2012). Measuring the distance between multiple sequence alignments. *Bioinformatics*. **28**(4): 495-502.

Chang D.T.H., Syu Y.T. and Lin P.C. (2010). Predicting the protein-protein interactions using primary structures with predicted protein surface. *BMC Bioinformatics*. **11**(1): S3.

Chen W., Hong Y.Q. and Meng Z.L. (2014). Bioinformatics analysis of molecular mechanisms of chronic obstructive pulmonary disease. *European Review for Medical and Pharmacological Sciences*. **18**(23): 3557-3563.

Choi Y.S., Han S.K., Kim J., Yang J.S., Jeon J., Ryu S.H. and Kim S. (2010). ConPlex: a server for the evolutionary conservation analysis of protein complex structures. *Nucleic Acids Research*. **38**(2): W450-W456.

Daugelaite J., O'Driscoll A. and Sleator R.D. (2013). An overview of multiple sequence alignments and cloud computing in bioinformatics. *ISRN Biomathematics*. **20**(13): 615-630.

Doğan T. and Karaçalı B. (2013). Automatic identification of highly conserved family regions and relationships in genome wide datasets including remote protein sequences. *PloS One*. **8**(9): e75458.

Drozdetskiy A., Cole C., Procter J. and Barton G.J. (2015). JPred4: a protein secondary structure prediction server. *Nucleic Acids Research*. **43** (W1): W389–W394.

Duhovny D., Nussinov R., Wolfson H.J., Guigo R., and Gusfield D. (2002). Efficient unbound docking of rigid molecules. In: editors. Proceedings of the Fourth

International Workshop on Algorithms in Bioinformatics; *Springer-Verlag GmbH Rome, Italy.* 185–200.

Ekins S., Mestres J. and Testa B. (2007). In silico pharmacology for drug discovery: methods for virtual ligand screening and profiling. *British Journal of Pharmacology*. **152**(1): 9-20.

Elumalai A. and Eswaraiah M.C. (2013). Review on application of bioinformatics. *Journal of Science Bioinformatics*. **3**(1): 21-27.

Gill S.K., Christopher A.F., Gupta V. and Bansal P. (2016). Emerging role of bioinformatics tools and software in evolution of clinical research. *Perspectives in Clinical Research*. **7**(3): 115.

Goswami A.M. (2017). Computational analysis, structural modeling and ligand binding site prediction of Plasmodium falciparum 1-deoxy-d-xylulose-5-phosphate synthase. *Computational Biology and Chemistry*. **66**: 1-10.

Grant M.A. (2011). Integrating computational protein function prediction into drug discovery initiatives. *Drug Development Research*. **72**(1): 4-16.

Hall T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*. **41**:95-98.

Heo L., Shin W.H., Lee M.S. and Seok C. (2014). GalaxySite: ligand-binding-site prediction by using molecular docking. *Nucleic Acids Research*. **42**(W1): W210-W214.

Hue M., Riffle M., Vert J.P. and Noble W.S. (2010). Large-scale prediction of protein-protein interactions from structures. *BMC Bioinformatics*. **11**(1): 144.

Källberg M., Margaryan G., Wang S., Ma J. and Xu J. (2014). RaptorX server: a resource for template-based protein structure modeling. *Protein Structure Prediction*. **1137**:17-27.

Kanehisa M. and Bork P. (2003). Bioinformatics in the post-sequence era. *Nature Genetics*. **33:** 305-310.

Kelley L.A., Mezulis S., Yates C.M., Wass M.N. and Sternberg M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*. **10**(6): 845-858.

Kemena C. and Notredame C. (2009). Upcoming challenges for multiple sequence alignment methods in the high-throughput era. *Bioinformatics*. **25**(19): 2455-2465.

Krieger E., Nabuurs S.B. and Vriend G. (2003). Homology modeling. *Methods of Biochemical Analysis*. **44**: 509-524.

Lovell S.C., Davis I.W., Arendall W.B., de Bakker P.I., Word J.M., Prisant M.G., Richardson J.S. and Richardson D.C. (2003). Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. *Proteins: Structure, Function, and Bioinformatics*. **50**(3): 437-450.

Meng X.Y., Zhang H.X., Mezei M. and Cui M. (2011). Molecular docking: a powerful approach for structure-based drug discovery. *Current Computer-aided Drug Design.* **7**(2): 146-157.

Mier P., Alanis-Lobato G. and Andrade-Navarro M.A. (2017). Protein-protein interactions can be predicted using coiled coil co-evolution patterns. *Journal of Theoretical Biology*. **412**: 198-203.

Mourão M.D.M., Grunau C., LoVerde P.T., Jones M.K. and Oliveira G. (2012). Recent advances in Schistosoma genomics. *Parasite Immunology*. **34**(2-3): 151-162.

Nadzirin N. and Firdaus-Raih M. (2012). Proteins of Unknown function in the Protein Data Bank (PDB): an inventory of true uncharacterized proteins and computational tools for their analysis. *International Journal of Molecular Sciences*. **13**(10): 12761-12772.

Nahum L.A., Mourao M.M. and Oliveira G. (2012). New frontiers in Schistosoma genomics and transcriptomics. *Journal of Parasitology Research*. **2012**: 849132.

Nisius B., Sha F. and Gohlke H. (2012). Structure-based computational analysis of protein binding sites for function and druggability prediction. *Journal of Biotechnology*. **159**(3): 123-134.

Oliveira K.C., Carvalho M.L., Venancio T.M., Miyasato P.A., Kawano T., DeMarco R. and Verjovski-Almeida S. (2009). Identification of the Schistosoma mansoni TNF-alpha receptor gene and the effect of human TNF-alpha on the parasite gene expression profile. *PLOS Neglected Tropical Diseases*. **3**(12): e556.

Pettersen E.F., Goddard T.D., Huang C.C., Couch G.S., Greenblatt D.M., Meng E.C. and Ferrin T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*. **25**(13): 1605-1612.

Qi L. and Cui J.A. (2013). A schistosomiasis model with praziquantel resistance. *Discrete Dynamics in Nature and Society.* **2013**: 945767.

Roche D.B., Brackenridge D.A. and McGuffin L.J. (2015). Proteins and their interacting partners: an introduction to protein–ligand binding site prediction methods. *International Journal of Molecular Sciences*. **16**(12): 29829-29842.

Schneidman-Duhovny D., Inbar Y., Nussinov R. and Wolfson H.J. (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Research*. **33**(S2): W363-W367.

Scholz S.W., Mhyre T., Ressom H., Shah S. and Federoff H.J. (2012). Genomics and bioinformatics of Parkinson's disease. *Cold Spring Harbor Perspectives in Medicine*. **2**(7): a009449.

Schwede T. (2013). Protein modeling: what happened to the “protein structure gap”?. *Structure*. **21**(9): 1531-1540.

Shekhar C. (2008). In silico pharmacology: computer-aided methods could transform drug development. *Chemistry and Biology*. **15**(5): 413-414.

Siderovski D.P. and Willard F. S. (2005). The GAP’s, GEF’s and GDI’s of heterotrimeric G-protein alpha subunits. *International Journal of Biological Sciences*. **1**: 51 – 66.

Snel B., Lehmann G., Bork P. and Huynen M.A. (2000). STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Research*. **28**(18): 3442-3444.

Thompson J.D., Higgins D.G. and Gibson T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*. **22**(22): 4673-4680.

Turanalp M.E. and Can T. (2008). Discovering functional interaction patterns in protein-protein interaction networks. *BMC Bioinformatics*. **9**(1): 276.

Vetter I.R. and Wittinghofer A. (2001). The guanine nucleotide-binding switch in three dimensions. *Science*. **294**(5545): 1299-1304.

Vyas V.K., Ukawala R.D., Ghate M. and Chintha C. (2012). Homology modeling a fast tool for drug discovery: current perspectives. *Indian Journal of Pharmaceutical Sciences*. **74**(1): 1–17.

Wadood A., Ahmed N., Shah L., Ahmad A., Hassan H. and Shams S. (2013). In-silico drug design: An approach which revolutionarised the drug discovery process. *OA Drug Design and Delivery*. **1**(1): 3.

Wang S., Peng J., Ma J. and Xu J. (2016). Protein secondary structure prediction using deep convolutional neural fields. *Scientific Reports*. **6**:18962.

Yang J., Roy A. and Zhang Y. (2013). Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinformatics*. **29**: 2588–2595.

You Z.H., Zhu L., Zheng C.H., Yu H.J., Deng S.P. and Ji Z. (2014). Prediction of protein-protein interactions from amino acid sequences using a novel multi-scale continuous and discontinuous feature set. *BMC Bioinformatics*. **15**(15): S9.

CHAPTER FOUR

RECOMBINANT EXPRESSION AND PURIFICATION OF THE GST-TAGGED SWITCH DOMAIN PROTEIN

4.1 Introduction

Protein expression and purification play an essential role in biochemistry. *Escherichia coli* are known as a cell factory in the production of recombinant proteins; owing to the fact that it is one of the most widely used microorganism species. Added to this, it has been identified as the organisms of choice, due to its low cost, ease of genetic manipulation and rapid growth for the production of recombinant proteins (Jia and Jeon, 2016; Rosano and Ceccarelli, 2014). Production of recombinant proteins can be challenging especially with proteins containing consecutive disulfide bond. This is because such proteins can end up in inclusion bodies formation and may require additional *in-vitro* refolding to obtain biologically active proteins. (Kong and Guo, 2014; Malik, 2016).

In recent years, numerous fusion tags have been developed in order to overcome this difficult, time and labour consuming challenges associated recombinant proteins production. To this end, the use of fusion tags has successfully enhanced the expression yields; improve solubility and promote proper folding of recombinant protein as well as facilitate protein purification and enhance structural and functional studies (Costa *et al*., 2014; Nallamsetty and Waugh, 2007; Young *et al*., 2012). Purified proteins have been successfully used in various biomedical interventions (development of novel therapeutic drugs, vaccine production, immunological studies just to mention a few) and other biochemical analysis (Harper and Speicher,

2011). Therefore, this chapter describes the recombinant expression and purification the GST-tagged switch domain protein.

4.2 Materials and Methods

4.2.1 General stock solutions, buffers and media

Ammonium persulphate (APS): 10 % stock solution of APS was prepared by dissolving 1 g of APS in 10 ml of distilled water; this was stored at -20 °C.

Ampicillin: 100 µg/ml (final concentration) stock solution was prepared by dissolving 1 g of ampicillin in 10 ml of distilled water. The solution was filter-sterilized and divided into 2000 µl aliquots and stored at -80 °C.

Ampicillin agar plates: 3.4 g Bacteriological agar, 3.2 g Tryptone, 2.0 g Yeast and 1.0 g NaCl was dissolved in 200 ml distilled water. This was autoclaved allowed to cool down before ampicillin was added.

Bis-Acrylamid: 29.2 g of acrylamid and 0.8 g of N,N-Methylenebisacrylamide was dissolved in 100 ml of distilled water.

Cell lysis buffer: 2.5 ml Tris (50 mM; pH 8.0), 100 µg/ml lysozyme, 250 µl PMSF (0.5 mM), 100 µg/ml DTT (5 mM) and 7.5 ml NaCl (100 mM). The buffer was made up to 50 ml by an addition of distilled water.

Coomassie staining solution: 1 g Coomassie® Brilliant Blue was dissolved in 450 ml (45 %) methanol and 100 ml (10 %) glacial acetic acid and the volume brought to 1000 ml by the addition of 450 ml distilled water.

Destaining solution: 500 ml (50 %) methanol and 70 ml (7 %) glacial acetic acid were mixed together and the volume was made up to 1000 ml by the addition of 430 ml distilled water.

Dithiothreitol (DTT): 1 M stock solution was prepared in 0.01 M sodium acetate, pH 5.2. The solution was filter-sterilized, divided into 1 ml aliquots and stored at -80 °C.

Isopropyl β-D-1-thiogalactopyranoside (IPTG): 1 M stock solution was prepared by dissolving 2.38 g of IPTG in 8 ml of double distilled water and brought to a final volume of 10 ml. This was divided into aliquots and kept at -20 °C.

PBS: this was prepared by the addition of 8 g/l NaCl, 0.2 g/l KCl, 1.44 g/l Na_2HPO_4 and 0.24 g/l KH_2PO_4, pH 7.4 and sterilized by autoclaving.

PMSF (100mM): 0.174 g of PMSF was dissolved in 10 ml of ethanol and stored at -20º C.

2 M Tris (pH 8.8): 242.2 g of Tris base was dissolved in 800 ml of ddH_2O. The pH was adjusted to 8.8 by adding concentrated HCl. The final volume was made up to 1 L with distilled water and the solution was sterilized by autoclaving.

2 M Tris (pH 6.8): 242.2 g of Tris base was dissolved in 800 ml of ddH_2O. The pH was adjusted to 6.8 by adding concentrated HCl. The final volume was made up to 1 L with distilled water and the solution was sterilized by autoclaving.

YT Broth: this was prepared by the addition of 16 g/l Tryptone powder, 10 g/l Yeast extract, 5 g/l NaCl and 4 g/l Glucose. Autoclaved and stored at room temperature, normally used within three days.

2X sample buffer: 100 mM Tris-HCl pH 6.8, 4 % SDS, 20 % Glycerol and 0.2 % Bromophenol Blue. The buffer was stored at room temperature, prior to use, 200 mM freshly prepared DTT was added

30 % glycerol: 30 ml of glycerol was mixed with 70 ml of distilled water

10 % Sodium Dodecyl Sulphate (SDS): 50 g SDS was dissolved in 400 ml of distilled water. The pH was adjusted to 7.2 and the volume adjusted to 500 ml with distilled water. Filter-sterilize, if needed.

5X running buffer: 15.1 g Tris-base, 94 g Glycine and 10 % SDS were combined and the volume brought to one liter with distilled water.

15 % SDS resolving/ separating gel recipe: 2812 µl Bis, 1425 µl Tris 8.8, 28.13 (10 %) µl SDS, 1350 µl distilled water, 75 µl APS and 15 µl TEMED.

SDS stacking gel recipe: 650 µl Bis, 1250 µl Tris 6.8, 25 (10 %) µl SDS, 3000 µl distilled water, 100 µl APS and 10 µl TEMED.

4.2.2 Bacterial strain used

E. coli JM109 e14^{-} (McrA^{-}) recA1 endA1 *gyr*A96 thi-1 hsdR17 (rK^{-} mK^{+}) supE44 *rel*A1 Δ (*lac*-proAB) (F´ traD36 proAB *lac*Iq ZΔM15) Thermo fisher Scientific, USA was used in this study.

4.2.3 Construction of plasmid expressing the switch domains of *S. mansoni* Smp_059340.1

Codon harmonized form of the gene for the switch domains of *S. mansoni* Smp_059340.1 was constructed by Genscript. The DNA segment encoding regions of the switch domains of *S. mansoni* Smp_059340.1 were PCR amplified using

BamHI and XhoI restriction site (due to their ability to use the same digestion buffer hence, restriction can be done once at 4 °C overnight). The gene was subsequently cloned in pGEX-6P-2 making pGEX-6P-2/*S. mansoni* Smp_059340.1 (switch domains) construct was used for expression in this study.

4.2.4 Preparation of Competent *E. coli* JM109 Cell for transformation

E. coli JM109 host cells were made competent by the calcium chloride method (Sambrook and Russel, 2001). Berifly, single colony of *E. coli* JM109 was inoculated overnight with shaking at 37 °C in 5 ml tube containing 2 X YT broth. The overnight culture was diluted 1:100 in 50 ml YT broth (that is the 5 ml overnight culture was transferred into a 45 ml broth) and allowed to grow with shaking to early log phase of absorbance 0.3-0.6, measured at 600 nm. The cells were harvested by centrifugation at 5,000 rpm for 10 minutes at 4 °C. From this point onwards, the cells were incubated on ice. The cells were gently re-suspended in 10 ml 0.1 M $MgCl_2$ and left on ice for 30 minutes. The suspension was centrifuged at 4000 rpm for 10 minutes at 4 °C. The pellets were re-suspended in 10 ml of 100 mM $CaCl_2$ solution by swirling and gentle vortexing. This was incubated over ice for 4 hours. Thereafter, the competent cells were pelleted by centrifugation at 4000 rpm for 10 minutes at 4 °C. The pellet was finally re-suspended in equal volumes of $CaCl_2$ and sterile 30 % glycerol. This was divided into 200 µl aliquots and stored at -80 °C until use.

4.2.5 Transformation of *E. coli* JM109 with pGEX-6P-2-RING expression constructs

Transformation was performed by heat shock protocol (Froger and Hall, 2007). The plasmid (pGEX-6P-2/*S. mansoni* Smp_059340.1) DNA was introduced into competent *E. coli* JM109 cells. The frozen competent cells which is stored at -80 °C was ice-thawed for 5 minutes before adding 100 µl of it to 2 µl of the recombinant pGEX-6P-2-Switch domain plasmid DNA) to form a transformation mixture, which was incubated on ice for 30 minutes. Thereafter, a heat activation step was set up at 42 °C for 45 seconds and immediately placed on ice for 10 minutes. Thereafter, 900 µl of 2YT broth was added and then incubated at 37 °C for 1- 2 hours with gentle agitation. The cells were harvested and 50 µl of the transformed cells was plated out on a 2YT agar plate containing the appropriate antibiotic (ampicillin, 100 µg/ml) and incubated overnight at 37 °C.

4.2.6 Small-scale expression screening of positive transformants

Small-scale expression screening was done in order to ascertain how soluble the recombinant protein will be expressed. This was done by randomly picking 5 colonies from *E. coli* JM109 (transformation product) cells which was used to inoculate five tubes containing 5 ml culture of 2YT with ampicillin (100 µg/ml), with vigorous shaking at 37 °C for 4 hours. Afterwards, 1 ml culture was taken from each tube and transferred into a clean 1.5 ml tube, designated as the un-induced sample. Thereafter, another 1 ml culture was taken from each tube and transferred into another clean 1.5 ml tube; at this point 0.5 mM Isopropyl β-D-thiogalactoside (IPTG) was added to the culture and this tube was designated as the induced culture. With vigorous shaking, both induced and un-induced cultures was incubated at 37 °C for 2

hours; cells was harvested by centrifugation at 13200 x g for 5 minutes. The supernatant was discarded, while the cell pellets was re-suspended in 50 µl PBS containing 1 % Triton X-100 (PBS-T). Twenty micro-liters (20 µl) of each sample were analyzed on SDS-PAGE gel to visualize expression of the protein. For future large-scale recombinant protein expression, glycerol stocks was made from the remaining 3 ml of the best expressing culture and these was kept at -80 °C until needed.

4.2.7 Large-scale recombinant protein expression of the switch domains of *S. mansoni* Smp_059340.1

4.2.7.1 Expression of the switch domains of S. mansoni Smp_059340.1

For large-scale expression of recombinant protein, 100 µl of the best recombinant protein expressing clone was used to inoculate 100 ml of 2YT broth; containing 100 µg/ml ampicillin overnight at 37 °C. The 100 ml primary culture was scaled up to 2 litres with 2YT broth containing 100 µg/ml ampicillin the next morning, while incubation was continuous with vigorous shaking at 37 °C; the optical density OD_{600} of the cell culture was periodically monitored until it reached between 0.4 - 0.6. Thereafter, 0.5 mM IPTG was added with vigorous shaking of the broth, and the temperature was reduced to 25 °C; protein induction was done overnight. The cells were harvested the next morning by centrifugation at 5000 x g for 20 mins at 4 °C; cell pellets was lysed immediately and frozen at -80 °C.

4.2.7.2 Extraction of the switch domains of S. mansoni Smp_059340.1

Frozen cell pellets (total bacterial cell lysates) containing the switch domain protein from the large scale expression stored in -80°C was taken out and subjected to three to five cycles of freeze-thawing (incubate at -80 °C for 5 minutes, followed by incubation at 37 °C for another 5 minutes) and three cycles of sonication to break open the cells in order to release GST-tagged switch domain protein. GST-tagged switch domain protein was extracted by centrifugation of the cell suspension at 5,000 x g for 20 minutes at 4 °C. Cleared total bacterial lysate without cell debris and particulate matter was poured into a clean 50 ml tube and placed on ice.

4.2.7.3 Affinity purification of the switch domains of S. mansoni Smp_059340.1

Glutathione-agarose beads was weighed out, allowed to swell overnight in deionised water at 4 °C; poured into a chromatography column and allowed to pack properly by allowing it to stand for an hour at 4 °C. Afterwards, the column was equilibrated with 5 column volumes (5 CV) of 50 mM Tris containing 20 mM NaCl at pH 8.0. After equilibration, the cleared total bacterial lysate containing the GST-tagged switch domain protein was poured onto the column and allowed to flow by gravity. The flow-through was collected and kept on ice. Prior to elution of GST-tagged switch domain protein; the affinity column was washed with 10 column volumes (10 CV) of equilibration buffer. The last three drops of the equilibration buffer was collected (wash 1) to examine clearness of the column before elution. Elution buffer (50 mM Tris pH 8 containing 15 mM reduced glutathione, 0.02 % NaN_3, 1 mM DTT and 1 mM PMSF) of seven column volumes (7 CV) was used to elute the GST-tagged

switch domain protein off the column. Thereafter, the column was washed with five column volumes (5 CV) of 2 M NaCl, wash 2 was collected. The rest of the salt solution was allowed to flow freely, when this was done; the column was then kept at 4°C until needed.

4.2.8 SDS Polyacrylamide gel electrophoresis

Various fractions collected were then subjected to electrophoresis on a 15 % SDS PAGE gel according to standard procedures (Laemmli, 1970). Briefly, separating gel (2812 µl Bis, 1425 µl Tris 8.8, 28.13 (10 %) µl SDS, 1350 µl distilled water, 75 µl APS and 15 µl TEMED) and stacking gel (650 µl Bis, 1250 µl Tris 6.8, 25 (10 %) µl SDS, 3000 µl distilled water, 100 µl APS and 10 µl TEMED) were prepared. Both were poured into the casting plate and allowed to stand for about 30 minutes in order to give room for solidification. The protein sample to be analyzed (20 µl of each) was mixed with 10 µl of 2x SDS PAGE sample buffer. The mixture was boiled for 5 - 10 minutes at 95 °C, and electrophoresed in 1x SDS electrophoresis buffer at a constant voltage of 100 Volts and 250A for 60 minutes. When the samples were done running, the gel was stained by incubation in Coomassie staining solution overnight. The following morning, the gel was destained in destaining solution for about two hours in order to view the bands.

4.2.9 Protein concentration determination

The protein concentration was determined using a NanoDrop ® ND 2000 spectrophotometer (Thermo Fisher Scientific) by measuring the absorbance at 280 nm (A280; Desjardins *et al*., 2009). Briefly, 0.5 - 2 µl sample of the extraction buffer was placed directly on top of the detection surface of the NanoDrop ® ND 2000

spectrophotometer (Thermo Scientific) to blank it. Thereafter, 0.5 - 2 µl of the protein sample (elution 1) was placed directly on top of the detection surface in order to determine the protein concentration of the GST-tagged switch domain protein.

4.3 Results and discussion

4.3.1 Small-scale expression screening of positive transformants

Based on the coding regions of the switch domains, a fusion gene (expression plasmid; pGEX-6P-2-*S. mansoni* Smp_059340.1 (switch domains), encoding residues 183 – 223 of the full length *S. mansoni* Smp_059340.1), fused to Glutathione-S-transferase (GST) was constructed by Genscript. The switch domains sequence was cloned into the *Bam*HI and *Xho*I sites of the pGEX-6P-2 multiple cloning cassette with a PreScission protease site (Figure 4.1).

The expected molecular weight of the switch domains, as calculated using the ExPASy PROTPARAM server (http://ca.expasy.org/cgi-bin/protparam), is approximately 5.8 kDa with an isoelectric point (pI) of 7.83 (Appendix 1). Given that the molecular weight of GST is 26 kDa, the expected molecular weight of the pGEX-6P-2-Switch domain protein (fusion gene; Figure 4.2) is approximately 31.84 kDa.

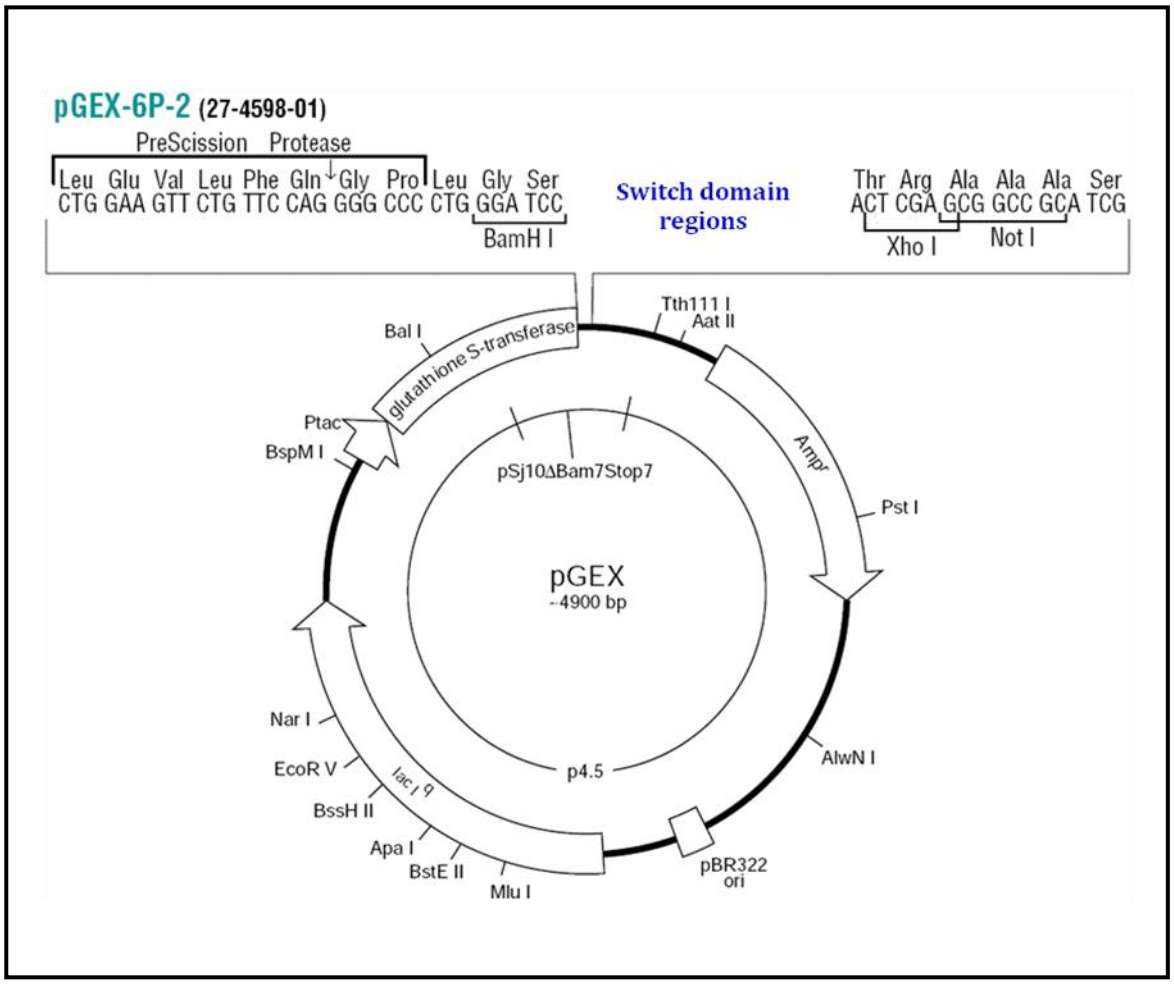


Figure 4.1: Map of pGEX-6P-2 expression vector showing the switch domains sequence cloned into the BamHI and XhoI sites of the pGEX-6P-2. The pGEX-6p-2 vector is a 4900 bp C-terminal GST expression vector with a tac promoter, allowing induction of expression of the target protein using IPTG.

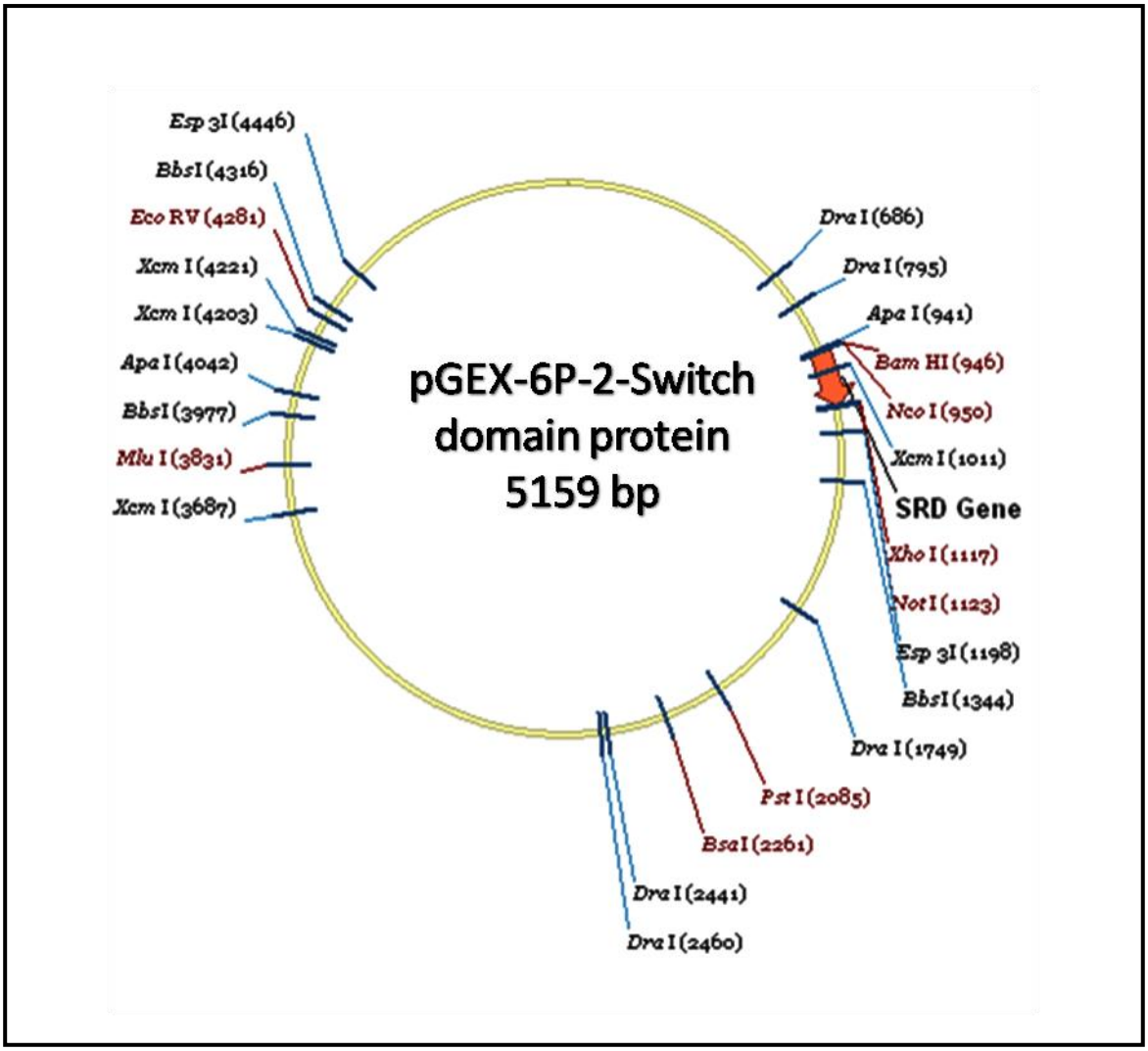


Figure 4.2: A diagrammatic representation of the pGEX-6P-2-Switch domain protein constructs.

For the first expression experiment, the pGEX-6P-2 vector containing the switch domain protein gene was transformed into *E. coli* JM109. Thereafter, expression screening of the positive transformants using SDS-PAGE was carried out following the protocol described in Section 4.2.6. SDS-PAGE analysis of the total bacterial lysates from the 4 colonies randomly picked is presented in Figure 4.3. Lanes 1-8 shows alternating pairs of un-induced and induced samples. Visible prominent bands (~31.84 kDa) corresponding to the GST-Switch domain protein (fusion protein) were observed in the region of 25 – 37 kDa. This bands are visible and prominent in the induced lanes (2, 4, 6 and 8) but not in the un-induced lanes (1, 3, 5 and 7). The result obtained is this study indicates a successful expression of the recombinant pGEX-6P-2-Switch domain protein.

Protein misfolding and inclusion body formation is one of the key challenges and drawbacks in the large scale production (high amounts) of recombinant proteins in a properly folded and soluble biological active form (Baneyx and Mujacic, 2004; Lebendiker and Danieli, 2014; Rosano and Ceccarelli, 2014). Expression screening of colonies may assists in resolving these serious challenges associated with recombinant protein production especially for structural and function determination of novel protein in anticipation of developing novel drug targets. Expression screening of colonies immediately after transformation in this study provides information on the best expressing colonies as presented in Figure 4.3.

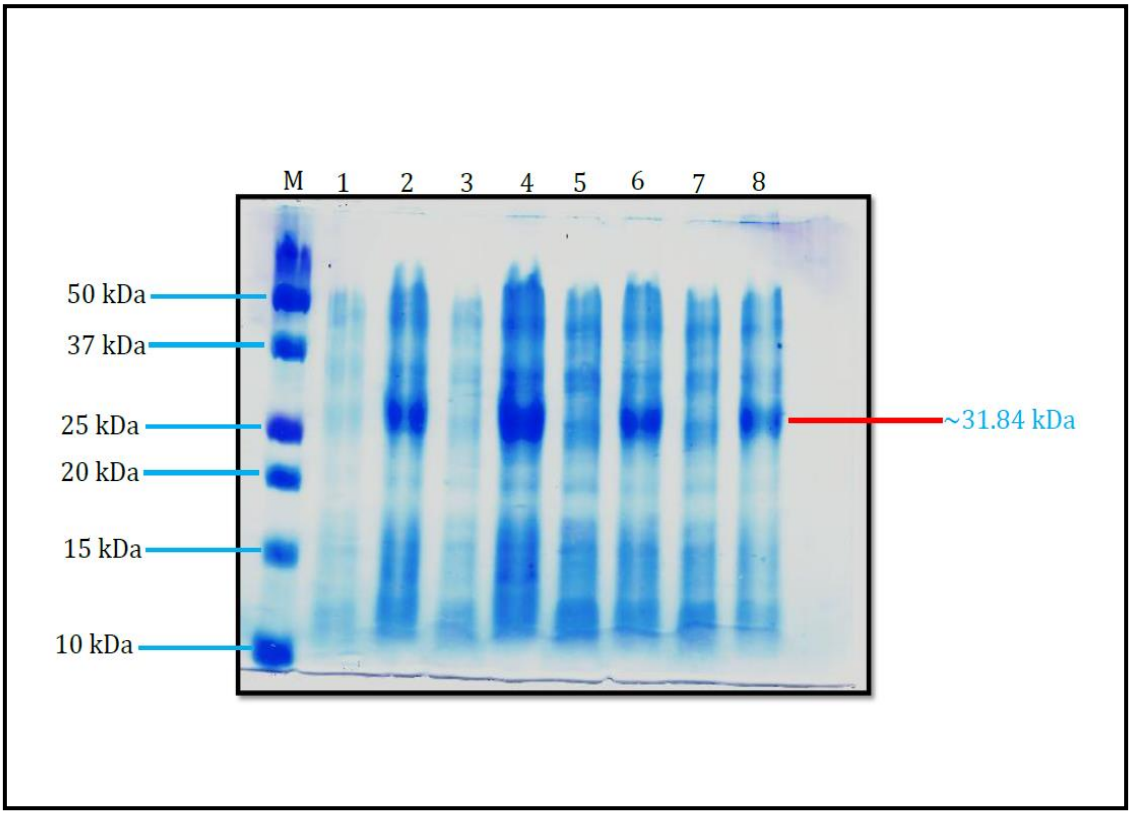


Figure 4.3: Expression screening results of colonies transformed with GST-Switch domain protein. Lane M shows protein molecular weight marker in kDa. Lanes 1, 3, 5 and 7 shows the un-induced total bacterial cell lysates while lanes 2, 4, 6 and 8 show the induced total bacterial cell lysates upon the addition of 0.5 mM IPTG. The band observed on each lane corresponds to the expected combined size of GST-Switch domain protein (~31.84 kDa).

4.3.2 Large scale expression of the switch domains of S. mansoni Smp_059340.1

Large-scale expression of GST-Switch domain protein (fusion protein) was carried out as described in Section 4.2.7.1, using the best expressing colony (lane 4) as shown on the SDS-PAGE analysis in Figure 4.3. From the induced culture in lane 4, 100 µl was used to inoculate a large-scale expression in this study. Following cell harvesting by centrifugation, cell lysis and extraction of the GST-Switch domain protein (fusion protein), SDS-PAGE analysis shows that a high yield of fusion protein was obtained. Based on the result presented in Figure 4.4, the expression of the recombinant protein (GST-Switch domain protein) reached a maximal level within 2 to 6 h after (post) induction. This is an indication that 0.5 mM IPTG is a sufficient concentration for induction.

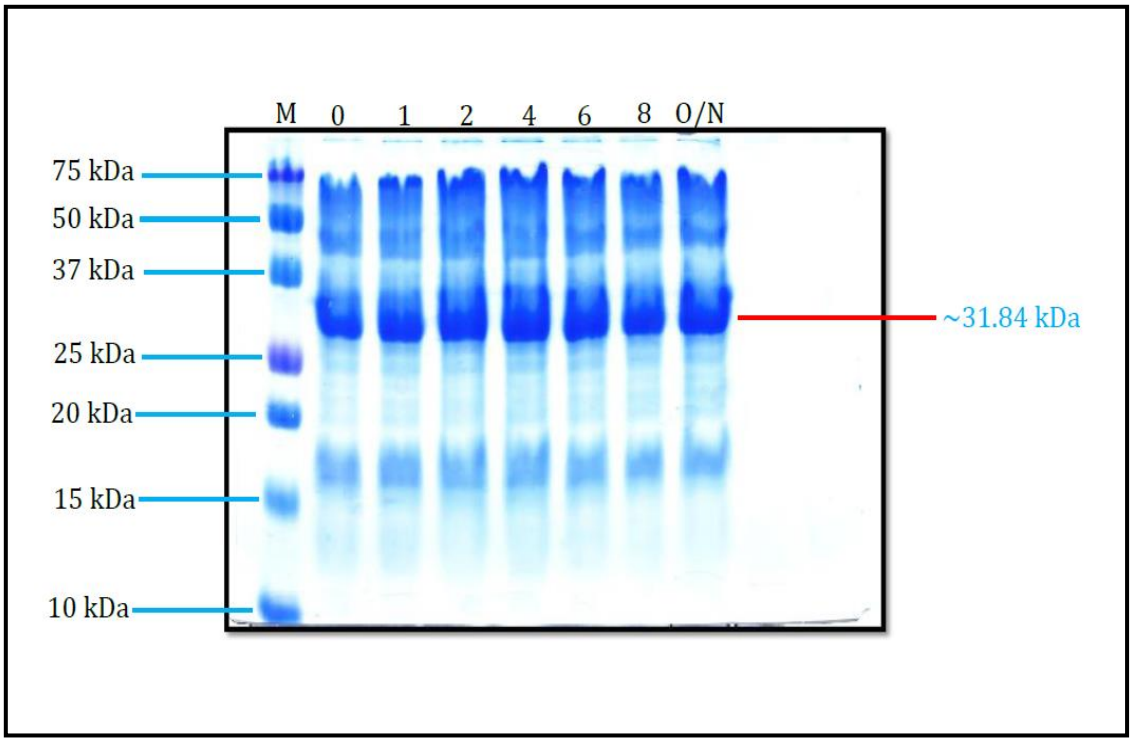


Figure 4.4: Large scale expression results of GST-Switch domain protein. Lane M shows protein molecular weight marker in kDa. Lanes 0 shows the un-induced total bacterial cell lysates before the addition of 0.5 mM IPTG while lanes 1, 2, 4, 6, 8 and O/N shows the total bacterial cell lysates obtained; 1, 2, 4, 6, 8 hours and overnight post induction (upon the addition of 0.5 mM IPTG). The band observed on each lane corresponds to the expected combined size of GST-Switch domain protein (~31.84 kDa).

4.3.3 Extraction and affinity purification of the switch domains of S. mansoni Smp_059340.1

Extraction and affinity purification of recombinant GST-Switch domain protein (fusion protein) were carried out as described in Sections 4.2.7.2 and 4.2.7.3. Figure 4.5 show the SDS-PAGE analysis of extraction and affinity purification of the recombinant GST-Switch domain protein. This result from the extraction of the recombinant GST-Switch domain protein analysed on SDS-PAGE (Figure 4.5) suggests that the fusion protein is soluble.

Structural and functional studies of proteins require expression and purification of recombinant proteins on a large scale. In the present study, all attempts to cleave the fusion protein (GST-Switch domain protein) using 3C protease was not successful as the expected 5.8 kDa (Switch domain protein) protein was not visible on the gel after SDS-PAGE analysis of the 3C protease cleavage of the GST-Switch domain protein. At the moment, no reason could be attributed for this abnormal behaviour of the GST-Switch domain protein on the SDS PAGE gel after 3C protease cleavage. Therefore, all preliminary characterization carried out in this study was done using the expressed GST-tagged switch domain protein.

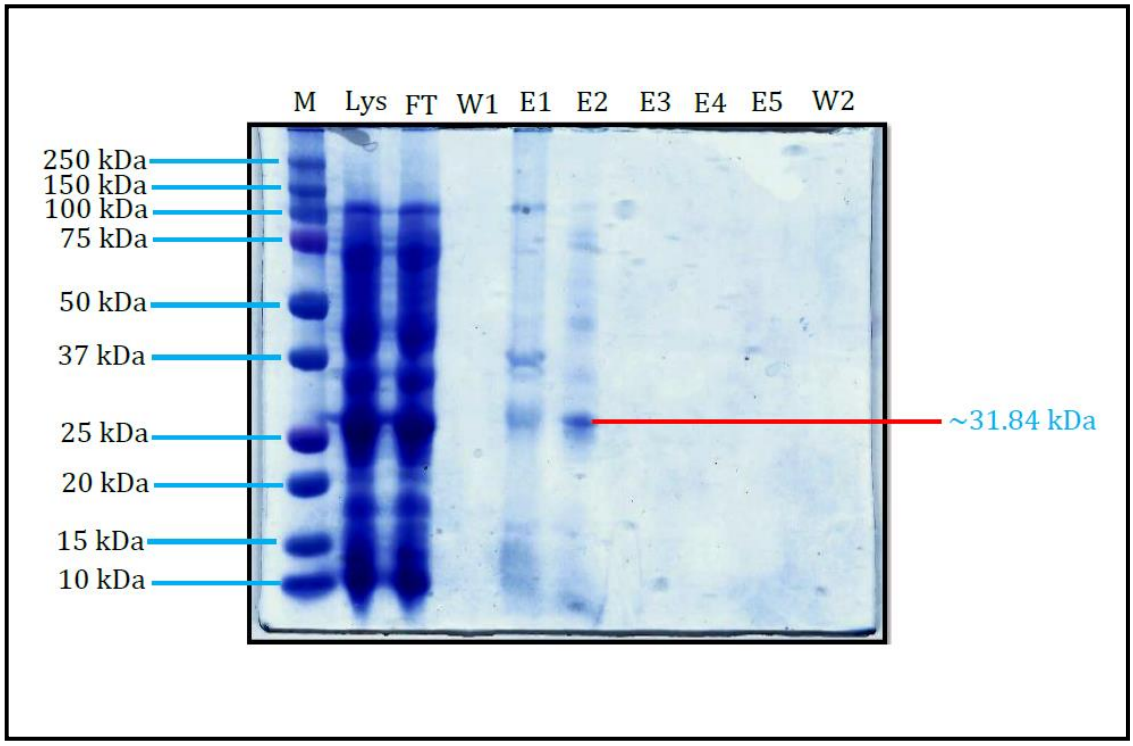


Figure 4.5: SDS-PAGE analysis of expression and affinity purification of GST-Switch domain protein (fusion protein). Lane M shows protein molecular weight marker in kDa. **Lys**: shows lysate obtained before purification, **FT:** shows flow through from glutathione affinity column, **W1:** shows the column wash with equilibration buffer, **E1, E2, E3, E4 and E5:** shows the eluted GST-fusion protein, standing for Elution. **W2:** shows the column wash with 2 M NaCl. The protein band observed at **E1** and **E2** corresponds to the expected combined size of GST-Switch domain protein (~31.84 kDa).

4.3.4 Protein concentration determination

The protein concentration of the GST-Switch domain protein (fusion protein) was determined using a NanoDrop ® ND 2000 spectrophotometer (Figure 5.6) as described in Section 3.9. The concentration of the soluble GST-Switch domain protein (fusion protein) measured by absorbance at 280 nm (A280) was 1.187 mg/ml, for a total protein of 20 ml. This concentration is a bit lower than what we expected, but beyond all doubt, we have been able to establish a purification scheme for a small and useful quantity of the fusion protein. The use of NanoDrop spectrophotometers (Thermo Scientific) for nucleic acid quantification is well established in the life sciences, and are also capable of quantifying proteins at 280 nm with high degree of accuracy and reproducibility (Desjardins *et al*., 2009).

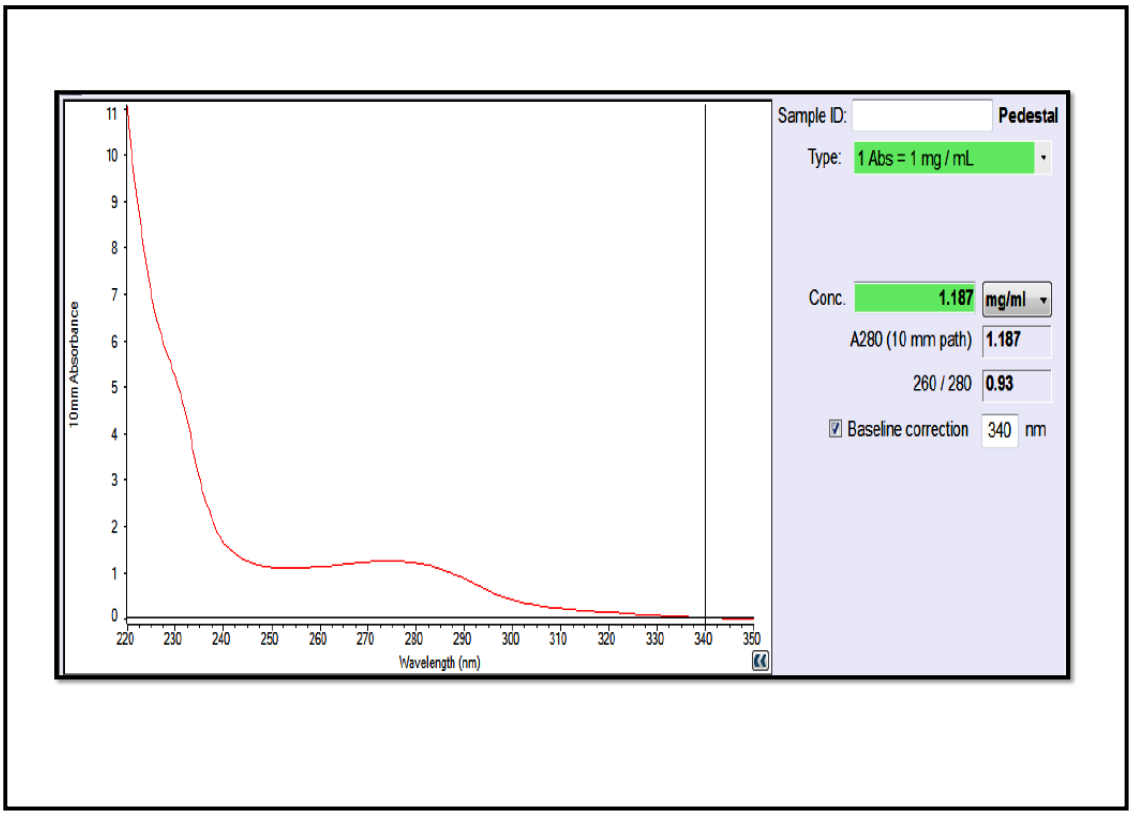


Figure 4.6: Estimation of the protein concentration of the GST-Switch domain protein (fusion protein) using NanoDrop 2000.The value obtained was 1.187 mg/ml.

4.4 Conclusion

In this chapter we reported the successful construction, expression, purification and concentration of a fusion protein (the GST-Switch domain protein) in *E. coli*. Added to this, it was also demonstrated that *E. coli* JM109 is an efficient expression system for the production of recombinant GST-Switch domain protein; with respect to the level of its expression and production of soluble active protein. In the next chapter, we present the preliminary characterization of the GST-Switch domain protein using different methods.

REFERENCES

Baneyx F. and Mujacic M. (2004). Recombinant protein folding and misfolding in Escherichia coli. *Nature biotechnology*. **22**(11): 1399-1408.

Costa S., Almeida, A., Castro, A. and Domingues L. (2014). Fusion tags for protein solubility, purification, and immunogenicity in Escherichia coli: the novel Fh8 system. *Frontiers in Microbiology*. **5**:63

Desjardins P., Hansen J.B. and Allen M. (2009). Microvolume protein concentration determination using the NanoDrop 2000c spectrophotometer. *Journal of Visualized Experiments*. **33:** e1610-e1610.

Froger A. and Hall J.E. (2007). Transformation of plasmid DNA into E. coli using the heat shock method. *Journal of Visualized Experiments*. **6**: 253.

Harper S. and Speicher D.W. (2011). Purification of proteins fused to glutathione S-transferase. *Protein Chromatography: Methods and Protocols*. **681**: 259–280.

Jia B. and Jeon C.O. (2016). High-throughput recombinant protein expression in Escherichia coli: current status and future perspectives. *Open Biology*. **6**(8): 160196.

Kong B. and Guo G.L. (2014). Soluble expression of disulfide bond containing proteins FGF15 and FGF19 in the cytoplasm of Escherichia coli. *PloS one*. **9**(1): e85890.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature **227***: 680-685.

Lebendiker M. and Danieli T. (2014). Production of prone-to-aggregate proteins. *FEBS letters*. **588**(2): 236-246.

Malik A. (2016). Protein fusion tags for efficient expression and purification of recombinant proteins in the periplasmic space of E. coli. *3 Biotech*. **6**:44.

Nallamsetty S. and Waugh D.S. (2007). A generic protocol for the expression and purification of recombinant proteins in Escherichia coli using a combinatorial His6-maltose binding protein fusion tag. *Nature Protocols*. **2**(2): 383-391.

Rosano G. and Ceccarelli E. (2014). Recombinant protein expression in Escherichia coli: advances and challenges. *Frontiers in Microbiology*. **5**: 172.

Sambrook J, Russell DW (2001) Molecular cloning: A laboratory manual. 3rd Ed, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; pp:132-150

Young C.L., Britton Z.T. and Robinson A.S. (2012). Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. *Biotechnology Journal*. **7**(5): 620-634.

CHAPTER FIVE

PRELIMINARY CHARACTERIZATION OF THE GST-TAGGED SWITCH DOMAIN PROTEIN

5.1 Introduction

At the moment, the exact mechanisms of assembly, function, and regulation of vast majority of protein complexes still remains unclear (Sharon, 2010). Proteins vary from one to another in terms of size, molecular structure and physiochemical properties. These variations permit the analysis and characterization of recombinant proteins, notably by separation and identification. The most important analysis performed on recombinant proteins are checks of purity and of biological or enzymatic activity. When a recombinant protein has been purified, it is essential to characterize the protein (Denslow *et al*., 2001). Several standard methods have been established in literature for protein identification and characterization (Mirza and Olivier, 2008). This chapter presents the preliminary characterization of the fusion protein (GST-tagged switch domain protein) using Fourier transform infrared spectroscopy (FTIR), ultraviolet-visible spectrophotometry, fluorescence spectroscopy, isothermal titration calorimetry (ITC), thermogravimetric analysis (TGA) and nuclear magnetic resonance spectroscopy (NMR).

5.2 Materials and Methods

5.2.1 Fourier Transform Infrared Spectroscopy (FTIR)

Infrared spectroscopy has the potential to provide information regarding protein secondary structure. In this study, data acquisition was performed using a Bruker, Tensor 27 platinum ATR-FTIR spectrometer. The spectral resolution was 4 cm^{-1}, with 96 scans, and an aperture of 4 mm. The optimum beam incidence angle was 45 °C and all the FTIR spectra were recorded in a range between 4000 and 600 cm^{-1}, at room temperature.

5.2.2 Ultraviolet-visible spectrophotometry

The identity of the expressed GST-tagged switch domain protein was determined by a Perkin Elmer UV–Vis spectrophotometer. The spectra were scanned in the region between 200-700 nm at room temperature using quartz cuvettes (1 cm path length). The cuvettes were cleaned before each use. UV–Vis differential spectra of the expressed GST-tagged switch domain protein were measured using double distilled water as a blank.

5.2.3 Fluorescence spectroscopy

The expressed GST-tagged switch domain protein fluorescence emission spectrum was monitored at room temperature using a Perkin Elmer LS45 Fluorescence Spectrophotometer. Quartz cuvette with 1 cm path length was used. 0.8 M KCl was added to the protein which was already prepared in 5 mM PBS (pH 7.4). Intrinsic fluorescence emission scanning was carried out in the range 280 – 500 nm (excitation value of 275.6 nm). PBS (pH 7.4) was used as a reference (blank) for all measurements.

5.2.4 Thermal characterisation of the GST-Switch domain protein

Thermal characterization of GST-Switch domain protein was performed by differential scanning calorimetry (DSC) and thermogravimetric analysis (TGA) using Universal V4.5A TA Instruments (simultaneous DSC and TGA analyzer) under nitrogen atmosphere at the heating rate of 10 °C min^{-1} from room temperature to 400 °C.

5.2.5 Nuclear magnetic resonance spectroscopy (NMR): (1D ^{1}H-NMR) spectroscopy

NMR spectra were obtained using the 600-MHz Varian NMR spectrometer of the NMR laboratory, University of Stellenbosch. The expressed GST-tagged switch domain protein used to obtained NMR data were concentrated and lyophilized in 50 mM PBS pH 7.4 to give a final concentration of 1.05 mg/ml. Prior to NMR experiments, this sample was dissolved in 90 % distilled water (DH_2O) and 10 % deuterium oxide (D_2O) (to serve as an internal reference to lock the spectrometer carrier frequency). The degree of resonance dispersion in a 1D spectrum was used to gauge the folding state of the protein.

5.2.6 Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry (ITC) is a powerful technique to study both protein-ligand and protein-protein interactions. Isothermal titration calorimetry experiments were carried out at 25 °C using a modular titration nanocalorimeter TAM III. The protein and the ligand were dialyzed against 5 mM phosphate-buffered saline (PBS), pH 7.4. All samples were filtered through a 0.22 µm sterile filter membrane, and thoroughly degassed at 25 °C before loading into the TAM III instrument; using a 1

ml glass ampoule with a gold propeller stirrer and for ligand injection; a 250 µl Hamilton syringe with a stainless steel canula was used. For ITC measurements, the stationary phase (protein; 1.02 mg/ml) was prepared in the ampoule (750 µl) and the syringe (250 µl) was filled with the ligand (5.10 mg/ml) while the reference cell contains 5 mM PBS. The protein-ligand titration was performed with an injection volume of 10 ml, a time spacing of 10 min between injections and a stirrer speed of 40 rpm. The ligand (guanosine triphosphate; GTP or Calcium (Ca) as Calcium chloride; $CaCl_2$)-protein (GST-tagged switch domain protein) binding equation is presented below (1) as:

$$M + nL \rightleftharpoons MLn \qquad (1)$$

M = molecule (the expressed GST-tagged switch domain protein),

L = ligand (GTP or Ca),

n = stoichiometry of interaction.

5.3 Results and discussion

5.3.1 Fourier Transform Infrared Spectroscopy (FTIR)

The FTIR spectrum of the GST-Switch domain protein, carried out as described in Section 5.2.1 is displayed in Figure 5.1. Of importance are the visible signal wave number (3349.53, 1624.19, 1460.48 and 1157.11) and their corresponding functional groups identified as (Amide A: N-H, O-H, Amide I: C=O, C=C, C=N, Amide II: C-N, N-H, C-H, C=C and Amide III: C-N, C-O) respectively, which are presented in Table

5.1 for a reference. The FTIR spectrum obtained for the fusion protein is in agreement with the ones documented in literature for similar related compounds. Sequel to this, the identified functional groups have been assigned by comparing them with other assignments for functional groups that have been published (Aboul-enein *et al*., 2014; Centeno and Shamir, 2008; Mbonyiryivuze *et al*., 2015).

Infrared spectroscopy (IR) is one of the earliest tools used to rapidly characterize protein structure, it was not until the advent of FTIR spectrometers that the efficacy of the technique for biological studies became generally recognized and accepted (Haris, 1999). Currently, FTIR spectroscopy is an attractive non-destructive tool for structural characterization of proteins and polypeptides. Although this technique may not provide the same level of structural details like X-ray crystallography and NMR spectroscopy; but it has be used to rapidly characterize protein secondary structure in aqueous solution in order to further understand how proteins function (Aboul-enein *et al*., 2014; Gallagher, 2009; Yang *et al*., 2015).

The most sensitive spectral (characteristic) region found in the infrared spectra of proteins and polypeptides consist of the Amide I and Amide II band. Typically, the stretching vibrations of the C=O bond of the Amide I and the bending vibrations of the N-H bond of the Amide II are responsible for the absorption attributed to the Amide I and Amide II bands. Interestingly, both the C=O and the N-H bonds are implicated in the hydrogen bonding that occurs among the different elements of secondary structure. Therefore, the position of both the Amide I and Amide II bands are important to the secondary structure content of a protein (Aboul-enein *et al*., 2014; Gallagher, 2009). Taken together, in this experiment, FTIR spectroscopy successfully revealed the identities of the secondary structure of the GST-Switch domain protein. The most important advantage of FTIR spectroscopy is the ability to

record spectra in a wide range of environments. Therefore, this technique is an exceptionally important technique in the field of protein biochemistry (Haris, 1999).

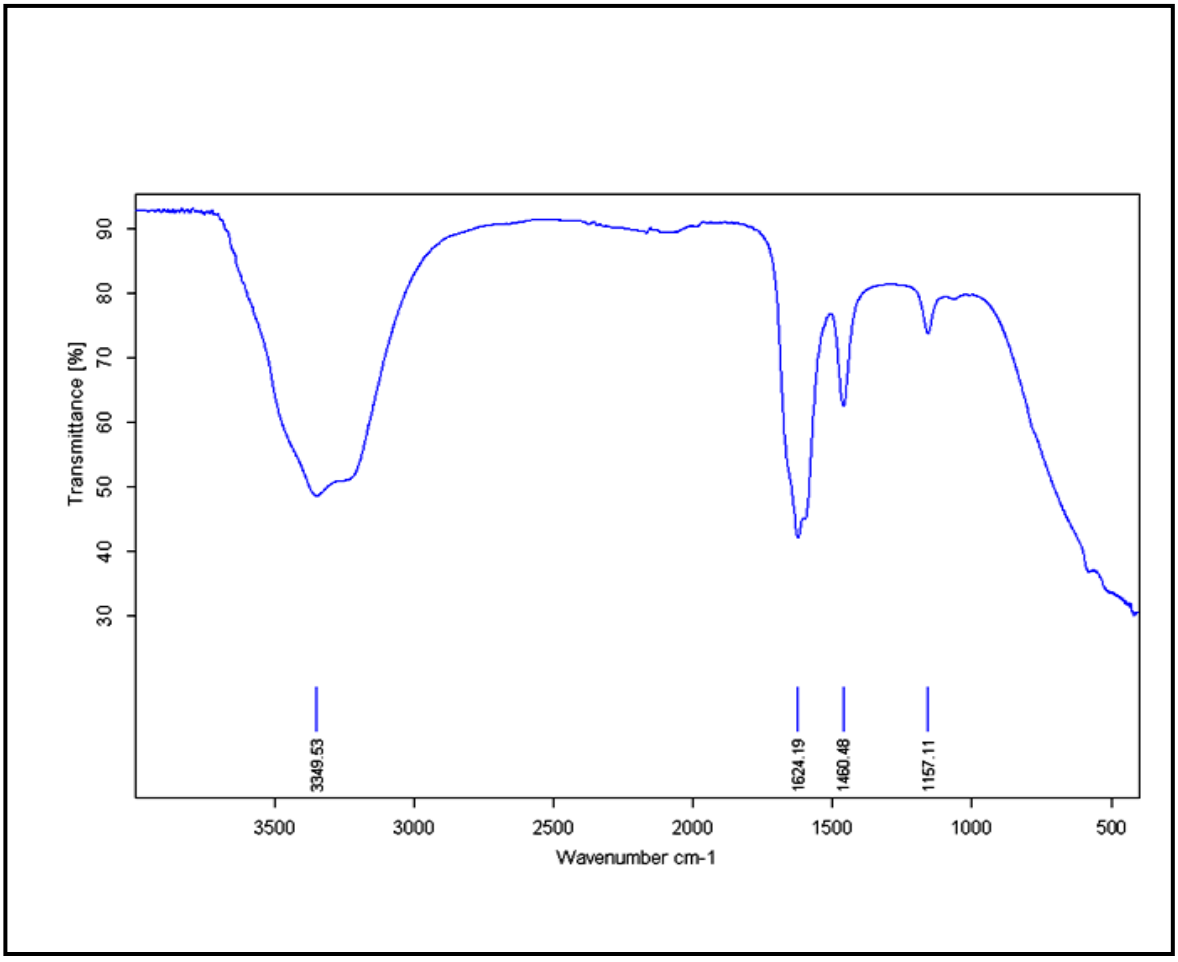


Figure 5.1: FTIR spectrum analysis of the GST-Switch domain protein

Table 5.1: Summary of the corresponding functional groups derived from the FTIR analysis of GST-tagged switch domain protein

Wave number Peaks (cm^{-1})	Vibration modes	Functional group	Intensity
3349.53	Stretching	Amide A: (N-H, O-H)	Medium
1624.19	Stretching	Amide I: (C=O, C=C, C=N)	Variable
1460.48	Bending, Stretch	Amide II: (C-N, N-H, C-H, C=C)	Variable, medium weak or multiple bands
1157.11	Stretch	Amide III: (C-N, C-O)	Medium weak, strong

5.3.2 Ultraviolet-visible spectrophotometry

The UV-visible wavelength scan of the GST-Switch domain protein presented in Figure 5.2 was carried out as described in Section 5.2.2, using a Perkin Elmer UV-Vis spectrophotometer. The UV-visible spectra displayed a peak around 281 nm and the absorbance decreased towards the visible region. Comparing this result with earlier documented reports, we observed that the same pattern was recorded for similar proteins (Apte *et al*., 2013; Deshmukh and Pethe, 2016). UV-Visible spectroscopy in this study was used to analyze the structural integrity of the GST-Switch domain protein. Although, it is really cumbersome to accurately determine the concentrations and purity of DNA, RNA and protein in complex mixtures. Nevertheless, measuring their absorbance between 260 nm and at 280 nm is a rapid and accurate technique that can provide validation of their purity (Teare *et al*., 1997).

The peak observed at the region of 281 nm is an indication of the purity of the GST-Switch domain protein. Additionally, the peak observed at this region (281 nm) can also be attributed to the π => π* and n => π* of the amino, carboxylic and aromatic moieties of the protein (Abdelhalim *et al*., 2013; Magarelli *et al*., 2010). Over the years, based on the Beer Lambert's law, UV-visible spectroscopy has been used as a conventional means of measuring protein absorbance in order to determine its concentration. This technique is also a suitable tool for the detection of non-protein contaminants, especially if the protein of interest contains aromatic residues and the absorption spectrum is recorded over a large range (240 nm - 350 nm). Reducing agents (notably DTT) can change the symmetry of the 280 nm absorbance peak by increasing the absorbance at 250 nm and below (Raynal *et al*., 2014; Teare *et al*., 1997).

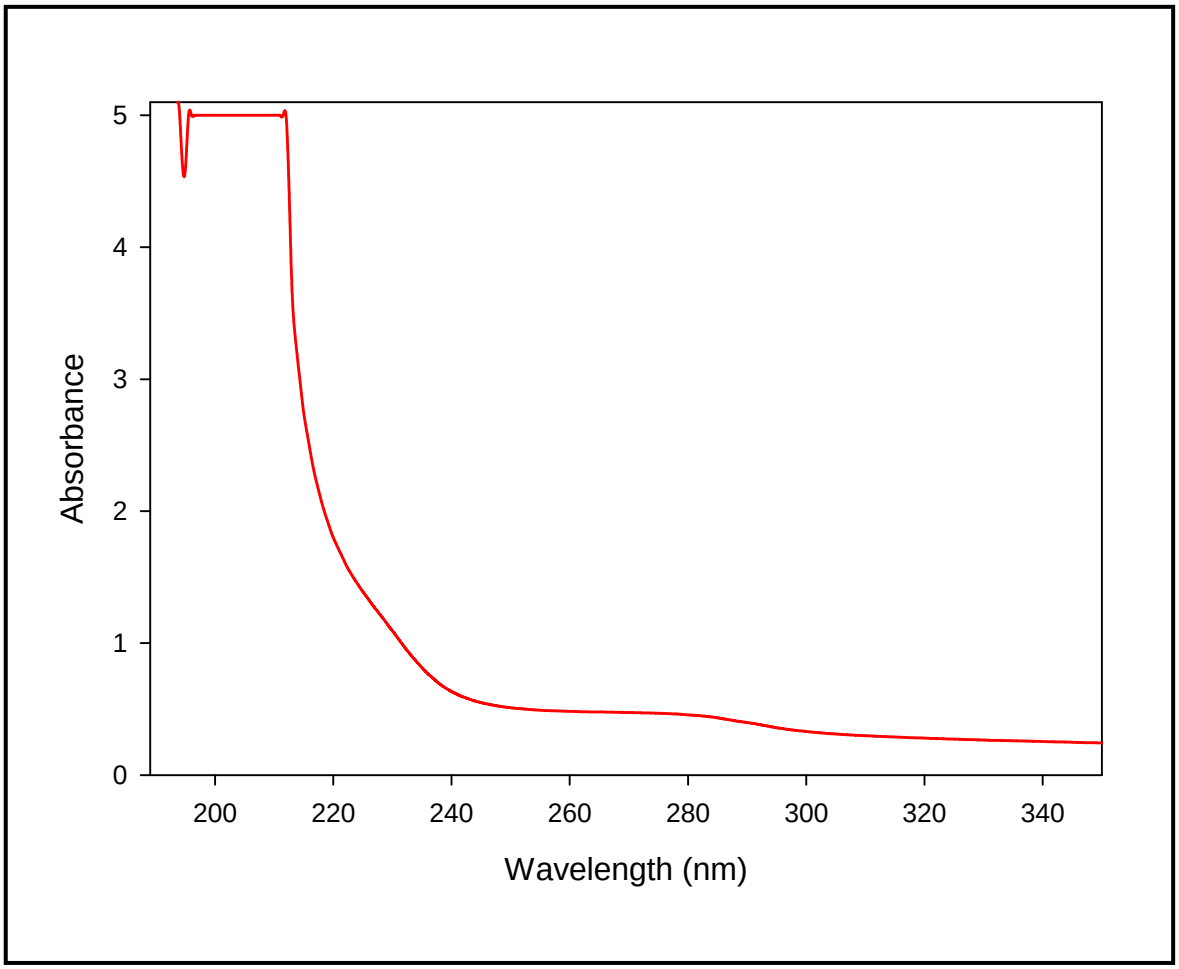


Figure 5.2: UV-Visible spectrum analysis of the GST-Switch domain protein

5.3.3 Probing the Structural Stability of the GST-Switch domain protein using fluorescence spectroscopy

This experiment was carried out to characterize the changes in stability at the tertiary structure level of the GST-Switch domain protein by recording the maximum fluorescence emission intensity upon excitation at 275.6 nm. The intrinsic fluorescence of GST-Switch domain protein was measured by the addition of 0.8 M KCl. As presented in Figure 5.3, when the GST-Switch domain protein was excited at 275.6 nm, it had an emission maximum at 345 nm. Basically, the intrinsic fluorescence of proteins mainly from their tryptophan residues, show an emission maximum at around 340 nm. The increase in the intrinsic fluorescence can be attributed to the presence of aromatic amino acid residue and tryptophan. The result reported here suggests that the GST-Switch domain protein is stable and this observation by fluorescence spectroscopy is consistent with reports from similar studies (El-Turk *et al*., 2012; Lee *et al*., 2006).

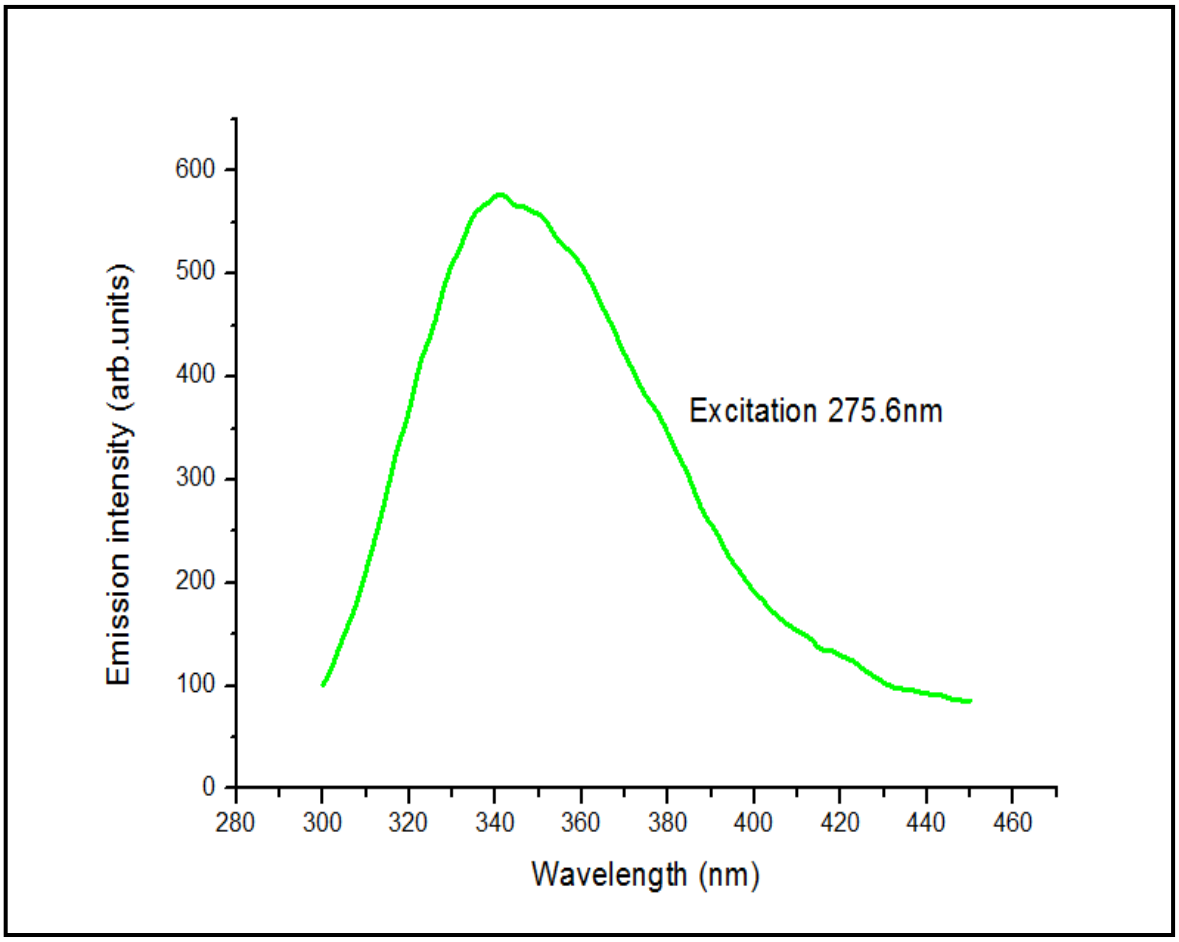


Figure 5.3: Fluorescence spectrum analysis of the GST-Switch domain protein

5.3.4 Thermal characterisation of the GST-Switch domain protein

The thermal stability of the GST-Switch domain protein was carried out to quantify its thermal degradation in a controlled atmosphere. Lyophilized GST-Switch domain protein (1.832 mg in weight) was heated starting from room temperature to 400 °C at a heating rate of 10 °C min^{-1}. The thermogravimetric analysis (TGA) profile for the GST-Switch domain protein shown in Figure 5.4 revealed a four-step degradation reaction (weight loss steps). The first major weight loss (1.412 mg corresponding to 77.01 %) in the temperature range of about 165 – 200 °C, could be attributed to the removal of moisture and physically adsorbed (physisorbed) water molecules (Jeyaraman *et al*., 2016; León-Mancilla *et al*., 2016). The second step of weight loss (0.1840 mg corresponding to 10.03 %) in the temperature range of about 215 – 260 °C, could be attributed to the exothermic degradation causing the partial collapse of functional groups in the GST-Switch domain protein. The third minor weight loss occurred (0.2293 mg corresponding to 12.50 %) in the temperature range of about 260 – 370 °C while the fourth and final weight loss is a steeper one leaving a residue of 0.006901 mg corresponding to 0.3763 % in the temperature range of about 315 – 400 °C.

Figure 5.5 shows the differential scanning calorimetry (DSC) profile of the GST-Switch domain protein. In the present study, we have investigated the thermal denaturation and stability of the GST-Switch domain protein by estimating the values of the onset temperature (T_{on}), the peak denaturation temperature (T_{d2}) and off-set temperature (T_{off}) as well as the enthalpy of denaturation (ΔH). As measured by DSC instrument (Universal V4.5A TA), the GST-Switch domain protein shows a T_{on} of 130.70 °C, T_{d2} of 133.39 °C, T_{off} of 145 °C and a ΔH of 89.16J/g (Table 6.2). Taken

together, the GST-Switch domain protein demonstrated thermal stability up to 130.70°C.

Differential scanning calorimetry (DSC) in one of the primary techniques frequently used in chemistry, biochemistry, cell biology, biotechnology, pharmacology and more recently in nanoscience (Gill *et al*., 2010). It is a very reliable and sensitive technique employed in evaluating the thermodynamic (thermotropic) properties of a wide range of biomolecules, nano-sized materials and extracts. In some biopharmaceutical industries across the globe, DSC has been used to characterize protein thermal stability, overall conformation, and domain folding integrity (Wen *et al*., 2012). DSC establishes a relationship between temperature and specific physical properties of material and is probably the only method for direct determination of the enthalpy associated with the process of interest (Chiu and Prenner, 2011; Gill *et al*., 2010). Understanding the relationship between the structure of proteins and the energetics of their stability is very important in biochemistry and biotechnology. (Bruylants *et al*., 2005).

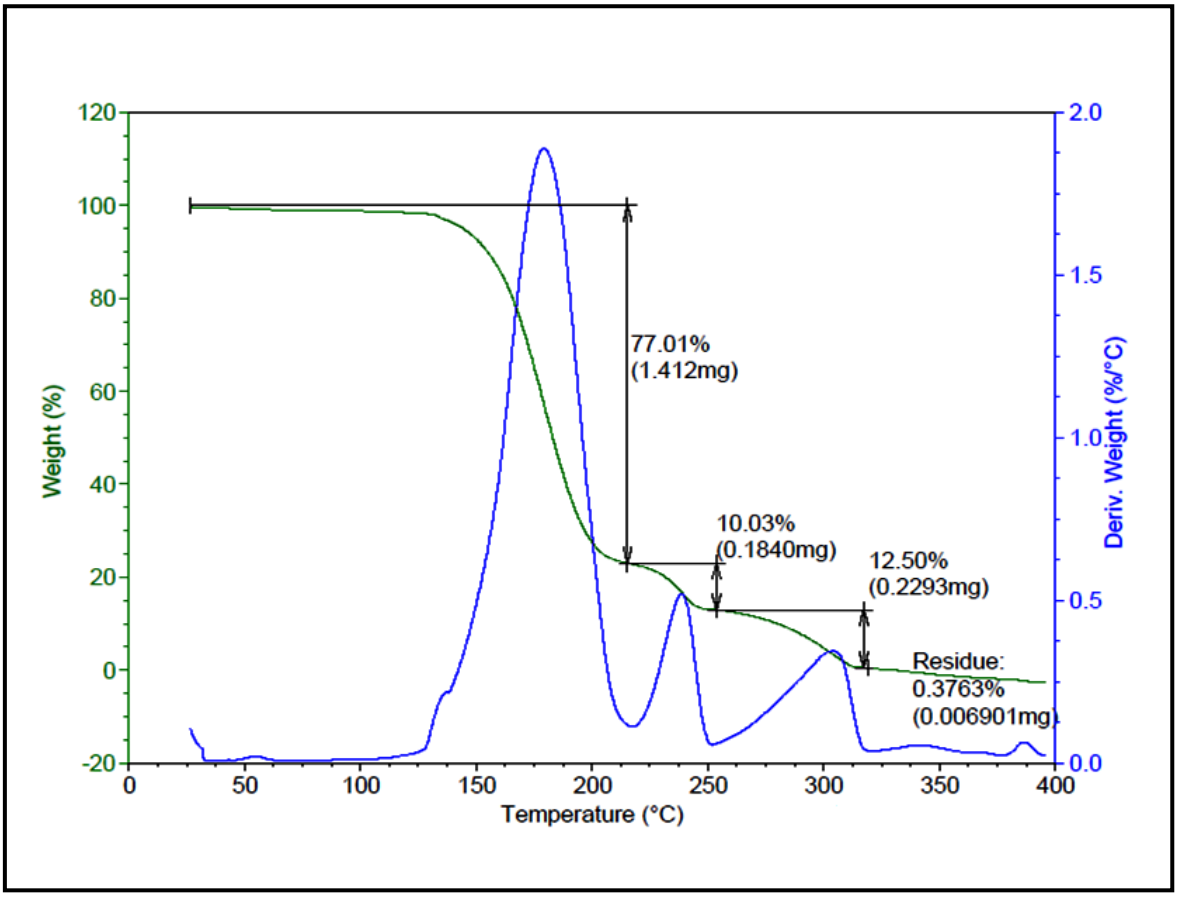


Figure 5.4: TGA profile of the GST-Switch domain protein

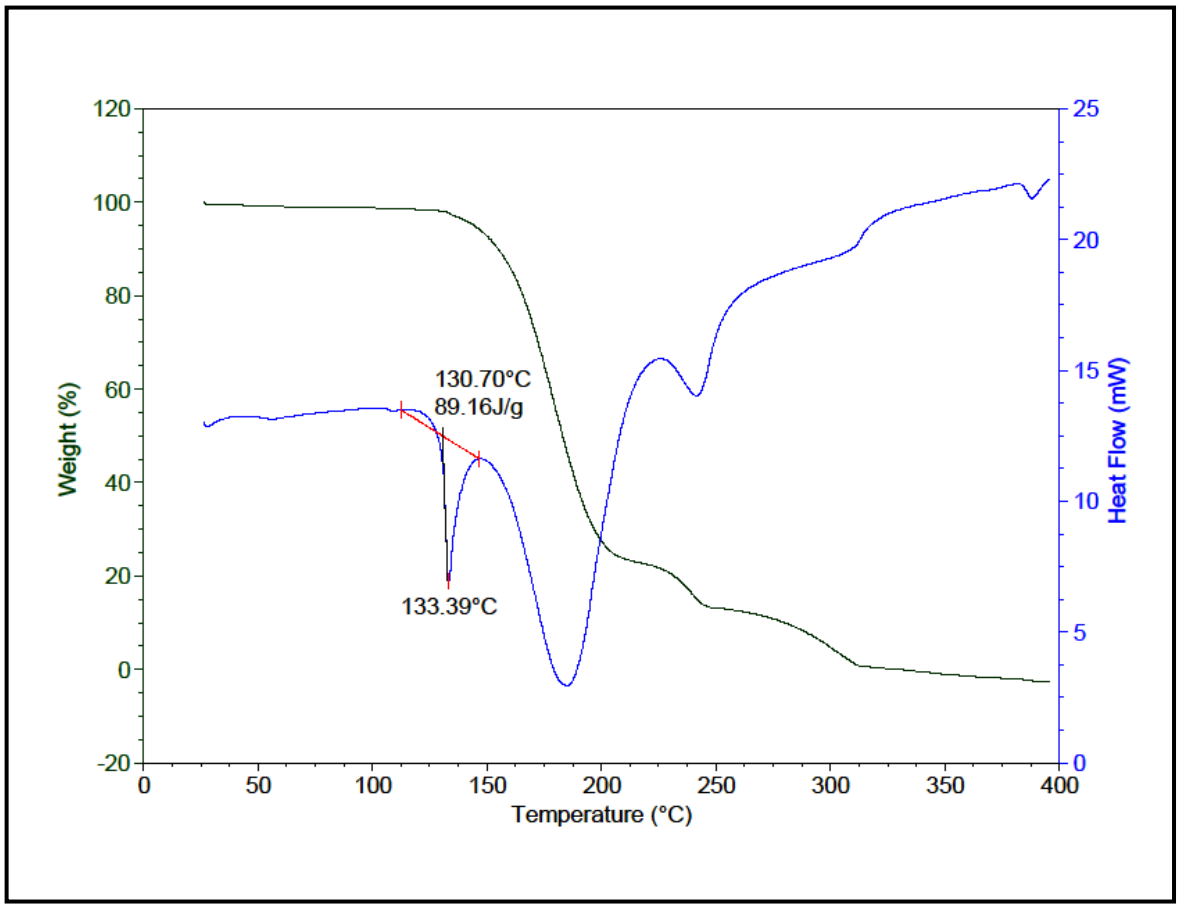


Figure 5.5: DSC profile of the GST-Switch domain protein

Table 5.2: Summary of the DSC profile of GST-tagged switch domain protein

Parameters	Values
T_{on}	130.70 °C
T_{d2}	133.39 °C
T_{off}	145 °C
ΔH	89.16J/g

Onset temperature: (T_{on}), peak denaturation temperature: (T_{d2}), off-set temperature: (T_{off}), and enthalpy of denaturation: (ΔH)

5.3.5 Nuclear magnetic resonance spectroscopy (NMR): (1D ^{1}H-NMR) spectroscopy

1D ^{1}H-NMR spectroscopy was carried out as described in Section 5.2.5. 1D ^{1}H homonuclear experiment was carried out to evaluate the integrity of the switch domains protein and also to ascertain whether the switch domains protein is folded, partially folded or unfolded. As presented in Figure 5.6, the presence of chemical shift dispersions in the methyl group region around 1.0 ppm as well as the presence of resonances upfield of 5.0 ppm are good indicators of the protein (the switch domains protein) 'foldedness' (Kwan *et al.,* 2011; Pugh, 2005).

Through protein structure elucidation, important information needed in understanding protein functions has been revealed. In recent years, this information has facilitated rational drug design and novel drugs development. Protein structural determination and characterization is important to identify suitable constructs; this is achieved by analyzing the expression clones (Geerlof *et al*., 2006). NMR is a standard method in determining protein structures in solution; the major drawback of this method is that of size limitation (Kielec *et al*., 2010; Pugh *et al.,* 2005). It is difficult to solve structures of protein that is above 30 kD using NMR due to broaden line-widths and intrinsic overlap of chemical shifts (Arora and Tamm, 2001). Protein folding state is another limitation of NMR; it is difficult to use NMR to solve protein structures if the protein of interest is unfolded (not properly folded) or denatured (Pugh *et al.,* 2005).

Recently, there are numerous solution-state NMR-based experiments for evaluation of protein integrity and folding state. Commonly, 1D or 2D NMR experimental methods are employed to provide information on the structure and dynamics of protein folding, partly folded and unfolded states (Kovermann *et al*., 2016. The rate

of folding of most proteins is faster than the time scale of the NMR experiment. Therefore, in real time it may be a little bit difficult to monitor the folding or unfolding events of proteins using NMR. In spite of this, NMR experiments can be used to acquire important information in favourable situation where folding is sufficiently slow (Dyson and Wright, 1996). 1D ^{1}H homonuclear experiments is the first choice and have been used extensively to obtain detailed information from several proteins following slow folding events immediately after initiation of the folding reaction. Large chemical shift dispersions of the proton signals, especially in the methyl group region and the amide region can be used as an indicator for protein folding (Alessandri *et al*., 2012).

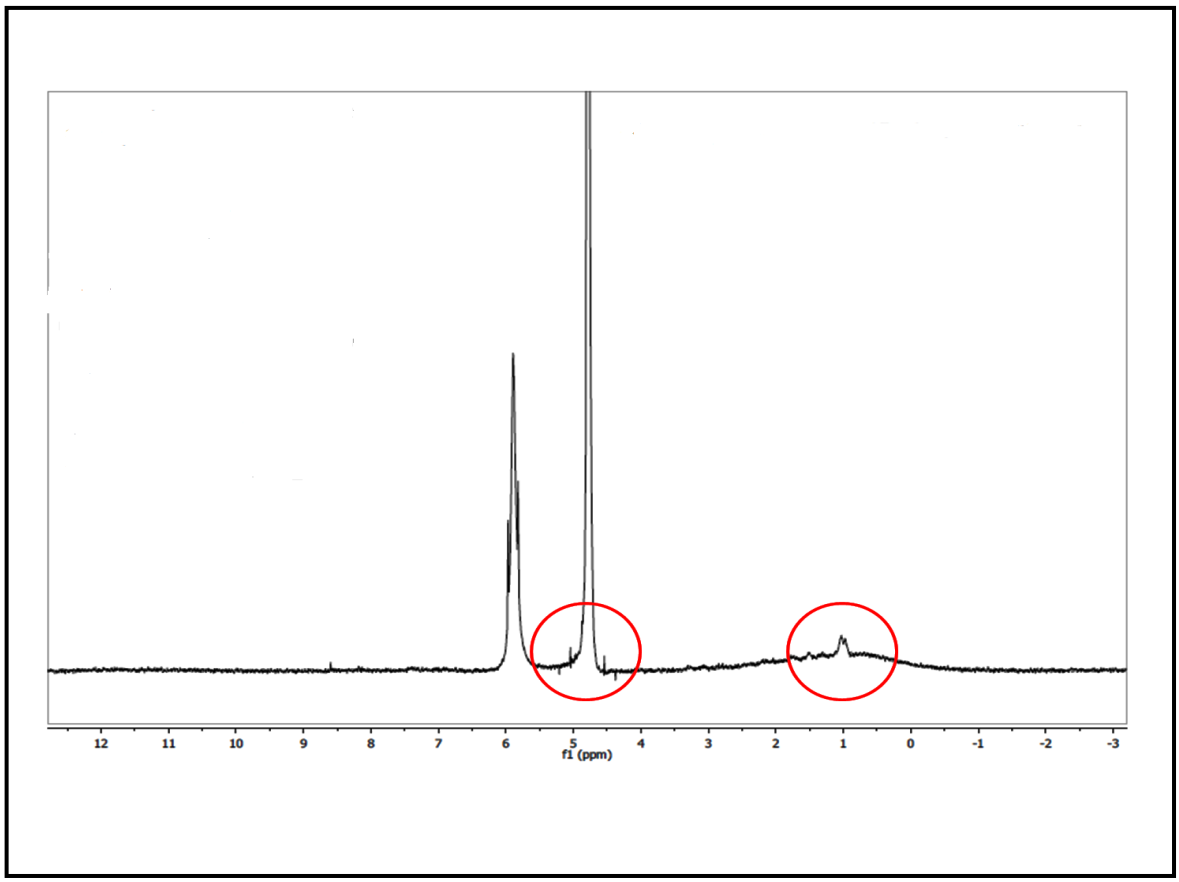


Figure 5.6: 1D proton spectra of the GST-Switch domain protein, recorded at 600 MHz at 25 °C.

5.3.6 Interaction studies using isothermal titration calorimetry (ITC)

Isothermal titration calorimetry analysis of the switch domains protein was carried out as described in Section 5.2.6. This experiment was carried out to characterize the interaction between two of the ligands (guanosine triphosphate and Calcium) that have been hypothesized in literature to interact with the switch domains protein. The same conditions and procedure was followed for the ligand (guanosine triphosphate; GTP or Calcium (Ca) as Calcium chloride; $CaCl_2$) titrations with protein (GST-tagged switch domain protein). The results of the ITC interaction titrations for both ligands (GTP or Ca) against the protein (GST-tagged switch domain protein) are presented in Figure 5.7 and Figure 5.8 respectively.

ITC studies of GTP or Ca binding to the GST-tagged switch domain protein showed that the binding were endothermic. The peaks displayed in Figure 5.7 and Figure 5.8 respectively shows the progress of binding of the analyte (GTP or Ca) to GST-tagged switch domain protein with time as well as the amount of heat released as the interaction progresses. The stoichiometry of interaction (n), association constant (K_a), free energy (ΔG), enthalpy (ΔH) and entropy (ΔS) are presented in Table 5.3. Appendix III shows the ligand binding report that was generated after each experiment. Taken together, the enthalpies and entropies for the binding of GST-tagged switch domain protein to GTP or Ca were all positive (Table 5.3). This is an indication that the binding is entropically favoured (Anbazhagan *et al*., 2011).

In recent times, Isothermal Titration Calorimetry (ITC) experiments have been widely used in investigating the association of biological macromolecules. Really, it is a dependable and quick method that can be used in the study of biochemical interactions (Chilom *et al*., 2006). At the moment, ITC is the only experimental

technique capable of providing thermodynamic parameters such as enthalpy, entropy, Gibbs free energy, binding constant and stoichiometry of the interaction (*n*) under isothermal and isobaric conditions in a single titration experiment (Baranauskienė *et al*., 2009; Rajarathnam and Rösgen, 2014).

Although, protein-ligand binding interactions can also be studied using other techniques such as surface plasmon resonance (SPR); fluorescence analytical ultracentrifugation (AUC) and NMR just to mention a few, ITC among others is gradually becoming an established method of choice for carrying out such studies, thereby contributing valuable information in new drug design and development (Holdgate, 2001; Hussain *et al*., 2016).

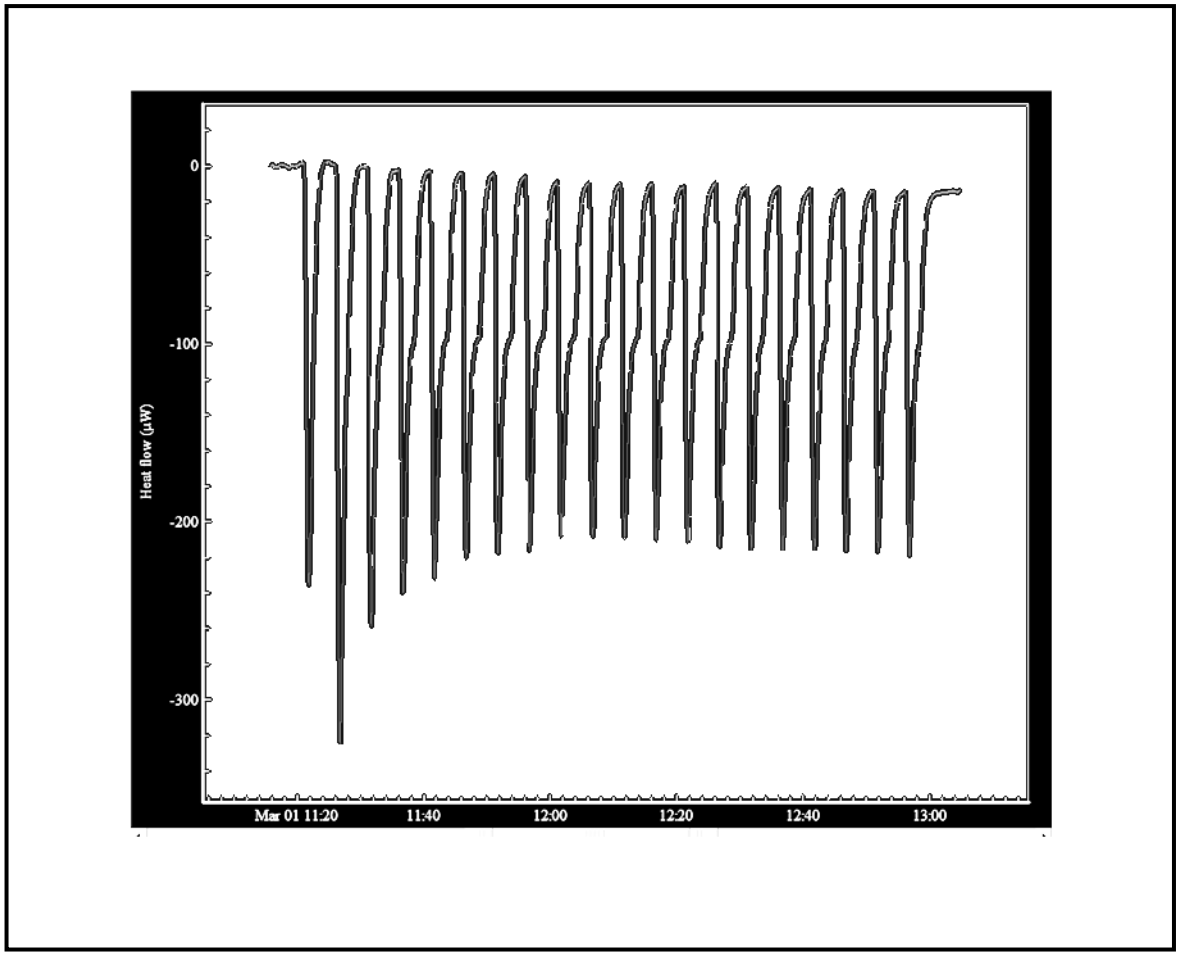


Figure 5.7: ITC thermogram showing the titration of GTP and the GST-Switch domain protein. A total of 20 injections were performed for interactions

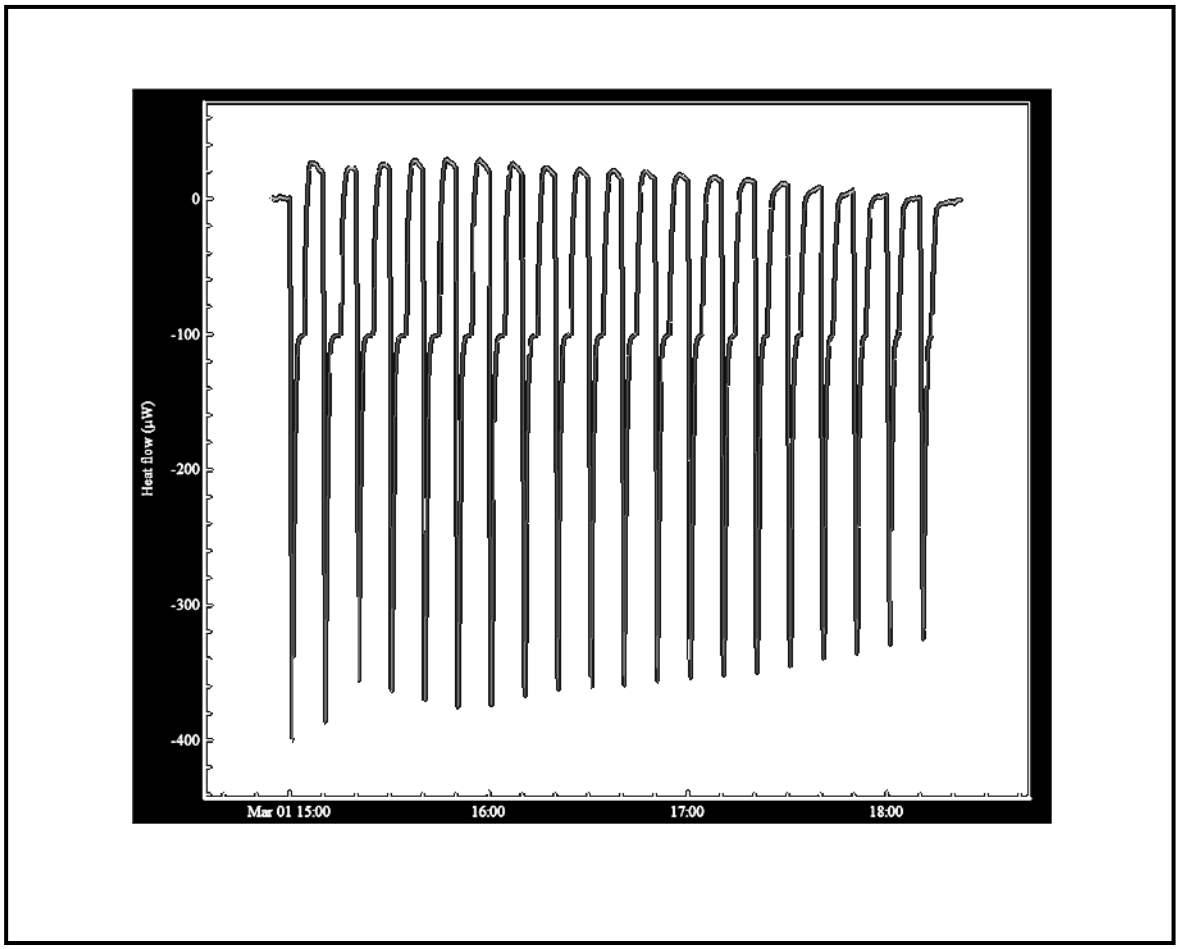


Figure 5.8: ITC thermogram showing the titration of Ca and the GST-Switch domain protein. A total of 20 injections were performed for interactions.

Table 5.3: Summary of the thermodynamic properties obtained from Isothermal titration calorimetry (ITC)

	Free energy (ΔG) (kJ/mol)	**Stoichiometry of the interaction (*n*)**	**Enthalpy (ΔH) (MJ/mol / GJ/mol)**	**Entropy (ΔS) (MJ/(K*mol))**	**Association constant (K_a)**
GTP/GST-Switch domain protein	-8.332	$1.11 \cdot 10^{0} \pm 3.7 \cdot 10^{-2}$	630 ± 1.3	2.13	$2.9 \cdot 10^{1} \pm 6.5 \cdot 10^{1}$
Ca/GST-Switch domain protein	-7.219	$1.13 \cdot 10^{0} \pm 1.3 \cdot 10^{-2}$	330 ± 44	1.093	$1.8 \cdot 10^{1} \pm 3.8 \cdot 10^{0}$

5.4 Conclusion

This chapter, have reported the preliminary characterization of the GST-tagged switch domain protein using various techniques. The results obtained from this study, has given a considerable insight into the nature of the protein in terms of its purity, functional groups, structural and thermal stability as well as its folding state and its binding interaction with two of the hypothesized ligands in literature.

REFERENCES

Abdelhalim M.A.K., Moussa S.A.A. and Al-Mohy Y.H. (2013). Ultraviolet-visible and fluorescence spectroscopy techniques are important diagnostic tools during the progression of atherosclerosis: diet zinc supplementation retarded or delayed atherosclerosis. *BioMed Research International*. **2013**: 652604.

Aboul-enein Y., Bunaciu A.A., and Fleschin S. (2014). Evaluation of the protein secondary structures using Fourier Transform Infrared Spectroscopy. *Gazi University Journal of Science*. 27(1): 637-644.

Alessandri S., Sancho A., Vieths S., Mills C.E., Wal J.M., Shewry P.R., Rigby, N. and Hoffmann-Sommergruber K. (2012). High-throughput NMR assessment of the tertiary structure of food allergens. *PloS One*. **7**(7): e39785.

Anbazhagan V., Sankhala R.S., Singh B.P. and Swamy M.J. (2011). Isothermal Titration Calorimetric Studies on the Interaction of the Major Bovine Seminal Plasma Protein, PDC-109 with Phospholipid Membranes. *PLoS One*. **6**(10): e25993.

Apte M., Girme G., Bankar A., RaviKumar A., and Zinjarde S. (2013). 3, 4-dihydroxy-L-phenylalanine-derived melanin from Yarrowia lipolytica mediates the synthesis of silver and gold nanostructures. *Journal of Nanobiotechnology*. **11**: 2.

Arora A. and Tamm L.K. (2001). Biophysical approaches to membrane protein structure determination. *Current opinion in Structural Biology*. **11**(5): 540-547.

Baranauskienė L., Petrikaitė V., Matulienė J. and Matulis D. (2009). Titration calorimetry standards and the precision of isothermal titration calorimetry data. *International Journal of Molecular Sciences*. **10**(6): 2752-2762.

Bruylants G., Wouters J. and Michaux C. (2005). Differential scanning calorimetry in life science: thermodynamics, stability, molecular recognition and application in drug design. *Current Medicinal Chemistry*. **12**(17): 2011-2020.

Centeno S.A. and Shamir J. (2008). Surface enhanced Raman scattering (SERS) and FTIR characterization of the sepia melanin pigment used in works of art. *Journal of Molecular Structure*. **873**(1): 149-159.

Chilom C.G., Craescu C.T. and Popescu A.I. (2006). Parameters of interaction between proteins and their specific ligands, deduced by isothermal titration calorimetry. *Romanian Journal of Physics*. **51**(3/4): 443-457.

Chiu M.H. and Prenner E.J. (2011). Differential scanning calorimetry: an invaluable tool for a detailed thermodynamic characterization of macromolecules and their interactions. *Journal of Pharmacy and Bioallied Sciences*. **3**(1): 39-59.

Denslow N.D., Wingfield P.T. and Rose K. (2001). Overview of the characterization of recombinant proteins. *Current Protocols in Protein Science*. **2001**: 7.0.1-7.0.4.

Deshmukh K.R. and Pethe A.S. (2016). Extraction and Analysis of Melanin Pigment Produced by Clostridium tertium Isolated from Water Sample of Saline Belt in West Vidardha Region. *International Journal of Science and Research*. **5**(8): 812 – 814.

Dyson H.J. and Wright P.E. (1996). Insights into protein folding from NMR. *Annual Review of Physical Chemistry*. **47**(1): 369-395.

El-Turk F., Fauvet B., Ashrafi A., Ouertatani-Sakouhi H., Cho M.K., Neri M., Cascella M, Rothlisberger U, Pojer F, Zweckstetter M. and Lashuel, H. (2012). Characterization of molecular determinants of the conformational stability of macrophage migration inhibitory factor: leucine 46 hydrophobic pocket. *PloS One*. **7**(9): e45024.

Gallagher W. (2009). FTIR analysis of protein structure. http://www.chem.uwec.edu/chem455_s05/pages/manuals/ftir_of_proteins.pdf [assessed on 14/03/17].

Geerlof A., Brown J., Coutard B., Egloff M.P., Enguita F.J., Fogg M.J., Gilbert R.J.C., Groves M.R., Haouz A., Nettleship N.P., Owens R.J., Ruff M., Sainsbury S., Svergun D.I. and Wilmanns M. (2006). The impact of protein characterization in structural proteomics. *Acta Crystallographica Section D: Biological Crystallography*. **62**(10): 1125-1136.

Gill P., Moghadam T.T. and Ranjbar B. (2010). Differential scanning calorimetry techniques: applications in biology and nanoscience. *Journal of Biomolecular Techniques*. **21**(4): 167-193.

Haris P.I. (1999). Characterization of protein structure and stability using Fourier transform infrared spectroscopy. *Pharmacy and Pharmacology Communications*. **5**(1): 15-25.

Holdgate G.A. (2001). Making cool drugs hot. *Biotechniques*. **31**: 164-184.

Hussain R., Harding S.E., Hughes C.S., Ma P., Patching S.G., Edara S., Siligardi G., Henderson P.J. and Phillips-Jones M.K. (2016). Purification of bacterial membrane sensor kinases and biophysical methods for determination of their ligand and inhibitor interactions. *Biochemical Society Transactions*. **44**(3): 810-823.

Jeyaraman P., Meenarathi B. and Anbarasan R. (2016). Synthesis, characterization and drug-delivery activity of rifampin anchored poly (vinyl alcohol). *Bulletin of Materials Science*. **39**(1): 201-207.

Kielec J.M., Valentine K.G. and Wand A.J. (2010). A method for solution NMR structural studies of large integral membrane proteins: reverse micelle

encapsulation. *Biochimica et Biophysica Acta (BBA)-Biomembranes*. **1798**(2): 150-160.

Kovermann M., Rogne P. and Wolf-Watz M. (2016). Protein dynamics and function from solution state NMR spectroscopy. *Quarterly Reviews of Biophysics*. **49**: e6.

Kwan A.H., Mobli M., Gooley P.R., King G.F. and Mackay J.P. (2011). Macromolecular NMR spectroscopy for the non-spectroscopist. *FEBS Journal*. **278**: 687-703.

Lee H.S., Kim J.S., Shim K., Kim J.W., Inouye K., Oneda H., Kim Y.W., Cheong K.A., Cha H., Woo E.J. and Auh J.H. (2006). Dissociation/association properties of a dodecameric cyclomaltodextrinase. *FEBS Journal*. **273**(1): 109-121.

León-Mancilla B.H., Araiza-Téllez M.A., Flores-Flores J.O. and Piña-Barba M.C. (2016). Physico-chemical characterization of collagen scaffolds for tissue engineering. *Journal of Applied Research and Technology*. **14**(1): 77-85.

Magarelli M., Passamonti P. and Renieri C. (2010). Purification, characterization and analysis of sepia melanin from commercial sepia ink (Sepia Officinalis). *Revista CES Medicina Veterinaria y Zootecnia*. **5**(2): 18-29.

Mbonyiryivuze A., Mwakikunga B., Dhlamini S.M. and Maaza M. (2015). Fourier transform infrared spectroscopy for sepia melanin. *Physics and Materials Chemistry*. **3**(2): 25-29.

Mirza S.P. and Olivier M. (2008). Methods and approaches for the comprehensive characterization and quantification of cellular proteomes using mass spectrometry. *Physiological Genomics*. **33**(1): 3-11.

Pugh D.J.R., Eiso AB., Faro A., Lutya P.T., Hoffmann E. and Rees D.J.G. (2006). DWNN, a novel ubiquitin-like domain, implicates RBBP6 in mRNA processing and ubiquitin-like pathways. *BMC Structural Biology*. **6**: 1.

Pugh DJR. (2005). Biomolecular NMR spectroscopy in South Africa: the first five years. *South African Journal of Science*. **101**: 421-429.

Rajarathnam K. and Rösgen J. (2014). Isothermal titration calorimetry of membrane proteins—Progress and challenges. *Biochimica et Biophysica Acta (BBA)-Biomembranes*. **1838**(1): 69-77.

Raynal B., Lenormand P., Baron B., Hoos S., and England P. (2014). Quality assessment and optimization of purified protein samples: why and how?. *Microbial Cell Factories*. **13**: 180.

Sharon M. (2010). How far can we go with structural mass spectrometry of protein complexes?. *Journal of the American Society for Mass Spectrometry*. **21**(4): 487-500.

Teare J.M., Islam R., Flanagan R., Gallagher S., Davies M.G. and Grabau, C. (1997). Measurement of nucleic acid concentrations using the DyNA Quant and the GeneQuant. *Biotechniques*. **22**(6): 1170-1174.

Wen J., Arthur K., Chemmalil L., Muzammil S., Gabrielson J. and Jiang Y. (2012). Applications of differential scanning calorimetry for thermal stability analysis of proteins: qualification of DSC. *Journal of Pharmaceutical Sciences*. **101**(3): 955-964.

Yang H., Yang S., Kong J., Dong A. and Yu S. (2015). Obtaining information about protein secondary structures in aqueous solution using Fourier transform IR spectroscopy. *Nature Protocols*. **10**(3): 382-396.

CHAPTER SIX

GENERAL DISCUSSION AND CONCLUSION

6.1 General discussion

Today it is very well established that *Schistosoma mansoni* is one of the two main widely distributed causative species responsible for the disease called schistosomiasis (Alebie *et al*., 2014; Mazigo *et al*., 2013). At the moment, human schistosomiasis remains one of the major public health issue facing over 74 developing countries where this disease is known to be endemic in many parts of their population (Mangal *et al*., 2010; Ossai *et al*., 2014). Globally, over 732 million people are at risk with over 235 million infections in known endemic countries (Oyinloye *et al.,* 2014). In sub-Saharan Africa, schistosomiasis and Human Immunodeficiency Virus-1/AIDS co-infection occurs frequently. In fact, a significant overlap between upstream causes of both infections has been documented in literature (Mazigo *et al*., 2013; Shim, 2015; Suttiprapa *et al*., 2016).

Additionally, it was reported that repeated exposure to some of the five schistosoma species parasitizing human may lead to hepatic damage and also result in anaemia, especially in children (Butler *et al*., 2012; Elbaz and Esmat, 2013; Mangal *et al*., 2010). The life cycles of the five documented schistosoma species known to parasitize humans are very similar, even though they are complex. Their life cycles are believed to alternate between two important hosts; the definitive host, comprising of a sexual phase in humans and an intermediate host, comprising of an asexual phase in snail (Bill, 2003; Colley *et al.,* 2014; Oyinloye *et al.,* 2014). In the last 30 years praziquantel (PZQ) has remain the only drug of choice in the treatment of

schistosomiasis due to its effectiveness and transient side effects when compared to other currently used antischistosomal drugs (Lee *et al.,* 2011; Mutapi *et al.,* 2011; Ojurongbe *et al.,* 2014).

Currently, some isolates of *S. mansoni* have been reported to exhibit some worrisome signs of resistance to treatment with PZQ, as a result of drug pressure (Doenhoff *et al*., 2002; Oyinloye *et al.,* 2014; Ross *et al*., 2002). Taken together, there is an urgent need to design and develop novel therapy for the treatment of human schistosomiasis and this has been an essential motivation for molecular research on schistosomiasis. However, novel drug discovery, design and development; is a high risk multifaceted time consuming adventure with potentially high rewarding opportunities (Katara, 2013). In recent years, bioinformatics approach has been recognized to play an increasingly significant role in nearly all areas of drug discovery, drug design and drug development as well as drug assessment. Various bioinformatics and computational tools have been used to predict, analyze, interpret and validate potential drug targets (Gill *et al*., 2016; Wishart, 2005).

Bioinformatics and computational approach play an important predictive role in designing wet-laboratory experiments (Dhanda *et al*., 2016; Emmert-Streib *et al*., 2014). In this study, various bioinformatics tools were employed to characterise and obtain information about the structural features of the switch domain regions from *Schistosoma mansoni*. As presented in chapter four, the functional residues corresponding to the switch domains of *S. mansoni* Smp_059340.1 (183 – 223 aa) consist of two promising functional motifs that are highly conserved in nature. The secondary structure analysis revealed that the switch domains of *S. mansoni* Smp_059340.1 contain 1 α-helices, 2 β-sheets and 4 random coils. Homology modeling of the switch domain regions and overall stereochemical parameters as

evaluated by the Ramachandran plot analysis in RAMPAGE reveals that the modelled structure is stable and of good quality.

Added to this, energy minimization with YASARA energy minimization server and GalaxyRefine server for protein structure refinement shows that the modelled structure is a thermodynamically stable model. On the other hand, protein network interaction analysis determined by STRING reveals that the protein interacts with ten other proteins. Consequently, ligand binding locations and residues were identified and molecular docking was carried out. Docking studies revealed the binding and interaction patterns of the switch domains protein with guanosine-5'-triphosphate, magnesium, calcium and zinc. Using bioinformatics approach, an unprecedented wealth of biological information can be generated and developed, leading to accelerated discovery of new drug targets (Bayat, 2002). The *in-silico* analysis performed in this study, shows that the switch domains protein is a potential candidate for novel antischistosomal drug design.

The roles of recombinant therapeutic proteins as prerequisite and essential component in small molecule drug development have become increasingly important over the years (Buckel, 1996). Recombinant GST-tagged switch domain was heterologously expressed in *Escherichia coli* (*E. coli*) JM109 and subsequently purified using glutathione-agarose beads. SDS-PAGE analysis of the recombinant GST-Switch domain protein shows that a high yield soluble fusion protein was obtained, as visible prominent bands (~31.84 kDa) corresponding to the GST-Switch domain protein were observed in the region of 25 – 37 kDa. These results suggest that the protein could be used for further structural studies in order to elucidate the biochemical mechanism and understand the molecular basis of how the *S. mansoni*

worm synthesize energy via the release of cAMP to transverse the hepatic portal system of the host.

The structural integrity (preliminary characterization) of the recombinant GST-Switch domain protein was assessed using Fourier transform infrared spectroscopy (FTIR), ultraviolet-visible spectrophotometry, fluorescence spectroscopy, differential scanning calorimeter (DSC), thermogravimetric analysis (TGA), nuclear magnetic resonance spectroscopy (NMR) and isothermal titration calorimetry (ITC). The results obtained from this study, has given us a considerable insight into the nature of the protein in terms of its purity, functional groups, structural and thermal stability as well as its folding state and its binding interaction with two of the hypothesized ligands in literature. Results from the preliminary characterization confirmed the druggability of the switch domain protein.

In recent times, thermodynamic characterization of biological macromolecules using ITC studies is an integral part of novel drug design and development (Rajarathnam and Rösgen, 2014). ITC among other techniques used for characterization is steadily becoming an established method of choice for carrying interaction studies. It has been reported that the current available information from biological, structural and thermodynamic studies are used in the design and development of the majority of the effective drugs available in the market (Garbett and Chaires, 2012). G protein-coupled receptors (GPCRs) represent one of the most popular drug targets; they present a great opportunity for new drug discovery and development because they are the site of action of 25-30% of current drugs (Congreve *et al*., 2015; Rajarathnam and Rösgen, 2014).

Although, various computational tools and modeling studies are been employed as valuable tools for drug discovery (Sliwoski *et al*., 2014); these tools may not be able to effectively capture the basics of the molecular foundations of binding interactions. As a result of this, thermodynamic information from ITC experiments is suitable to provide useful data that will pave way in designing and developing new therapeutic drugs with high-affinity and specificity (Rajarathnam and Rösgen, 2014).

6.2 Conclusion

This work describes for the first time the recombinant expression and purification, bioinformatics and computational analysis as well as structural characterization of the switch domain protein of *S. mansoni*. Since these molecular switch domains are critical for the survival of *Schistosoma* worm, determining the structure of these domains will serve as the committed step (since the structure of the protein will serve as the lead compound) in search of new anti-schistosoma agents. In view of the aforementioned, there is an urgent necessity for the design of novel structure based drugs with better efficacy and minimal transient side-effects that could combat human schistosomiasis as an alternative treatment to praziquantel (PZQ). Taken together, the overall results from this study are impressive and provide evidence that affirm the druggability of the switch domain protein of *S. mansoni* for the development of novel anti-schistosomal drug.

6.3 Recommendation for further studies

The findings from this study suggest that the switch domain protein is soluble, well folded and stable. Therefore, this protein can be explored as a new drug target for schistosomiasis. Although, one of the major challenges faced during this study was cleaving the fusion protein (GST-Switch domain protein) using 3C protease, and

since all attempts were not successful, as the expected 5.8 kDa (Switch domain protein) protein was not visible on the gel after SDS-PAGE analysis of the 3C protease cleavage of the GST-Switch domain protein, it is highly important to conduct further studies in order to develop a purification strategy, whereby the GST-tag will be completely cleaved off the Switch domain protein. Secondly, the expression of the switch domains of protein should be optimized in minimal media and biophysical assessment using mass spectrometry and circular dichroism should be carried out. Additionally, 3-dimensional structure elucidation of the protein using X-ray crystallography and/or heteronuclear NMR spectroscopy should be carried out. Added to this, one of the limitations of this study is that we did not investigate the interactions of antimicrobial peptides with the expressed Gα protein from the switch domains of *S. mansoni*. Therefore, it is also recommended that future exploration of antimicrobial peptides with anti-schistosomal activity be carried out.

REFERENCES

Alebie G., Erko B., Aemero M. and Petros B. (2014). Epidemiological study on Schistosoma mansoni infection in Sanja area, Amhara region, Ethiopia. *Parasites Vectors* **7**:15.

Bayat A. (2002). Bioinformatics. *British Medical Journal.* **324**(7344): 1018–1022.

Bill P. (2003). Schistosomiasis and the nervous system. *Practical Neurology*. **3**(1): 12-21.

Buckel P. (1996). Recombinant proteins for therapy. *Trends in Pharmacological Sciences*. **17**(12): 450-456.

Butler S.E., Muok E.M., Montgomery S.P., Odhiambo K., Mwinzi P.M., Secor W.E. and Karanja D.M. (2012). Mechanism of Anemia in Schistosoma mansoni–Infected School Children in Western Kenya. *American journal of Tropical Medicine and Hygiene*. ***87***(5): 862-867.

Colley D.G., Bustinduy A.L., Secor W.E. and King C.H. (2014). Human schistosomiasis. *Lancet*. **383**(9936): 2253-2264.

Congreve M., Doré, A.S., Jazayeri, A. and Nonoo R. (2015). Engineering G Protein-Coupled Receptors for Drug Design. In *Multifaceted Roles of Crystallography in Modern Drug Discovery*. (pp. 1-18). Springer Netherlands.

Dhanda S.K., Usmani S.S., Agrawal P., Nagpal G., Gautam A. and Raghava G.P. (2016). Novel in silico tools for designing peptide-based subunit vaccines and immunotherapeutics. *Briefings in Bioinformatics*. **doi**: (2016).10.1093/bib/bbw025

Doenhoff M.J., Kusel J.R., Coles G.C. and Cioli D. (2002) Resistance of Schistosoma mansoni to praziquantel: is there a problem? *Transactions of the Royal Society of Tropical Medicine and Hygiene*. **96**: 465-69.

Elbaz T. and Esmat G. (2013). Hepatic and intestinal schistosomiasis: review. *Journal of Advanced Research*. **4**(5): 445-452.

Emmert-Streib F., Zhang S.D. and Hamilton P. (2014). Dry computational approaches for wet medical problems. *Journal of Translational Medicine*. **12**(1): 26.

Garbett N.C. and Chaires J.B. (2012). Thermodynamic studies for drug design and screening. *Expert opinion on drug discovery*. **7**(4): 299-314.

Gill S.K., Christopher A.F., Gupta V. and Bansal P. (2016). Emerging role of bioinformatics tools and software in evolution of clinical research. *Perspectives in Clinical Research*. **7**(3): 115-122.

Katara P. (2013). Role of bioinformatics and pharmacogenomics in drug discovery and development process. *Network Modeling Analysis in Health Informatics and Bioinformatics*. **2**(4): 225-230.

Lee J.M., Lim H.S. and Hong S.T. (2011). Hypersensitive reaction to praziquantel in a clonorchiasis patient. *Korean Journal of Parasitology*. **49**(3): 273-275.

Mangal T.D., Paterson S. and Fenton A. (2010). Effects of snail density on growth, reproduction and survival of Biomphalaria alexandrina exposed to Schistosoma mansoni. *Journal of Parasitology Research*. **2010**: 186792.

Mazigo H.D., Nuwaha, F., Wilson, S., Kinung'hi, S.M., Morona, D., Waihenya, R., Heukelbach J. and Dunne D.W. (2013). Epidemiology and interactions of Human Immunodeficiency Virus–1 and Schistosoma mansoni in sub-Saharan Africa. *Infectious Diseases of Poverty*. **2**(1): 2.

Mutapi F., Rujeni N., Bourke C., Mitchell K., Appleby L., Nausch N., Midzi N. and Mduluza T. (2011). Schistosoma haematobium treatment in 1–5 year old children: safety and efficacy of the antihelminthic drug praziquantel. *PLoS Neglected Tropical Diseases*. **5**(5): e1143.

Ojurongbe O., Sina-Agbaje O.R., Busari A., Okorie P.N., Ojurongbe T.A. and Akindele A.A. (2014). Efficacy of praziquantel in the treatment of Schistosoma haematobium infection among school-age children in rural communities of Abeokuta, Nigeria. *Infectious diseases of poverty*, *3*(1), 30.

Ossai O.P., Dankoli R., Nwodo C., Tukur D., Nsubuga P., Ogbuabor D., Ekwueme O, Abonyi G, Ezeanolue E, Nguku P. and Nwagbo D. (2014). Bacteriuria and urinary schistosomiasis in primary school children in rural communities in Enugu State, Nigeria, 2012. *Pan African Medical Journal*. **18**:15.

Oyinloye B., Adenowo F., Gxaba N., and Kappo A. (2014). The Promise of Antimicrobial Peptides for Treatment of Human Schistosomiasis. *Current Drug Targets* **15**(9): 852-859.

Rajarathnam K. and Rösgen J. (2014). Isothermal titration calorimetry of membrane proteins—Progress and challenges. *Biochimica et Biophysica Acta (BBA)-Biomembranes*. **1838**(1): 69-77.

Ross A.G., Bartley P.B., Sleigh A.C., Olds G.R., Li Y., Williams G.M. and McManus D.P. (2002). Schistosomiasis. *New England Journal of Medicine***. 346**: 1212-20.

Shim K. (2015). Expanding the Definition of Infectious Disease. *Journal of Global Health*. http://www.ghjournal.org/expanding-the-definition-of-infectious-disease-2/ [Accessed 20 March 2014].

Sliwoski G., Kothiwale S., Meiler J. and Lowe E.W. (2014). Computational methods in drug discovery. *Pharmacological reviews*. **66**(1): 334-395.

Suttiprapa S., Rinaldi G., Tsai I.J., Mann V.H., Dubrovsky L., Yan H.B., Holroyd N., Huckvale T., Durrant C., Protasio A.V. and Pushkarsky (2016). HIV-1 Integrates Widely throughout the Genome of the Human Blood Fluke Schistosoma mansoni. *PLoS Pathogens*. **12**(10): e1005931.

Wishart D.S. (2005). Bioinformatics in drug development and assessment. *Drug metabolism reviews*. **37**(2): 279-310.

APPENDIX I: SUPPLEMENTARY DATA

NCBI Reference Sequence: XP_018644269.1

Guanine nucleotide-binding protein G(s) subunit alpha (Adenylate cyclase-stimulating G alpha protein), putative [Schistosoma mansoni].

ATGGAACAGGATCGCTGTCGTGTACTAACTTCCGGAATATTTGAAACCAAGTTC
AGCGTGGATAAAGTTAATTTCCATATGTTTGACGTCGGCGGGCAGAGGGAAGA
GCGCCGAAAGTGGATACAATGCTTTAACATCATTTTTTGA

Figure A1: DNA sequence corresponding to the switch domains of the *S. mansoni* Smp_059340.1

```
          10          20          30          40
MEQDRCRVLT SGIFETKFSV DKVNFHMFDV GGQREERRKW IQCFNIIF
```

Figure A2: Amino acid sequence corresponding to the switch domains of the *S. mansoni* Smp_059340.1

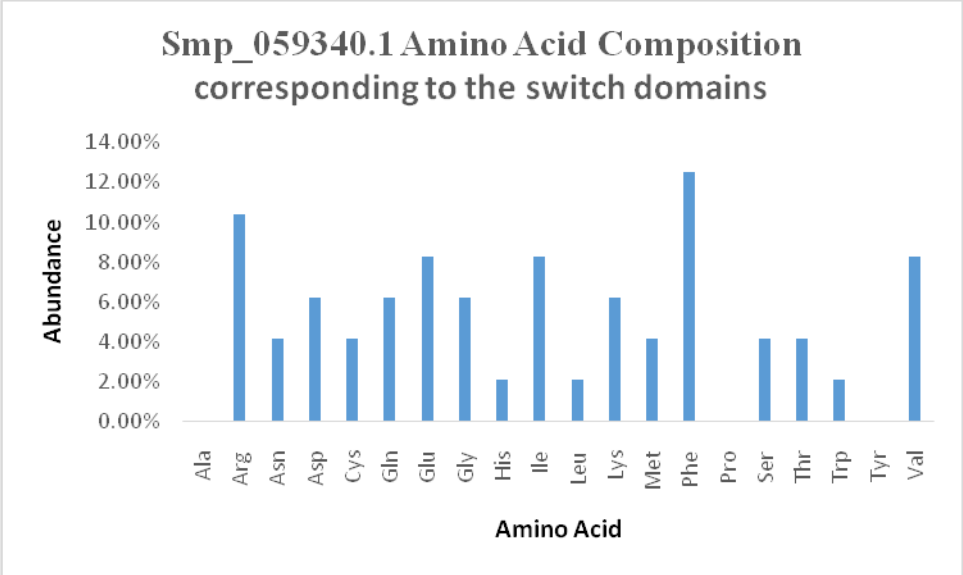


Figure A3: Amino acid sequence corresponding to the switch domains of the *S. mansoni* Smp_059340.1

Table A1: Physicochemical properties of the Switch domain protein calculated from the amino acid sequence

ProtParam Parameters	Values
Number of amino acids	48
Molecular weight	5842.73
Theoretical pI	7.83
Atomic composition	Carbon C 260 Hydrogen H 400 Nitrogen N 74 Oxygen O 72 Sulfur S 4
Formula	$C_{260}H_{400}N_{74}O_{72}S_4$
Number of negatively charged residues	7
Number of positively charged residues	8
Extinction coefficient	5625 Abs 0.1% (=1 g/l) 0.963, assuming all pairs of Cys residues form cystines 5500 Abs 0.1% (=1 g/l) 0.941, assuming all Cys residues are reduced
Total number of atoms	810
Estimated Half-life	The N-terminal of the sequence considered is M (Met) 30 hours (mammalian reticulocytes, *in vitro*). >20 hours (yeast, *in vivo*). >10 hours (Escherichia coli, *in vivo*).
Aliphatic index	64.79
Grand average of hydropathicity (GRAVY)	-0.423
Instability index	34.36

NCBI Reference Sequence: XP_018644269.1

Guanine nucleotide-binding protein G(s) subunit alpha (Adenylate cyclase-stimulating G alpha protein), putative [Schistosoma mansoni].

ATGGTTATTGGCTGTTGCACACTAAATAATATTTCGGAAG ATGCCAAAAC
GCGAAGCGATGCAAATAAACAGATTGAGAAATTAATTGAAAAGGAGAAAA
AGAATTTTAAGTCTACACATAGACTGCTCCTTCTGGGTGCTGGAGAGTCT
GGGAAGTCGACCATTGTTAAACAAATGAGGATATTGCACATAGATGGCTT
TTCAGAACGAGAAAAAAAGGAAAAGATTGATGCAATCCGAAAAAACCTCC
GTGACGCTATTTGTTCGATAGCTGGCGCTATGGGTTCTTTGAAACCACCA
GTGAAGCTTGAACTCAGTGAAAACAGAAAGCTAAGGGATTATATACTGGA
AACTGCATCTAAACCTGACTTCGACTATCC GCCGGAGTTT TTCACATATT
GCGCCAAATTATGGAAGGATGGAGGTATACAGGAAACATTCGAAAGATC
AAATGAGTATCAACTAATCGACTGTGCAAAGTATTTCCTTGATAAGGCTTT
AGAGGTCGGCGCTCCAAACTACATCCCATCTGAACAGGATATTCTGCGC
TGTCGTGTACTAACTTCCGGAATATTTGAAACCAAGTTCAGCGTGGATAA
AGTTAATTTCCATATGTTTGACGTCGGCGGGCAGAGGGAAGAGCGCCGA
AAGTGGATACAATGCTTTAACGATGTCACTGCCATCATTTTTGTCGCTGC
CTGTTCTTCGTACAACATGGTCTTACGAGAAGATCCTAGCCAAAACCGTG
TAAAGGAGTCTTTAGAACTTCTTGCTTCCATATGGAACAACAGGTGGTTG
CGAAACATATCTGTGATTCTCTTTCTTAACAAGCAAGATTTACTTACGGAG
AAGGTTTTAGCTGGCAAGTCCAAGATTGAAGTTTATTTCCCGCATTATGC
TACTTACCAAGCTCCAGCTGATACACTAGCAGAGTATCGTCATGAAAACT
CAGAAGTTGTACGTGCCCGATTTTTCTTCCGTGATGAATTTCTTAAAGTAA
CATCAAATAACAATGGTGGACGTCATTATT GTTATCCTCA CTTAACGTGT
GCAGTTGATACGGAAAATATCCGACGCGTATTCAATGATT GTCGTGATAT
CATTCAGCGAATGCACCTGCGGCAGTATGAATTACTCTAG

Figure A4: Full DNA sequence of *S. mansoni* Smp_059340.1

```
        10         20         30         40         50         60
MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID

        70         80         90        100        110        120
GDVKLTQSMA IIRYIADKHN MLGGCPKERA EISMLEGAVL DIRYGVSRIA YSKDFETLKV

       130        140        150        160        170        180
DFLSKLPEML KMFEDRLCHK TYLNGDHVTH PDFMLYDALD VVLYMDPMCL DAFPKLVCFK

       190        200        210        220        230        240
KRIEAIPQID KYLKSSKYIA WPLQGWQATF GGGDHPPKSD LEVLFQGPLG SGSEQDILRC

       250        260        270        280
RVLTSGIFET KFSVDKVNFH MFDVGGQREE RRKWIQCFND VTAIIF
```

Figure A5: Full amino acid sequence of *S. mansoni* Smp_059340.1

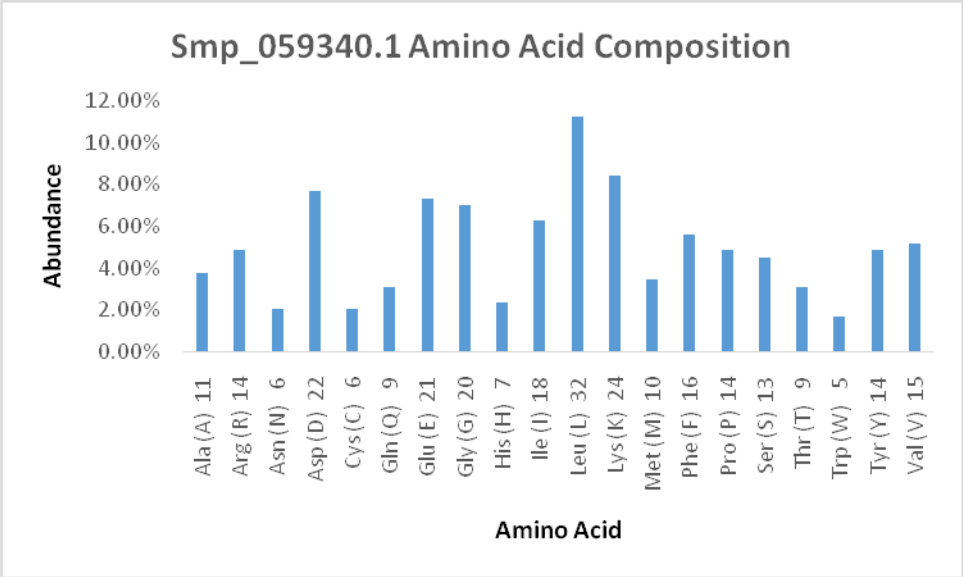


Figure A6: Full amino acid sequence of the *S. mansoni* Smp_059340.1

Table A2: Physicochemical properties of the GST-Switch domain protein (Codon harmonized) calculated from the amino acid sequence

ProtParam Parameters	**Values**
Number of amino acids	286
Molecular weight	33292.5 Da
Theoretical pI	5.86
Atomic composition	Carbon C 1518 Hydrogen H 2338 Nitrogen N 386 Oxygen O 424 Sulfur S 16
Formula	$C_{1518}H_{2338}N_{386}O_{424}S_{16}$
Number of negatively charged residues	43
Number of positively charged residues	38
Extinction coefficient	48735 Abs 0.1% (=1 g/l) 1.464, assuming all pairs of Cys residues form cystines 48360 Abs 0.1% (=1 g/l) 1.453, assuming all Cys residues are reduced
Total number of atoms	4682
Estimated Half-life	30 hours (mammalian reticulocytes, *in vitro*). >20 hours (yeast, *in vivo*). >10 hours (Escherichia coli, *in vivo*).
Aliphatic index	87.24
Grand average of hydropathicity (GRAVY)	-0.306
Instability index	35.01

Table A3: Names and accession numbers of organisms retrieved used in the multiple sequence alignment and phylogenetic tree of full length sequence *S. mansoni* Smp_059340.1

Accession number	Organisms/Source	Protein length	BLAST E-value
XP_018644269.1	Schistosoma mansoni	379	0.0
XP_014470426.1	Dinoponera quadriceps	381	0.0
CDS32166.1	Hymenolepis microstoma	379	0.0
EEC10034.1	Ixodes scapularis	380	0.0
XP_003743481.1	Galendromus occidentalis	378	0.0
XP_005092762.1	Aplysia californica	376	0.0
ETN62507.1	Anopheles darlingi	383	0.0
XP_019769995.1	Dendroctonus ponderosae	379	0.0
XP_015794619.1	Tetranychus urticae	381	0.0

Table A4: Names and accession numbers of organisms retrieved used in the multiple sequence alignment and phylogenetic tree of the switch domains of *S. mansoni* Smp_059340.1

Accession number	Organisms/Source	Protein length	BLAST E-value
XP_018644269.1	Schistosoma mansoni	379	3e-21
GAA53581.1	Clonorchis sinensis	394	3e-21
XP_018596464.1	Scleropages formosus	332	5e-21
KFQ38841.1	Mesitornis unicolor	263	5e-21
XP_019651163.1	Ailuropoda melanoleuca	357	5e-21
ALE66358.1	Ovis aries	211	5e-21
AGJ70283.1	Terebratalia transversa	382	6e-21
XP_011385811.1	Pteropus vampyrus	181	4e-22
AAR04851.1	Bos taurus	211	3e-22

Table A5: Sequence identities and similarities Table (full length sequence of *S. mansoni* Smp_059340.1 protein and its homologues)

-	XP_018644269.1	XP_014470426.1	CDS32166.1	EEC10034.1	XP_003743481.1	XP_005092762.1	ETN62507.1	XP_019769995.1	XP_015794619.1
XP_018644269.1	-	80.5 (90.0)	80.2 (89.2)	69.4 (82.7)	69.4 (81.9)	69.4 (81.0)	66.6 (81.3)	67.6 (80.3)	68.9 (82.0)
XP_014470426.1	80.5 (90.0)	-	93.4 (96.3)	69.4 (80.9)	69.4 (80.9)	68.6 (80.7)	65.8 (79.3)	64.5 (77.7)	69.5 (81.2)
CDS32166.1	80.2 (89.2)	93.4 (96.3)	-	68.3 (80.6)	68.3 (80.4)	68.3 (79.4)	64.5 (78.5)	63.0 (76.9)	67.9 (80.7)
EEC10034.1	69.4 (82.7)	69.4 (80.9)	68.3 (80.6)	-	92.4 (97.4)	83.7 (90.8)	84.6 (90.4)	81.8 (89.3)	89.8 (94.0)
XP_003743481.1	69.4 (81.9)	69.4 (80.9)	68.3 (80.4)	92.4 (97.4)	-	82.8 (91.3)	84.9 (89.8)	81.4 (89.5)	88.5 (93.2)
XP_005092762.1	69.4 (81.0)	68.6 (80.7)	68.3 (79.4)	83.7 (90.8)	82.8 (91.3)	-	80.2 (86.9)	77.0 (87.7)	84.3 (91.1)
ETN62507.1	66.6 (81.3)	65.8 (79.3)	64.5 (78.5)	84.6 (90.4)	84.9 (89.8)	80.2 (86.9)	-	85.1 (90.9)	83.6 (88.8)
XP_019769995.1	67.6 (80.3)	64.5 (77.7)	63.0 (76.9)	81.8 (89.3)	81.4 (89.5)	77.0 (87.7)	85.1 (90.9)	-	80.8 (87.8)
XP_015794619.1	68.9 (82.0)	69.5 (81.2)	67.9 (80.7)	89.8 (94.0)	88.5 (93.2)	84.3 (91.1)	83.6 (88.8)	80.8 (87.8)	-

Table A6: Sequence identities and similarities Table (switch domains of the *S. mansoni* Smp_059340.1 and its homologues)

-	XP_018644269.1	GAA53581.1	XP_018596464.1	KFQ38841.1	XP_019651163.1	ALE66358.1	AGJ70283.1	XP_011385811.1	AAR04851.1
XP_018644269.1	-	95.3 (97.7)	95.3 (97.7)	90.7 (95.3)	90.7 (95.3)	90.7 (95.3)	93.0 (97.7)	90.7 (95.3)	90.7 (95.3)
GAA53581.1	95.3 (97.7)	-	100.0 (100.0)	95.3 (97.7)	95.3 (97.7)	95.3 (97.7)	97.7 (100.0)	95.3 (97.7)	95.3 (97.7)
XP_018596464.1	95.3 (97.7)	100.0 (100.0)	-	95.3 (97.7)	95.3 (97.7)	95.3 (97.7)	97.7 (100.0)	95.3 (97.7)	95.3 (97.7)
KFQ38841.1	90.7 (95.3)	95.3 (97.7)	95.3 (97.7)	-	100.0 (100.0)	100.0 (100.0)	97.7 (97.7)	100.0 (100.0)	100.0 (100.0)
XP_019651163.1	90.7 (95.3)	95.3 (97.7)	95.3 (97.7)	100.0 (100.0)	-	100.0 (100.0)	97.7 (97.7)	100.0 (100.0)	100.0 (100.0)
ALE66358.1	90.7 (95.3)	95.3 (97.7)	95.3 (97.7)	100.0 (100.0)	100.0 (100.0)	-	97.7 (97.7)	100.0 (100.0)	100.0 (100.0)
AGJ70283.1	93.0 (97.7)	97.7 (100.0)	97.7 (100.0)	97.7 (97.7)	97.7 (97.7)	97.7 (97.7)	-	97.7 (97.7)	97.7 (97.7)
XP_011385811.1	90.7 (95.3)	95.3 (97.7)	95.3 (97.7)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	97.7 (97.7)	-	100.0 (100.0)
AAR04851.1	90.7 (95.3)	95.3 (97.7)	95.3 (97.7)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	97.7 (97.7)	100.0 (100.0)	-

Ligand Binding Report

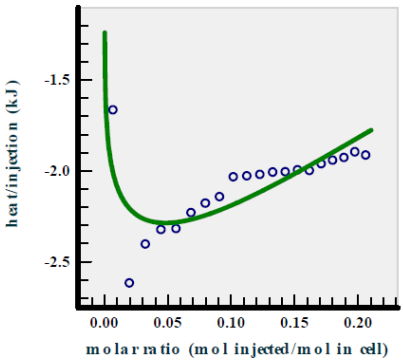
Mar 01, 2017 16:46

Model

$M + nL \leftrightarrow MLn$ $\quad \beta_1, \Delta H_1$

Thermodynamics Parameters

k_1	$2.9 \cdot 10^1$ ±	$6.5 \cdot 10^1$
ΔH_1	630MJ/mol ±	1.3GJ/mol
n	$1.11 \cdot 10^0$ ±	$3.7 \cdot 10^{-2}$
ΔG_1	-8.332kJ/mol	
ΔS_1	2.13MJ/(K*mol)	



Statistics

Degrees of freedom	17
ξ^2	$3.65 \cdot 10^{-5}$
Reduced ξ^2	$2.15 \cdot 10^{-6}$
Standard error of point	$1.47 \cdot 10^{-3}$

Input Data

SRDGTPT1Titration (3-1-17)

Volume in vessel	750µl
Number of datapoints	20
Concentration of [M] in vessel	30.6µM
Concentration of [L] in syringe	9.75mM

No	Volume (µl)	Q_{exp} (mJ)	Q_{calc} (mJ)	Residual (mJ)
1	10.0	-16.59	-19.46	2.86
2	10.0	-26.09	-21.96	-4.13
3	10.0	-23.96	-22.60	-1.36
4	10.0	-23.16	-22.79	-0.37
5	10.0	-23.11	-22.76	-0.35
6	10.0	-22.23	-22.61	0.37
7	10.0	-21.71	-22.38	0.67
8	10.0	-21.36	-22.10	0.74
9	10.0	-20.27	-21.79	1.52
10	10.0	-20.22	-21.45	1.24
11	10.0	-20.14	-21.10	0.96
12	10.0	-20.01	-20.74	0.73
13	10.0	-19.99	-20.38	0.39
14	10.0	-19.89	-20.02	0.13
15	10.0	-19.92	-19.65	-0.27
16	10.0	-19.55	-19.29	-0.26
17	10.0	-19.35	-18.93	-0.42
18	10.0	-19.21	-18.58	-0.63
19	10.0	-18.90	-18.23	-0.67
20	10.0	-19.08	-17.88	-1.19

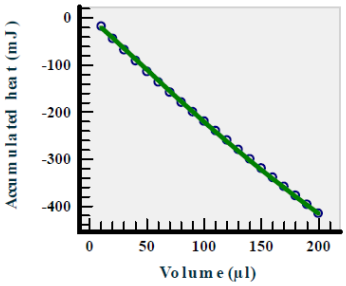


Figure A7: Ligand Binding Report (GTP and the GST-Switch domain protein).

Ligand Binding Report

Mar 02, 2017 7:31

Model

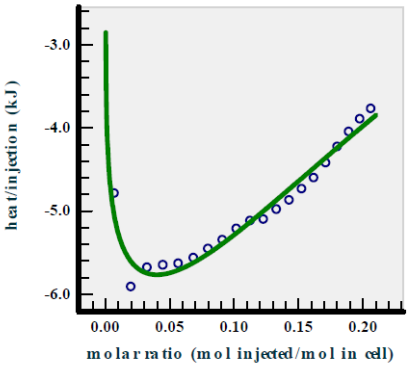
$M + nL \leftrightarrow MLn$ $\quad \beta_1, \Delta H_1$

Thermodynamics Parameters

k_1	$1.8 \cdot 10^1$	±	$3.8 \cdot 10^0$
ΔH_1	330MJ/mol	±	44MJ/mol
n	$1.13 \cdot 10^0$	±	$1.3 \cdot 10^{-2}$
ΔG_1	-7.219kJ/mol		
ΔS_1	1.093MJ/(K*mol)		

Statistics

Degrees of freedom	17
ξ^2	$2.51 \cdot 10^{-5}$
Reduced ξ^2	$1.48 \cdot 10^{-6}$
Standard error of point	$1.22 \cdot 10^{-3}$



Input Data

SRDCITJTitration (3-1-17)

Volume in vessel	750µl
Number of datapoints	20
Concentration of [M] in vessel	30.6µM
Concentration of [L] in syringe	69.4mM

No	Volume (µl)	Q_{exp} (mJ)	Q_{calc} (mJ)	Residual (mJ)
1	10.0	-47.71	-48.64	0.93
2	10.0	-58.92	-55.77	-3.15
3	10.0	-56.62	-57.30	0.68
4	10.0	-56.32	-57.44	1.12
5	10.0	-56.15	-56.94	0.79
6	10.0	-55.47	-56.08	0.61
7	10.0	-54.37	-55.00	0.63
8	10.0	-53.31	-53.80	0.49
9	10.0	-51.97	-52.53	0.56
10	10.0	-51.02	-51.22	0.20
11	10.0	-50.83	-49.89	-0.94
12	10.0	-49.64	-48.57	-1.07
13	10.0	-48.54	-47.26	-1.27
14	10.0	-47.19	-45.98	-1.21
15	10.0	-45.86	-44.72	-1.14
16	10.0	-44.08	-43.49	-0.59
17	10.0	-42.13	-42.29	0.16
18	10.0	-40.34	-41.13	0.79
19	10.0	-38.82	-40.00	1.19
20	10.0	-37.57	-38.91	1.34

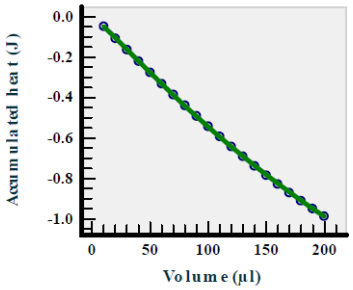


Figure A8: Ligand Binding Report (Ca and the GST-Switch domain protein).

APPENDIX II: GENERAL CHEMICALS AND ENZYMES

PRODUCT	COMPANY
Acetic Acid	Merck
Ammonium Persulphate (APS)	Biorad
Ampicillin sodium salt	Inqaba Biotech
Calcium Chloride	Merck Coomasie
Brilliant Blue	Sigma
Dithiothreitol (DTT)	Merck
Ethanol	Merck
Fastcast	Inqaba Biotech
Glucose	Merck
Glutathione Agarose	Sigma
Glycerol	Merck
Glycine	Merck
Imidazole	Sigma
IPTG	Inqaba
Lysozyme	Merck
Magnesium Chloride	Merck
TEMED	Promega
Tris	Merck
Tryptone powder	Merck