Rehana Masood

Production, characterization and structural analysis of proteins from *Corynebacterium pseudotuberculosis* and snake venoms

SÃO JOSÉ DO RIO PRETO
2015

Rehana Masood

Production, characterization and structural analysis of proteins from *Corynebacterium pseudotuberculosis* and snake venoms

Tese apresentada para obtenção do título de Doutor em Biofísica Molecular, área de concentração Biofísica Molecular, junto ao programa de Pós-Graduação em Biofísica Molecular do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Campus de São José do Rio Preto, São Paulo, Brasil.

Supervisor: Prof. Dr. Raghuvir Krishnaswamy Arni

SÃO JOSÉ DO RIO PRETO
2015

REHANA MASOOD

Production, characterization and structural analysis of proteins from *Corynebacterium pseudotuberculosis* and snake venoms

Tese apresentada como parte dos requisitos para obtenção do título de Doutor em Biofísica Molecular, junto ao Programa de Pós-Graduação em Biofísica Molecular do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Campus de São José do Rio Preto.

COMISSÃO EXAMINADORA

Prof. Dr. Raghuvir Krishnaswamy Arni
Dep. Física – UNESP/São José do Rio Preto
Orientador

Prof. Dr. Patrick Jack Spencer
IPEN/CNEN/ São Paulo

Prof. Dr. Marcos Roberto de Mattos Fontes
Dep. Física e Biofísica/ UNESP/ Botucatu

Prof. Dr. João Ruggiero Neto
Dep. Física – UNESP/São José do Rio Preto

Prof. Dr. Gustavo Orlando Bonilla Rodriguez
Dep. Física – UNESP/São José do Rio Preto

São José do Rio Preto
19 de Outubro de 2015

Dedicated to my fantastic four

Ajmal (late) (brother), Noor Zaman (brother), Anwar (husband) and Imaad (son)

Acknowledgments

First and foremost, praise and thanks to Allah (God), The Almighty and merciful, for giving me life and all that blessings that many don't have.

There are many people, who contributed in my educational career directly or indirectly, without their support and contribution this journey was without its destination.

I would like to thank Professor Dr. Marinonio Lopes Cornelio for inviting me and made me a part of this prestigious institution UNESP and department of Physics.

I would like to thank my supervisor Dr. Raghuvir Krishnaswamy Arni that gave me a place in his group and provided all the facilities that were required for the completion of my research work.

I would like to convey my thanks to all the faculty and administration members of Physics department for their help and support.

Special thanks to my lab fellows and friends, Ricardo, Vinicius, Hevila, Livia, Carolina, Raphael, Andre and Fernanda.

Special thanks to my class fellows during my course work, Marcelo, Icaro, Josimar, Daniel, Antonio, Karina, Natalia, Taisa, Gabriella, Thiago, Pedro.

I have no words to thank my best and nearest friends, they supported me at every step during my studies, research and also in personal life, Aziz Ullah, Sadia Manzoor, Miriam Guerra, Luciane, Eliana, Peterson, Diana, Harvey, Anna Lopes and her family.

I have no words to express my deep sense of gratitude to my partner, colleague and husband Dr. Anwar Ullah, who have gave me endless love, care and support. Many times, I have tried to step back but that was his encouragement and confidence on me that I have made it through all the steps to reach this point in life. My son Imaad Ullah, instead of thanks I would like to say sorry to him that due to my studies and research work, I could not manage to give that sufficient time that he deserves. His smiles and hugs make me more focused and strong towards hardships of life.

The last but not the least, to thank my parents and siblings for their love care and prayers. Special thanks to my brothers, Muhammad Ajmal (late) and Dr. Noor Zaman who fought for my education against the system and family. They raised me loved me and supported me to the extent that I can proudly say that I will be the first Ph.D. in my family.

At the end, I would like to acknowledge **CAPES-Toxinology** for the financial support throughout my Ph.D. work.

عقل مه غواړه بخت د غواړه اے رحمانه

عقلمند د بختورو غلامان يي

"O Lord don't make me among the Talented Ones but make me among the Blessed Ones
because I have seen Many Talented Ones Serving the Blessed Ones"

Rahman Baba

RESUMO

Corynebacterium pseudotuberculosis (*C. pseudotuberculosis*) é o agente etiológico da linfadenite caseosa (CLA), uma doença que afeta diferentes grupos de animais em todo o mundo, especificamente áreas de produção de ovinos e caprinos, que causam perdas econômicas significativas. Estas bactérias podem infectar humanos, e 25 casos de infecções em humanos foram relatados na literatura. Atualmente, não há nenhum tratamento eficiente para esta doença e também não existem estruturas cristalográficas de proteínas de *C. pseudotuberculosis* que possam ajudar a compreender melhor o mecanismo de ação desse patógeno. No presente estudo, foram expressas e purificadas proteínas-chave que desempenham papéis importantes no metabolismo deste patógeno. Três proteínas de *C. pseudotuberculosis* foram expressas e purificadas. Triose fosfato isomerase (TIM) é uma enzima ativa da via glicolítica, sendo uma das principais enzimas envolvidas no fornecimento de energia para o organismo. Esta enzima foi expressa purificada e cristalizada. Tiorredoxina redutase (TrxR) e Tioredoxina (Trx) participam do sistema tioredoxina (redox), em que o Trx atua como um substrato para o TrxR. Estas enzimas têm diversas funções nos organismos. Estas enzimas foram expressas, purificadas e caracterizadas. A estrutura cristalográfica da TIM foi determinada a 2,5 Å. Proteínas de veneno de serpente, também foram investigados neste estudo, e cinco proteínas diferentes do gênero Bothrops foram purificados e cristalizadas. A estrutura cristalina da Atroxlysin-I foi determinada a 1,8 Å de resolução e foi comparada com diferentes metaloproteinases-I que foram depositados no PDB.

Palavras-Chave: *Corynebacterium pseudotuberculosis*. Triosefosfato Isomerase. Tiorredoxina redutase. Tioredoxina. Expressão. Purificação. Caracterização. Cristalização. Venenos de serpente. *Bothrops moojeni*. *Bothrops atrox*. Metaloproteinases.

ABSTRACT

Corynebacterium pseudotuberculosis (*C. pseudotuberculosis*) is the etiological agent of Caseous Lymphadenitis (CLA), a disease that affects different groups of animals worldwide, specifically sheep and goat production areas, which result a significant economic losses. These bacteria even infect humans; to date 25 different cases of infections in humans are reported in the literature. Currently, no efficient treatment for this disease is available neither a crystallographic structure of any protein from *C. pseudotuberculosis* that could assist in better understanding of the mechanism of action of this pathogen. In the present study, we have expressed and purified key proteins that play a major role in the metabolism of this pathogen. Plasmids of three different proteins from *C. pseudotuberculosis* were designed. Triose phosphate isomerase (TIM) is an active enzyme of glycolytic pathway, which is one of the main energy supplier enzymes for the organisms. This enzyme was expressed, purified and crystallized, the crystal structure was determined at 2.5 Å. Thioredoxin reductase (TrxR), and Thioredoxin (Trx) are the part of thioredoxin (redox) system, in which the thioredoxin act as a substrate for the TrxR. These enzymes have a diverse function in organisms. These enzymes were expressed, purified and characterized.

Snake venom proteins were also investigated in this study; five different proteins from genus Bothrops were purified and crystallized. The crystal structure of Atroxlysin-I was determined at 1.8 Å resolutions and compared with different metalloproteinases-I, deposited to PDB.

Keywords: *Corynebacterium pseudotuberculosis*, Triose phosphate isomerase, Thioredoxin reductase, Thioredoxin, Expression, Purification, Characterization, Crystallization, Snake venoms, *Bothrops moojeni, Bothrops atrox*, Metalloproteinases.

CONTENTS

List of Figures

List of Tables

List of Abbreviations

TIM	Triose phosphate isomerase
CpTIM	*Corynebacterium pseudotuberculosis* Triose phosphate isomerase
TrxR	Thioredoxin Reductase
Trx	Thioredoxin
FAD	Flavin Adenine Dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
EC	Enzyme Commission
SVMP	Snake Venom Metalloproteinase
BmMP-III	*Bothrops moojeni* metalloproteinase-III
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Electrophoresis
HCl	Hydrochloric acid
NaCl	Sodium Chloride
NaH$_2$PO$_4$	Sodium Dihydrogen Phosphate
CH$_3$COONa	Sodium Acetate
Na$_2$CO$_3$	Sodium Carbonate
PDB	Protein Data Bank
PEG	Polyethylene Glycol
Å	Angstrom
K$_a$	Acid dissociation constant
kDa	kilo Dalton
K	Kelvin
M	Molar
MW	Molecular weight
pI	Isoelectric Point

Chapter 1

Proteins from *Corynebacterium pseudotuberculosis*

Corynebacterium spp (https://www.flickr.com/photos/ajc.com)

1) Introduction

A French bacteriologist Edward Nocard isolated for the first time an unusual organism from a case of lymphangitis in a cow in 1898. Three years later, a Bulgarian bacteriologist Hugo von Preisz identified a similar bacterium in cultures from a renal abscess in a ewe. This similar organism becomes known as "Perisz Norcardo" bacillus for decades (Baird and Fontaine, 2007). By the end of the 19th century, German bacteriologists Lehmann and Neumann described the bacterium in the first edition of their bacteriological atlas and renamed the bacterium as *Bacillus tuberculosis* derived from Greek word as *pseudotuberculosis* or "false tuberculosis" because of the clinical similarity of the lesions to caseous nodules of *Mycobacterium tuberculosis*. Later in 1923, the organism was placed in the Corynebacterium genus, which was created as a category for the human pathogen *Corynebacterium diphteriae* (Baird and Fontaine, 2007).

The Corynebacterium genus is located in a supergeneric group of actinomycetes, which also contain the genera Mycobacterium, Rhodococcus and Nocardio. This group has considerable biological, medical and veterinary importance. *Corynebacterium pseudotuberculosis* (*C. pseudotuberculosis*) is a Gram positive, motile, non-spore, facultative and intracellular pathogen, which has dimensions ranging from 0.5 to 0.6 μm in diameter and from 1.0 to 3 μm in length. It may be present individually, in pairs or palisade and assume coccoid, bacillary or filamentous forms (Sood *et al.*, 2012; Osman *et al.*, 2014).

C. pseudotuberculosis is a cosmopolitan microorganism found predominantly in soil, skin or mucous membranes of animals (Pugh and Bair, 2004). In conditions where, these microorganisms are not directly exposed to sunlight, can survive for long periods (Radostits *et al.*, 2007) and in purulent secretions, for 6 to 12 months (Kennelly, 2002). This disease is prevalent in several countries, such as Australia, New Zealand, South Africa, USA, Canada, England, France, Italy, Argentina, Chile, Uruguay and Brazil (Alves and Olander, 1998; Eggleton *et al.*, 1991; Walburgur *et al.*, 2004). In Brazil, it is estimated that most flocks are infected and that clinical prevalence can reach 30% of the hosts. *C. pseudotuberculosis* is the etiological agent of Caseous lymphadenitis (CLA) a disease that affects different groups of animals worldwide, specifically sheep and goat production areas, which cause significant economic losses. *C. pseudotuberculosis* has two main biotypes, the type called equi infects horses, while the type, which affects small ruminants named as ovis (Peel *et al.*, 1997). This bacteria even infect human, to date 25 different cases of infections in human are reported in the literature (Habuš *et al.*, 2015).

CLA is a chronic infectious disease, which is responsible for the reduction of wool production, meat and milk, a reduction of reproductive performance of animals and condemnation of carcasses in slaughterhouses and skins. The disease is evidenced by the presence of necrosis in the lymph glands, appearing in two forms such as external and internal surface CLA or visceral (Salim, 2001). The external CLA has abscess formation in superficial lymph nodes, especially parotid, submandibular, popliteal, pre-crural, supramammary and prescapular and subcutaneous tissues. (Fig 1).

Figure 1: CLA abscess formation; **(A)** Submandibular region; **(B)** Prescapular region; **(C)** Prefermoral region; **(D)** Supramammary region (http://www.goatworld.com/articles/cl/cl.shtmL)
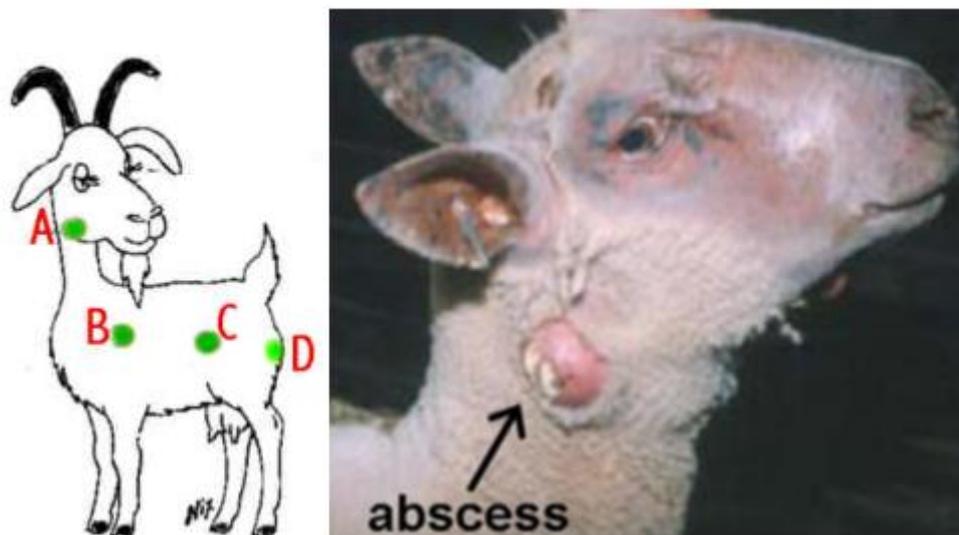
This type of CLA is the most reported or frequent clinical form. In visceral CLA, abscesses develop in bronchial and mediastinal lymph nodes reported the presence of abscesses in internal organs such, as lungs, kidneys, liver and spleen (Merchant and Packer 1967). *C. pseudotuberculosis* contains lipids in bacterial cell wall, similar to mycolic acids present in the Nocardia, Rhodococcus and Mycobacterium (Baird and Fontaine, 2007) that complicates the process of phagocytosis by preventing the enzymatic hydrolysis of lysosomes and enhances the cytotoxic effects in the host. *C. pseudotuberculosis* has different susceptibility pattern towards antimicrobial agents, depending upon the source of their isolation (Connor *et al.*, 2000; Foley *et al.*, 2004; Songer *et al.*, 1988). Treatment of animals with systemic antibiotics is generally ineffective; the majority of cases require surgical excision of the affected node (Baird and Fontaine, 2007), that is not feasible commercially.

Although the bacterium is sensitive to various types of antibiotics *in vitro*, they are not able to penetrate the capsule of abscesses, which makes therapy inefficient (Olson *et al.*, 2002). Olson et al. (2002) cultivated *C. pseudotuberculosis* as a biofilm, to reproduce a naturally infected environment, where they observed the resistance pattern of this bacterium and considered it as highly resistant to all drugs tested.

Generally, the removal of affected lymph node or internal organs do not guarantee 100% elimination of these pathogens, and this technique may result in environmental contamination, where a single bacterium from an abscess may contaminate the whole herd. Therefore, the best strategy to combat the disease is prophylaxis (Piontkowaski and Shivvers, 1998). The identification of infected animals and their removal from the herd are the most efficient methods for the control of CLA in small ruminants (Binns *et al.*, 2002). Thus, CLA control should be based on the prevention of invasion and spreading of the microorganisms in the herds. Currently the best strategy is to vaccinate the healthy animals, along with the identification and elimination of infected animals, but it still needs the development of effective vaccines and necessary diagnostic tests, since there is currently no ideal immune prophylaxis to combat this disease (Hard, 1969).

The secretion of effector molecules is one of the major mechanisms by which intracellular pathogens interact with their host cells. Thus, secreted proteins from pathogenic bacteria represent a large fraction of the molecular determinants of their virulence (Hansmeier *et al.*, 2006; Jankovic *et al* 2007; Stavrinides *et al* 2008; zarraga *et al.*, 2006). The secretome comprises a wide range of proteins, which mediate interactions with the environment, such as receptors, adhesions, carriers, enzymes, toxins and virulence factors. The bacteria colonizing mammalian cells secrete proteins able to mediate adhesion to their host, destruction of the host tissue or interference with the immune system. Conserved secreted proteins that are able to induce protective immune responses are potent targets to produce a potential vaccine. Bacterial secretome study is essential for a variety of applications, from the identification of useful enzymes to the understanding of virulence (Hodgson *et al.*, 1990). The main methods for determining the three-dimensional structures of proteins are X-ray crystallography, Nuclear magnetic resonance spectroscopy (NMR), and computational methods. Currently no crystallographic structure of a protein is determined from *C. pseudotuberculosis*

A number of proteins from *C. pseudotuberculosis* are involved in the mechanisms of various diseases caused by this pathogen. Their study is of great importance for their cure and brings information for drug discovery. This study describes purification, characterization and

structural determination of some of the key proteins involved in the metabolism of *C. pseudotuberculosis*.

1.1) Triose phosphate isomerase/ Triose isomerase monophosphate

Triose phosphate isomerase (TIM) (EC 5.3.1.1) is an enzyme of the glycolytic pathway that take part in the isomerization or interconversion of glyceraldehyde-3 phosphate (GAP) and dihydroxyacetone phosphate (DHAP). The major function of this enzyme is energy production. In glycolysis, TIM provides these two products for the formation of pyruvate, while in glycogenesis both substrates are supplied to aldolase. TIM is present in almost every organism, which performs glycolysis such as humans, plants, insects, fungi and bacteria. This enzyme is called as a "catalytically perfect enzyme" as its performance in a reaction is about 10^8 times faster than it occurs naturally in a solution. The reaction takes place without any cofactor or ion.

TIM is an essential enzyme of the glycolytic pathway that is responsible for the central carbon metabolism, which is required for glycolysis and gluconeogenesis. In glycolytic pathway, it is considered as a "near-perfect enzyme" as the rate of the reaction *in vitro* depends only on the rate of diffusion of its substrates and *in vivo* to prevent the accumulation of DHAP the enzyme interconverts the dihydroxyacetone phosphate (DHAP) and glycerlaldehyde-3-phosphate (GAP) (Maria Gruning *et al.*, 2014). The TIM mechanism illustrates an example of acid-base catalysis, in which the Glutamate acts like a base (B⁻) and the Histidine acts like an acid (HA) that is involved in the cleavage of a C-H bond. During this process, Glutamate extracts the hydrogen ion or the proton from the C-2 position of the substrate and the Histidine imidazole group donates a proton to C-1 position. In general, the Glu-COOH removes the proton from C-2 position and His accepts the proton from C-1 position (Trujillo *et al.*, 2014).

Glyceraldehyde-3-phosphate   *cis*-enediol   Dihydroxyacetone phosphate
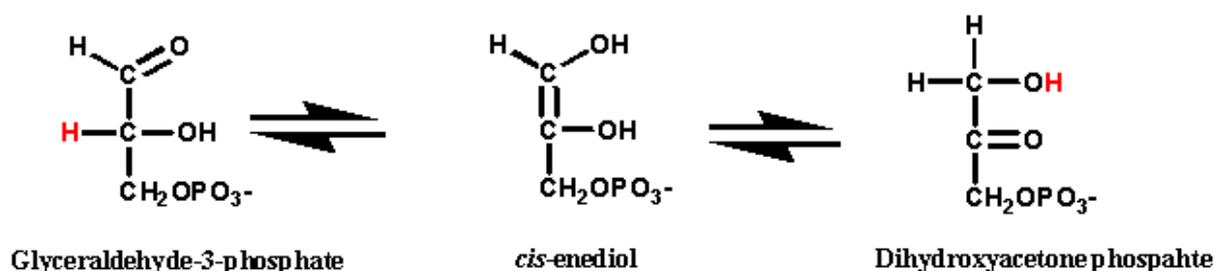
TIM is a homodimer and a well-studied enzyme, which represents a prototypical and extensively occurring TIM barrel structure. The typical structure of each monomer of TIM has eight parallel β strands/ sheets that are wrapped on the outside by eight α-helices. Such pattern is called as $(α/β)_8$ fold which is the most common fold of enzymes in nature and the active site of each subunit is located at the C-terminal side. The typical active site of a TIM consists of four main residues such as, Asparagine (Asn11), Lysine (Lys13), Histidine (His95) and Glutamate (Glu167) that take an active part in the process of catalysis (Wierenga *et al.*, 2010). These enzymes have a high structural similarity and are responsible for the energy production as well as the growth of organisms (Trujillo *et al.*, 2014).

TIM is considered as a multifunctional protein. Its deficiency in human is the only glycolytic deficiency that is lethal and its symptoms in early childhood, such as hemolytic anemia, neurological disorders. Its activity is very important for the proper functioning of a cell, as the glycolytic pathway is interconnected with the metabolism of pentose phosphate, lipid and also with gluconeogenesis. All these pathways disordered with the deficiency of TIM that leads to the symptoms of serious diseases (Hewitson *et al.*, 2014).

1.2) Thioredoxin reductase

Thioredoxin reductases (TrxRs) (EC 1.8.1.9) are dimeric flavoproteins that belong to a family of pyridine nucleotide-disulphide oxidoreductases that have some other members such as lipoamide dehydrogenase, glutathione reductase, and mercuric ion reductase. TrxRs have a distinct yellow color due to the presence of FAD (flavonoid adenine dinucleotide) a prosthetic group. All the proteins of this family are homodimers; each subunit has an FAD domain, an NADPH binding site, and an active site that has a redox-active disulphide (Mustacich and Powis, 2000). These enzymes are divided into two types according to their molecular weight and structure. The thioredoxin reductases with higher molecular weight (55 kDa) have an additional C-terminal domain (interface domain) which take part in dimerization as well as in catalysis, are known as higher molecular weight type and identified as H-TrxRs, they have been isolated from higher eukaryotes. While the TrxRs isolated from prokaryotes, archaea, and lower eukaryotes have low molecular weight such as 35 kDa and lack a separate interface domain, and identified as L-TrxRs (Williams *et al.*, 2000). TrxRs are responsible for the catalysis of the NADPH-dependent reduction of thioredoxin and some other endogenous and exogenous substrates such as lipoic acid, lipid hydro peroxidases, vitamin K_3, dehydroascorbic acid, the cytotoxic peptide NK-lysin, the ascorbyl free radical

and the tumor- suppressor protein p[53] (Mustacich and Powis, 2000) . The TrxRs reduces the oxidized thiol redoxins, which are small proteins (10-12 kDa). These small proteins have a highly conserved catalytic site (-Trp-Cys-Gly-Pro-Cys-Lys) in which, the two cysteine residues are responsible for the reversible oxidation/reduction process. The active site of TrxR takes electrons from NADPH through FAD and then reduces the oxidized thioredoxin or other substrates (Williams, 1995a). The overall reaction of the system is as follows,

$$\textbf{NADPH} + \textbf{H}^+ + \textbf{Tr(S}_2\textbf{)} \;\rightleftharpoons\; \textbf{NADP}^+ + \textbf{Tr(SH)}_2$$
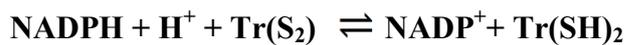
Where, $Tr(S_2)$ is the oxidized thioredoxin and $Tr(SH)_2$ is the reduced thioredoxin (Williams, 1995b).

The thioredoxin system (thioredoxin reductase, thioredoxin, and NADPH) is well known for its cellular defense property against reactive oxygen species (ROC) and electrophilic species, which are the usual products of metabolism. These enzymes have specific function in each organism such as the thioredoxin and thioredoxin reductase system play a key role against oxidative stress in Archaea, bacteria, and eukaryote, while, in bacteria, yeast and mammals it also take part in regulating DNA synthesis, gene transcription, cell growth and apoptosis (Hirt *et al.*, 2002).

The overall functions of TrxRs are shown in a flow scheme (Fig 2). The TrxR utilizes the NADPH for the conversion of oxidized Trx into reduced form and recycles the ascorbate from its reduced form. In the same process, the reduced Trx functioned further as, it provides equivalents to Trx peroxidase that break down into water molecules, it also reduce ribonucleotides to deoxyribonucleotide for DNA synthesis. It provides transcription factors for gene transcription. Besides these functions, it also increases the cell growth and inhibits apoptosis.

Figure 2: Schematic of the reaction and function of TrxR inside the cell (Mustacich and Powis, 2000)
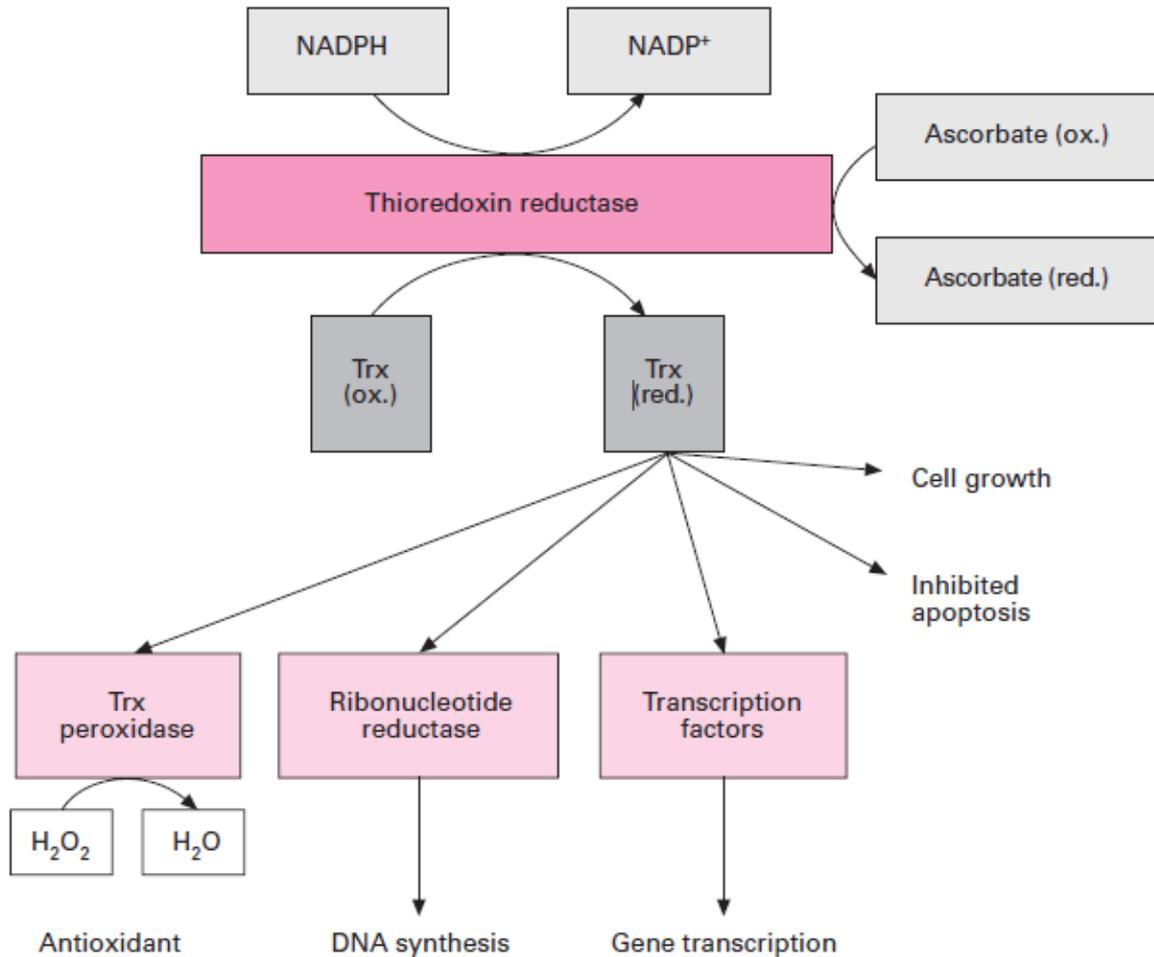
The study and occurrence of TrxRs in different groups of organisms are of utmost importance in the cells. The property of TrxRs that they reduce substrates other than Trx lead to novel biological functions of TrxRs. The parasites such as aerotolerant, anaerobic parasitic protozoa have L-TrxR, which have different structures and mechanisms as their host H-TrxR may represent the best targets for chemotherapeutic attack. The inhibition of TrxR through different inhibitors may lead to new discoveries for the treatments of human diseases such as cancer, AIDS and other autoimmune diseases (Mustacich and Powis, 2000). Some of the H-TrxRs such as from cancer cells were inhibited by aurofins and aurothioglucose, which are considered as efficient inhibitors for human enzyme. Currently, the H-TrxRs of *Plasmodium falciparum* is a major focus of the researchers for the discovery of selective inhibitors that inhibit the plasmodium enzymes (Hirt *et al.*, 2002). As *Plasmodium falciparum* is the major

cause of malaria, inhibition of plasmodium enzymes will likely form the next step towards the discovery of antimalarial drugs.

1.3) Thioredoxin

Thioredoxin (Trx) is a small oxidoreductase enzyme from a class of redox proteins that are present in all living organisms. The molecular weight of this protein is approximately 12 kDa. It has been purified and characterized from both prokaryotes and eukaryotes. Trx was recognized for the first time in *Escherichia coli* as a hydrogen donor for ribonucleotide reductase and deoxyribonucleotide synthesis and repair (Vizuete *et al.*, 1997). Trx is well known for its function in thioredoxin system (thioredoxin, thioredoxin reductase and NADPH as a proton donor). In this system, the oxidized thioredoxin is reduced by the flavoprotein thioredoxin reductase (TrxR), which takes proton from NADPH through FAD. The reduced thioredoxin as an end product of the reaction of the thioredoxin system has two thiol groups that serves for the catalysis of other exposed disulfides and reduce them from oxidized forms. The overall process of reduction of Trx and its further reaction is shown in two-step equations below:

Reaction 1: $Trx\text{-}S_2 + NADPH + H^+ \rightleftharpoons Trx\text{-}(SH)_2 + NADP^+$ (catalyst: TrxR)

Reaction 2: $Trx\text{-}(SH)_2 + protein\text{-}S_2 \rightleftharpoons Trx\text{-}S_2 + protein(SH)_2$
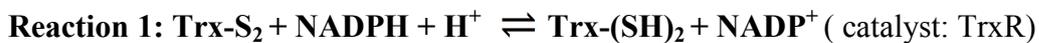
These reactions are reversible. The second reaction is very important for many biological functions. For example, the Trx in *E. coli*, besides its function in DNA synthesis and repair, also known as a hydrogen donor for the enzymes involved in the sulfate assimilation pathway. An essential component for the life cycle of bacteriophages, such as T7 and M13 and f. Trxs act as donors of Hydrogen or proton in eukaryotic cells, and take part in refolding of disulfide-containing proteins. Taking part in initiation or activation of interleukin-2 receptors, it also contributes in regulating the DNA binding activity of some transcription factors for example NF-kB (Vizuete *et al.*, 1997). Trxs also stimulate the rapid multiplication of lymphoid cells (Wollman *et al.*, 1988) and also the rapid growth of solid tumor cells, this function of Trx that taking part in cell multiplication and cell growth leads to the evidence of its major role (as a part of redox system) in the development of multicellular organisms (Foyer and Noctor, 2005).

Plants have six different types of Trxs, named as *f, m, x, y, h* and *o* that take part in an array of processes, such as taking part in photosynthesis, respond to fungal infections in plants etc. The specific function of some of these Trxs (*f, m* in chloroplasts and *h* in cytosol and endoplasmic reticulum) in photosynthetic organisms is their involvement in the regulatory systems in oxygen photosynthesis (Menga *et al.,* 2010).

Trxs fight oxidative stress, in which they act as antioxidants to reduce hydrogen peroxides, convert it into water molecules, and remove free radicals from the cells (Saitoh *et al.*, 1998).

2) Materials and methods

2.1) Materials

2.1.1) Chemicals

All the chemicals used were of analytical grade and were purchased from Sigma, Merck, GE Healthcare Life Science, Hampton research and Dinamica.

2.1.2) Plasmid design

The plasmids (pD441) were designed in our laboratory and contained T5 promotor, Kanamycin resistance and a gene of the desired protein and were synthesized by DNA 2.0 Company. The genes were designed according to UniProt code D9Q8J4 and XtalPred predictor (http://services.mbi.ucla.edu/SER/) was used to check for the existence of the signal peptide. The sequence nucleotides encode the target gene followed by a TEV protease cleavage site and a C-terminal His tag as shown in the plasmid map. ExPASy (Expert Protein Analysis System, http://ca.expasy.org) ProtParam tool, analyzed the complete amino acid sequence encoded by this construct.

Figure 3: Plasmid map of TIM. The TEV protease cleavage site facilitates the His-tag removal

2.2) Methods

2.2.1) Triose phosphate isomerase (TIM)

2.2.1.1) Transformation of plasmid

The transformation of the recombinant plasmid was done by following the heat shock method. A DNA plasmid (50 ng) was mixed in a 50 µL of different competent strains of *Escherichia coli* (*E. coli*) cells in a micro centrifuge tube. The mixture was placed on ice for 30 min. For heat shock, the water bath was adjusted at 42 °C. The micro centrifuge tube

containing the mixture of DNA plasmid and competent bacterial cells were heated at 42 °C for 45 seconds or "heat shocked". After heat shock the tube containing the mixture was placed on ice for 2 min. SOC medium, (LB medium with 0.02 M Magnesium Chloride) was prepared. The mixture of competent bacterial cells and DNA plasmid was mixed in a 500 μL SOC medium. The cells were incubated for 1 hour at 37 °C in an agitating incubator. Agar Petri dishes were prepared by using a selective solid culture medium (LB + kanamycin 50 μg/mL). The cells were plated in the agar Petri dishes and incubated at 37 °C overnight for colony screening. A single bacterial colony was selected for the preparation of the pre-inoculum.

2.2.1.2) Bacterial cell lines for expression

E.coli of different cell lines (table. 1) was used and tested a small scale expression. After analysis of the intensity of bands in SDS-PAGE, strain BL21 Star™ (DE3) (Novagen) was selected for the expression process of this protein at large scale.

Table 1: *E. coli* cell lines tested

E. coli cell lines
BL21 Star^{TM} (DE3)
BL21-CodonPlus (DE3) *RIL*
C41 (DE3)
C43 (DE3)
BL21 (DE3) *pLysS*

2.2.1.3) Test for the expression of TIM at different conditions

For optimum expression process, different tests at small scale (5 mL of LB culture) were performed. Different temperatures (20, 25, 30 and 37 °C) were tested and after sonication, the soluble fractions were analyzed on SDS-PAGE and stained with Comassie brilliant blue. A temperature of 30 °C was selected for the expression of this protein at large scale.

2.2.1.4) Expression of TIM at large scale

For a large scale protein production, 1 liter LB medium with appropriate antibiotics (kanamycin 50 µg/mL) and pH 7.7 was inoculated with 20 mL of overnight pre-inoculum/ preculture, in 3 liters flask (no more than 1 liter per flask). The cells grew at 30 °C by continuous stirring at 120 rpm in shaking incubator until OD_{600} reached 0.5-0.8. At this point, 0.3 mM IPTG was added for induction of expression. The cells were harvested after 4 hours, centrifuged at 5000 × g for 15 min at 10 °C.

2.2.1.5) Cell lysis

The pellet was suspended in a 20 mL lysis buffer (0.02 M Tris-HCl, 0.5 M NaCl, 0.005 M Imidazole and 5% glycerol pH 7.7). After incubation on ice up to 20 minutes, the cells were sonicated (ultrasonic processor Marconi - MA 103) at a pulse rate of 30 seconds at 300 W as many times as necessary to disrupt the cells completely (3 to 4 times). After sonication, the cell suspension was centrifuged at 17000 × g for 45 min at 10 °C. The supernatant was stored for further analysis.

2.2.1.6) Protein purification

The protein was purified through three different chromatographic steps, explained as follows.

2.2.1.6.1) Ni-NTA affinity chromatography

Immobilized metal affinity chromatography (IMC) is specific for the purification of proteins by binding it with the metal; Ni-sepharose resin (GE) contains nickel, which has the ability to bind specifically with the stretches of polyhistadines in the proteins. Many expression systems have tags 6xHis either on N-terminal or C-terminal end or on both. One mL of resin is sufficient to bind 20 mg of tagged protein. The supernatant of the cell lysis process was filtered on a syringe filter (0.22 µm). Two mL of the Ni-sepharose resin (Invitrogen) was loaded on the BioRad disposable plastic column and was equilibrated with the lysis buffer (0.02 M Tris-HCl, 0.5 M NaCl, 0.005 M Imidazole, 5% glycerol pH7.7). Once equilibrated, the filtered sample was loaded on the column. After the loading, the column was washed with 10 to 20 mL loaded buffer. Further washing was done by changing the imidazole concentration in the loading buffer such as 0.04 M imidazole, 0.06 M imidazole and the final elution buffer having 0.4 M imidazole. Elution was collected drop-wise.

2.2.1.6.2) Ion exchange chromatography

For the second step of the protein purification, anion exchange chromatography was used. A Mono Q 5/50 GL column and an empty loop of 50 mL was used. The solvent of the elution sample from Ni-NTA column was exchanged to buffer A (0.02 M Tris-HCl. pH 7.7) by using desalting column. The Mono Q column was equilibrated with buffer A and the sample was loaded. The elution was achieved by a stepwise gradient from zero to 100% of buffer B (0.02 M Tris-HCl, 1M NaCl pH 7.7). The elution fractions containing TIM were pooled.

2.2.1.6.3) Size exclusion/ Gel filtration chromatography

Gel filtration chromatography is based on the sieving properties of porous, inert matrices through which the macromolecules move. The matrices are characterized in such a way that the small molecules are delayed in the column. The Superdex 75 GL column was equilibrated with a buffer (0.02 M Tris-HCl, 0.5 M NaCl, pH 7.7).The pooled fraction that was obtained from the anion exchange chromatographic step, was concentrated to 0.5 mL in a micro-concentrator (Falcon at 30 kDa cutoff). The sample (0.5 mL) was injected into super loop with a flow rate of 0.5 mL/min. The eluted fractions containing that have the protein were collected for further analysis.

2.2.1.7) Electrophoresis SDS-PAGE

All the samples collected during the expression and purification steps of the protein were analyzed by electrophoresis in the presence of SDS and the denaturing agent β-mercaptoethanol following the methodology of Laemmli (1970). The samples were prepared by mixing 10 µL of the protein and 5µL sample buffer (0.125 M Tris-HCl, pH 6.8, containing 4% SDS, 20% glycerol, 10% β- mercaptaethanol and 0.001 M bromophenol blue). The samples were heated for 5 min at 100 °C and centrifuged at 1000 × g for 5 min. Ten microliters of the supernatant of each sample was applied in the gel. The electrophoresis was carried out by using the Hoefer mini (VE Amersham Biosciences GE). The gels were stained in a mixture of methanol, water and acetic acid (50:40:10) and 0.25% Comassie brilliant blue R-250, for 1 hour and destained with methanol and water (1:1).

2.2.1.8) Concentration of the purified protein

After analysis of the protein through SDS-PAGE and confirmed its purity, all the purified fractions that containing the desired protein were pooled and concentrated in the ultrafiltration concentrators AMICON (3 kDa) (Millipore) at $5000 \times g$ and at 10 °C.

2.2.1.9) Determination of protein concentration

The protein concentration was determined by using the spectrophotometer (Spectronic Genesys 2), at the wavelength 280 nm; the extinction molar coefficient value (ε) of the protein was obtained through a programme ProtParam (http://web.expasy.org). The concentration was also measured on a Nanodrop/NanoVaue (GE).

2.2.1.10) Circular dichroism

The circular dichroism experiments were carried out by using a Jasco J-815 spectropolarimeter (Jasco, USA) equipped with a Peltier-type temperature control system and a 0.1 cm path-length quartz cuvette that was coupled a 0.9 cm metal spacer block. The far UV-CD spectra of the protein were collected from 260 to 190 nm with a scan speed of 50 nm/min, response time of 1.0 second, spectral bandwidth of 1.0 nm and spectral resolution of 0.1 nm. The signal was averaged over 10 scans. CD spectra of TIM were acquired in pH 3, 7.4 and 10, and at the temperatures of 25 and 75 °C for each pH. The protein was diluted to 4.8 μM in sodium acetate (0.02 M CH_3COONa, 0.05 M NaF, and pH 3), sodium phosphate (0.02 M NaH_2PO_4, 0.05 M NaF, pH 7.4) and sodium carbonate buffer (0.02 M Na_2CO_3, 0.05 M NaF, pH 10). In the thermal unfolding experiments, the protein sample was heated from 25 to 75 °C at a scan rate of 1.0 °C/min with the ellipticity being monitored at 222 nm every 2.0 °C. The thermal denaturation of TIM was performed in pH 3, 7.4 and 10. The contribution of the buffer was subtracted from the protein spectra. Percentages of secondary structure of TIM in solution were calculated with the CONTINLL software of the CDPro package, using the reference set of proteins SMP56 (Sreerama and Woody, 2000). All the experiments were performed twice.

2.2.1.11) Crystallization

Crystallization is one of the key techniques used for the structural determination of macromolecules (Protein, nucleic acid, protein and protein complex, protein and nucleic acid complex, protein and ligand complex), which further leads to the better understanding of the

structure and function of the macromolecules in a living cell (Dessau and Modis, 2011). The major factors involved in a crystallization experiment are concentration of the protein, pH, temperature, buffer, precipitants, and sufficient homogeneity. These factors play a key role in the success of the crystallization process. The pH conditions of the experiment are very important, as different pHs can result in different packing orientations of the solute, for crystallization, a supersaturated solution is required therefore a solution should be highly concentrated (2-50 mg/mL) without aggregation or precipitation of the protein (Dessau and Modis, 2011). Precipitants, like ammonium sulfate and PEG (polyethylene glycol), are compounds that play a role in the protein precipitation out of a solution (Gale, 1993). There are different techniques used for protein crystallization such as Vapor-diffusion techniques, Batch-crystallization, Dialysis and Free-interface diffusion. The most popular technique now used for crystallization, is vapor-diffusion. It consists of two types, the hanging-drop vapor diffusion, in which siliconized cover slides are used for a drop of protein and precipitant (1:1), 24-well hanging-drop plates, and high vacuum grease for providing a close system for the crystallization process. The sitting-drop vapor diffusion method required 96-well plates and automated setups (robot) for the application of the sample for crystallization. The hanging-drop method is mainly used for manual setups; while the sitting-drop method is most prevalent method in automated crystallization plate setups. Both the methods are illustrated in a figure below.

Figure 4: Protein crystallization by vapor-diffusion method. **(A)** Hanging-drop vapor diffusion method; **(B)** Sitting-drop vapor diffusion method. (http://www.bio.davidson.edu/)
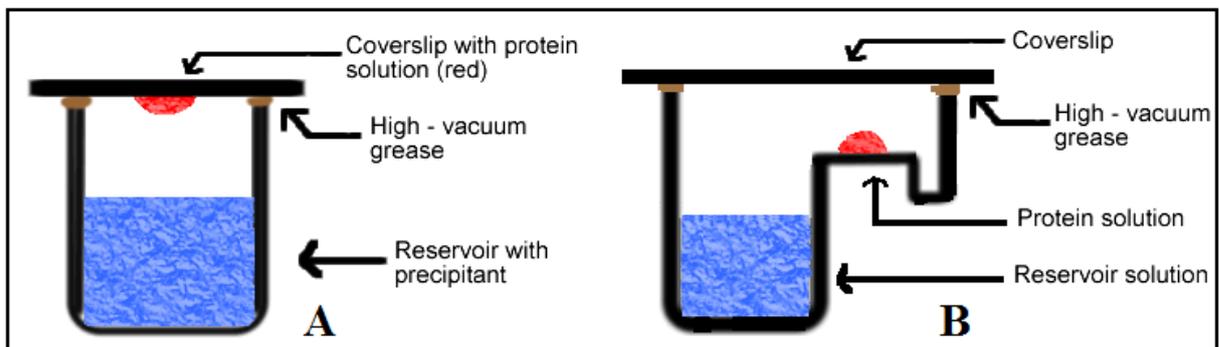
The crystallization process of proteins can be explained by crystallization (phase) diagram, which consists of four zones (Asherie, 2004). Initially a drop consists of one microliter of protein and one microliter of precipitant are sealed in a closed system. The vapor

diffuses from the two microliters drop into the reservoir and concentration of both protein and the precipitant increases, this leads to nucleation and growth of the crystals. During the growth of the crystals, the protein concentration will increases continuously and reaches to the growth zone where some of the crystals will grow further and finally the process of enlargement of crystals will terminates (Fig 5).

Figure 5: Phase diagram divided into four phases; under saturated, crystal growth, nucleation and precipitation zone. (http://www.dolomite-microfluidics.com)
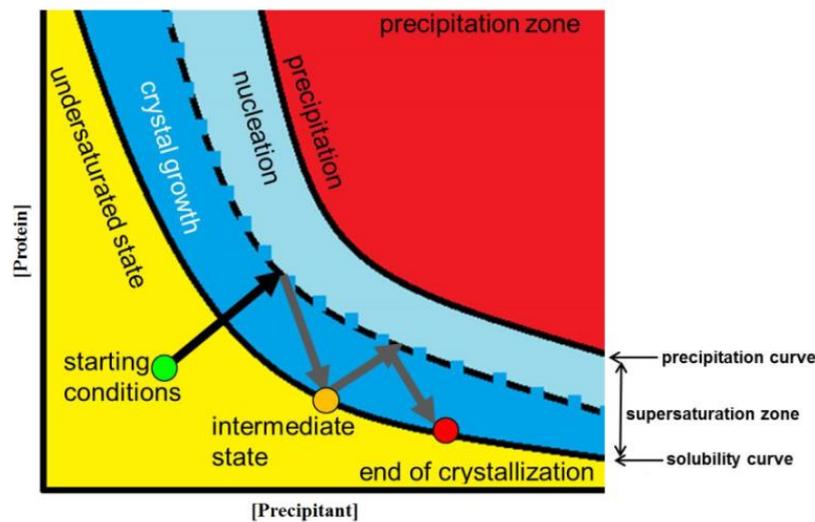
For crystallization of the protein, a manual hanging-drop vapor diffusion method was used. The purified TIM in a buffer 0.02 M Tris-HCl pH 7.7 at a concentration of 17 mg/mL was used. Crystallization experiments were performed by the hanging-drop vapor-diffusion method using 24-well tissue-culture plates and commercially available crystallization screens such as crystal screens 1 and 2, polyethylene glycol (PEG) 6000, ammonium sulfate kits Small crystals were obtained in the condition containing 0.2 M ammonium sulfate, 0.1 M Na-cacodylate and 20% *(w/v)* PEG 8000, pH 6.5. Small crystals were obtained in the condition containing 0.2 M ammonium sulfate, 0.1 M Na-cacodylate and 20% *(w/v)* PEG 8000, pH 6.5. This condition was optimized by variation in pH and concentration of PEG and crystals suitable for X-ray diffraction experiments were obtained with the reservoir contained 0.2 M ammonium sulfate, 0.1 M Na-cacodylate and 23% *(w/v)* PEG 8000, pH 6.0.

2.2.1.12) Data collection, processing and structure determination

For X-ray diffraction data collection, the TIM crystal was directly flash-cooled in a 100 K nitrogen-gas stream at the W01B-MX2 beamLine at the Brazilian Synchrotron Light Laboratory (LNLS-Campinas, Brazil). The wavelength of the radiation source was set to 1.4 Å and a MarCCD 165 mm CCD detector was used to record the X-ray diffraction intensities. The TIM crystal was exposed for 30 s per degree of rotation around φ, a total of 380 images were collected and the detector distance was set at 100 mm. Molecular replacement was carried out using the program MOLREP (Vagin and Teplyakov, 1997) and a model based on the atomic coordinates of 3AT6 (PDB). This model was chosen for its higher percentage of sequence identity.

2.2.2) Thioredoxin reductase (TrxR)

2.2.2.1) Plasmid design

Plasmid (pD441) was constructed and analyzed as described in item 2.1.2 Thioredoxin reductase gene was designed according to UniProt code D9QDF7.

Figure 6: Full plasmid map of TrxR. The TEV protease cleavage site allows the His tag removal

2.2.2.2) Transformation of plasmid

The plasmid containing the TrxR gene was transformed by following the "heat shock method" as described in the item number 2.2.1.1.

2.2.2.3) Bacterial cell lines for the expression of TrxR

E. coli of different cell lines (table 1) had been given prominent colonies and was tested in expression at small scales. After analysis of the intensity of bands in SDS-PAGE,

strain BL 21-star (Novagen) and C43 (DE3) were selected for the expression process of this protein at large scale. Both the strains were used parallel for the expression of TrxR.

2.2.2.4) Tests for the expression of TrxR at different conditions

Different tests of expression were performed as described in item number 2.2.1.3.

2.2.2.5) Expression of TrxR at large scale

Expression of the protein on a large scale was performed according to item number 2.2.1.4.

2.2.2.6) Cell lysis and protein purification

The cells were lysed by sonication with 30 s sonication pulses four times repeatedly on ice (ultrasonic processor Marconi - MA 103) and centrifuged at 17000 x g for 45 min. The supernatant containing recombinant protein was passed through further process of purification such as Ni-NTA affinity, Ion exchange and size exclusion chromatography as explained in item number 2.2.1.5 and 2.2.1.6.

2.2.2.7) Electrophoresis SDS-PAGE

The purity of the protein was evaluated by 15% SDS-PAGE gel as explained in item number 2.2.1.7.

2.2.2.8) Determination of protein concentration

The concentration of the protein was measured on UV-visible spectrometer at 280 nm using a formula A= $\varepsilon \times l \times [C]$. Where, A is absorbance, ε is the extinction coefficient, l is the path length in centimeters and C is the concentration of the protein.

2.2.2.9) UV-Vis absorbance spectroscopy

The UV-Vis experiments were performed on a BioMate 3S UV-Vis spectrophotometer (Thermo Scientific) equipped with a quartz cell of 10 mm path length. The UV-Vis spectra were obtained at room temperature. The concentration of TrxR was determined by measuring the absorption at 280 nm and using the extinction coefficient of 16.055 M^{-1} cm^{-1} (ExPASy-ProtParam) (Gasteiger *et al.*, 2005). The interaction studies between TrxR and metal salts ($AgNO_3$, Li_2SO_4, $ZnSO_4$, $NiSO_4.6H_2O$, $CuCl_2$, and $CsSO_4$) were recorded in the range of 310 − 550 nm with the increment of 1.0 nm. In these

experiments, the concentration of protein and each metal salt was equal to 27.3 μM. All measurements were performed in a phosphate buffer (0.02 M Na_2HPO_4/NaH_2PO_4, 0.1 M NaF, pH 7.4).

2.2.2.10) Circular dichroism

CD measurements were conducted in a Jasco J-701 spectropolarimeter (Jasco, USA), Using a cell of path-length 1 cm. A protein stock solution of 0.26 mM was prepared in 0.02 M sodium phosphate, 0.1 M sodium fluoride pH 7.7 as explained in item number 2.2.1.10.

2.2.2.11) Differential scanning calorimetry

Differential scanning calorimetry is based on the principle of measuring the heat input and output of the sample and heat of the sample relative to a reference. This technique enables the tendency of the stability of the sample at different temperatures. DSC for the protein (TrxR) was performed on N-DSC III (TA Instruments, USA) in the temperature range 20 °C to 90 °C and at a scan rate of 1 °C/min. The protein concentration for this experiment was 30 μM (0.9 mg/mL) and a buffer of 0.02 M sodium phosphate and 0.1 M sodium fluoride pH 7.7. The calorimeter cells (reference cell and sample cell) were loaded with the appropriate buffer and equilibrated at 20 °C for 10 min. The buffer scan was repeated three times to ensure a stable baseline. The protein sample was loaded in a clean and rinsed sample cell and scanned in the desired scan range. The heat capacity versus temperature scan of the protein was determined by subtracting the baseline or scan of 'buffer vs buffer' from the 'protein vs buffer' scan.

2.2.2.12) Crystallization

Pure TrxR of concentration 16 mg/mL was applied for crystallization as described in item number 2.2.1.11.

2.2.3) Thioredoxin (Trx)

2.2.3.1) Plasmid design

Plasmid (pD441) was constructed and analyzed as described in item number 2.1.2. Thioredoxin gene was designed according to UniProt code D9Q5F4.

Figure 7: Full plasmid map of Trx. The TEV protease cleavage site allows the His tag removal

2.2.3.2) Transformation of plasmid

The plasmid containing the Thioredoxin gene was transformed by following the " heat shock method" as described in the item number 2.2.1.1

2.2.3.3) Bacterial cell lines for the expression of Trx

Different cell lines were tested as described in item number 2.2.1.2 and table 1.

C43 (DE3) (Novagen) was selected for the expression process of this protein at large scale.

2.2.3.4) Tests for the expression of Trx at different conditions

Different tests of expression were performed as described in item number 2.2.1.3.

2.2.3.5) Expression of Trx at large scale

Expression of the protein on a large scale was performed according to item number 2.2.1.4.

2.2.3.6) Cell lysis and protein purification

The cells were lysed by sonication with 30 s sonication pulses four times repeatedly on ice (ultrasonic processor Marconi - MA 103) and centrifuged at 17000 x g for 45 min. The supernatant containing recombinant protein was passed through further process of purification as explained in item number 2.2.1.5 and 2.2.1.6.

2.2.3.7) Electrophoresis SDS-PAGE

The purity of the protein was evaluated by 15% SDS-PAGE as explained in item number 2.2.1.7.

2.2.3.8) Determination of protein concentration

The concentration of the protein was measured on UV-visible spectrometer at 280 nm using a formula A= $\varepsilon \times l \times [C]$. Where, A is absorbance, ε is the extinction coefficient (get through ExPASY/Protoparam) l is the path length in centimeters and C is the concentration of the protein. The protein concentration was also determined on nanodrop (GE).

2.2.3.9) Circular dichroism

CD measurements were conducted in a Jasco J-815 spectropolarimeter (Jasco, USA), Using a cell of path-length 1 cm. A protein stock solution of 0.346 mM was prepared in 0.02 M sodium phosphate, 0.05 M sodium fluoride pH 7.4 as explained in item number 2.2.1.10.

2.2.3.10) Crystallization

Pure Trx of concentration 17.33 mg/mL was applied for crystallization as described in item number 2.2.1.11.

3) Result and discussion

3.1) Triose phosphate isomerase

3.1.1) Sequence analysis of TIM through the ExPASy- ProtParam tool

ExPASy-ProtParam tool (Gasteiger *et al.*, 2005) has been used for the analysis of different physical and chemical parameters of Triose phosphate isomerase. It shows the total number of amino acids that are 274 and molecular weight 29.1 kDa. The theoretical isoelectric point is 5.37. The predicted instability index is 21.17, which classified the protein as stable. The aliphatic index is 95.11. At physiological pH, the protein is negatively charged. The total number of negatively charged residues (Asp + Glu) is 37 while, the total number of positively charged residues (Arg + Lys) is 25. The dominant residues in the sequence of this protein are Alanine (13.9%), Glycine (9.9%) and Valine (9.9%) as shown in the table number 2.

MERKPLIAGNWKMNLDHMQAVATVQKLAFALPAEYYEKVDVAVTVPFTDLRSVQT
VIDGDKLQITYGAQDVSEHESGAYTGEISAAMLAKLGCWVVVGHSERREYHGESSK
LVAAKAKAALSKDISPIVCVGEPLTIREAGTHVDFVVEQTRESLAGLTEDELAKTVIA
YEPVWAIGTGKVASAADAQEVCKAIRGLIKELANEKIAAGIRILYGGSVKEETVAEIV
GQPDVDGGLVGGASLDGEAFAKLAANAATGL<u>GGENLYFQG</u>HHHHHH

The amino acids in red and underlined are the cleavage site of TEV protease following by His tag of the vector pD441 that the letters in black corresponds to the amino acids sequence of TIM from *C. Pseudotuberculosis*.

Table 2: Number of each amino acid residue and its percentage in TIM of *C. pseudotuberculosis*

Ala (A)	38	13.9%	Arg (R)	8	2.9%
Asn (N)	5	1.8%	Asp (D)	13	4.7%
Cys (C)	3	1.1%	Gln (Q)	9	3.3%
Glu (E)	24	8.8%	Gly (G)	27	9.9%
His (H)	11	4.0%	Ile (I)	15	5.5%
Leu (L)	22	8.0%	Lys (K)	17	6.2%
Met (M)	4	1.5%	Phe (F)	5	1.8%
Pro (P)	7	2.6%	Ser (S)	13	4.7%
Thr (T)	15	5.5%	Trp (W)	3	1.1%
Tyr (Y)	8	2.9%	Val (V)	27	9.9%

3.1.2) Expression and purification of TIM

The His-tagged protein was expressed overnight at 30 °C in an agitating incubator. As explained in item number 2.2.1.4, the supernatant was passed on Ni-sepharose resin (GE). Lysis buffer (0.02 M Tris-HCl, 0.5 M NaCl, 0.005 M Imidazole, 5% glycerol pH 7.7) was used for the equilibration of the column. The bound fractions were washed and eluted with the lysis buffer, containing three different concentrations of imidazole (0.04 M, 0.06 M and 0.4 M imidazole). After Ni-NTA affinity chromatography, the purity of the protein was evaluated by 15% SDS-PAGE (fig 8) and the gel was stained with Comassie brilliant blue R-250. The solvent of the fraction that eluted with 0.4 M of imidazole was changed to buffer A (0.02 M Tris-HCl pH 7.7) and then passed through a Mono Q 5/50 GL column (previously equilibrated with a buffer A (0.02 M Tris-HCl pH 7.7). The fraction separated into two peaks and eluted with buffer B (0.02 M Tris-HCl, 1 M NaCl pH 7.7) as shown in the figure 9. The purity of the protein was evaluated by 15% SDS-PAGE gel as shown in figure 10. The eluted fraction (peak 1b) from Mono Q 5/50 GL column was then concentrated up to 500 µL and passed on size exclusion chromatography using Superdex 75 GL column (fig 11).

Figure 8: SDS-PAGE 15%, expressed protein TIM in *E. coli* passed through Ni-NTA. Stained with Comassie brilliant blue. Band M represents the molecular weight markers, 1a and 2a corresponds to wash with 0.04 M and 0.06 M imidazole respectively. 3a shows the elution of TIM with 0.4 M imidazole.
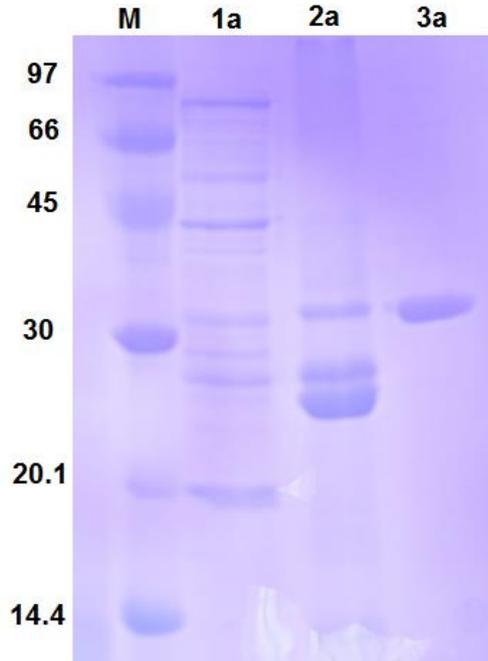
Figure 9: Anion exchange chromatographic profile. Blue line indicates the absorbance at 280 nm; the green line presents the concentration of buffer B.
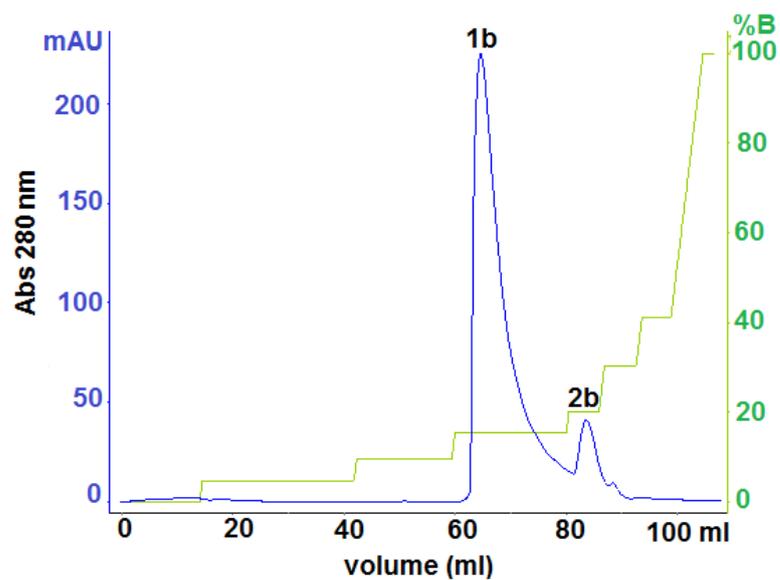
Figure 10: SDS-PAGE 15%; lane M represents molecular weight markers. 1b and 2b correspond to the peak 1b and 2b of the chromatogram of anion exchange chromatography.
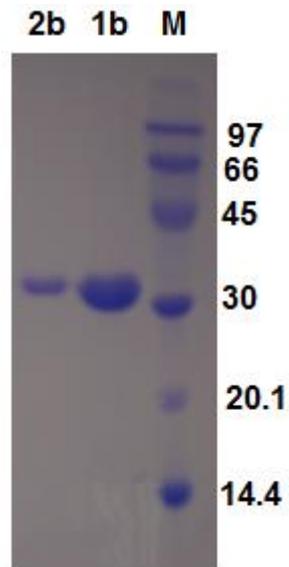
Figure 11: Size exclusion chromatography; peak 1c shows the eluted pure protein. The inserted figure shows the 15% SDS-PAGE with a band M represents the molecular weight markers (kDa) and a band 1c shows the pure 29.1 kDa TIM protein.
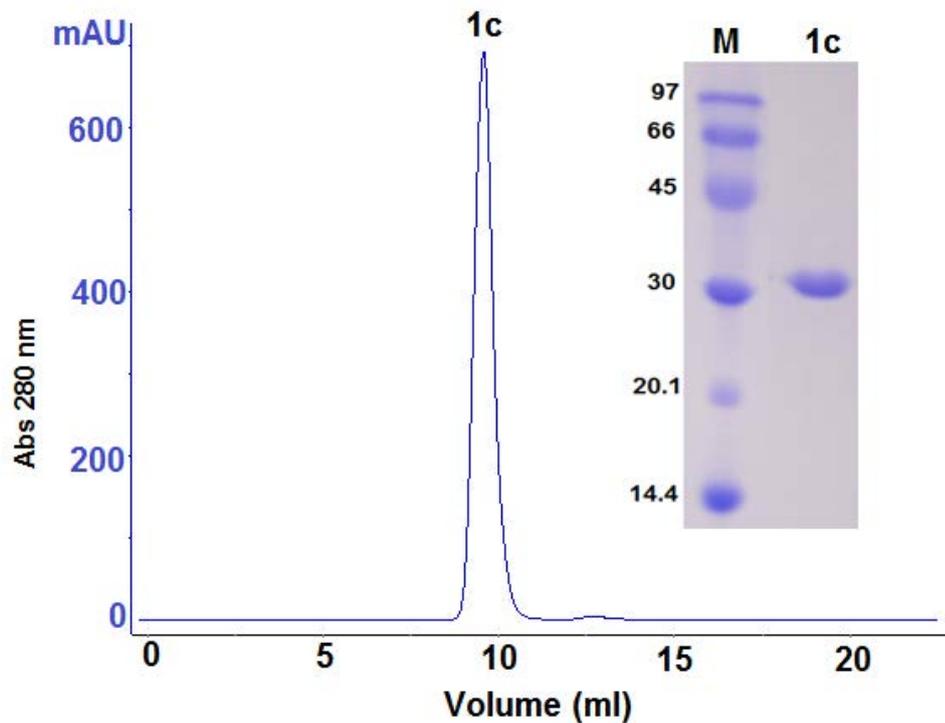
3.1.3) Circular dichroism

Circular dichroism (CD) spectroscopy was applied to study the thermal stability of TIM at different pH values (pH 3, 7.4 and pH 10) and temperature (25 to 75 °C). The CD spectrum of TIM (fig 12) presents, a significant change at temperature 25 °C and pH 3 while at other pH (7.4 and 10) the CD spectra are similar. This aspect is confirmed by calculating of the secondary structure percentages of TIM using the CD data (Table 3). At pH 3, the unfolding of TIM was induced. Since, CD spectrum reveals characteristics of a denatured protein and their secondary structure percentages exhibit significant changes with respect to the values found at pH 7.4 and 10. This pH-dependent behavior of TIM is very similar to PfuTIM (*Pyrococcus furiosus* triose phosphate isomerase) (Chandrayan and Guptasarma, 2009). The CD spectra of TIM showed different behaviors to pH as the temperature changed from 25 to 75 °C (Fig 12). The CD spectra of the protein are similar for pH 7.4 and 10 at 25 °C. Heating of the protein up to 75 °C and at pH 10 induces a larger unfolding and secondary structure percentages are characteristics of proteins in a higher level of denaturation. These results suggest that TIM from *C. pseudotuberculosis* presents a lower thermal stability at pH 10 than at pH 7.4. On the other hand, the temperature change (25 $\rightarrow$ 75 °C) at pH 3 did not provide any further information as the protein was already unfolded (not shown in graph). Figure 13 shows the molar ellipticity at 222 nm as a function of the temperature at different pH conditions for the thermal denaturation of TIM. It is clear from the graph that the protein exhibits a lower melting temperature (T_M) at pH 10 (39 °C) than at pH 7.4 (48 °C) and that there was no significant changes in molar ellipticity (at 222 nm) at pH 3.

Figure 12: Far-UV CD spectra of TIM from *C. pseudotuberculosis* at different temperatures (25 and 75 °C) and pH values
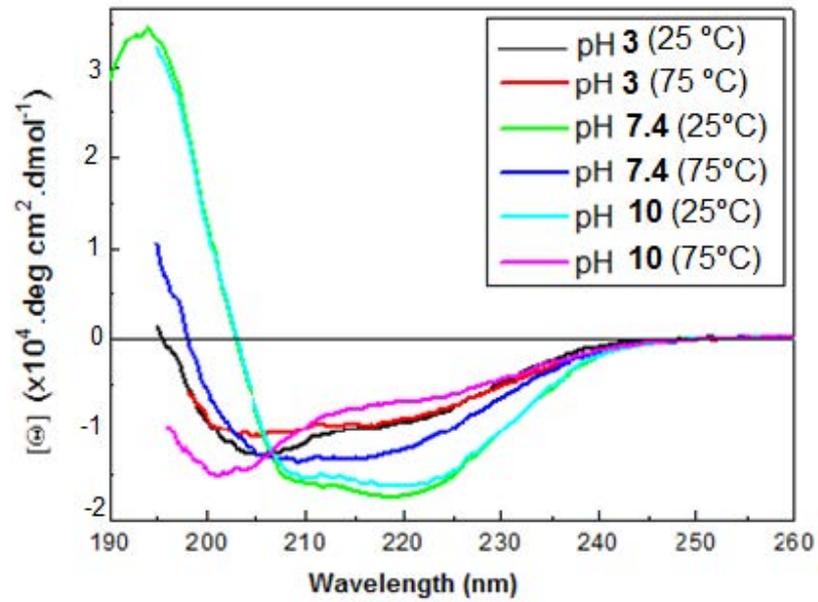
Table 3: This table shows the secondary structure (α-helices, β-sheets, turns and random coils) percentage values of TIM at different temperature and pH values

Secondary structures (%)	pH 3		pH 7.4		pH 10	
	25 °C	75 °C	25 °C	75 °C	25 °C	75 °C
α-helice	22	17	50	34	51	15
β-sheet	22	26	11	16	11	24
Turn	22	25	15	22	14	25
Random coil	34	32	24	28	24	36

Figure 13: Temperature dependence of the molar ellipticity at 222 nm for the thermal unfolding of TIM. The TIM spectra in a NaH_2PO_4 buffer pH 7.4 (green circles), the TIM spectra in a CH_3COONa buffer, pH 10 (blue triangle), TM is melting temperature
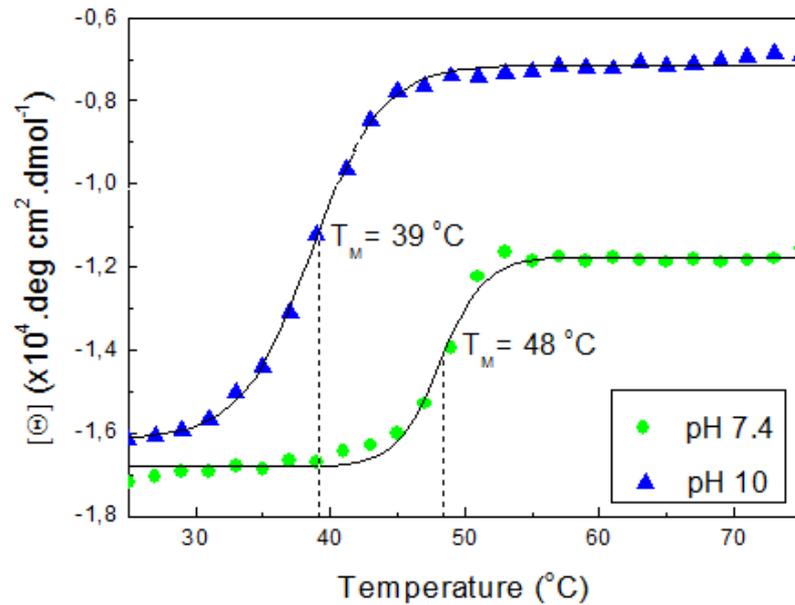
3.1.4) Crystallization, data collection and structure determination

The purified TIM in a buffer 0.02 M Tris–HCl pH 7.7 at a concentration of 17 mg/mL was used for the crystallization experiments, 24-well tissue-culture plates and commercially available crystallization screens were used in the experiments. A manual hanging-drop method was used for the crystallization. The details of the crystallization conditions are given in table 4.

Table 4: Crystallization conditions

	Initial conditions	Optimization
Method	Hanging-drop vapor-diffusion	Hanging-drop vapor-diffusion
Plate type	24 well tissue culture plates	24 well tissue culture plates
Temperature (K)	298	298
Protein concentration (mg/mL)	17	17
Buffer composition of protein solution	0.02 M Tris–HCl pH 7.7	0.02 M Tris–HCl pH 7.7
Composition of reservoir solution	0.2 M ammonium sulfate, 0.1 M Na-cacodylate and 20% *(w/v)* PEG 8000, pH 6.5	0.2 M ammonium sulfate, 0.1 M Na-cacodylate and 20% *(w/v)* PEG 8000, pH 6
Volume and ratio of drop (µL)	0.8 : 0.8	2 : 2
Volume of reservoir (µL)	500	500

Figure 14: (A) Microphotograph of TIM crystals. **(B)** X-diffraction pattern of TIM

A   B
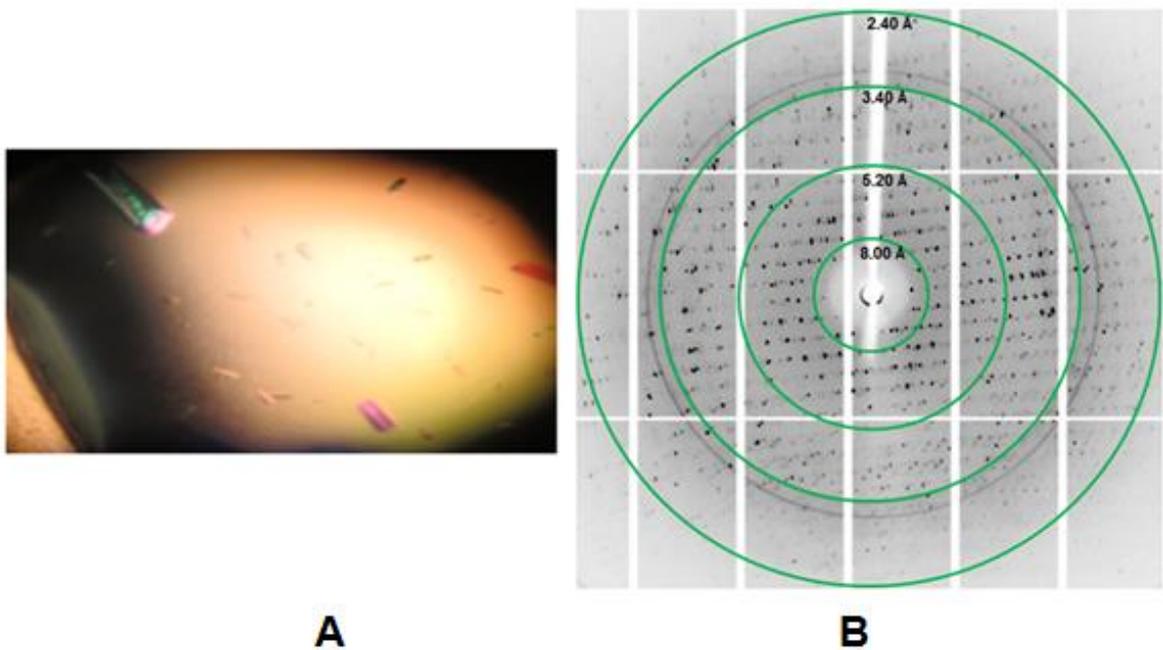
For the diffraction of TIM crystal and data collection, a single crystal was isolated and was directly flash-cooled in a 100 K nitrogen-gas stream at W01B-MX2 beamLine at Brazilian Synchrotron Light   Laboratory (LNLS-Campinas, Brazil). The wavelength of the radiation was set to 1.4 Å and X-ray diffraction intensities were recorded on MarCCD 165 mm CCD detector. The crystal diffraction data was processed in the space group $P2_12_12_1$. The Matthews coefficient calculated for the crystal is 2.24 Å3 Da-1 that corresponds to a solvent content of 45.6 % (Mathews, 1968). Considering the molecular weight of the protein (~29,100 Da), the asymmetric unit contains two molecules. This information is given by molecular replacement by CCP4 Program. A final model of the TIM structure was refined up to 2.40 Å with Rfree/Rfactor = 21/28%.

The statistics of data collection, refinement and parameters for the model quality are given in table 5.

Table 5: Data collection and refinement statistics

Diffraction source:	Brazilian Synchrotron Light Lab
Wavelength (Å):	1.458
Temperature (K):	100
Detector:	MarMosaic 225 mm
Space group:	$P2_12_12_1$
Unit-cell parameters (Å)	a = 40.40, b = 101.40, c = 124.03
Resolution range (Å)	20.00-2.40 (2.50-2.40)
R_{merge}† (%)	17.1 (51.5)
$\langle I/\sigma(I) \rangle$	4.2 (2.0)
Data completeness (%)	96.7 (95.1)
No. of measured reflections:	97.095 (6795)
No. of unique reflactions:	19.984 (2017)
Redundancy (Multiplicity):	4.9 (3.7)
V_M (Å^3 Da^{-1})	2.39
Slovent content (%)	48.56
Molecules per asymmetric unit	2
Model refinement	
R_{model} (%)	21.00
R_{free} (%)	28.50
R.m.s.d from ideal bond lengths (Å)	0.012
R.m.s.d from ideal angles (°)	1.65
Average B-factor (Å^2):	27

* Values in parentheses are for the last resolution shell

† $R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith observation of reflection *hkl* and $\langle I(hkl) \rangle$ is the weighted average intensity for all observations *I* of reflection *hkl*.

The Ramachandran plot analysis for TIM shows that more than 90% of the amino acid residues were in the favored region, about 7.8% in additional allowed region and only 0.2% in disallowed region (Fig 15).

Figure 15: Ramachandran plot of TIM from *C. pseudotuberculosis*
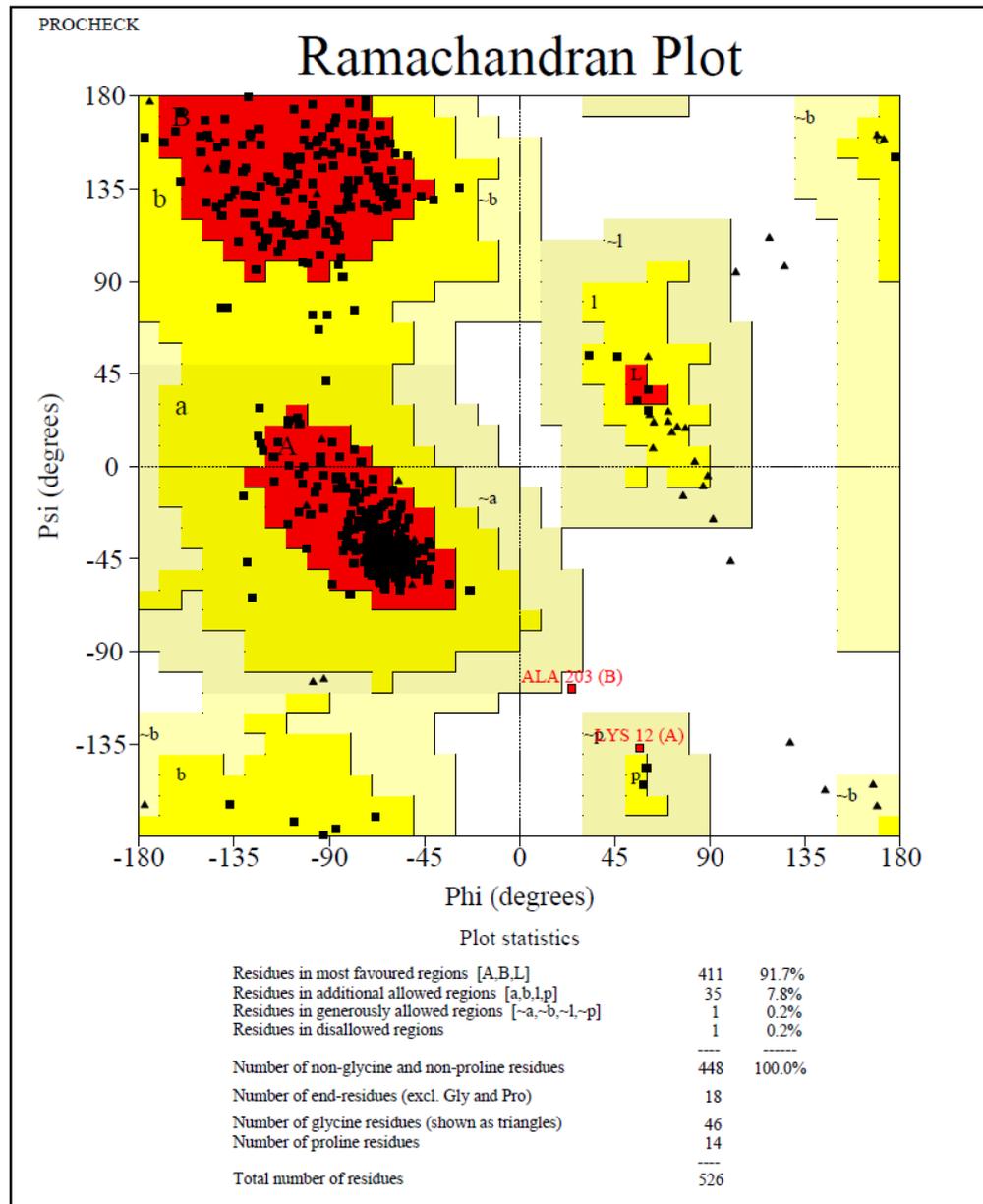
The overall structure of CpTIM (TIM from *Corynebacterium pseudotuberculosis*) conserved all the structural properties as presented by TIM structure from other organisms. Its

structure contains eight alpha helices and eight beta sheets $(\alpha/\beta)_8$ with additionally four 3_{10} alpha helices. The helices are arranged around the sheet, producing a special type of barrel, which is called TIM barrel (Fig 16). On one end of the barrel, the N and C termini are located and on the opposite end, the active site is located. The N-terminal starts with beta sheet and the C- terminal ends with an α-helix (Fig 17). The highly conserved active site residues include Asn11, Lys13, His95 and Glu167 (Fig 16 B).

Figure 16: (A) Overall structure of CpTIM-cartoon representation, the helices, sheets and loops are colored in red, yellow and green respectively. The helices are labelled as H1-H8 and the sheets as S1-S8. **(B)** Active site residues highlighted
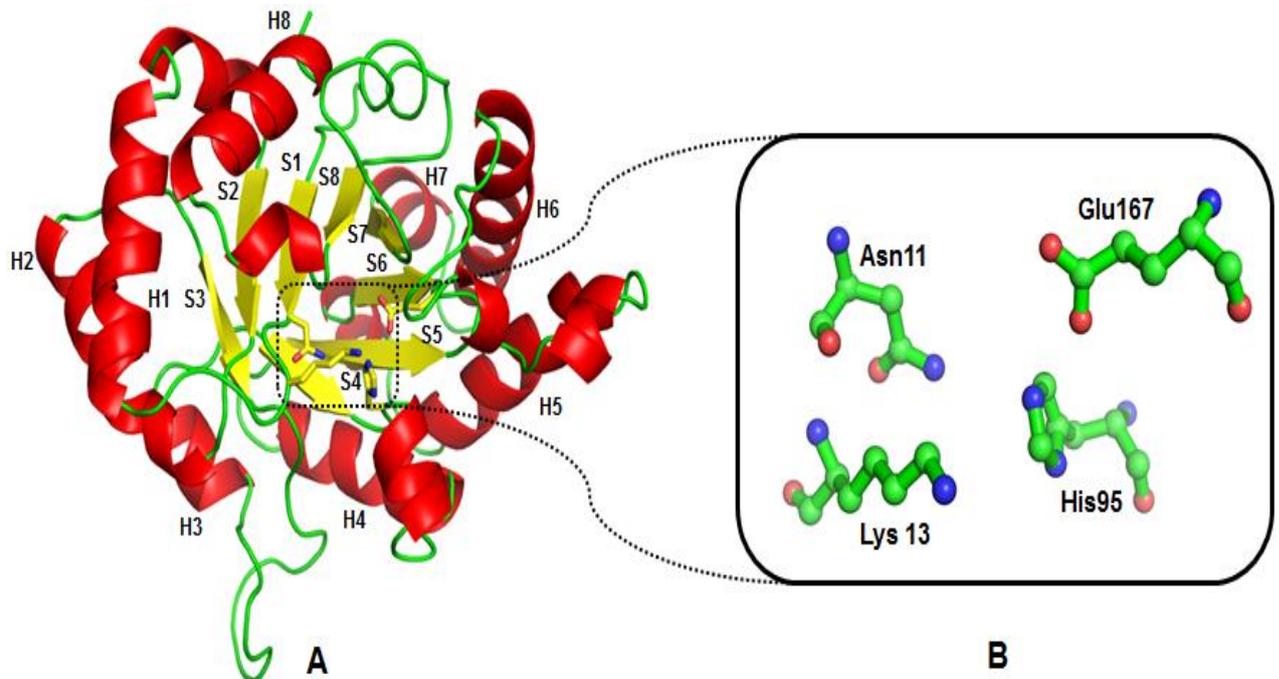
Figure 17: Topology diagram of CpTIM. The helices (H1-H8) are represented by cylinders and the sheets (S1-S8) by arrows
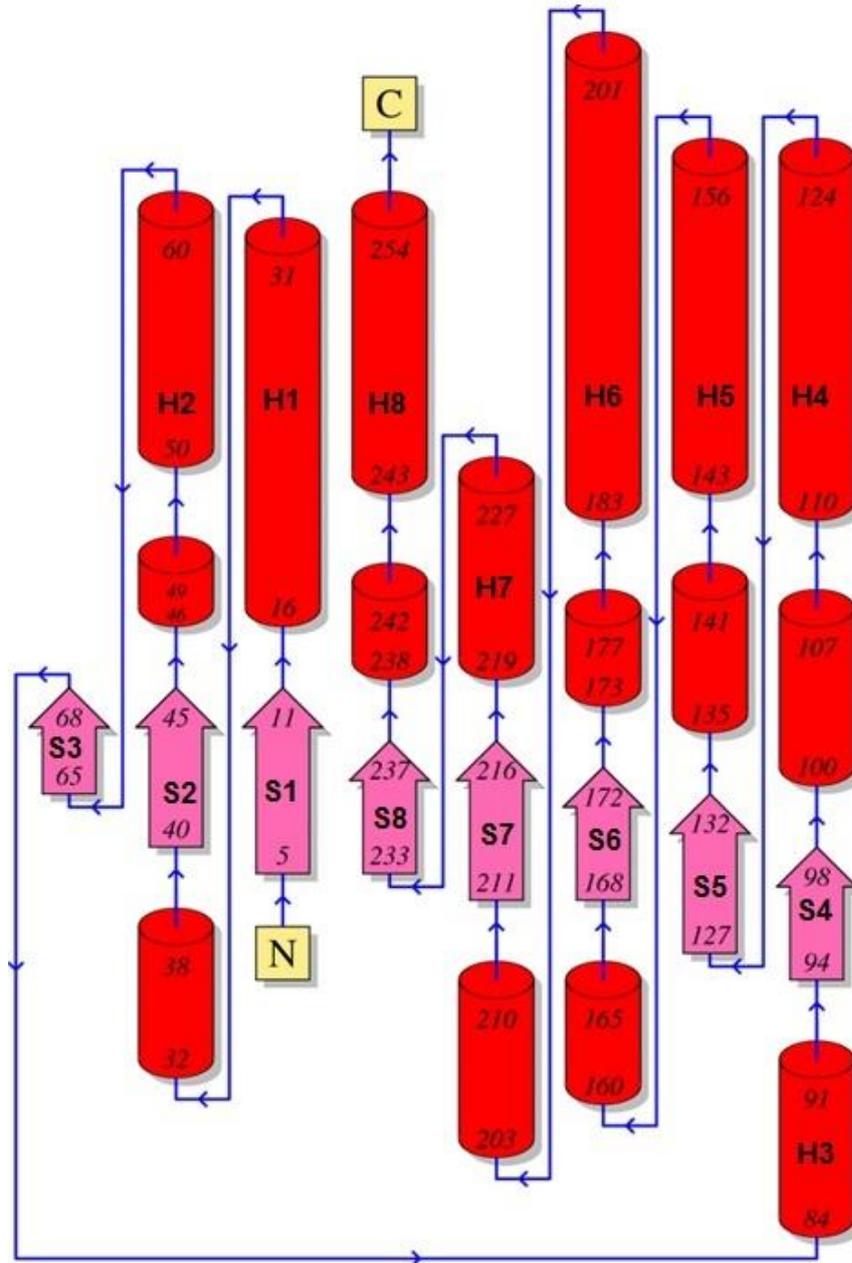
3.1.5) Sequence and structure comparison of CpTIM with TIMs from various organisms

          Sequence alignment among TIMs from various organisms has been done and it indicates mean sequence identity of 40% (Table 6). The highest sequence identity was observed with 3TA6 and 67% and the lowest with 3QST, 1AW1 and 4OHQ (42%). All the catalytic residues are conserved among these TIMs.

Table 6: Sequence identity of TIM from *C. pseudotuberculosis* with other TIMs

Species	PDB Code	Sequence Identity (%)
Mycobacterium tuberculosis	3TA6	67
Thermus thermophiles	1YYA	49
Thermotoga maritima	1B9B	45
Bacillus stearothermophilus	2BTM	44
Staphylococcus aureus	3M9Y	43
Burkholderia thiandensis	4G1K	49
Vibrio marinus	1AW1	42
Tenebrio molitor	2I9E	43
Rabbit muscle	1R2R	43
Arabidopsis thaliana	4OHQ	42
Trichomonas vaginalis	3QST	42

Figure 18: Sequence alignment; **CpTIM**: TIM from *C. pseudotuberculosis*, **3TA6**: *M. tuberculosis*, **1YYA**: *Thermus thermophilus* HB8, **1B9B**: *Thermotoga maritima*, **2BTM**: *Bacillus stearothermophilus*, **3M9Y**: *Staphylococcus aureus*, **4G1K**: *Burkholderia thailandensis*, **1AW1**: *Vibrio marinus*, **2I9E**: *Tenebrio molitor*, **1R2R**: Rabbit muscle, **4OHQ**: *Arabidopsis thaliana*, **3QST**: *Trichomonas vaginalis*. The conserved active site residues are underlined in blue.

Structural alignment among different TIMs (3TA6, 3GVG, 1YYA, 1B9B, 4G1K and 2BTM) indicate that they are very similar to each other (Fig. 19). Alignment of *C. pseudotuberculosis* TIM with TIM (PDB code 3TA6) from *Mycobacterium tuberculosis* shows a very slight difference in the loop region as shown in the figure 19 A as they share 67% sequence identity and 0.4 RMSD value. The higher difference observed among CBTIM and TIMs from *Thermus thermophiles HB8*, *thermotoga maritima*, *Burkholderia thailandensis* and *Bacillus stearothermophilus* with RMSD values 0.42, 0.79, 0.76 and 0.76 respectively. The percentage sequence identity of these TIMs is given in table 6. The overall fold is similar as well as the active site the only difference is in the loop region and with the above TIM structures, the CpTIM is not completely superimposed as shown in figure 19.

Figure 18: Structural alignment among TIMs from various organisms; **(A)** CpTIM: TIM from *C. pseudotuberculosis* (green- in all cases) aligned to 3TA6: *M. tuberculosis* (cyan) **(B)** CpTIM aligned 3GVG: *M. tuberculosis* (yellow) **(C)** CpTIM aligned with 1YYA *T. thermophilus* HB8 (Blue)**(D)** CpTIM aligned 1B9B: *T. maritima* (Light gray) **(E)** CpTIM aligned 4G1K: *B. thailandensis* (Magenta) **(F)** CpTIM aligned 2BTM: *B. stearothermophilus* (Light brown). Orange indicates the helix that is different in all structures. RMSD = Root mean square deviation.
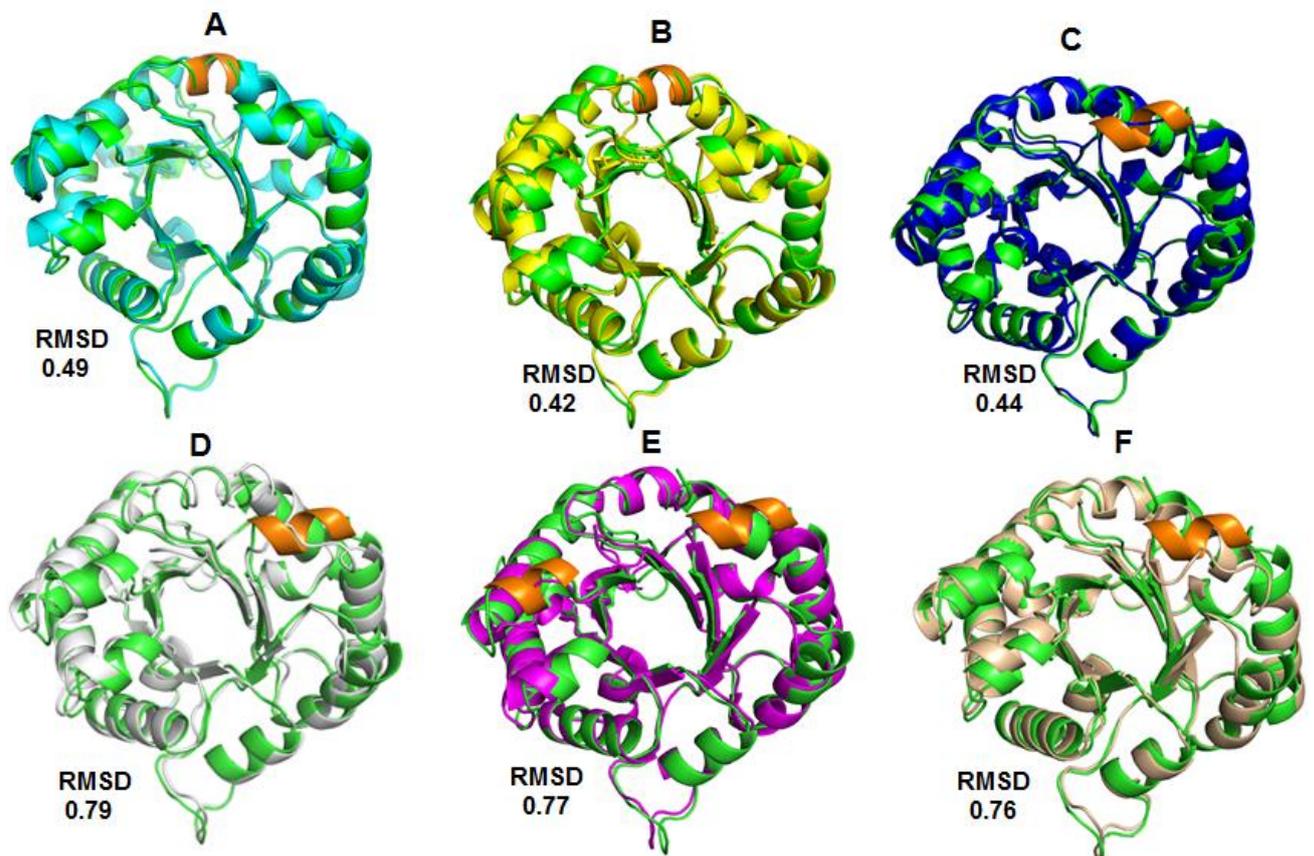
3.1.6) Dimerization and stability

The CpTIM like from other organisms is a functional dimer. There are about 68 (33 residues from chain A/monomer and 35 residues from chain B/monomer) amino acid residues participating in the dimerization. Electrostatic forces of attraction or nonbonding contacts along with three salt bridges/ionic bonds and fifteen Hydrogen bonding stabilize the two monomers or chains (A and B). The interface area of chain A is 1701 Å and that of B is about 1714 Å. The electrostatic forces of attraction between the two chains/monomers, is the attraction between them due to the charge (positive and negative) which stabilize the dimer (Fig 20).

Figure 19: Dimerization of TIM and the amino acid residues participating in the dimerization
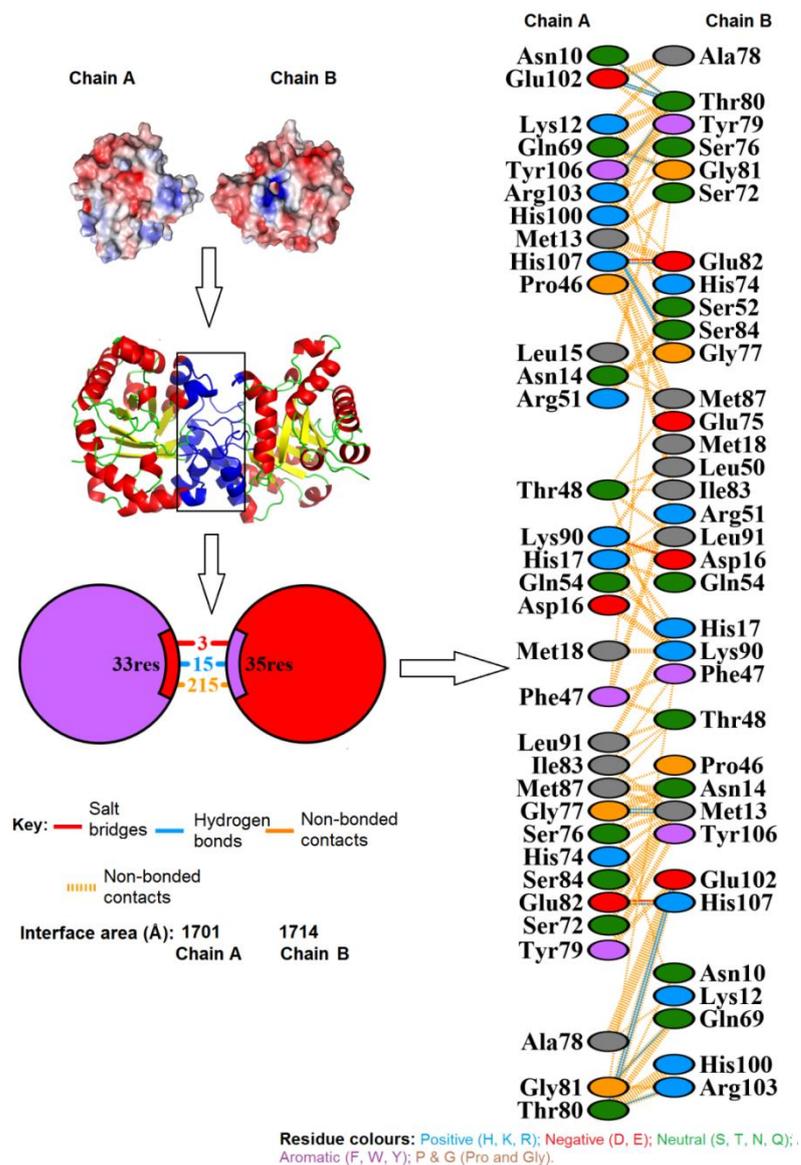
Although the TIMs from all organisms are very similar in sequence as well as structural fold but each TIM is specific towards its substrate, there substrate specificity can be observed by the surface charge distribution. The difference in face A and face B of CpTIM and TIMs from other organisms are given in table 7 and figure 21.

Table 7: Surface charge analysis among different TIMs

Species	PDB code	Difference of face A	Difference of face A
C. pseudotuberculosis	-	Negative with an active site cavity partially neutral and partially positive	Partially neutral and partially negative
Mycobacterium tuberculosis	3TA6	Partially neutral and partially negative	Partially neutral and partially negative
Thermos thermophilus	1YYA	More positive and partially neutral and negative	Partially neutral and equally positive and negative
Staphylococcus aurus	3M9Y	Positive and somewhat neutral and partially negative	Negative and partially neutral and less positive
Burkholderia thailandesis	4G1K	Partially neutral and partially positive	Partially neutral and partially positive

Figure 20: Surface charge distribution of various TIM structures. CpTIM: TIM from *C. pseudotuberculosis, M. tuberculosis,* 3TA6: TIM from *M. tuberculosis*, 1YYA: TIM from *Thermus thermophilus HB8*, 3M9Y: *Staphylococcus aureus,* 4G1K*: Burkholderia thailandensis.*
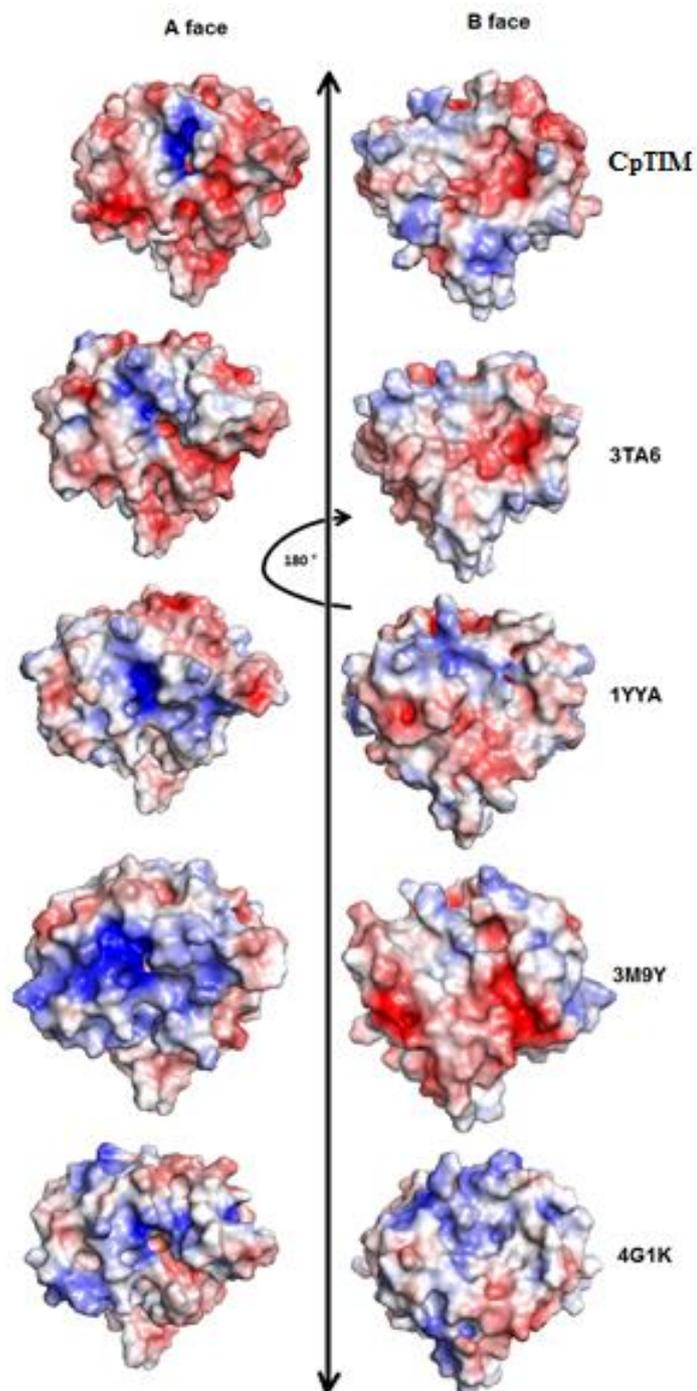
3.2) Thioredoxin reductase (TrxR)

3.2.1) Sequence analysis of TrxR through the ExPASy- ProtParam tool

Thioredoxin reductase has 330 amino acid residues in its sequence and molecular weight 35.5 kDa. The sequence analysis of thioredoxin reductase through ExPASy-ProtParam tool (Gasteiger *et al.*, 2005) gave the following information. The protein has a theoretical pI 4.94. The protein is negatively charged at physiological pH. The total number of negatively charged amino acids (Asp+Glu) is 48, whiles those of positively charged amino acids (Arg+Lys), is 28. The instability index value is 24.82 that classified the protein as stable, while the aliphatic index is 80.76. The dominant amino acids in in the sequence are Alanine (11.5%) and Glycine (10.3%) as shown in table 8.

MSEAPNATTVHDVAIIGSGPAGYTAALYAARAELKPIVFEGIEFGGSLMTTTEVENFP
GFPEGIMGPDLMDNMRSQAERFGADLRMELVTKVELEGEIKKIWVDDQEFHARSVIL
ATGSAPRYIGAEGEQTLLGRGVSACATCDGFFFRDHDIAVIGGGDSAMEEATFLTKFA
KSVTIVHRREEFRASAIMLERAKNNDKIKFLTNKTVSKVLGENTVSGLELTDTATGET
SILDVTAMFVAIGHDPRSDLFRGVVDTNDAGYVRVEEPSTRTNVPGVFAVGDLVDA
HYQQAITAAGSGCRGAIDAENYLAALDA<u>GGENLYFQG</u>HHHHHH

The amino acids in red and underlined corresponds to the cleavage site of TEV protease following by His tag of vector pD441 while the letters in black are the amino acids sequence of TrxR (Thioredoxin reductase) from *C. Pseudotuberculosis.*

Table 8: Number of each amino acid residue and its percentage in a TrxR from *C. pseudotuberculosis* sequence

Ala (A)	38	11.5%	Arg (R)	17	5.2%
Asn (N)	11	3.3%	Asp (D)	21	6.4%
Cys (C)	3	0.9%	Gln (Q)	6	1.8%
Glu (E)	27	8.2%	Gly (G)	34	10.3%
His (H)	12	3.6%	Ile (I)	18	5.5%
Leu (L)	22	6.7%	Lys (K)	11	3.3%
Met (M)	9	2.7%	Phe (F)	17	5.2%
Pro (P)	10	3.0%	Ser (S)	16	4.8%
Thr (T)	25	7.6%	Trp (W)	1	0.3%
Tyr (Y)	7	2.1%	Val (V)	25	7.6%

3.2.2) Expression and purification of TrxR

The protein was expressed overnight, after centrifugation at 5000 $\times$ g the pellet was suspended in a lysis buffer (0.02 M NaH_2PO_4, 0.5 M NaCl, 0.005 M Imidazole, 5% glycerol pH 7.7) and subjected to sonication. The filtered supernatant was passed on Ni-sepharose resin (GE). Lysis buffer was used for the equilibration of a column. The bound fraction was washed and eluted through three different concentrations of imidazole (0.04 M, 0.06 M and 0.4 M imidazole) in the lysis buffer. The protein was eluted in a large quantity with 0.4 M imidazole. After Ni-NTA affinity chromatography, the purity of the protein was evaluated by 15% SDS-PAGE shown in figure 22.

Figure 21: SDS-PAGE 15%, of first step (Ni-NTA) protein purification. M: molecular weight markers (kDa), lane 1a and 2a show the wash with 0.04 M and 0.06 M imidazole concentration in the lysis buffer, while 3a represents the elution of TrxR with 0.4 M imidazole. Stained with Comassie brilliant blue.
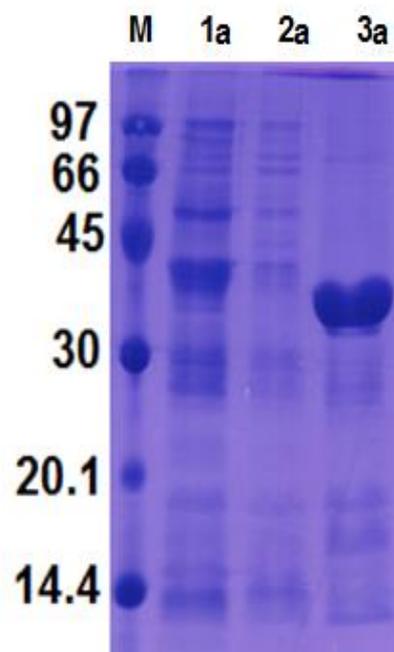
3.2.3) Ion (cation and anion) exchange chromatography

The 40 mL elution fraction of protein from Ni-NTA chromatography was subjected to anion exchange chromatography by passing through a column Mono Q 5/50 GL, using Akta purifier. The sample separated and eluted in two major peaks 1a and 1b as indicated in the figure 23. The purity of the protein was evaluated by 15% SDS-PAGE (Fig 24). The protein was with contamination and the two peaks (1a and 1b) from anion exchange chromatography were pooled and subjected to cation exchange chromatography by using Mono S 5/50 GL column. The pH was changed from 7.7 to 5.6. The pooled fraction was separated into four peaks (Fig 25). The purity of each fraction was evaluated by 15% SDS-PAGE (26).

Figure 22: Anion-exchange chromatographic profile; a fraction 3a from Ni.NTA chromatography. Blue and green lines represent UV at 280 nm and concentration of buffer B respectively.
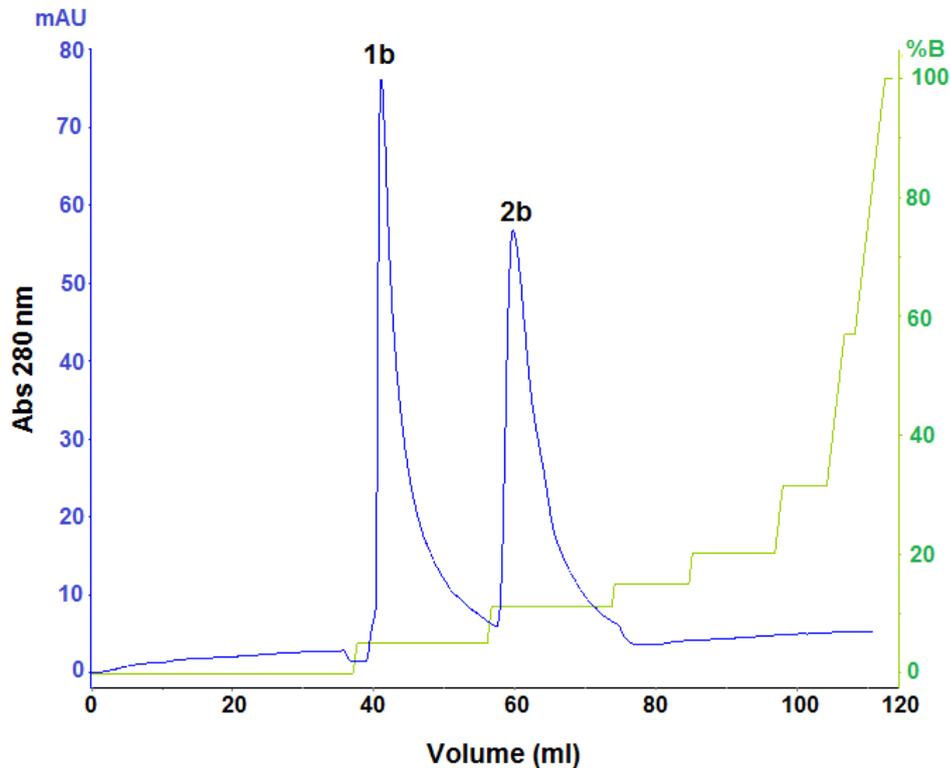
Figure 23: SDS-PAGE 15%; Peak fractions from anion-exchange chromatography. Lane M, molecular weight markers (labelled in kDa); lanes labelled 1b and 2b correspond to chromatographic peak fractions with the same label.
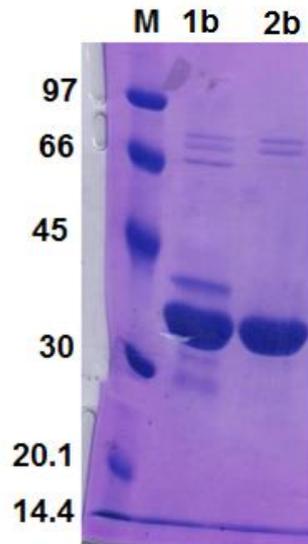
Figure 24: Cation-exchange chromatographic profile; peaks 1b and 2b from Mono Q 5/50 GL column. Blue and green lines represent UV at 280 nm and concentration of buffer B respectively.
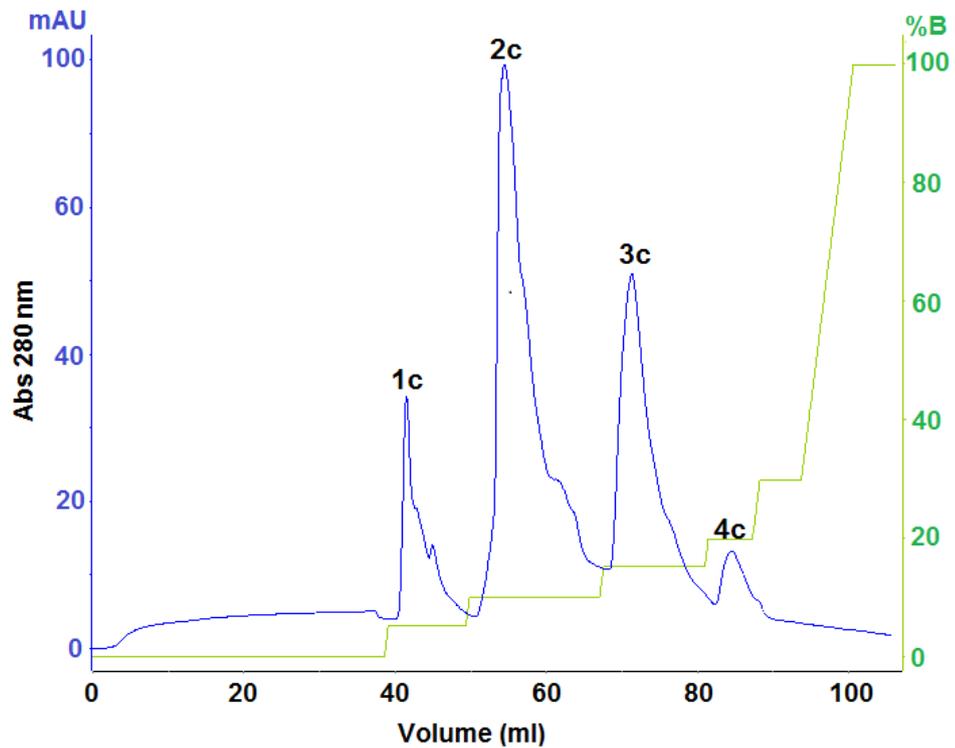
Figure 25: SDS-PAGE 15%; Peak fractions from cation-exchange chromatography. Lane M, molecular weight markers (labelled in kDa); lanes labelled 1c, 2c, 3c and 4c correspond to chromatographic peak fractions with the same label. Lane 2c and 3c represent TrxR.
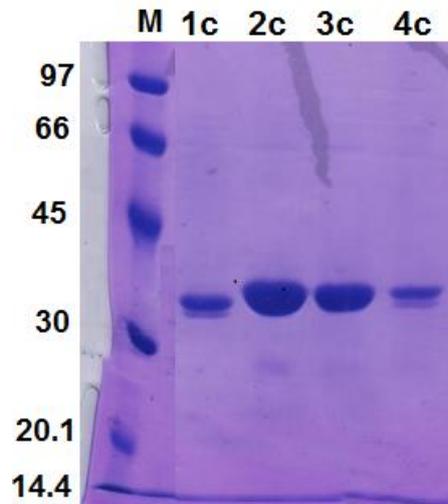
3.2.4) Size exclusion chromatography

The protein eluted in fraction 2c and 3c from cation exchange chromatography were pooled and concentrated up to 500 µL. The sample was then applied on Superdex 75 GL column, following the size exclusion chromatography. The peak 1c in figure 27 represents pure TrxR. The purity was evaluated through SDS-PAGE gel.

Figure 26: Size exclusion chromatographic profile. Insert 15% SDS-PAGE gel, lane M representing the molecular mass markers in kDa, lane 1d represents pure TrxR from *C. pseudotuberculosis*.
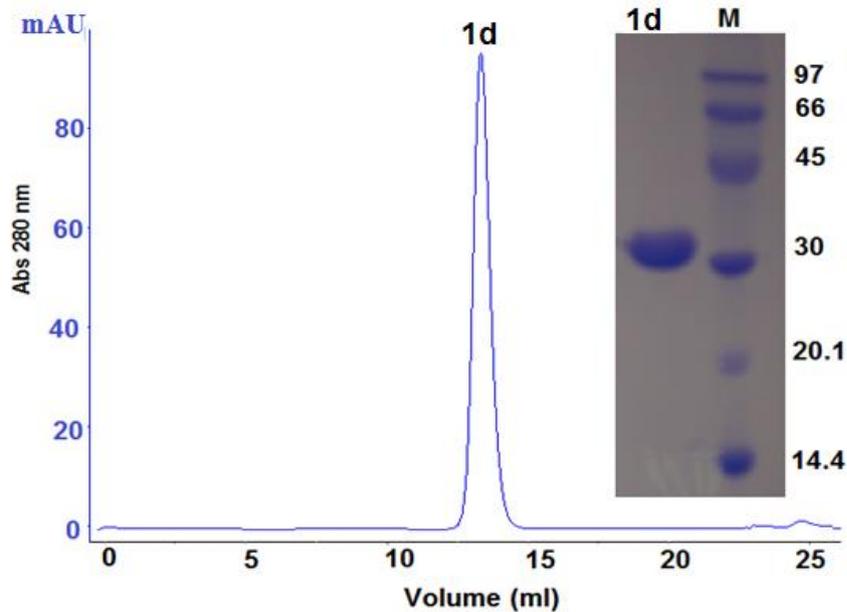
3.2.5) Thermal stability of TrxR by differential scanning calorimetry (DSC)

The thermal stability of TrxR from *C. pseudotuberculosis* was investigated by using the DSC technique. Figure 28 shows the DSC thermogram of the unfolding process (heating scan) of the protein. From this thermogram, it can be seen that TrxR presents a melting temperature (T_M) at approximately 61 °C, corresponding to the heat capacity maximum of the endothermic transition (Privalov, 1979), and the beginning of transition occurs at 47 °C. In the literature, thermal denaturation studies through CD spectroscopy shows, that human TrxR presents a melting temperature at 63 °C and its conformational changes begins above 45 °C (Oblong *et al.*, 1993), which is similar to the thermal characteristic of TrxR from *C. pseudotuberculosis*. The cooling scan after the denaturation of the protein did not show any transition in the DSC thermogram, which indicating that its thermal unfolding is irreversible, similar to that of human TrxR (Oblong *et al.*, 1993).

Figure 27: DSC thermogram of the unfolding process of TrxR. The apparent excess heat capacity curve was obtained for TrxR (0.9 mg/mL) in phosphate buffer (0.02 M Na_2HPO_4/NaH_2PO_4, 0.05 M NaF, pH 7.4) at a scan rate of 1°C/min. The dotted lines (- - -) indicates the melting temperature (T_M) of the protein.
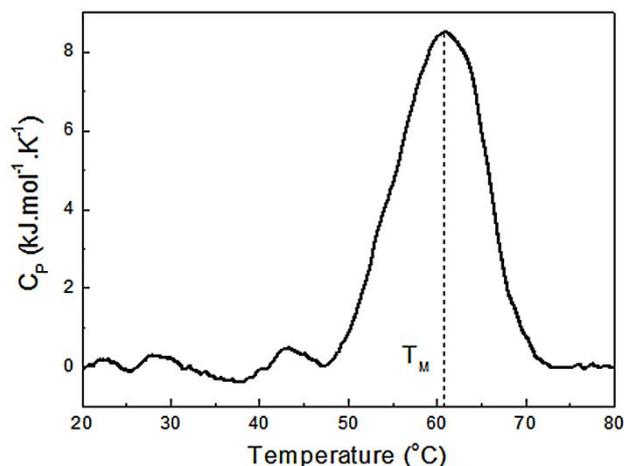
3.2.6) Metal ion interaction study of TrxR

The UV-Vis spectroscopy was used to investigate FAD in *C. pseudotuberculosis* TrxR structure and the interaction of this protein with metal salts ($AgNO_3$, Li_2SO_4, $ZnSO_4$, $NiSO_4.6H_2O$, $CuCl_2$, and $CsSO_4$). The presence of the non-covalently bound FAD as a prosthetic group in TrxR was confirmed with the absorption peaks at 374 and 452 nm along with a shoulder at 475 nm (Fig.29), which is a characteristic of flavoproteins (Brown *et al.*, 1996). The interaction of TrxR with metal salts revealed that among the studied compounds only silver salt (Ag^{1+}) presented significant changes in FAD spectrum of the protein. The presence of Ag cation caused a blue shift of the absorption peaks of FAD, from 374 to 368 and from 452 to 447 nm, and vanished with the shoulder at 475 nm (Fig 29). The features of FAD spectrum after the interaction of TrxR with the silver ions are similar to a FAD free solution (Thelander, 1967), suggesting the FAD exposure to the solvent probably due to the conformational changes.

Figure 28: UV-Visible spectra of TrxR in absence and presence of metal salts and UV-Vis spectra of the protein were obtained in the absorption region of FAD. The black line indicates the spectrum of TrxR, the red, green, blue, cyne, magenta and yellow indicate the TrxR in the presence of Ag^+, Ni^{+2}, Cu^{+2}, Zn^{+2}, Cs^{+2}, and Li^+ respectively.
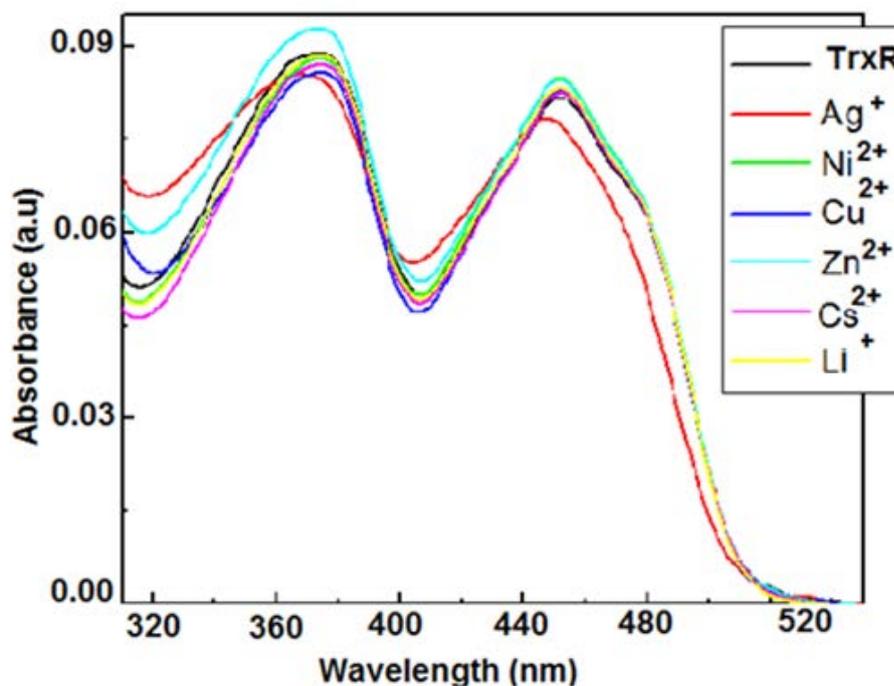
3.2.7) Circular dichroism

Far-UV CD experiments of the interaction between TrxR and metal salts were performed in order to investigate the possible conformational changes in the protein structure. Fig. 30 shows that only the interaction with the silver salt provided significant conformational changes in the TrxR structure. CD spectrum in the absence of Ag^+ displayed characteristics of a structured protein; its observed minimum was near 208 nm with a positive ellipticity at 190 nm (Fig. 30). The addition of Ag^+ ion to the TrxR solution induced unfolding of the protein, which can be observed by the negative ellipticity at 190 nm, the shift of the minimum from 208 to 201 nm, and decreasing of the negative ellipticity around of 222 nm. CD results collaborate with UV-Vis findings previously described in item 3.2.6.

Figure 29: Far-UV CD spectra of TrxR from *C. pseudotuberculosis* in absence and presence of metal salts (AgNO₃, Li₂SO₄, ZnSO₄, NiSO₄.6H₂O, CuCl₂ and CsSO₄)
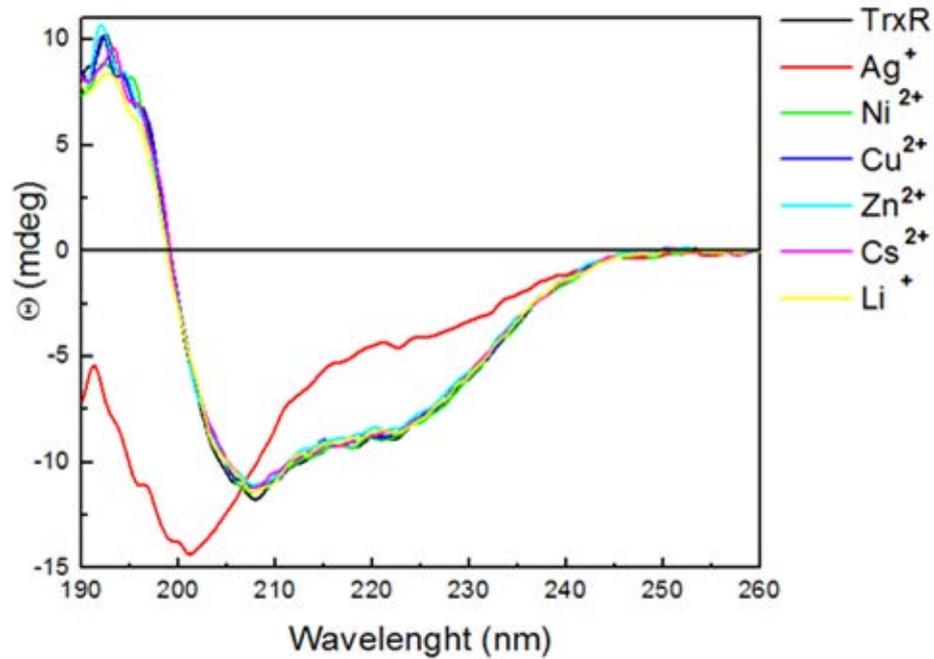
3.2.8) Crystallization

Pure 16 mg/mL TrxR was applied manually for crystallization through hanging drop method. The protein gave yellow crystals in a condition given in table 9. The crystals were needle like (Fig 31A) and were not suitable for collecting diffraction data.

Figure 30: Microphotograph of TrxR crystals **(A)** Needle like **(B)** Single crystal
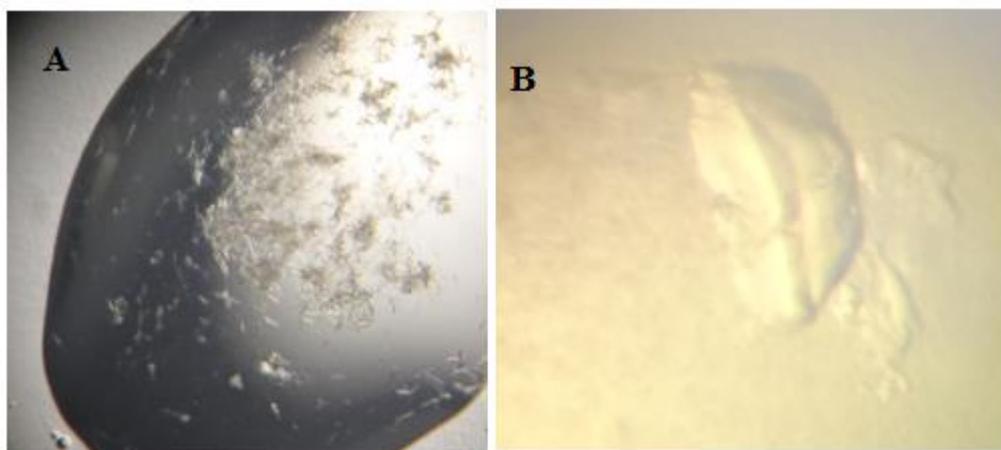
Table 9: Crystallization conditions

	Initial conditions	Optimization
Method	Hanging-drop vapor-diffusion	Hanging-drop vapor-diffusion
Plate type	24 well tissue culture plates	24 well tissue culture plates
Temperature (K)	298	298
Protein concentration (mg/mL)	16	16
Buffer composition of protein solution	0.02 M Tris–HCl pH 7.	0.02 M Tris–HCl pH 7.7
Composition of reservoir solution	0.1 M Tri sodium citrate dihydrate pH 5.6, 20% *(v/v)* isopropanol and 20% PEG 4000	0.1 M Tri sodium citrate dihydrate pH 5.2, 20% *(v/v)* isopropanol and 20% PEG 4000
Volume and ratio of drop (µL)	0.8 : 0.8	2 : 2
Volume of resorvour (µL)	500	500

3.3) Thioredoxin (Trx)

3.3.1) Sequence analysis of Trx through the ExPASy- ProtParam tool

Thioredoxin from *C. pseudotuberculosis* is a small protein having molecular weight 13.5 kDa and 122 amino acid residues. Through ExPASy-ProtParam tool (Gasteiger *et al.*, 2005) the following chemical and physical parameters has been obtained. The protein has a theoretical pI 5.74. The protein is negatively charged at a physiological pH. As the number of negatively charged proteins (Asp+ Glu) are 17 while that of positively charged amino acids (Arg +Lys) are 12. The instability index is computed to be 12.07 and classified the protein as stable. The Aliphatic index is 90.25.

MNAPIALTEATFKTTVIDSDKPVLVDFWAEWCGPCKKLGPIIDEIAAEMGDEVVVGK VDVDAERNLGAMFQIMSIPTVLIFKDGQKVAEFVGVRPKSEIVAKLRSHQGGENLYF QGHHHHHH

The amino acids in red and underlined are correspond to the cleavage site of TEV protease of vector pD441 while the letters in black corresponds to the amino acids sequence of Trx (thioredoxin) from *C. pseudotuberculosis*.

Table 10: Number of each amino acid residue and their percentage in a Trx sequence (from *C. pseudotuberculosis*).

Ala (A)	10	8.2%	Arg (R)	3	2.5%
Asn (N)	3	2.5%	Asp (D)	8	6.6%
Cys (C)	2	1.6%	Gln (Q)	4	3.3%
Glu (E)	9	7.4%	Gly (G)	10	8.2%
His (H)	7	5.7%	Ile (I)	9	7.4%
Leu (L)	7	5.7%	Lys (K)	9	7.4%
Met (M)	4	3.3%	Phe (F)	6	4.9%
Pro (P)	6	4.9%	Ser (S)	4	3.3%
Thr (T)	5	4.1%	Trp (W)	2	1.6%
Tyr (Y)	1	0.8%	Val (V)	13	10.7%

3.3.2) Expression and purification of Trx

Thioredoxin (Trx) was expressed in *E. coli* BL21 (C43) overnight and was harvested by centrifuging the expressed culture at $5000 \times$ g for 15 minutes. The pellet was suspended in a lysis buffer (0.02 M NaH_2PO_4, 0.5 M NaCl, 0.005 M Imidazole, 5% glycerol pH 7.7) and subjected to sonication. The supernatant obtained from the sonication and centrifugation process, was filtered and passed through Ni-sepharose resin.

The column was equilibrated with the lysis buffer. For purification and elution of the desired protein, the lysis buffer was divided further according to the concentration of

imidazole. After the protein mixture passed through the column the bounded proteins were washed through lysis buffer having different imidazole concentration such as 0.04 M, 0.06 M and 0.4 M. The evaluation of the purity of the protein was observed through 15% SDS-PAGE gel as shown in figure 32. The protein eluted with higher concentration of imidazole (0.4 M) in the lysis buffer was almost 40 mL was passed through anion exchange chromatography by using Mono Q 5/50 GL column (Fig 33). The purity of the eluted fraction from anion exchange chromatography was evaluated by 15% SDS-PAGE gel (Fig 34). A fraction 1b was pooled and concentrated up to 500 μL and passed through size exclusion chromatography by using Superdex 75 GL column (Fig 35). The purity of the protein was analyzed by 15% SDS-PAGE gel (insert fig 35).

Figure 31: SDS-PAGE 15%, Lane M is the molecular weight markers. Lane 1a, 2a and 3a represents the elution of protein with 0.04 M, 0.06 M and 0.4 M imidazole concentration.
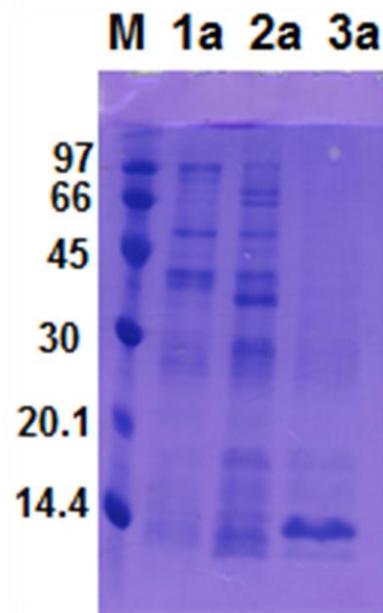
Figure 32: Anion-exchange chromatographic profile. The fraction 3a from Ni-NTA chromatography was passed through Mono Q 5/50 GL column. The fraction separated into two peaks, 1b and 2b. The blue and green lines present the absorbance at 280 nm and concentration of buffer B respectively.
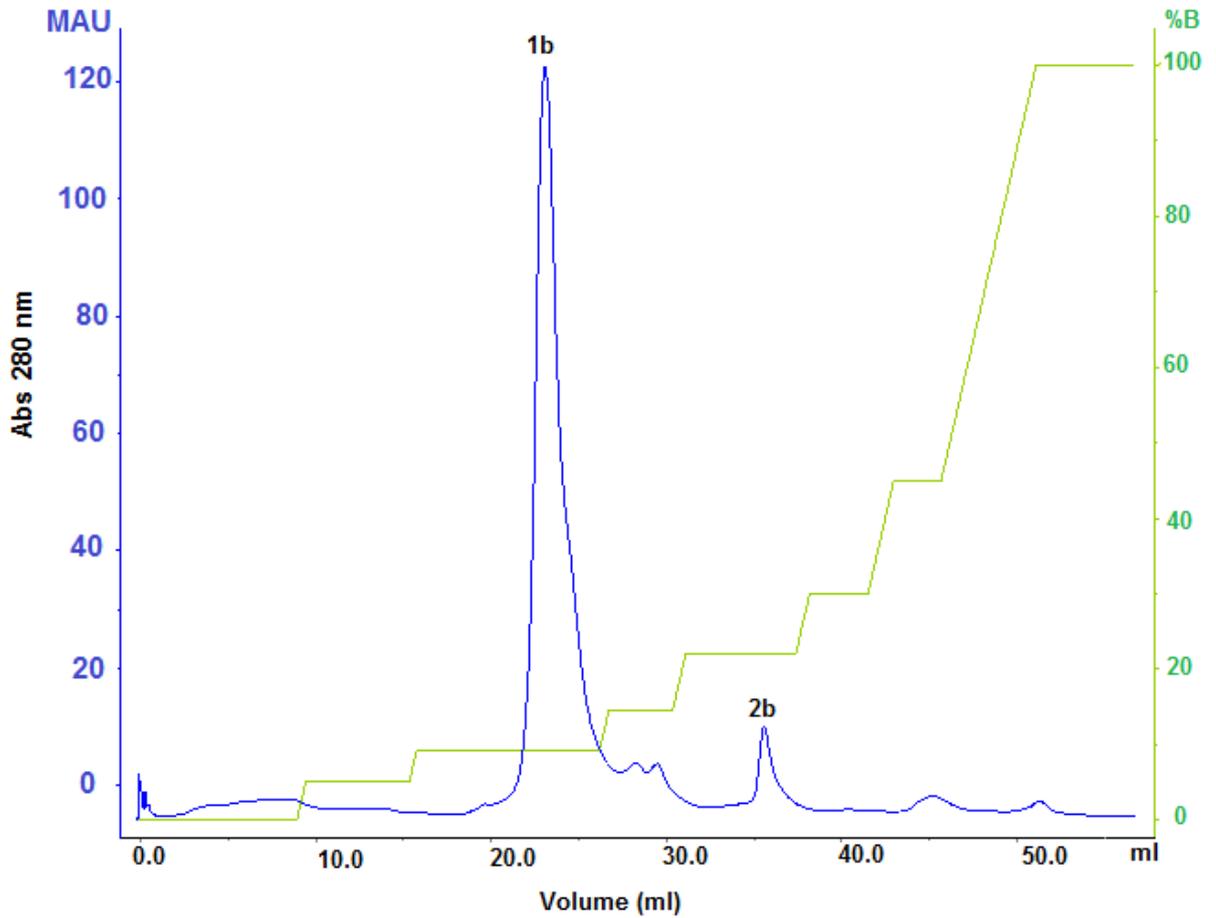
Figure 33: SDS-PAGE 15%; Lane M represents molecular weight markers in kDa, lane 1b and 2b corresponds to the peaks labelled in the chromogram of anion exchange chromatography
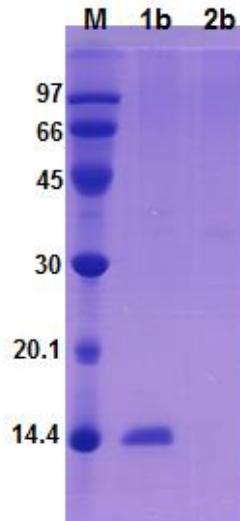
Figure 34: Size exclusion chromatographic profile of peak 1b from anion exchange chromatography. **Insert** 15% SDS-PAGE gel, lane M: molecular weight markers (kDa), 1c represents the corresponding peak 1c protein
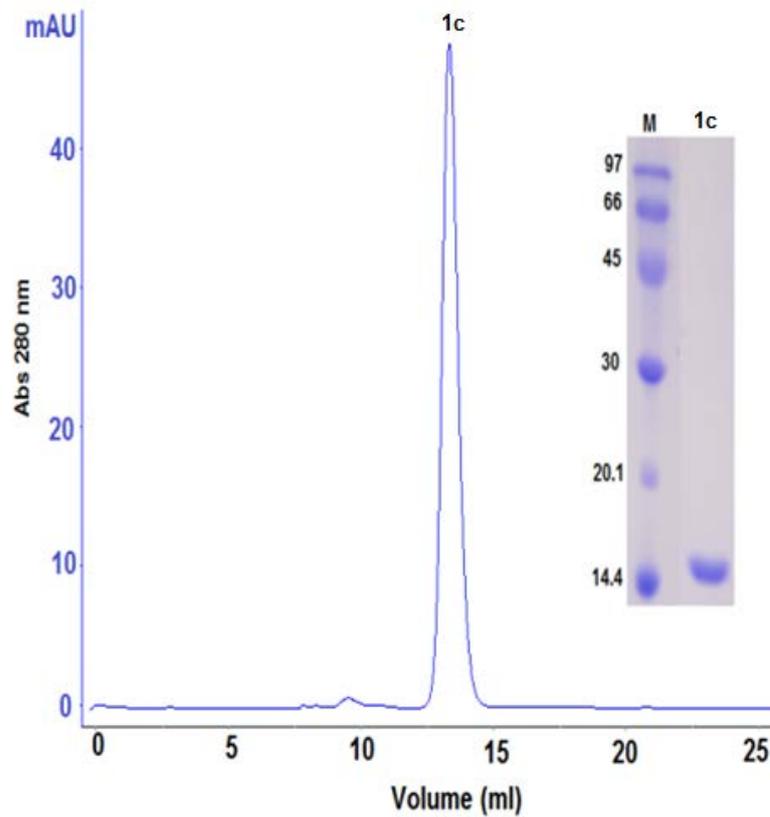
3.3.3) Circular dichroism

The circular dichroism experiments were performed on Jasco J-815 spectropolarimeter (Jasco, USA) equipped with a Peltier-type temperature control system and a 2 mm path-length quartz cuvette. The far UV-CD spectra of the protein were collected from 260 to 198 nm for the secondary structure and quality of the protein at 25 °C and pH 7.4. The CD scan and percentage content of secondary structure is shown in the figure 36.

Figure 35: Far UV-CD spectra of Trx. Insert: Percentage values of each content of the secondary structure.
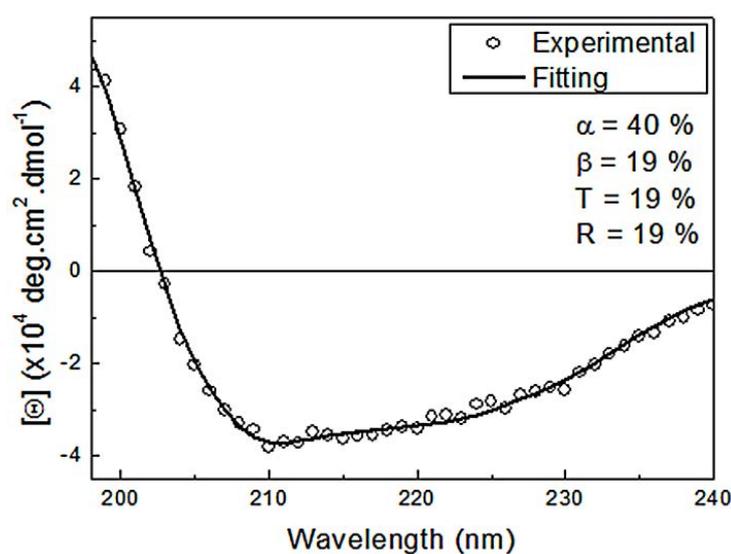
3.3.4) Crystallization

Pure protein (Trx) obtained from size exclusion chromatography was concentrated to 17.33 mg/mL and was applied to crystallization by a manual hanging drop method. The crystals (Fig 37) did not diffract. Crystallization conditions are given in table 11.

Table 11: Crystallization conditions

	Initial conditions	**Optimization**
Method	Hanging-drop vapor-diffusion	Hanging-drop vapor-diffusion
Plate type	24 well tissue culture plates	24 well tissue culture plates
Temperature (K)	298	298
Protein concentration (mg/mL)	17.33	17.33
Buffer composition of protein solution	0.02 M Tris–HCl pH 7	0.02 M Tris–HCl pH 7.7
Composition of reservoir solution	0.02 M Tris–HCl pH 8.5, 1.8 M ammonium sulfate	0.02 M Tris–HCl pH 8.2, 1.8 M ammonium sulfate
Volume and ratio of drop (μL)	0.8 : 0.8	2 : 2
Volume of reservoir (μL)	500	500

Figure 36: Microphotograph of Trx crystal approx 300 μm)

Chapter 2

Snake Venom

https://ask.extension.org/questions.com

1) Introduction

1.1) Snake venoms

From centuries, snakes are the center of attraction for human beings, by playing significant roles in religions and cultures both as good and evil (Stanley, 2008).

Venom is a modified saliva and complex mixture of chemical and biological compounds, (Bottrall *et al.*, 2010) which are pharmacologically active. Snake venoms affect hemostasis, neuromuscular transmission, cardiovascular system, (Eggertsen *et al.*, 1980). Venoms contain poisonous glycosides, which has a carbohydrate molecule with cardiac effects (Koh *et al.*, 2006). It effect the blood clotting mechanism to such an extent that the victim can die of internal bleeding.. Some other common symptoms are lethargy, headaches, nausea, vomiting, etc. (Jain and Kumar, 2012). The venom of *Crotalus atrox*, commonly known as the western diamondback rattlesnake is one of the example of such type.

They interact with ion channels and affect the central nervous system and brain. It causes a very fast degeneration of the synaptic nerves, and this is the reason for the blockage of nerve impulses sent to and from the brain to the muscles. It also affects the respiratory system and heart failure (Hodgson and Wickramaratna, 2002). The victim has trouble in speech, swallowing, drooling, breathing, convulsions or sometimes-prolonged unconsciousness. Some milder symptoms are dizziness, blurred vision and increased sweating (He *et al.*, 2004) The venom from Cobras, mambas, sea snakes, kraits, and coral snakes has such property.

The Cytotoxic property of venoms damage blood vessels, cells, and tissues (Omran *et al.*, 2004). The symptoms are generally a localized pain accompanied by severe swelling and bleeding, formation of red blisters near the bite area, blue spots appears on the skin due to limited blood circulation, nausea and vomiting. The neurotoxic property of venoms causes muscular necrosis. The symptoms started with drooping eyelid and then thickened-tongue sensation, dry throat, thirst, muscular spasms, loss of breath, brown urine discharge convulsions and kidney failure (Gasanov *et al.*, 2014). The muscle proteins enter the blood stream and kidney overworks, which lead to kidneys failure (Sitprija, 2006). It also causes a stiffness of the jaw, neck, trunk, and limbs, along with severe pain in movement. Venom from *Bothrops moojeni* commonly called Brazilian lancehead snake has such properties.

1.2) Snakebites and their symptoms

Snakebites are a type of injury due to a bite of snake, which are categorized by World Health Organization as a neglected disease. It mainly consists of two punctures of the snake fangs. The most common accidents occur on hands, arms and legs. Snakes bite mainly in two conditions, either for hunt or for protection.

Figure 37: **Most common symptoms of Snakebites (Häggström, Mikael 2014)**
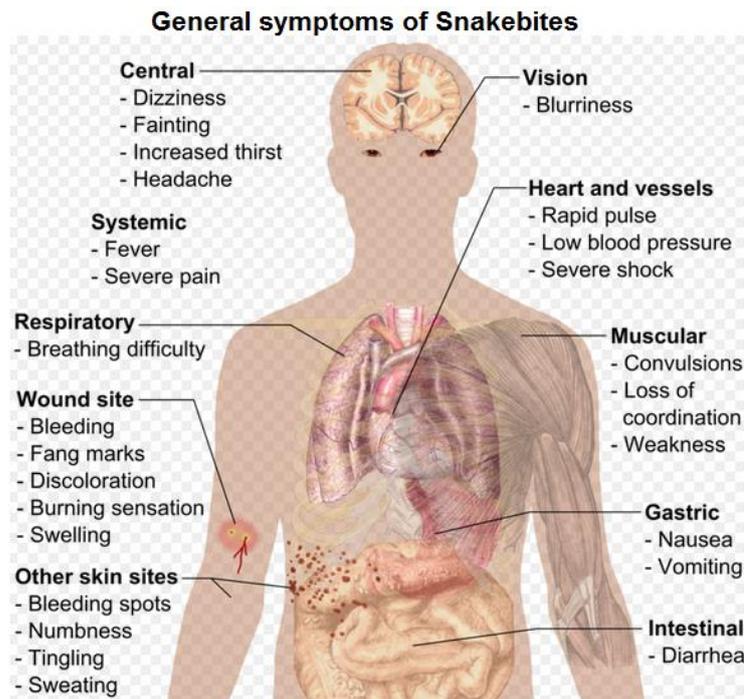
1.3) Compositions of snake venoms

Snake venom is a complex mixture of different proteins and peptides that has active biological and enzymatic activities. The composition of snake venom depends upon the genus, species, subspecies, population and the most important factor that changes the composition from specie to specie is the diet, sex, age and biogeography (Vonk *et al.*, 2011). It consists of proteins, peptides, lipids, carbohydrates, organic compounds and some metal ions such as calcium, magnesium, zinc etc., which contribute to the effects of different proteins and enzymes of the snake venom (Eggertsen *et al.*, 1980) (Vyas *et al.*, 2013). To date more than 100 different proteins are purified from snake venom. These proteins are from some of the families that have enzymes such as, Metalloproteinases, Serine proteinases, L-Amino acid oxidases, group II Phospholipases A_2 and proteins with non-enzymatic activities

such as, ohanin, disintegrins, c-type lectins, natriuretic peptides, myotoxins, cystein-rich secretory protein (CRISP) toxins, nerve and vascular endothelium growth factors, cystatin, and Kunitz-type protease inhibitors etc. respectively (Calvete et al., 2007; Fox and serrano 2005). The proteins with enzymatic activities are lead to the fatal and infirmity effect of the prey and those of nonenzymatic help in immobilization of the prey. Their interaction with different plasma proteins, ion channels and specific receptors also disturb the physiological process of the pray (Mccleary and manjuntha Kini, 2013). Venoms of *Bothrops atrox* and *Bothrops moojeni* were used in the present study of different proteins.

Table 12: Classification of *Bothrops moojeni* and *Bothrops atrox*

Kingdom	Animalia	Kingdom	Animalia
Phylum	Chordata	Phylum	Chordata
Subphylum	Reptilia	Subphylum	Reptilia
Class	Reptilia	Class	Reptilia
Order	Squamata	Order	Squamata
Suborder	Serpents	Suborder	Serpents
Family	Vipride	Family	Vipride
Subfamily	Crotalidae	Subfamily	Crotalidae
Genus	*Bothrops*	Genus	*Bothrops*
Specie	*moojeni*	Specie	*atrox*

Figure 38: Snakes (*B. moojeni, B. atrox*)
(http://animaisdnatureza.blogspot.com.br/2011/11/serpentes.htmL)

(A) *B. moojeni*

(B) *B. atrox*

1.4) Snake venom metalloproteinases

Snake venom metalloproteinases (SVMPs) occupied the major portion in the total mixture of snake venom and are present in both crotalid and viperid species. They are almost 50% of the total proteins present in viperidae species venom (Sousa *et al*., 2015). They initially synthesized as inactive precursors in cytoplasm of secretory cells, later they are subjected to proteolysis in which the cleavage of peptide bond occurs and the N-terminal propeptide is separated from the precursor. The metal ion, certainly the zinc ion is present in metalloproteinases, which coordinates with three Histidine residues, and one or two solvent molecules (Coronado *et al.*, 2014).

The SVMPs are divided into different families and classes according to their primary structure, catalytic site and domain organization. SVMPs belong to the reprolysin family and consider the most abundant protein among different proteins of snake venom (Gutiérrez *et al.*, 2005). SVMPs are mainly responsible for hemorrhagic and proteolytic activities. Due to the disruption of the blood vessels and inhibition of platelet aggregation, they induce both local bleeding and also disturb the entire blood circulation so that in some cases bleeding occurs through every possible opening of a body (Kang *et al.*, 2011)

SVMPs are classified into four major groups (P-I to P-IV) according to their molecular weight, domain structure (Fox and Serrano, 2008). Their molecular weight ranges from 20 to 100 kDa. P-I SVMPS having molecular weight from 20-30 kDa is the simplest protein among SVMPS. They have only a metalloproteinase domain. P-II SVMPS comprised 30-60 kDa proteins having metalloproteinase and followed by disintegrin like domain. P-III group have a molecular weight ranges from 60-80 kDa and along with metalloproteinase and disintegrin domains, they have a cysteine rich domain and the P-IV group have molecular weight from 80-100 kDa and they have an additional lectin-like domain in addition to the metalloproteinase, disintegrin and cysteine rich domain (Cintra *et al.*, 2012).

1.4.1) Snake venom metalloproteinases-PI (PI-SVMPs)

MPIs are the smallest proteins among SVMPs both in size and in domain composition (Fox and Serrano, 2008). It has molecular weight ranging from 20 to 30 kDa and has a single domain which is called metalloproteinase doamin or zinc-dependent catalytic domain (Oyama and Takahashi, 2015). Several PI-SVMPs are (Walburger *et al.*, 2004) purified and characterized from different snake venoms. All of the MPI have some typical characteristics such as fibrinogenelatic activities, disturbance in hemostasis, hemorrhagic and

non-hemorrhagic activities, platelets aggregation etc. (Teixeira *et al.*, 2005). There is a variation in the intensity of the hemorrhagic activities of PI-SVMPs. As they have, only a metalloproteinase domain so the variation in its hemorrhagic activity has been related to the differences in side the domain from one to another PI-SVMP. Beside a vast study on PI-SVMPs, the hemorrhagic and non-hemorrhagic activities and their key component for the variation is still unclear (Escalante *et al.*, 2006).

1.4.2) Snake venom metalloproteinases-III (PIII-SVMPs)

These SVMPs are ranging in size from 60 to 80 kDa and are multi domain proteins (Oyama and Takahashi, 2015). MPIII-SVMPs have a zinc-dependent catalytic domain, disintegrin-like and cysteine-rich domains. All the SVMPs are identical in their catalytic domain structure, but they very in their hemorrhagic activities as well as other characteristics. The MPIII-SVMPs are more hemorrhagic than PI-SVMPs mostly because of the complexity in their domains beside catalytic domain. As well as studies reveling that, MPIII-SVMPs are resistant to inhibition (plasma proteinase inhibitor α_2-macroglobulin), unlike MPI-SVMPs and the reason is the additional domains in MPIII-SVMPS (Escalante *et al.*, 2011).

2) Materials and methods

2.1) Materials

Sephacryl S-100 or S-200 Hiprep 16/60 (GE Healthcare), AKTA purifier (GE Healthcare), Amicon concentrator MWCO 30,000 Da (Millipore), Mono Q^{TM} 5/50 GL (GE Healthcare), molecular weight markers (97 kDa Phosphorylase I, 66 kDa Albumin, 45 kDa Ovalbumin, 30 kDa Carbonic Anhydrase, 20.1 kDa Trypsin inhibitors, 14.4 kDa α-lactalbumin) from GE Healthcare). All the chemicals were of analytical grade and purchased from Sigma Aldrich.

2.1.1) Snake venom extraction

Lyophilized crude *B. atrox* and *B. moojeni* venom was purchased from the Sanmaru Serpentarium (SANMARU Ltda, Taquaral, São Paulo, Brazil.

2.2) Methods

2.2.1) Purification of class P-I metalloproteinase (Atroxlysin-I) from *Bothrops atrox* venom

This protein was purified through two different chromatographic techniques using Akta purifier and Sephacryl S-200 (GE life science, USA) and Mono Q 5/50 GL column.

2.2.1.1) Size exclusion chromatography

Crude venom from *Bothrops atrox* (200 mg) was dissolved in 1 mL of a 0.02 M Tris–HCl buffer, pH 8 containing 0.1 M NaCl and was centrifuged at 10,000 x g for 10 min. The clear supernatant was subjected to size-exclusion chromatography on a Sephacryl S-200 column (GE Life science, USA) previously equilibrated with the same buffer. Elution was performed at a flow rate of 0.2 mL/min, the absorbance was monitored at 280 nm.

2.2.1.2) Anion exchange chromatography

The factions corresponding to peak 3a from size-exclusion chromatography were pooled and fractionated on a Mono Q 5/50 GL Ion exchange column,using a 0.02 M Tris–HCl, pH 8 as buffer A and 0.02 M Tris–HCl, pH 8 containing 1 M NaCl as buffer B. The column was previously equilibrated and washed with a buffer A till the baseline returned to the initial value. The bound protein fractions were then eluted with a non-linear gradient of buffer B with a flow rate of 0.6 mL/min, the absorbance was monitored at 280 nm.

2.2.1.3) Electrophoresis SDS-PAGE

All the fractions collected from Size exclusion and Ion exchange chromatography were analyzed by electrophoresis in the presence of SDS and denaturing agent (β-mercaptoethanol) following the methodology described by (LaemmLi, 1970). The samples were prepared by mixing 10 µL of the protein and 5µL sample buffer (0.125 M Tris-HCl, pH 6.8, containing 4% SDS, 20% glycerol, 10% β- mercaptaethanol and 0.001 M bromophenol blue). The samples were heated for 5 min at 100 °C and centrifuged at 1000 × g for 5 min. Ten microliter of the supernatant of each sample was applied in gel. The electrophoresis was done on Hoefer mini VE Amersham Biosciences GE. The gels were stained in a mixture of methanol, water and acetic acid (50:40:10) and 0.25% comassie brilliant blue R-250, for 1 hour and distained with methanol and water (1:1).

2.2.1.4) Fluorescence spectroscopy

The fluorescence measurements were determined by using an ISS PC1 steady-state spectrofluorimeter (Champaign, IL, USA) equipped with a quartz cell of 1 cm path length and a Neslab RTE-221 thermostat bath. Both excitation and emission bandwidths were set at 8 nm. The excitation wavelength at 295 nm was chosen to excite only tryptophan residues of Atroxlysin-I. The emission spectrum was collected in the range of 305 – 500 nm with the increment of 1.0 nm, which was corrected for the background fluorescence of the buffer and for inner filter effects (Lakowicz, 1999). Each point in the emission spectrum is the average of 10 accumulations.

2.2.1.5) Fluorescence quenching of Atroxlysin-I by suramin

In the fluorescence quenching experiments, the titrations were performed by adding small aliquots from Suramin stock solution (23.6 mM) to enzyme solution (2 mL) at constant concentration of 2.5 µM. Suramin concentration varied from 0 to 37.7 µM at 298 K (25 °C). The interaction between Atroxlysin-I and Suramin was investigated by using the fluorescence quenching method. The concentration of the inhibitor (suramin) was constantly increased to observe the fluorescence intensity of Atroxlysin-I.

2.2.1.6) Crystallization

The atroxlysin-I sample was dialyzed against 0.02 M Tris–HCl pH 8.0 and concentrated to 16 mg/mL in microconcentrators with a 10 kDa cut off (AMICON). Crystallization experiments were performed by the hanging-drop vapor-diffusion method using 24-well tissue-culture plates (Jancarik and Kim, 1991) using commercially available crystallization screens such as crystal screens 1 and 2, polyethylene glycol (PEG) 6000, ammonium sulfate kits, PEG/ion screen (Hampton Research) and the PEG suite (Qiagen). Typically, 0.8 µl drop of the protein solution was mixed with equal volume of the screening solution and equilibrated over a reservoir containing 0.5 mL of the latter solution. Small crystals were obtained in the condition containing 0.2 M ammonium sulfate, 0.1 M Na-cacodylate and 20% *(w/v)* PEG 8000, pH 6.5. This condition was optimized by variation in pH and concentration of PEG. Crystals suitable for X-ray diffraction experiments were obtained with the reservoir contained 0.2 M ammonium sulfate, 0.1 M Na-cacodylate and 23% *(w/v)* PEG 8000, pH 6.0.

Table 13: Crystallization conditions

	Initial conditions	**Optimization**
Method	Hanging-drop vapor-diffusion	Hanging-drop vapor-diffusion
Plate type	24 well tissue culture plates	24 well tissue culture plates
Temperatura (K)	298	298
Protein concentration (mg/ml)	16	16
Buffer composition of protein solution	0.02 M Tris–HCl pH 8.0	0.02 M Tris–HCl pH 8.0
Composition of reservoir solution	0.2 M ammonium sulfate, 0.1 M Na-cacodylate and 20% (*w/v*) PEG 8000, pH 6.5	0.2 M ammonium sulfate, 0.1 M Na-cacodylate and 20% (*w/v*) PEG 8000, pH 6.2
Volume and ratio of drop (μL)	0.8 : 0.8	2 : 2
Volume of resorvour (μL)	500	500

2.2.1.7) Data collection, processing and structure determination

For X-ray diffraction and data collection, an atroxlysine-1 crystal was directly flash-frozen in a 100 K nitrogen-gas stream at the W01B-MX2 beamLine at the Brazilian Synchrotron Light Laboratory (LNLS-Campinas, Brazil). The wavelength of the radiation source was set to 1.458 Å and a MarCCD 165 mm CCD detector was used to record the X-ray diffraction intensities. The crystal was exposed for 30 s per degree of rotation around φ, a total of 161 images were collected and the detector distance was set at 70 mm. The data were indexed, integrated and scaled using the DENZO and SCALEPACK programs from the HKL-2000 package (Otwinowski and Minor, 1997). Data collection and processing statistics are summarized in Table 14. Molecular replacement was carried out using the program MOLREP (Vagin and Teplyakov, 1997) and a model based on the atomic coordinates of BmooMPalpha-I (PDB Code: 3GBO) (Akao *et al.*, 2010).

2.2.2) Purification of P-III SVMP (BmMPIII) from *Bothrops moojeni* venom

This protein was purified through two chromatographic techniques described in item number 2.2.1

2.2.3) Electrophoresis SDS-PAGE

The evaluation of the purity of a protein was done by 12% SDS-PAGE gel as described in item number 2.2.1.3.

2.2.4) Fibrinogenlytic activity in the absence and presence of suramin

Fibrinogenlytic activity was performed by using the method described by Rodrigues et al. (2000). Samples of 10 µL of bovine fibrinogen (1 mg/mL) was dissolved in a buffer (0.05 M Tris-HCl, 0.01 M NaCl pH 7.4) were incubated with different amounts of MPIII (0.01 – 1.0 µg) at 37 °C for 30 min. The reaction was terminated with 10 µL of 0.05 M Tris-HCl pH 8.8 buffer containing 10% (v/v) 2-mercaptoethanol, 2% (w/v) SDS and 0.05% (w/v) bromophenol blue. The samples were then analyzed by SDS-PAGE (12.5%, w/v). The fibrinogenolytic activity in the presence of suramin was also done by using the same method as mentioned above. The enzyme and suramin were mixed in a 1 to 1 ratio [1:1].

2.2.5) Proteolytic activity on casein

The experiment for proteolytic activity was carried out by using the method described by Santana et al. (2008) with minor modifications. We used different pHs instead of using different concentrations of the proteins. Casein solution (1% w/v) was prepared in different pHs (4.6, 5.4, 6.2, 7.0, 8.0, 8.6 and 10.2). The BmMP-III (1 mg/mL (10 µL) were added to this casein solution (100 µL) and was incubated at 37 ˚C for 30 min. The reaction was stopped by adding 5% TCA (Trichloro acetic acid) (300 µL) to this solution. The samples were kept at rest for 30 min and then centrifuged at 10,000 x g for 10 min. The absorbance of supernatant was measured specrophotometrically at 280 nm. The control experiment was done by using only casein solution without metalloproteinase. This experiment was repeated in triplicates.

2.2.6) Dynamic light scattering

Dynamic light scattering (DLS) measurements were performed on Dynapro Molecular Dimension instrument at room temperature. The BmMP-III samples were

centrifuged at 15000 x g for 20 min and illuminated in quartz cell of 1 cm. Data were collected for every 20 s, with at least 20 acquisitions. The diffusion coefficient (DT) was determined from the decay rate distribution of the correlation of intensity profiles and were used to calculate the hydrodynamic radius (Rh) of the protein using the Stokes-Einstein equation (DT = kBT $(6\pi\eta Rh)^{-1}$ where T is the temperature in Kelvin, kb Boltzmann constant and ŋ is the solvent viscosity. The analysis was performed by using the software Dynamics V6.3.40.

2.2.7) Circular dichroism

Far Ultra Violet- Circular dichroism (UV-CD) spectra were recorded at room temperature on a Jasco J-710 spectropolarimeter (Jasco, Tokyo, Japan) equipped with 0.5 mm path length quartz cells, in the range from 185 – 260 nm at a scan rate of 50 nm/min, response time of 1 second, spectral bandwidth of 1 nm and spectral resolution of 0.1 nm. For each spectrum 10 accumulations were performed. CD spectra of BmMP-III (0.7 μM) were performed in the absence and presence of suramin (66 μM). The contribution of CD spectrum of buffer and suramin was subtracted from free and complexed protein respectively. Secondary structure percentages for each tested condition were calculated with CONTINLL software of CD Pro package, using the reference set of protein SMP50 (Sreerama and Woody, 2000).

2.2.8) Fluorescence spectroscopy

The fluorescence spectroscopy measurements were done on ISS PC1 steady-state spectrofluorimeter (Champaign, IL, USA) as explained in item number 2.2.1.4.

2.2.9) Differential scanning calorimetry

Differential scanning calorimetry experiments were performed by using N-DSC III (TA Instruments, USA) in the temperature range of 10 – 90 °C with heating and cooling scan rate of 1 °C/min. The protein was diluted in acetate buffer (0.02 M $NaC_2H_3O_2$, 0.1 M NaCl, pH 5.6) to a final concentration of 1.3 mg/mL. Both calorimeter cells were loaded with buffer solution and equilibrated at 10 °C for 10 min. The scanned were repeated as described above until the baseline was reproducible. Afterwards the sample cell was loaded with BmMP-III and scanned in the same way. Baseline correction was conducted by subtracting the 'buffer vs

buffer' scan from corresponding 'protein vs buffer' scan. The measurements were performed in duplicate.

3) Results and discussions

Results and discussions of experimental and computational work.

Figure 39: Flow sheet of experimental work
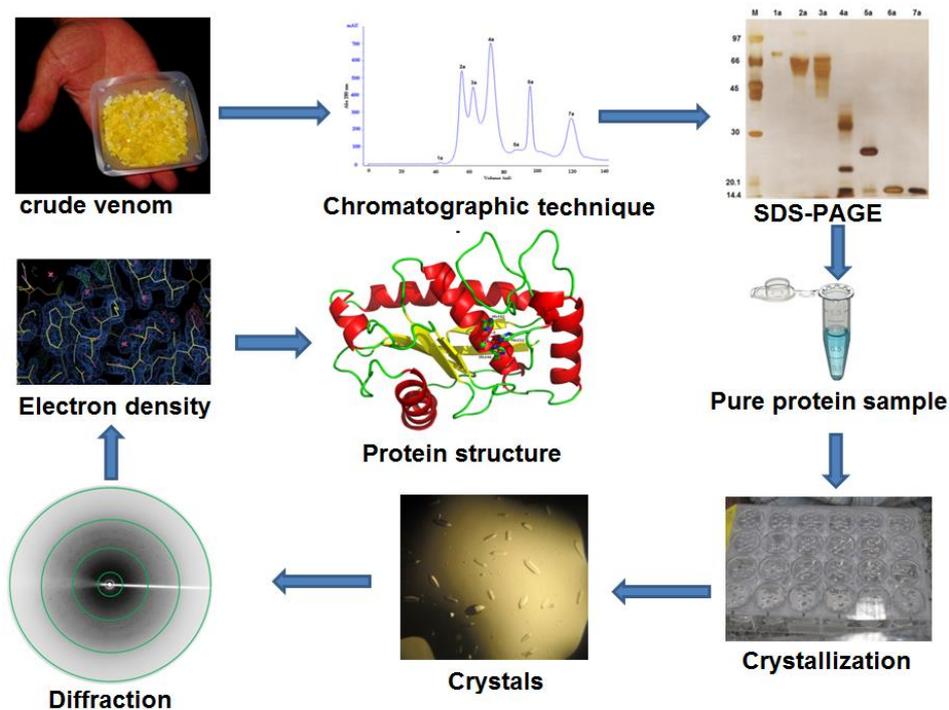
3.1) PI-SVMP (Atroxlysin-I)

3.1.1) Purification of Atroxlysin-I

Class-I metalloproteinase (Atroxlysin-I) from *B. atrox* venom was purified by simple two chromatographic techniques as explained in item numbers 2.2.1.1 and 2.2.1.2. The size-exclusion chromatography resulted in four peaks that were analyzed by SDS-PAGE (Fig 40). The 3rd peak containing Atroxlysin-I (~23 kDa) was applied onto an anion-exchange

column that resulted in further separation of two peaks and the presence of purified atroxlysin-I (> 95%) was confirmed in peak 2b (Fig 41). The total yield of purified protein was 2 mg from 200 mg of crude venom.

Figure 40: Size-exclusion chromatographic profile of *B. atrox* venom on Sephacryl S-200. **Inset.** SDS–PAGE gel of peaks from size-exclusion chromatography. Lane M, molecular weight markers (labelled in kDa); T.P; total soluble protein from crude venom; lanes 1a, 2a, 3a and 4a represent proteins obtained from peak fractions 1a, 2a, 3a and 4a.
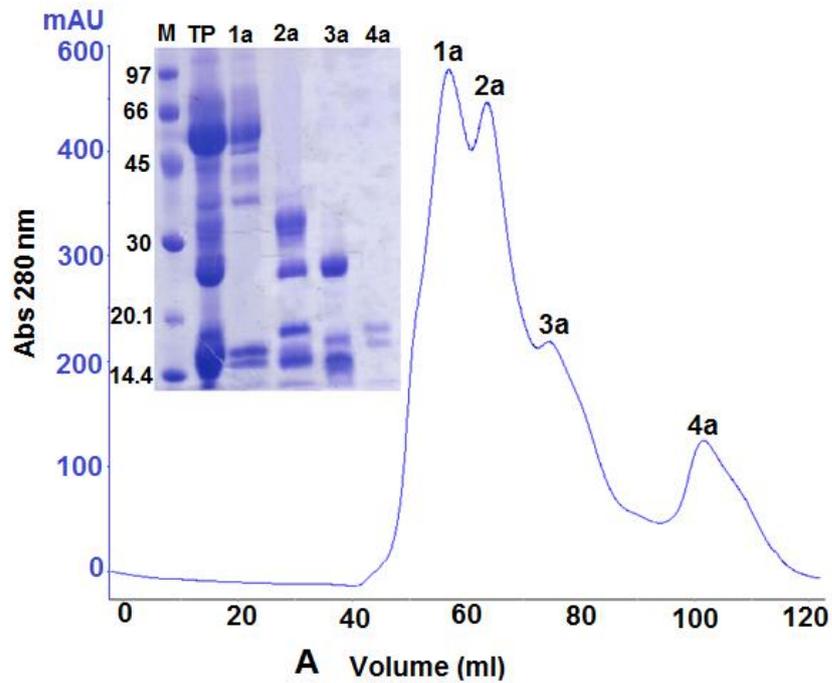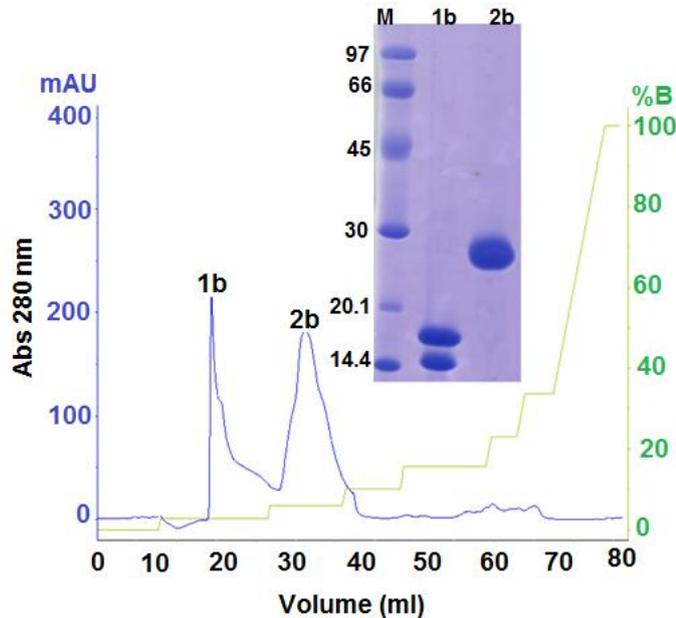
Figure 41: Ion-exchange chromatographic profile of peak 3a from Sephacryl S-200 on a Mono Q 5/50 GL column. **Inset.** SDS–PAGE gel of protein peak fractions from ion-exchange chromatography. Lane M, molecular weight markers (labelled in kDa); lanes labelled 1b and 2b correspond to chromatographic peak fractions with the same label. Lane 2b represents pure Atroxlysin-I.

3.1.2) Fluorescence spectroscopy and fluorescence quenching of Atroxlysin-I by suramin

Fluorescence quenching data were fitted with the Eq. (1) which describes ligand binding to a single protein site (1:1 complex) (Beckett, 2011).

$$(F_0 - F)/(F_0 - F_S) = \left((P_T + L_T + K_{d1}) - \sqrt{(P_T + L_T + K_{d1})^2 - 4P_T L_T}\right)/2P_T \quad (1)$$

Where F is the measured fluorescence intensity in the presence of suramin, F_0 is the starting fluorescence in the absence of Suramin, F_S is the fluorescence of the fully complexed protein (saturated Atroxlysin-I with suramin), K_{d1} is the dissociation constant, P_T is the total concentration of the enzyme, and L_T is the total concentration of suramin. The fluorescence intensity of Atroxlysin-I was decreased with increasing of the suramin concentration, suggesting that the microenvironment of the tryptophan residues of the protein are affected by the presence of suramin shown in Figure 43. The maximum emission wavelength of the enzyme presents a red shift with the increment of suramin concentration, from 340 to 363 nm (insert in Figure. 43A). This red shift indicates that the polarity around the tryptophan

residues of Atroxlysin-I is increased (Lakowicz, 1999). Figure 43B shows the fluorescence quenching changes $((F_0 - F)/(F_0 - F_S))$ at 337 nm in function of the Suramin concentration.

Figure 43: (A) Fluorescence emission spectra of Atroxlysin-I in absence and presence of suramin (λ_{ex} = 295 nm, pH 8.0 and 298 K). [Atroxlysin-I] = 2.5 μM; (A − N) [Suramin] = 0, 2.4, 4.7, 7.1, 9.4, 11.8, 14.2, 16.5, 18.9, 21.2, 23.6, 25.9, 28.3, 33 and 37.7 μM. The insert corresponds to the red shift of the maximum emission wavelength (λ_{max}) of Atroxlysin-I. **(B)** Fluorescence quenching change of Atroxlysin-I (λ_{em} = 337 nm) in function of suramin concentration. Line shows fit of the experimental data using Eq. (1) and R is the correlation coefficient.
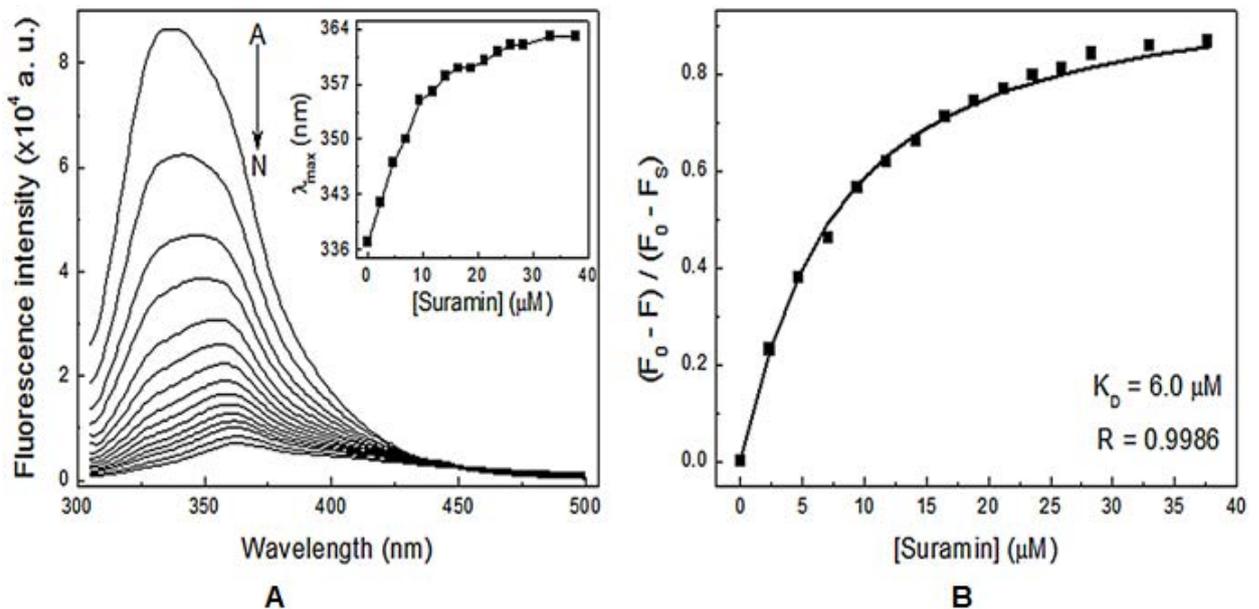
3.2) Crystallization, data collection and structure determination of Atroxlysin-I

Crystals of Atroxlysin-I (Fig 43) were obtained when 16 mg/mL solution of the purified protein was equilibrated with 0.2 M ammonium sulfate, 0.1 M Na-cacodylate and 23% *(w/v)* PEG 8000, pH 6.0.(Table 12).

Figure 44: (A) Microphotograph of Atroxlysin-I crystals, **(B)** X-ray diffraction pattern of an Atroxlysin-I crystal. The concentric circles indicate resolutions of 4.5, 2.2 and 1.8 Å respectively
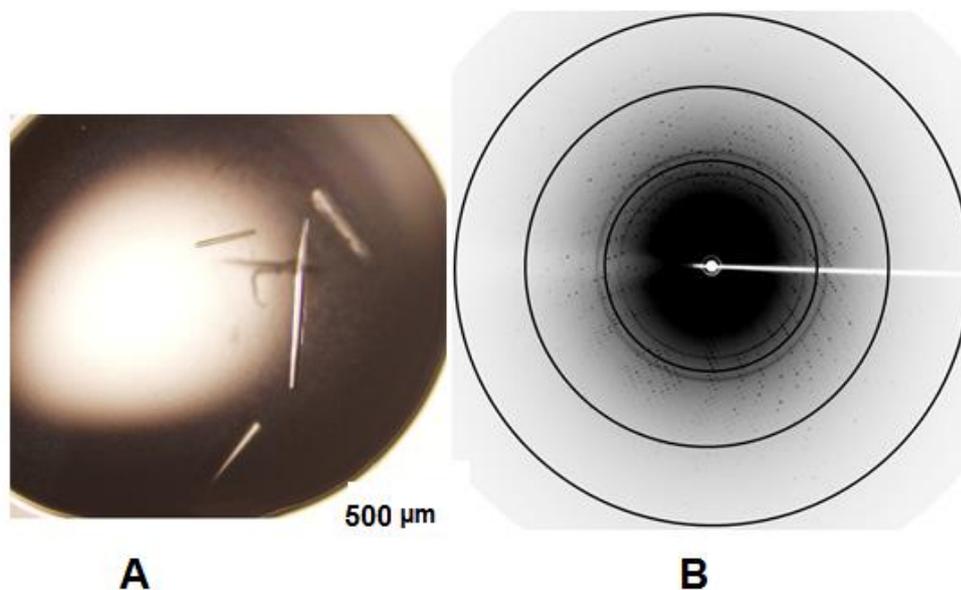
The Atroxlysin-I crystal was diffracted to a maximum resolution of 2.0 Å and the reflections were indexed in the space group $P2_12_12_1$. Taking into consideration the molecular weight (~23 kDa) and the presence of two molecules in the asymmetric unit, it results in a Matthews coefficient (Matthews, 1968) of 2.30 $Å^3$ Da^{-1}, which corresponds to a solvent content of 46.6%. Data collection and processing statistics are presented in Table 14.

The atomic coordinates of the BmooMPalpha-I (PDB Code: 3GBO) (Akao *et al.*, 2010) which shares 67% sequence identity with atroxlysin-I, were used to generate a search model for the molecular-replacement calculations and carried out by using the programme MOLREP (Vagin and Teplyakov, 1997). A clear solution was obtained for the two molecules in the asymmetric unit in space group $P2_12_12_1$. REFMAC5 (Murshudov *et al.*, 1997) was used for the rigid-body refinement and resulted in an R_{factor} of 18.% and R_{free} of 20%.

Table 14: Data collection and processing

Diffraction source:	**Brazilian Synchrotron Light Lab**
Wavelength (Å):	**1.458**
Temperature (K):	**100**
Detector:	**MarCCD 165 mm**
Crystal-detector distance (mm):	**70 mm**
Rotation range per image (°):	**1**
Total rotation range (°):	**180**
Exposure time per image (s):	**30**
Space group:	**$P2_12_12_1$**
a, b, c (Å) :	**61.33, 67.58, 94.52**
α, β, γ (°):	**90**
Mosaicity (°):	**1.04**
Resolution range (Å):	**1.8 (1.90)**
Total No. of reflections:	**28268 (31941)**
No. of unique reflections:	**39956 (5719)**
Completeness (%):	**99.2 (98.8)**
Redundancy(Multiplicity):	**5.7 (5.6)**
$\langle I/\sigma(I)\rangle$:	**9.2 (2.1)**
$R_{\text{r.i.m.}}$:	**4.4 (28.4)**
Overall B factor from Wilson plot :	**19.3 (Å^2)**

Values for the outer shell are given in parentheses.

The 3D structure of atroxlysin-I was determined at 1.8 Å by molecular replacement (3GBO). The final structure contains 201 amino acid residues, one Zn^{+2}, one $Ca^{+}2$, 304 water molecules, eleven sulfate ions, seven ethylene glycols and five PEG molecules. The quality of the final model was checked by PROCHECK (Laskowski *et al.*, 1993). The Ramachandran plot shows that >95% of the amino acid residues are in allowed region (Fig 45). The overall structure of Atroxlysin-I pretends all the properties as shown by the other snake venom metalloproteinases I. Its structure contains four long alpha helices and five beta strands (Fig 46 and 47). The three disulfide bridges (Cys117–Cys197, Cys157–Cys181 and Cys159–Cys164) stabilized the atroxlysin-I structure. The three Histidine residues that make the catalytic site together with Zn^{+2} are located near the helix three. Zinc ion is pentahedrally coordinated by three Histidine residues and two water molecules. The calcium ion, which is catalytically not important for the enzyme, is heptacoordinated by Asn200, Cys197, Asp193, Asp193, Aps9, and two water molecules.

Figure 45: Ramachandran plot

Figure 46: (A) Overall structure of Atroxlysin-I, **(B)** active site
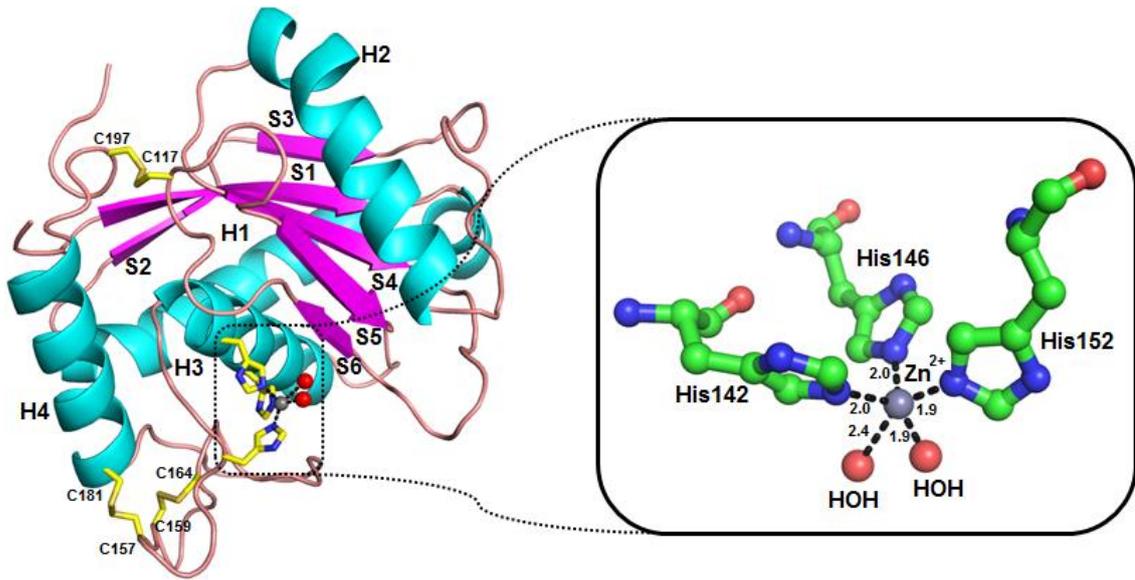
Figure 47: Topology schematic of Atroxlysin-I
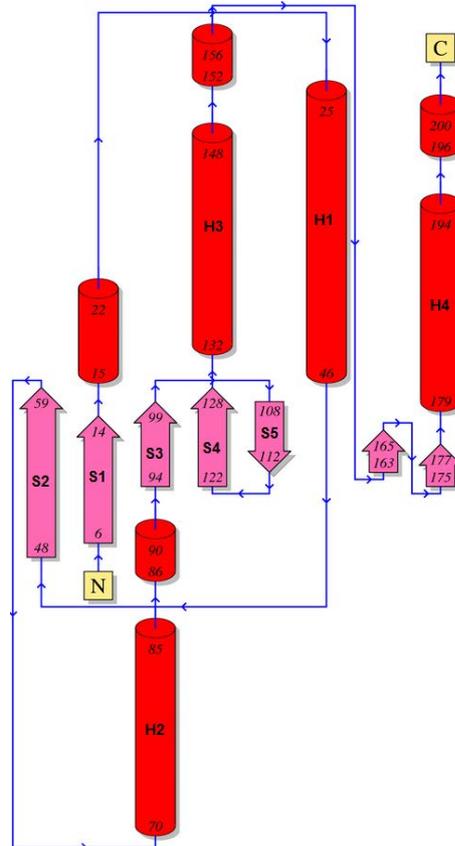
3.3) Sequence and structural alignment and active site comparison

Sequence alignment among various metalloproteinase-I from snake venom (Fig 48) displays an average identity of 60% (Table 15). The three His residues are highly conserved among these enzymes.

Table 15: Sequence identity of different SVMPs-I with Atroxlysin-I

Snake venom metalloproteinases-I	Species	PDB code	Sequence identity (%)
Acutolysin-A	*Agkistrodon acutus*	1BSW	67
BaPI	*B. asper*	2W12	57
BmooMPalpha-I	*B. moojeni*	3GBO	55
Leucurolysin-A	*B. leucurus*	4Q1L	53
Atrolysin-C	*C. atrox*	1HTD	53
Adamlysin II	*C. adamenteus*	2AIG	51
TM-I	*Trimeresurus mucrosquamatus*	4J4M	56
Astacin	crayfish	3LQ0	38
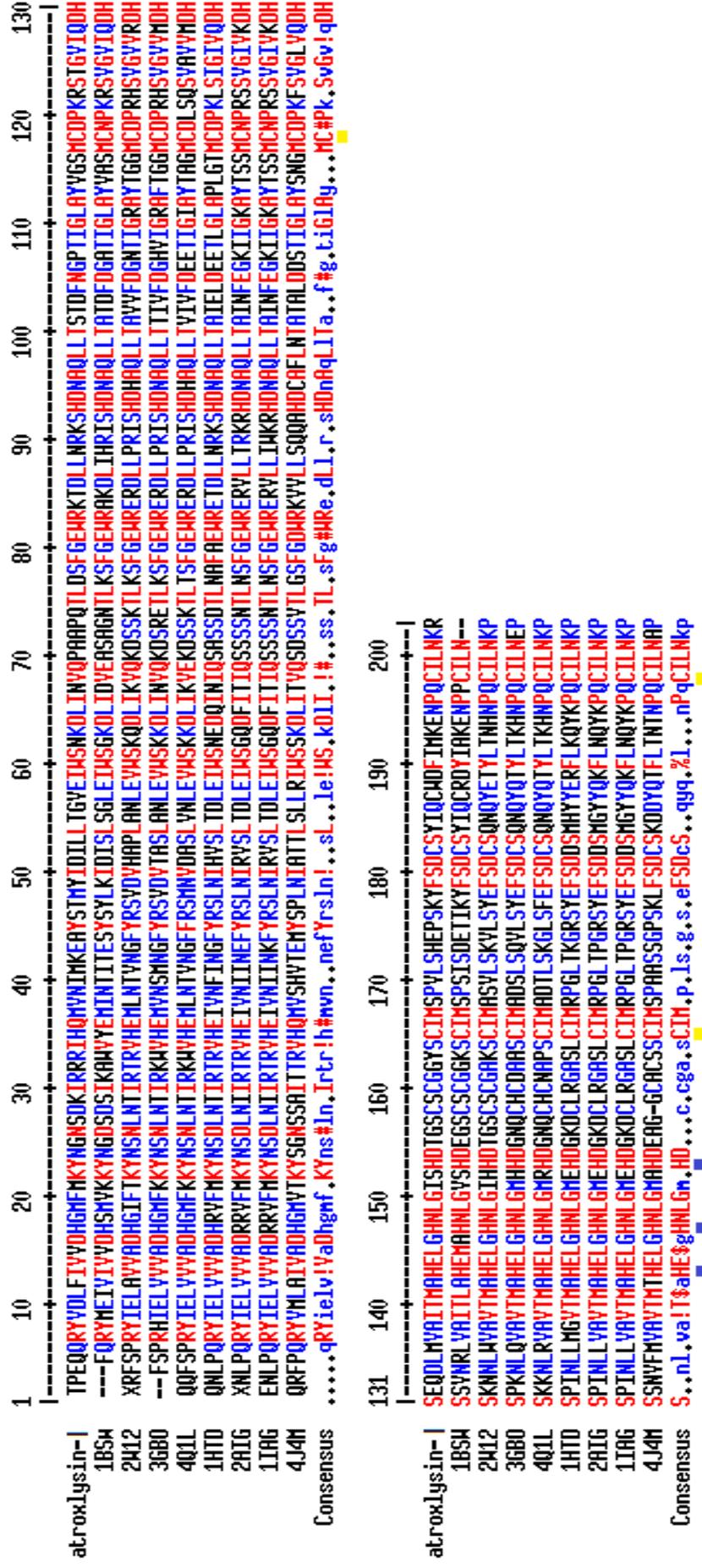
Figure 48: Multiple sequence alignment among various class I metalloproteinases from snake venoms

The 3D structure folds of class I metalloproteinases are very similar, they have very small differences in loop regions and some parts of the loops showing flexibility (Fig 49).

Figure 49: Structural alignment among various SVMPs-I **(A)** Atroxlysin-I (green) and BmooMPalpha-I (cyan), **(B)** Atroxlysin-I (green) and AdamLysin-II (yellow) **(C)** Atroxlysin-I (green) and Astacin (magenta). Difference is shown in orange.
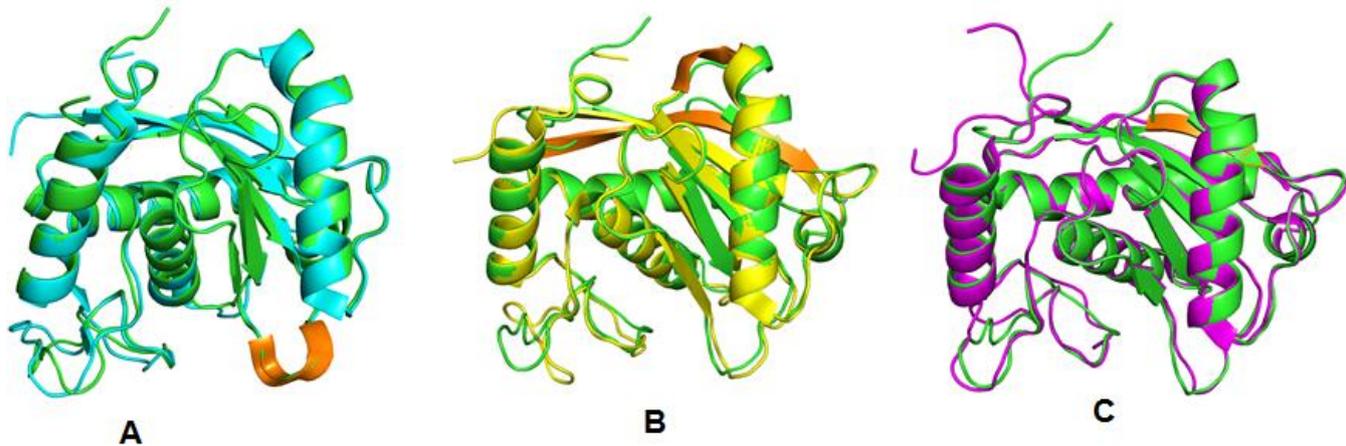
The coordination of Zn^{+2} is completed in different ways in the snake venom metalloproteinases (Fig 50). For example in Atroxlysin-I it is penta-coordinated by three Histidine residues and two water molecules (Fig 50A). In BmooMPalpha-I (Akao *et al.*, 2010) it is hexa-coordinated by three Histidine residues and three water molecules (Fig 50B) while in Adamlysin-II (Gomis-Ruth et al., 1993) it is tetra-coordinated by three Histidine residues and one water molecules (Fig 50C). A very different case has been observed in Astacin (Bode *et al.*, 1992) in which the zinc ion is penta-coordinated by three Histidine residues, one water and one tyrosine residue (Fig 50D).

Figure 50: Active site comparison of different SVMPs-I, **(A)** Atroxlysin-I, **(B)** BmooMPalpha-I, **(C)** AdamLysin-II, **(D)** Astacin
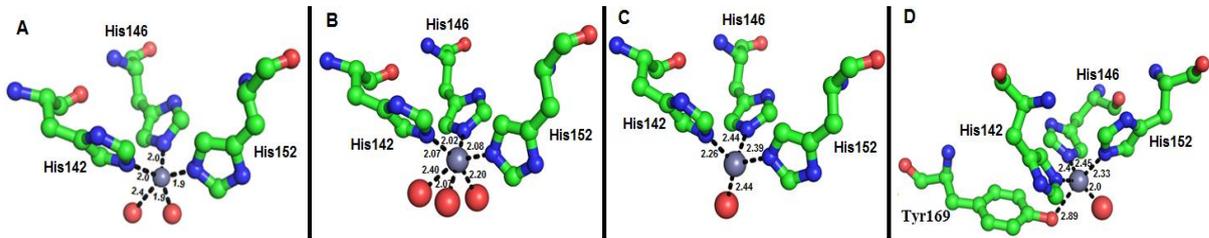
3.4) Hemorrhagic activity in consideration of structure and surface charge analysis

The Atroxlysin-I from *B. atrox* is hemorrhagic while BmooMPalpha-I (PDB; 3GBO) from *B. moojeni* is non- hemorrhagic. Sequence and structural analysis indicate that these two metalloproteinases are identical to each other. Analysis of the loop containing amino acid residues 153–164 and 167–176 in both structures, indicate that the hemorrhagic metalloproteinase in this region contains amino acid residues with small side chain that facilitate the binding of substrate, (Escalante *et al.*, 2011) while in non- hemorrhagic, these regions are dominated by amino acids residues with large side chain that may hinder the binding of substrate to this region (Fig 51). Surface charge distribution also indicates that the A face of Atroxlysin-I is negative as compared to BmooMPalpha-I (Fig 52), which may be relevant to the hemorrhagic activity of this enzyme.

Figure 51: Hemorrhagic activity
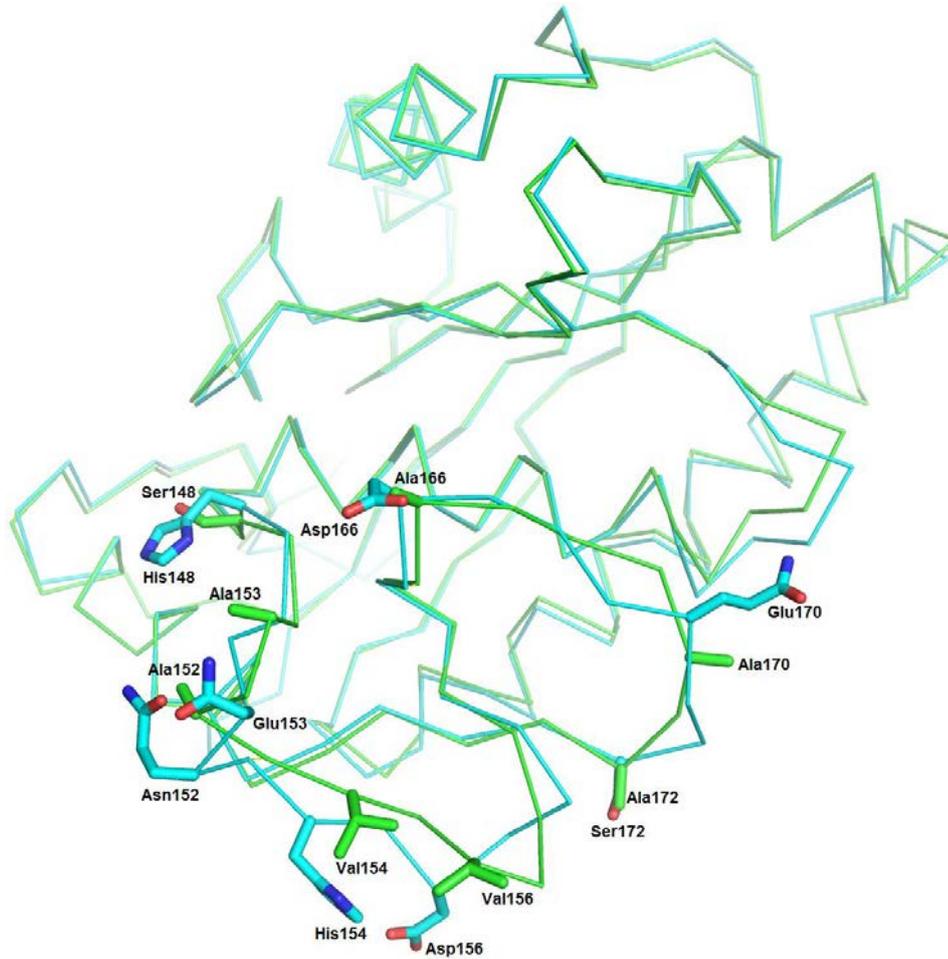
Table 16: Displaying the difference between face A and B between different P-I SVMPs

SVMP-I	PDB code	Face A	Face B
Atroxlysin-I	-	Partially positive, partially neutral	Neutral
BmmoMPalpha-I	3GBO	Partially negative, partially neutral	Partially negative, partially neutral
Adamlysin II	1IAG	Partially positive, partially neutral	Positive
Bap-I	2W12	Partially negative, partially positive	Neutral
Astacin	1AST	Neutral	Partially negative, partially neutral

Figure 52: Surface charge distribution analysis
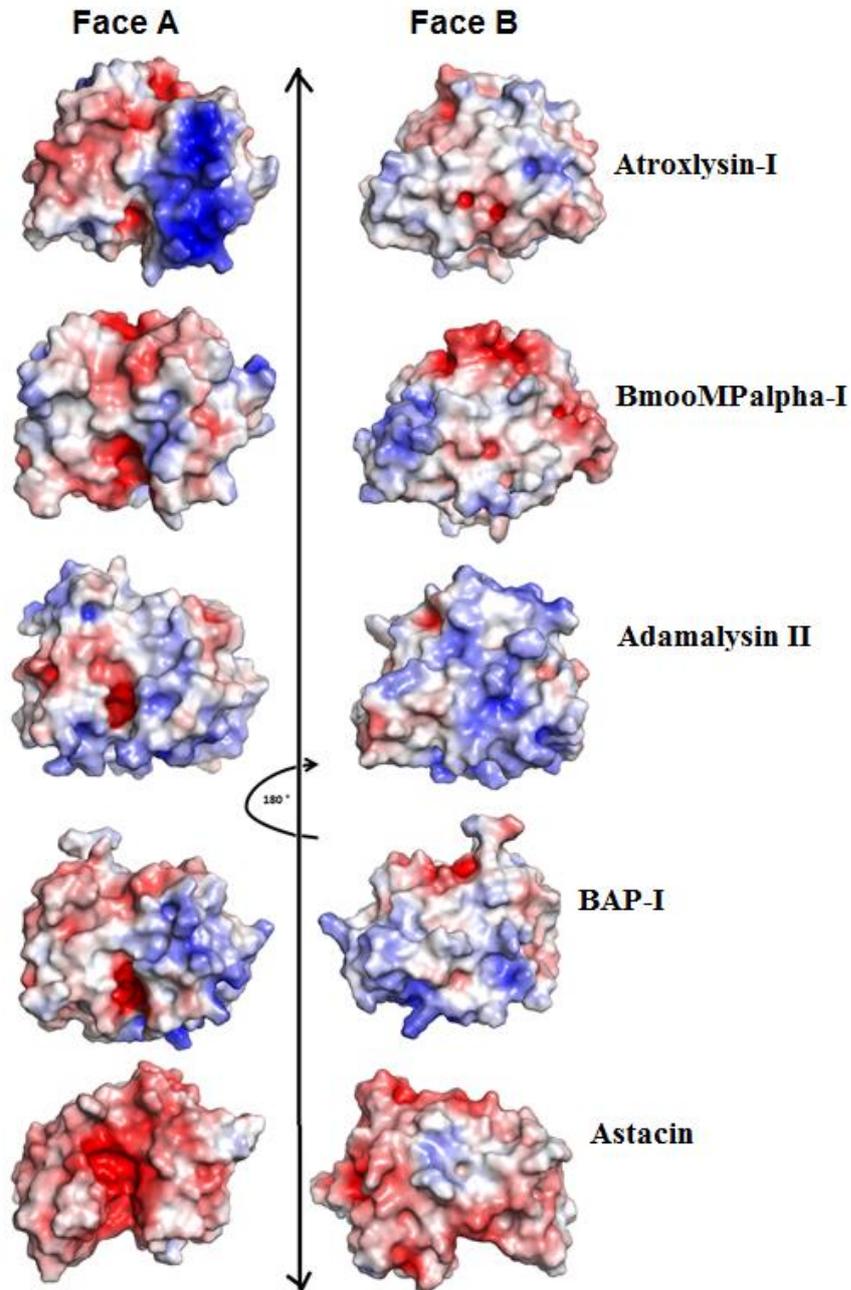
Atroxlysin-I exists as dimer in crystal, which is stabilized by electrostatic attraction between the surfaces of two monomers (Fig 53). Six amino acid residues from each monomer are involved in the stabilization; these residues are in contact through 3 Hydrogen bonding and 44 nonbonding or electrostatic forces of attraction.

Figure 53: Dimerization and stability
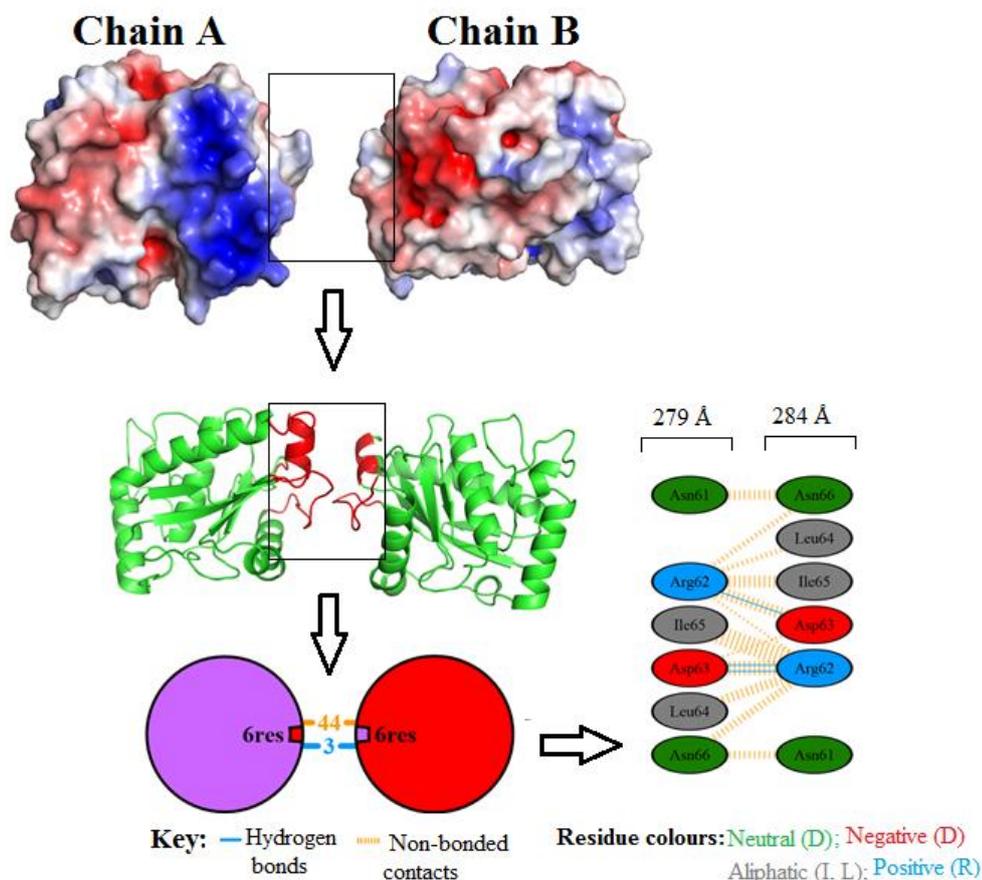
3.5) Metalloproteinase-III from *Bothrops moojeni* (BmMP-III)

3.5.1) Purification and Identification

BmMP-III was purified from the crude venom of *B. moojeni* by two chromatographic steps. The first step consisting of molecular size exclusion chromatography which fractionates the crude venom into six peaks numbered as 1a to 6a, all of them were analyzed by SDS-PAGE Figure 54. The fractions, containing BmMP-III showing M.wt around 60 kDa (corresponding to peak 2a from the molecular size exclusion chromatography) were pooled and further purified on Mono Q™ 5/50 GL column that also result in further separation of five peaks numbered 1b to 5b, all these peaks were analyzed by SDS-PAGE as shown in Figure 55 (A & B). The BmMP-III was detected in the first peak, which was pure as shown by SDS-PAGE. This protein was submitted to MS/MS and fragments found by fingerprint analysis share sequence similarity with snake venom metalloproteinases,

indicating the isolation of a high molecular weight metalloproteinases from the crude venom of *B. moojeni*, which probably corresponds to a MP-III protein, named here as BmMP-III (data not shown).

Figure 54: Size-exclusion chromatographic profile on Sephacryl S-100 crude venom of *B .moojeni* indicating the elution of six peaks. **Insert** SDS-PAGE analysis of peaks from molecular size exclusion chromatography. M; molecular weight markers (kDa), TP; total protein (crude venom). 1a to 6a corresponding peaks protein. Gel stained with silver nitrate.
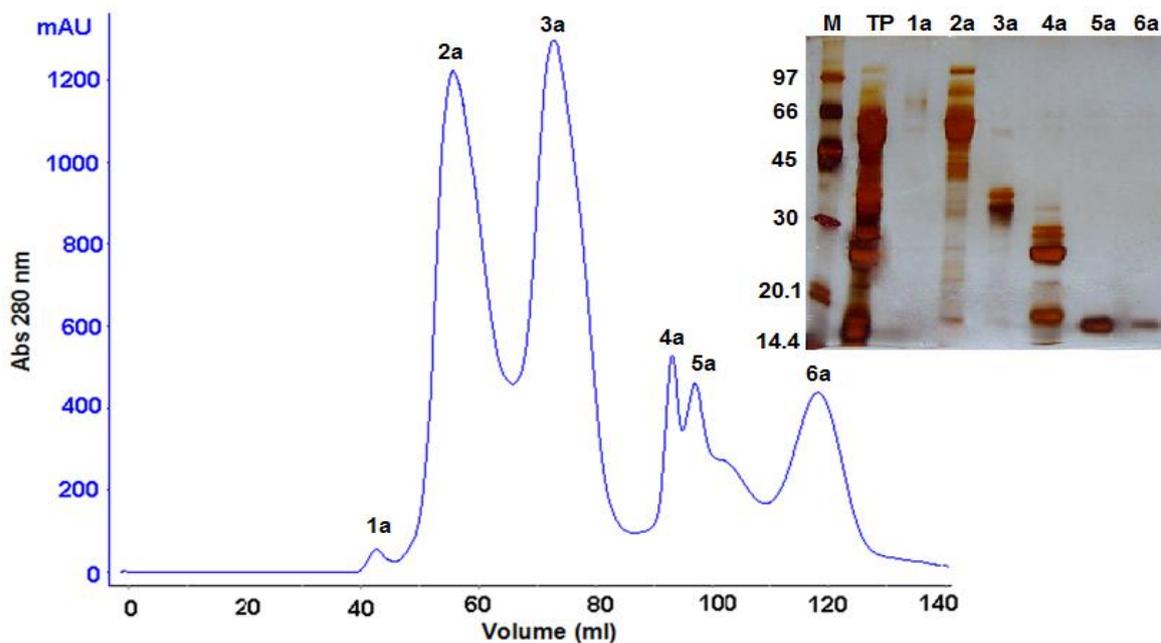
Figure 55: (A): Ion exchange chromatographic profile; Mono Q 5/50 GL (peak 2a MSEC) showing elution of five peaks **(B):** SDS-PAGE analysis of peaks from Ion exchange chromatography TP; total proteins (Peak 1a molecular size exclusion chromatography) 1b to 5b corresponding peaks proteins Lane 1b represent purified BmMP-III.

(A)

(B)
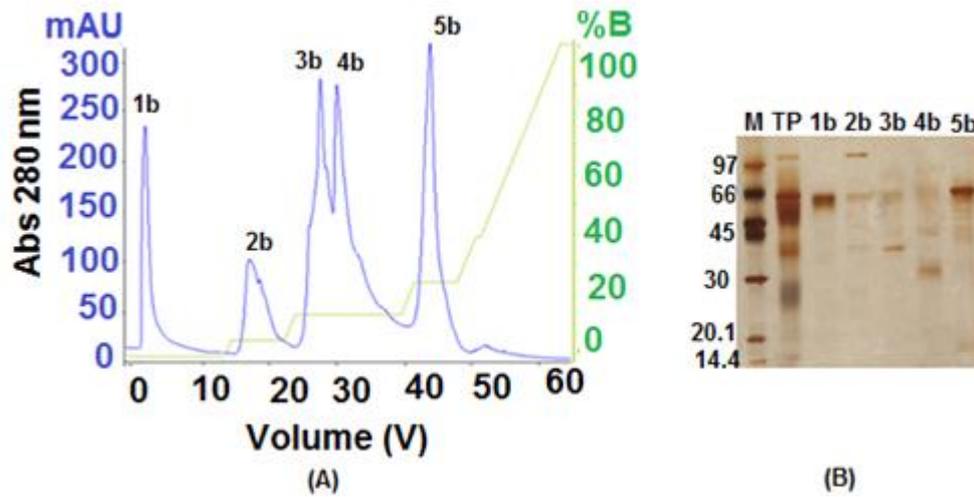
3.5.2) Fibrinogenolytic activity of BmMP-III in the absence and presence of suramin

The purified BmMP-III alone degraded the Aα and Bβ chains of bovine fibrinogen (Fig 56 A) remaining the γ chain unaffected. In the presence of suramin, BmMP-III partially cleaved the Aα and Bβ chains showing that suramin partially inhibit this enzyme (Fig 56 B).

Figure 426: (A) Fibrinogenolytic activity of BmMP-III. 12% SDS-PAGE analysis of bovine fibrinogen degradation. 1a; fibrinogen control (10 µg) without BmMP-III 2a; fibrinogen (10 µg) + BmMP-III (0.05 µg) 3a; fibrinogen (10 µg) + BmMP-III (0.1 µg) 4a; fibrinogen (10 µg) + BmMP-III (0.2 µg) 5a; fibrinogen (10 µg) + BmMP-III (0.4 µg) 6a; fibrinogen (10 µg) + BmMP-III (0.6 µg) 7a; fibrinogen (10 µg) + BmMP-III (0.8 µg) 8a; fibrinogen (10 µg) + BmMP-III (1 µg) 9a; BmMP-III control without fibrinogen (0.6 µg). **(B)** Fibrinogenolytic activity in the precence of suramin 1b; fibrinogen control (10 µg) without BmMP-III, 2b; fibrinogen (10 µL) + BmMP-III + suramin (0.05 µg), 3b; fibrinogen (10 µg) + BmMP-III + suramin (0.1 µg) 4b; fibrinogen (10 µg) + BmMP-III +suramin (0.2 µg) 5b; fibrinogen (10 µg) +BmMP-III + suramin (0.4 µg) 6b; fibtinogen (10 µg) + BmMP-III + suramin (0.6 µg) 7b; fibrinogen (10 µg) + BmMP-III + suramin (0.8 µg) 8b; BmMP-III (0.6 µg).
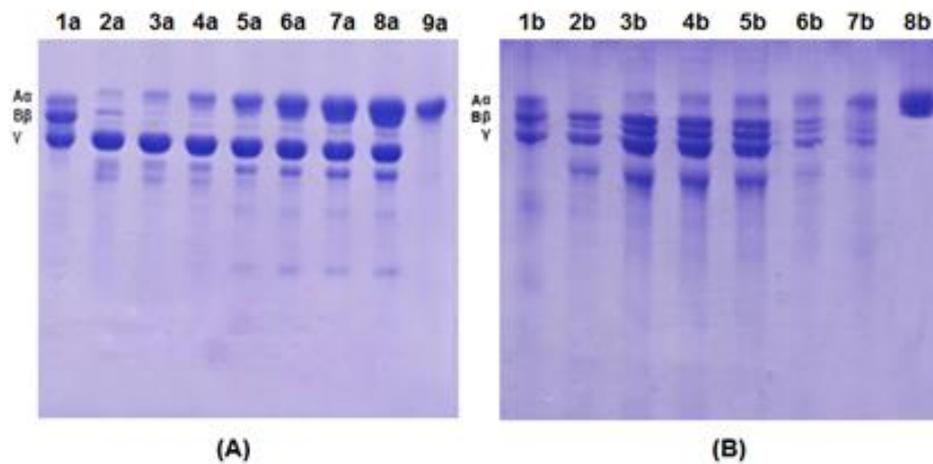
3.5.3) Proteolytic activity on casein

The proteolytic activity of BmMP-III was done by following the method described by Santana et al, (2008) with minor modifications such as using different pH instead of using different concentration to find out the pH at which BmMP-III is highly active. The casein solution was prepared at different pHs (4.6, 5.4, 6.2, 7.0, 8.0, 8.6 and 10.2) and a BmMP-III solution (1 mg/mL and 10 µL). BmMP-III displayed highly proteolytic activity at pH 8.6.

Figure 57: Proteolytic activity histogram of BmMP-III on casein
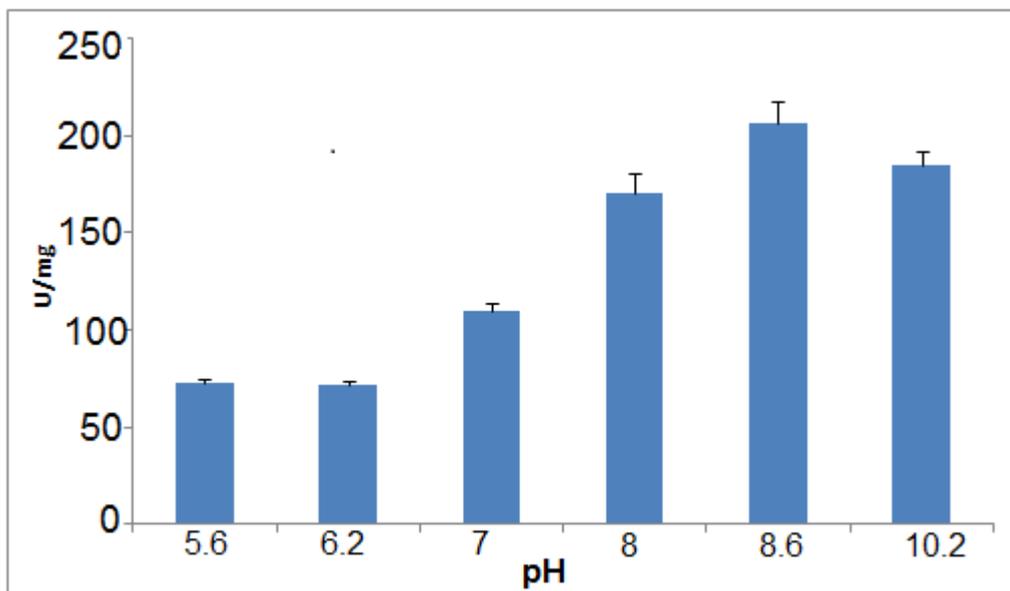
3.5.4) Dynamic light scattering (DLS)

Dynamic Light Scattering experiment shows that BmMP-III is homogenous and monomodal in solution.

Figure 58: (A) Dynamic Light Scattering Correlogram; **(B)** Radius distribution histogram

(A)

(B)
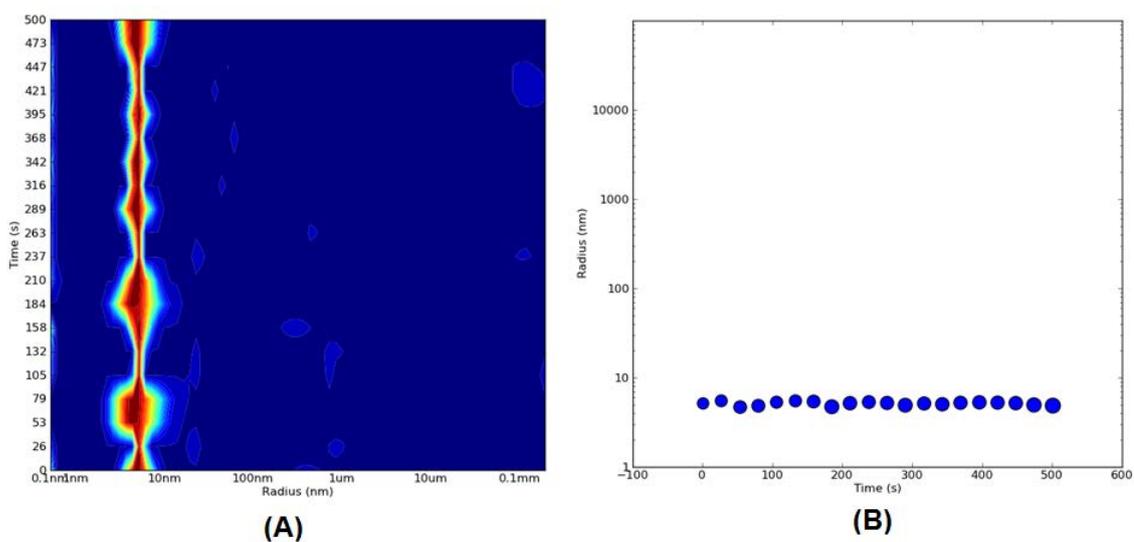
3.5.5) Circular dichroism

Far UV-CD spectra of BmMP-III (0.7 µM) were taken in the absence and presence of suramin (66 µM). The CD spectrum of BmMP-III in the absence of suramin presented structural characteristics of a structured protein since its minimum at 210 nm and positive ellipticity at 193 nm (Fig 59). The addition of suramin to the BmMP-III solution induced folding into an α-helices conformation. The folding of α-helices in BmMP-III, as shown based on an increasing of the positive ellipticity at 193 nm, the minimum shift from 210 to 209 nm, and increasing of the negative ellipticity at 222 nm. Table 17 shows the secondary structure percentages of BmMP-III carried out with the CONTINLL program, indicating that the α-helice propensity of BmMP-III in the presence of suramin, from 25 to 32%. Because of stabilization of α-helice, β-sheet percentage decreased from 26 to 15%. Other secondary structures presented slight fluctuations or decreases with the addition of suramin.

Figure 59: Far UV-CD spectra of BmMP-III (0.7 µM) in the absence (□) or in the presence (◊) suramin (66 µM). The black line represents the best result for curve fit performed by CONTINLL program.

Table 17: Secondary structure percentages of BmMP-III in the absence (1) and presence (2) of Suramin
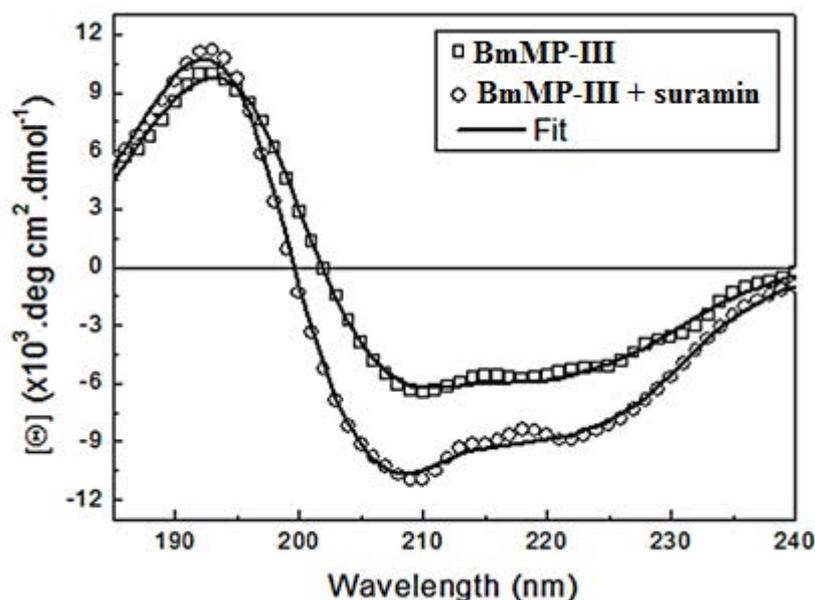
	α-helice (%)	β-sheet (%)	turn (%)	random coil (%)
1	25	26	20	29
2	32	15	22	30

3.5.6) Fluorescence quenching

The fluorescence quenching of the tryptophan residues of BmMP-III by suramin was used to investigate the binding sites. The fluorescence titration data for binding of suramin to BmMP-III are shown in Figure 60. The maximum emission wavelength of the tryptophan residues of BmMP-III presented a red shift with the increment of suramin concentration, from 340 to 363 nm (Fig 60A). Such red shift suggests that the interaction of BmMP-III with suramin resulted in a polar environment for tryptophan residues of the protein (Lakowicz, 1999) Figure 60B shows that $(F_0 - F)/F_0$ change with respect to increasing of Suramin concentration. The binding isotherm for the BmMP-III-suramin complex presented a biophysical profile. This profile corroborates the fact that BmMP-III has different domains in its tertiary structure and consequently, these can become independent and non-equivalent binding sites for suramin.

$$(F_0 - F)/F_0 = f_1 \left[\left((P_T + L_T + K_{d1}) - \sqrt{(P_T + L_T + K_{d1})^2 - 4P_T L_T} \right)/2P_T \right] + f_2 [L_T/(K_{d2} + L_T)] \dots (1)$$

The fluorescence quenching data was fitted with the Equation (1), which is composed of two terms that refers to the site 1 and 2. The first term corresponds to a simple 1:1 binding process with $P_T > K_{d1}$ (high-affinity) and the second term characterized by $P_T \ll K_{d2}$ (low-affinity). For the higher and lower affinity sites, the calculated dissociation constants (K_{d1} and K_{d2}) were 2.24×10^{-8} and 2.95×10^{-5} M, and maximum fluorescence changes (f_1 and f_2) were 0.22 and 0.45, respectively. The maximum fluorescence change associated with Suramin binding at the site 1 and 2 corresponds to 67 and 33 % of total fluorescence change, respectively.

Figure 60: (A) Normalized fluorescence emission spectra of BmMP-III in absence and presence of Suramin (λ_{ex} = 295 nm, pH 8.0 and 298 K). (1) [BmMP-III] = 0.7 µM; (2 – 16) [suramin] = 1.0, 1.9, 3.8, 5.8, 7.7, 9.6, 11.5, 13.4, 15.3, 19.1, 24.8, 32.4, 41.8, 53 and 66 µM. **(B)** Normalized fluorescence change in function of Suramin concentration. Line shows fit of the data using Equation (1). R is the correlation coefficient.
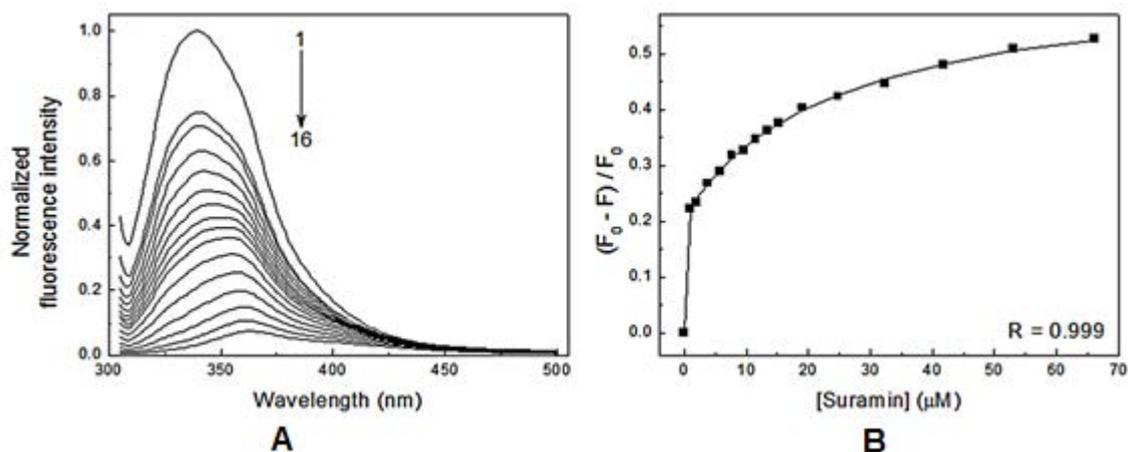
3.5.7) Differential scanning calorimetry

DSC thermogram of thermally induced unfolding of BmMP-III. It was found that the shape of the excess heat capacity function of BmMP-III in the transition temperature range is complex, showing a shoulder on the main transition peak (68 °C). The literature shows that proteins (e.g. immunoglobulins and ovomucoid) containing multiple domains with different thermal stabilities and/or domains that interact with each other usually may present complex thermograms, requiring a multiple curves fitting of the excess heat capacity data for a better understanding of the denaturation process (Privalov, 1979; Privalov and Potekhin, 1986). From the deconvolution analysis, the DSC thermogram of BmMP-III was fitted using three Gaussian curves centered at 36, 46 and 68 °C (Fig 61). This results indicate that BmMP-III can present a three-dimensional structure organized in domains, which is in agreement with the crystallographic data mentioned previously (zinc-dependent catalytic domain, disintegrin-like and cysteine-rich domain). The three overlapping endothermic transitions in the BmMP-III thermogram can suggest differences in the thermal stability of the domains and/or interactions between the domains of this protein.

Figure 61: Deconvolution of the excess heat capacity function of BmMP-III (1.3 mg/mL) in acetate buffer (0.02 M NaC2H3O2, 0.1 M NaCl, pH 5.6) at a scan rate of 1°C/min. The experimentally obtained function is shown by open circles (○), dotted lines (- - -) represent Gaussian adjustment curves and solid line (——) corresponds to the overlap of the summed Gaussians curves.
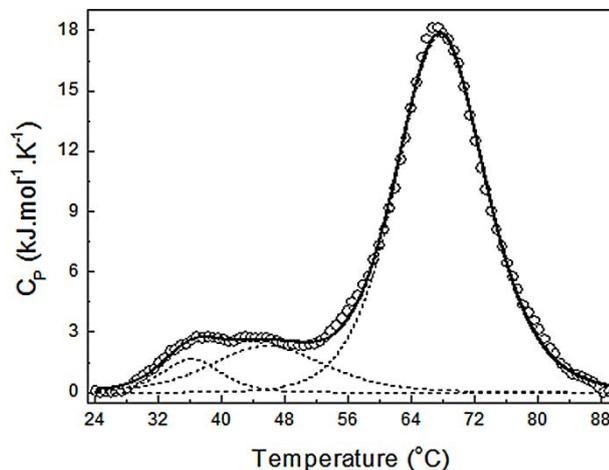
3.6) Purification and crystallization of proteins from other snake venoms

Table 18: Purification and crystallization of proteins from other snake venoms

Species	Proteins
B. atrox	L-amino acid oxidase
B. brazili	Metalloproteinase-I
B. alternatus	Serine proteinase inhibitor

For the purification of these proteins two chromatographic techniques were used.

a) Size exclusion chromatography:

- Akta purifier (GE)

- Column: Sephacryl S-100

- Empty loop: (1 mL)

- Buffer: 0.02 M Tris-HCl, 0.15 M NaCl, pH 8

b) Ionexhcange chromatography:

- Column: Mono Q 5/50 GL

- Empty loop: 50 mL

- Buffer A: 0.02 M Tris-HCl, pH 8

- Buffer B: 0.02 M Tris- HCl, 1 M NaCl, pH 8

The evaluation of the purity of these proteins were checked through SDS-PAGE, crystallized and crystals were diffracted in LNLS (Campinas, Brazil).

Figure 62: (A) SDS-PAGE 12%, displaying pure 60 kDa Laao, **(B)** microphotograph of laao crystal, **(C)** X-ray diffraction pattern.
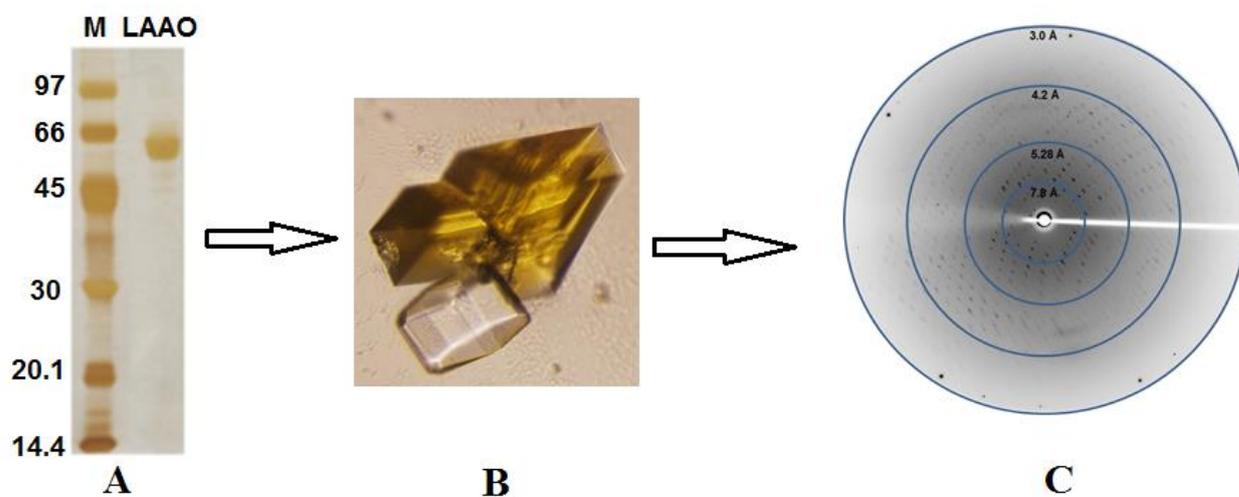
Figure 63: (A) SDS-PAGE 12% displaying a pure 25 kDa metalloproteinase-I, **(B)** microphotograph of MPI crystal, **(C)** X-ray crystallographic pattern.
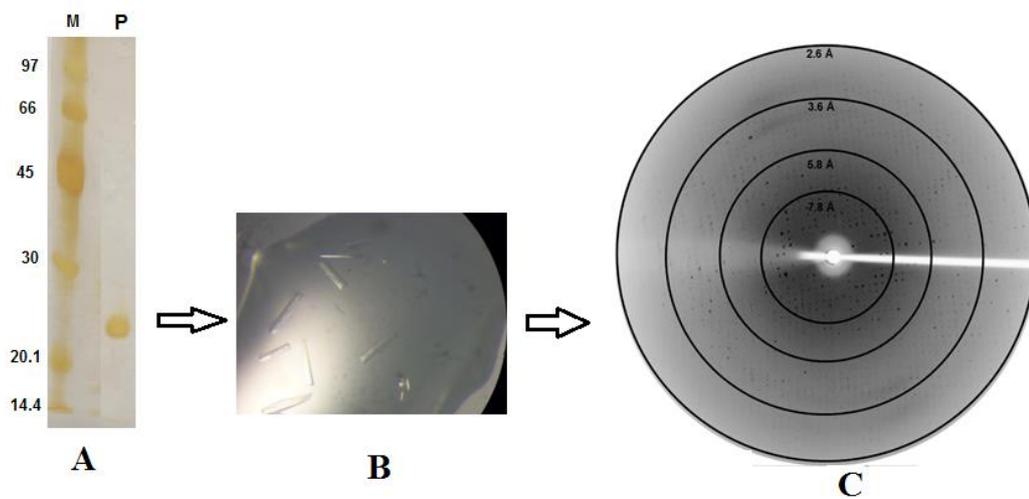
Figure 64: (A) SDS-PAGE 12%, displaying an 18 kDa Serine proteinase inhibitor (SPI), **(B)** microphotograph of SPI crystal, **(C) X**-ray crystallographic pattern.
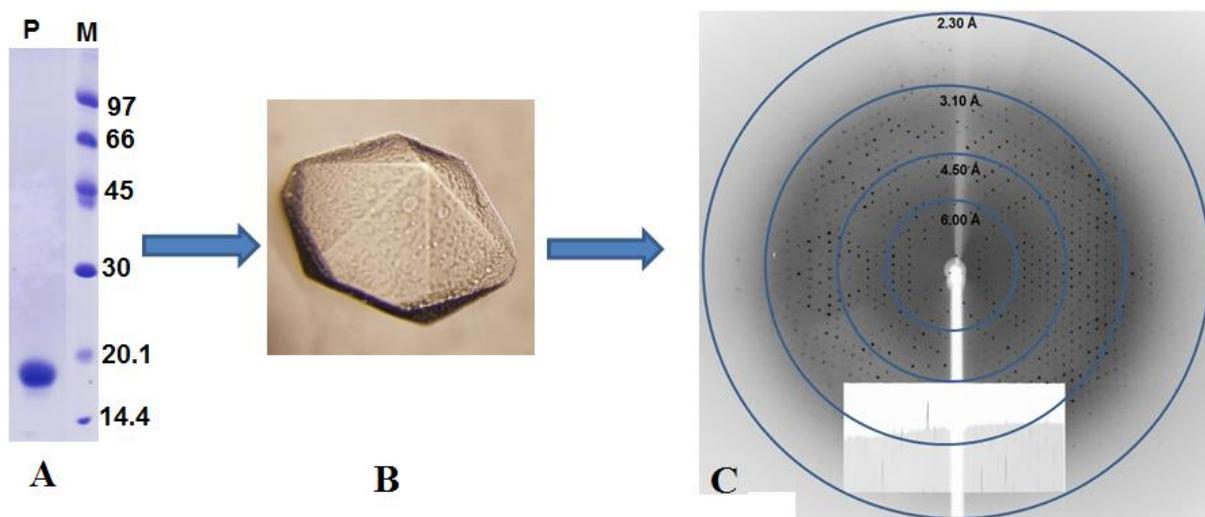
4) Conclusion

1.   In the first step of this project, different DNA plasmids were designed and constructed.

2.   Screening involved different cell lines, expression and purification conditions were performed that allowed to establish specific protocols for the production of recombinant proteins at large scale and with high purity.

3.   The circular dichroism analysis showed that all the proteins were structurally folded and suitable for crystallization trails.

4.   All the proteins were applied for crystallization trials, but only TIM crystals diffracted and its structure was determined at 2.4 Å that is the first crystal structure of any protein from *C. pseudotuberculosis*. Overall structure of TIM was analyzed and its sequence and structural comparison was done with TIMs of other organisms.

5.   The characterization of TrxR showed that this protein is stable up to 60 ˚C at physiological pH, its interaction with silver ions and the CD spectra shows that heavy metal salts can inhibit it. This work will help in further investigation of inhibitors for TrxR and related enzymes.

6.   Five different proteins from snake venoms were purified, crystallized and their diffraction data were collected.

7.   Metalloproteinases (SVMP-I and SVMP-III) from snake venoms (*B. atrox* and *B. moojeni*) were characterized.

8.   Suramin was used as an inhibitor for metalloproteinases in biophysical experiments and it was observed that it partially inhibit BmMP-III.

9.   The overall structure of Atroxlysin-I was resolved and compared with other class P-I metalloproteinases. Its hemorrhagic activity was compared with BmooMPalpha-I, which is a non-hemorrhagic metalloproteinase-I.

5) References

AKAO, P. K. et al. Structural studies of BmooMPα-I, a non-hemorrhagic metalloproteinase from *Bothrops moojeni* venom. **Toxicon,** v. 55, n. 2 - 3, p. 361 - 368, August 23 2010.

Beckett, D. Measurement and analysis of equilibrium binding titrations: A beginner's guide. **Methods in Enzymology,** v. 488, p. 1-16, 2011.

BINNS, S. H.; BAILEY, M.; GREEN, L. E. Postal survey of ovine caseous lymphadenitis in the United Kingdom between 1990 and 1999. **Journal of Veterinary Research,** v. 150, p. 263 - 268, 2002.

BODE, W. et al. Structure of Astacin and implications for activation of Astacins and zinc-ligation of collagenases. **Nature,** v. 358, n. 6382, p. 164 - 167, July 9 1992.

BOTTRALL, J. L. et al. Proteolytic activity of Elapid and Viperid Snake venoms and its implication to digestion. **Journal of Venom Research,** v. 1, p. 18-28, September 30 2010.

BROWN, D. M.; UPCROFT, J. A.; P, U. A thioredoxin reductase-class of disulphide reductase in the protozoan parasite Giardia duodenalis. **Molecular and Biochemical Parasitology,** v. 83, n. 2, p. 211-220, december 20 1996.

CINTRA, A. C. et al. Batroxase, a new metalloproteinase from *B. atrox* snake venom with strong fibrinolytic activity. **Toxicon,** v. 60, n. 1, p. 70 - 82, March 30 2012.

CORONADO, M. A. et al. Three-Dimensional Structures and Mechanisms of Snake Venom Serine Proteinases, Metalloproteinases, and Phospholipase A2s. In: GOPALAKRISHNAKONE, P.and CALVETE, J. J. (Ed.). **Venom Genomics and Proteomics**, 2014. p. 1-25.

EGGERTSEN, G.; FOHLMAN, J.; SJOQUIST, J. *in vitro* studies on complement inactivation by snake venoms. **Toxicon,** v. 18, n. 1, p. 87 - 96, 1980.

ESCALANTE, T. et al. Key events in microvascular damage induced by snake venom hemorrhagic metalloproteinases. **Journal of Proteomics,** v. 74, n. 9, p. 1781 - 1794, April 6 2011.

ESCALANTE, T. et al. Novel insights into capillary vessel basement membrane damage by snake venom hemorrhagic metalloproteinases: a biochemical and immunohistochemical study. **Archives of Biochemistry and Biophysics,** v. 455, n. 2, p. 144 - 153, Nov 15 2006.

GALE, R. **Crystallography made Crystal Clear**. san Diego: Academic press, 1993.

GASANOV, S. E.; Dagda, R. K.; RAEL, E. D. Snake Venom Cytotoxins, Phospholipase A2s, and Zn2+-dependent Metalloproteinases: Mechanisms of Action and Pharmacological Relevance. **Journal of Clinical Toxicology,** v. 4, n. 1, Jan 25 2014.

GASTEIGER, E. et al. Protein Identification and Analysis Tools on the ExPASy Server. In: WALKER, J. M. (Ed.). **The proteomics Protocols handbook**. Totowa. NJ, 2005. p.571-607.

GRÜNING, N. M. et al. Inhibition of triosephosphate isomerase by phosphoenolpyruvate in the feedback-regulation of glycolysis. **Open biology,** v. 4, n. 3, 2014.

GUTIÉRREZ, J. M. et al. Hemorrhage induced by snake venom metalloproteinases: biochemical and biophysical mechanisms involved in microvessel damage. **Toxicon,** v. 45, n. 8, p. 997 - 1011, June 15 2005.

HABUŠ, J. et al. Comparision of the epizootiological and clinical features of caseous lymphadenitisis and morel's disease in goats. **Veterinarski Arhiv** v. 85, n. 2, p. 163-173, 2015.

HARD, G. C. Electron microscopy examination of *Corynebacterium ovis*. **Journal of Bacteriology,** v. 97, p. 1480 - 1485, 1969.

HE, Y. Y.; LEE, W. H.; ZHANG, Y. Cloning and purification of α-neurotoxins from king cobra (*Ophiophagus hannah*). **Toxicon,** v. 44, n. 3, p. 295-303, september 1 2004.

HEWITSON, J. P. et al. The Secreted Triose phosphate isomerase of *Brugia malayi* Is Required to Sustain Microfilaria Production *in vivo*. **PloS Pathogens,** v. 10, n. 2, 2014.

HIRT, R. P. et al. The diversity and evolution of thioredoxin reductase: new perspectives. **Trends in Parasitology,** v. 18, n. 7, p. 302-308, Jul 01 2002.

HODGSON, A. L.; BIRD, P.; NISBET, I. T. Cloning, nucleotide sequence, and expression in Escherichia coliof the phospholipase D gene from *Corynebacterium pseudotuberculosis*. **Journal of Bacteriology,** v. 172, p. 1256 - 1261, 1990.

KANG, T. S. et al. Enzymatic toxins from snake venom: structural characterization and mechanism of catalysis. **The FEBS Journal,** v. 278, n. 23, p. 4544 - 4576, May 11 2011.

KENNELLY, P. J. Protein kinases and protein phosphatases in prokaryotes: a genomic perspective. **FEMS Microbiology Letters,** v. 2002, n. 1, p. 1 - 8, Jan 2 2002.

KOH, D. I. C.; A. ARMUGAM, A.; JEYASEELAN, K. Snake venom components and their applications in biomedicine. **Cellular and Molecular Life Sciences,** v. 63, n. 24, p. 3030-3041, 2006.

LAEMMLI, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. **Nature,** v. 227, n. 5259, p. 680 - 685, Agust 15 1970.

LAKOWICZ, J. R. **Principles of Fluorescence Spectroscopy**. 2nd. New York: Kluwer Academic Publishers/Plenum Press, 1999.

LASKOWSKI, R. A. et al. PROCHECK: A program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography,** v. 26, p. 283-291, 1993.

MURSHUDOV, G. N.; VAGIN, A. A.; DODSON, E. J. Refinement of macromolecular structures by the maximum-likelihood method. **Acta Crystallographica Section D, Biological Crystallography,** v. 53, n. 3, p. 240 - 255, May 1 1997.

OBLONG, J. E. et al. Purification of human thioredoxin reductase: Properties and characterization by absorption and circular dichroism spectroscopy **Biochemistry,** v. 32, n. 28, p. 7271-7277, July 1993.

OLSON, M. E. et al. Biofilm bacteria: formation and comparative susceptibility to antibiotics. **Canadian Journal of Veterinary,** v. 66, p. 86 - 92, 2002.

OMRAN, M. A. A.; FABBS, A.; DICKSON, G. Biochemical and Morphological analysis of cell deathinduced by Egyptian cobra (*naja haje*) venom on cultured cells **Journal of Venomous Animals and Toxins including Tropical Diseases,** v. 10, n. 3, p. 219 -241, Sep 10 2004.

OSMAN, A. Y. et al. The study of caseous of lymphadenitis: dose dependent infection of *C. pseudotuberculosis* in mouse model via oral. **Research Opinions in Animal & Veterinary Sciences,** v. 4, n. 3, p. 163-169, 2014.

PEEL, M. M. et al. Human lymphadenitis due to *Corynebacterium pseudotuberculosis*: report of ten cases from Australia and review. **Journal of Clinical Infectious Diseases** v. 24, n. 2, p. 185-191, Feb 1997.

PRIVALOV, P. L. Advance Protein chemistry. In: GLICK, D. (Ed.). **Methods of Biochemical Analysis**, v.33, 1979. cap. 35, p.01-104.

RADOSTITS, O. M.; BLOOD, D. C.; GAY, C. C. **Veterinary medicine: a textbook of the diseases of catlee, sheep, pigs, goats and horses**. 9. Philadelphia Bailliere Tindall, 2007.

SALIM, S. A. Oedematous Skin Disease of Buffalo in Egypt. **Journal of Veterinary Medicine, Series B,** v. 48, n. 4, p. 241-258, May 2001.

SITPRIJA, V. Sbakebite nephropathy. **Journal of Nephrology,** v. 11, n. 5, p. 442 - 448, 2006.

SOOD, N. K. et al. Mesenteric caseous lymphadentis in a cow calf caused by *Corynebacterium Pseudotuberculosis*: a case report. **Veterinarni medicina,** v. 57, n. 7, p. 371-375, 2012.

STANLEY, J. W. Snakes: Objects of Religion, Fear, and Myth **http://altweb.astate.edu/electronicjournal/Articles/Stanley_snake_myth_paper_r...pdf**, March 4 2008.

TEIXEIRA, C. F. P. et al. Inflammatory effects of snake venom metalloproteinases. **Mem. Inst. Oswaldo Cruz,** v. 100, n. 1, p. 181 - 184, June 14 2005.

THELANDER, L. Thioredoxin reductase. Characterization of a homogenous preparation from *Escherichia coli* B. **The Journal of Biological Chemistry,** v. 242, n. 5, p. 852-859, March 10 1967.

TRUJILLO, C. et al. Triosephosphate Isomerase Is Dispensable *in vitro* yet Essential for *Mycobacterium tuberculosis* To Establish Infection. **American Society for Microbiology,** v. 5, n. 2, p. 00085-14, 2014.

VIZUETE, A. M. et al. Cloning, Expression, and Characterization of a Novel *Escherichia coli* Thioredoxin. **The Journal of Biological Chemistry,** v. 272, n. 49, p. 30841-30847, December 5 1997.

VONK, F. J. et al. Snake venom: From fieldwork to the clinic: Recent insights into snake biology, together with new technology allowing high-throughput screening of venom, bring new hope for drug discovery. **BioEssays,** v. 33, n. 4, p. 269 - 79, Jan 27 2011.

VYAS, V. K. et al. Therapeutic potential of snake venom in cancer therapy: current perspectives. **Asian Pacific Journal of Tropical Biomedicine,** v. 3, n. 2, p. 156 - 162, Feb 28 2013.

WALBURGER, A. et al. Protein kinase G from pathogenic mycobacteria promotes survival within macrophages. **Science,** v. 304, p. 1800 - 1804, 2004.

WIERENGA, R. K.; KAPETANIOU, E. G.; VENKATESAN, R. Triosephosphate isomerase: a highly evolved biocatalyst. **Cellular and Molecular Life Sciences,** v. 67, n. 23, p. 3961-3982, 2010.

WILLIAMS, C. H. **Chemistry and Biochemistry of Flavoenzymes**. CRC Press. 1995a.

_____. Mechanism and structure of Thioredoxin reductase from *Escherichia Coli*. **Federation of American Societies for Experimental Biology, Federation of American Society of Experimental Biology,** v. 9, n. 13, p. 1267-1276, 1995b.

WILLIAMS, C. H. et al. Thioredoxin reductase two modes of catalysis have evolved. **European Journal of Biochemistry,** v. 267, p. 6111-6117, 2000.

WOLLMAN, E. E. et al. Cloning and expression of a cDNA for human thioredoxin. **Journal of Biochemistry,** v. 263, n. 30, p. 15506-15512, october 25 1988.

crystallization communications

Anwar Ullah,[a] Tatiana de Arruda Campos Brasil de Souza,[b] Rehana Masood,[a] Mario Tyago Murakami[b] and Raghuvir Krishnaswamy Arni[a]*

[a]Centro Multiusuário de Inovação Biomolecular, Departamento de Física, Universidade Estadual Paulista (UNESP), São Jose do Rio Preto-SP, 15054-000, Brazil, and [b]Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais, Campinas-SP, 13083-970, Brazil

Correspondence e-mail: arni@sjrp.unesp.br

Purification, crystallization and preliminary X-ray diffraction analysis of a class P-III metalloproteinase (BmMP-III) from the venom of *Bothrops moojeni*

Snake-venom metalloproteinases (SVMPs) comprise a family of haemostatically active toxins which can cause haemorrhage, coagulopathy, inhibition of platelet aggregation and inflammatory response. These effects are attributed to the proteolytic action of SVMPs on extracellular matrix components, plasma proteins and cell-surface proteins. SVMPs are classified into four classes (P-I to P-IV) based on their domain structures. In order to understand the multiple roles played by the domains of P-III SVMPs, a P-III SVMP (BmMP-III) from the venom of *Bothrops moojeni* was purified, characterized and crystallized. The crystals belonged to space group $I4_122$, with unit-cell parameters $a = b = 108.16$, $c = 196.09$ Å. Initially, flash-cooled crystals diffracted poorly to a resolution of about 10 Å. However, a significant improvement in the diffraction resolution was observed upon annealing and a complete data set was collected to 3.3 Å resolution. The asymmetric unit contained one molecule and the structure was determined and partially refined to an R factor of 34%. Structural comparisons indicated that the cysteine-rich domain can adopt different conformations in relation to the catalytic domain, which may modulate the enzyme activity.

1. Introduction

Snake venoms are rich and complex mixtures of biologically active peptides and proteins, principally acetylcholinesterases, L-amino-acid oxidases (LAAOs), serine proteinases (SVSPs), metalloproteinases (SVMPs) and phospholipases A_2 (PLA$_2$s) that perturb physiological processes and hence serve as models for biomedical investigations and the development of specific inhibitors (Calvete *et al.*, 2007; Fox & Serrano, 2008; Kang *et al.*, 2011). The proteinase content of snake venoms varies between genera and species, but is generally the most abundant component. It has been estimated that most viperid venoms contain at least 32% SVMPs and 18% SVSPs (Calvete *et al.*, 2007; Fox & Serrano, 2008; Serrano & Maroun, 2005).

SVMPs belong to a subgroup whose leading members are the reprolysins, which participate in the haemorrhagic process by the proteolytic degradation of endothelial cell-surface proteins and extracellular matrix components involved in the maintenance of capillary structure and integrity, leading to the disruption of capillary networks and consequently resulting in oedema and haemorrhage (Escalante *et al.*, 2006; Fox & Serrano, 2005).

The molecular weights of SVMPs usually range from 20 to 100 kDa and are divided into four classes (P-I to P-IV) based on their molecular weights and domain organization. The P-I class of SVMPs (20–30 kDa) are the simplest and contain only a metalloproteinase (M) domain, P-II SVMPs (30–50 kDa) contain metalloproteinase (M) and disintegrin-like (D) domains, P-III SVMPs (50–80 kDa) consist of metalloproteinase (M), disintegrin-like (D) and cysteine-rich (C) domains, and P-IV SVMPs (80–100 kDa) additionally enclose two C-type lectin domains connected by disulfide bonds to the cysteine-rich domain (Bjarnason & Fox, 1995; Fox & Serrano, 2008). The P-III SVMPs are more haemorrhagic than the P-I SVMPs and this has been attributed to the presence of the D and C domains (Bjarnason & Fox, 1995; Jeon & Kim, 1999).
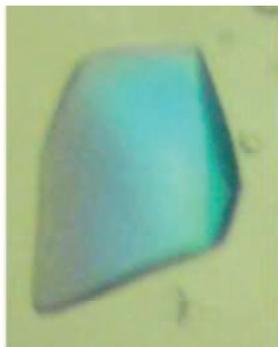
Rapid purification of serine proteinases from *Bothrops alternatus* and *Bothrops moojeni* venoms

Liliane Maria Fernandes de Oliveira [a,1], Anwar Ullah [a,1], Rehana Masood [a], André Zelanis [b], Patrick J. Spencer [c], Solange M.T. Serrano [b], Raghuvir K. Arni [a,*]

[a] *Centro Multiusuário de Inovação Biomolecular, Departamento de Física, Universidade Estadual Paulista (UNESP), São José do Rio Preto, SP 15054-000, Brazil*
[b] *Laboratório Especial de Toxinologia Aplicada-CeTICS, Instituto Butantan, Av. Vital Brasil 1500, São Paulo, SP 05503000, Brazil*
[c] *Centro de Biotecnologia, Instituto de Pesquisas Energéticas e Nucleares-CNEN/SP, Avenida Professor Lineu Prestes 2242, São Paulo, SP 05508-000, Brazil*

ARTICLE INFO

ABSTRACT

Envenomation by Bothrops species results, among other symptoms, in hemostatic disturbances. These changes can be ascribed to the presence of enzymes, primarily serine proteinases some of which are structurally similar to thrombin and specifically cleave fibrinogen releasing fibrinopeptides. A rapid, three-step, chromatographic procedure was developed to routinely purify serine proteinases from the venoms of *Bothrops alternatus* and *Bothrops moojeni*. The serine proteinase from *B. alternatus* displays an apparent molecular mass of ∼32 kDa whereas the two closely related serine proteinases from *B. moojeni* display apparent molecular masses of ∼32 kDa and ∼35 kDa in SDS–PAGE gels. The partial sequences indicated that these enzymes share high identity with serine proteinases from the venoms of other Bothrops species. These proteins coagulate plasma and possess fibrinogenolytic activity but lack fibrinolytic activity.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Snake venoms are especially interesting since they contain high concentrations of proteins and peptides that are chemically and structurally similar to their mammalian counterparts and which, upon envenomation, trigger a wide spectrum of secondary effects that interfere with the maintenance and functioning of essential biological functions such as hemostasis, platelet aggregation and lipid digestion (Lewis and Gutmann, 2004) and thus, some of these proteins have been commercialized as diagnostic and clinical tools (Lewis and Garcia, 2003).

Crotalidae and Viperidae proteinases (Kang et al., 2011; Serrano, 2013; Takeda et al., 2012) are synthesized by the exocrine venom glands and are either metalloproteinases or serine proteinases and catalyze the cleavage of covalent peptide bonds in proteins. Snake venom serine proteinases (SVSPs) likely originated as digestive enzymes and subsequently evolved by gene duplication and sequence modifications to serve other functions. SVSPs encountered in Bothrops venoms are in many aspects functionally similar to endogenous blood clotting enzymes and they interfere with the maintenance and regulation of the blood coagulation cascade by proteolytically cleaving specific bonds and activating proteins involved in blood coagulation, fibrinolysis, and platelet aggregation and also in the proteolytic degradation of cells resulting in an imbalance of the hemostatic system (Kini, 2005; Serrano and Maroun, 2005).

crystallization communications

Crystallization and preliminary X-ray diffraction studies of an L-amino-acid oxidase from *Lachesis muta* venom

Anwar Ullah,[a] Rehana Masood,[a] Patrick Jack Spencer,[b] Mário Tyago Murakami[c] and Raghuvir Krishnaswamy Arni[a]*

[a]Department of Physics, UNESP/IBILCE, Rua Cristovão Colombo 2265, São José Do Rio Preto, São Paulo 15054-000, Brazil, [b]Comissão Nacional de Energia Nuclear, Instituto de Pesquisas Energéticas e Nucleares, IPEN, São Paulo 05508-900, Brazil, and [c]Laboratório Nacional de Biociências (LNBio), Centro Nacional de Pesquisa em Energia e Materiais, Campinas 13083-970, Brazil

Correspondence e-mail: arni@sjrp.unesp.br

Snake-venom proteins form multi-component defence systems by the recruitment and rapid evolution of nonvenomous proteins and hence serve as model systems to understand the structural modifications that result in toxicity. L-Amino-acid oxidases (LAAOs) are encountered in a number of snake venoms and have been implicated in the inhibition of platelet aggregation, cytotoxicity, haemolysis, apoptosis and haemorrhage. An L-amino-acid oxidase from *Lachesis muta* venom has been purified and crystallized. The crystals belonged to space group $P2_1$, with unit-cell parameters $a = 66.05$, $b = 79.41$, $c = 100.52$ Å, $\beta = 96.55°$. The asymmetric unit contained two molecules and the structure has been determined and partially refined at 3.0 Å resolution.

1. Introduction

Snake venoms, which are used for attack, defence and digestion, are complex mixtures of proteins, enzymes and biologically active peptides that interfere with key physiological processes, triggering a wide spectrum of secondary effects such as blood coagulation, myotoxicity, neurotoxicity, platelet aggregation and lipid digestion (Lewis & Garcia, 2003). Structural studies of these proteins are relevant since it has been suggested that mammalian proteins involved in key regulatory processes have been recruited, modified and expressed in the venom gland to function as toxins (Fry, 2005).

Lachesis muta envenomation causes extensive tissue damage, hypotension and shock, and can result in death owing to the neurotoxic, haemorrhagic and coagulant activities of the venom (Jorge *et al.*, 1997)

L-Amino-acid oxidases (LAAOs; EC 1.4.3.2) are homodimeric enzymes which, together with the noncovalently bound cofactor flavin adenine dinucleotide (FAD), catalyse the oxidative deamination of L-amino acids to the corresponding α-keto acids with the concomitant liberation of ammonia (NH_3) and hydrogen peroxide (H_2O_2) and the reduction of FAD (Zhang *et al.*, 2003; Sun *et al.*, 2010). In addition to snake-gland secretions, LAAOs are encountered in many other organisms such as fungi, bacteria, fish skin mucus and plants (Arima *et al.*, 2009; Yang *et al.*, 2009; Kitani *et al.*, 2007; Nagashima *et al.*, 2009; Du & Clemetson, 2002; Kasai *et al.*, 2010; Stábeli *et al.*, 2007), where they are involved in the utilization of nitrogen sources (Du & Clemetson, 2002). LAAOs from snake venoms are cytotoxic, apoptotic, inhibit platelet agreggation and display bactericidal and antiviral activities (Rodrigues *et al.*, 2009; Alves *et al.*, 2008; Li *et al.*, 1994; Stábeli *et al.*, 2007; Zhang *et al.*, 2003).

LAAOs are homodimeric glycosylated proteins with molecular weights ranging from 110 to 150 kDa (Stábeli *et al.*, 2007; Du & Clemetson, 2002). Each subunit is composed of three domains: a FAD-binding domain, a substrate-binding domain and a helical domain (Georgieva *et al.*, 2011; Zhang *et al.*, 2004; Faust *et al.*, 2007; Kang *et al.*, 2011; Ullah, Souza *et al.*, 2012).

The crystal structures of LAAOs from *Calloselasma rhodostoma* (PDB entry 1f8r; 2.0 Å resolution; Moustafa *et al.*, 2006), *Agkistrodon halys pallas* (PDB entry 1reo; 2.3 Å resolution; Zhang *et al.*, 2004), *Vipera ammodytes ammodytes* (PDB entry 3kve; 2.6 Å resolution; Georgieva *et al.*, 2011) and *Bothrops jararacussu* (PDB entry 4e0v; 3.1 Å resolution; Ullah *et al.*, 2012) have been determined. The crystal
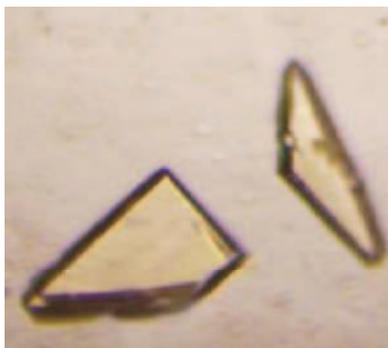
crystallization communications

Crystallization and preliminary X-ray diffraction analysis of a novel sphingomyelinase D from *Loxosceles gaucho* venom

Anwar Ullah,[a] Geraldo Santana Magalhães,[b] Rehana Masood,[a] Ricardo Barros Mariutti,[a] Monika Aparecida Coronado,[a] Mário Tyago Murakami,[c] Katia Cristina Barbaro[b] and Raghuvir Krishnaswamy Arni[a]*

[a]Multiuser Center for Biomolecular Innovation, Department of Physics, UNESP/IBILCE, Rua Cristovão Colombo 2265, São José Do Rio Preto, São Paulo 15054-000, Brazil, [b]Laboratorio de Imunopatologia Divisão de Desenvolvimento Científico, Instituto Butantan, Avenida Dr Vital Brasil 1500, São Paulo 05503-900, Brazil, and [c]Laboratório Nacional de Biociências (LNBio) Centro Nacional de Pesquisa em Energia e Materiais, Campinas 13083-970, Brazil

Correspondence e-mail: arni@sjrp.unesp.br
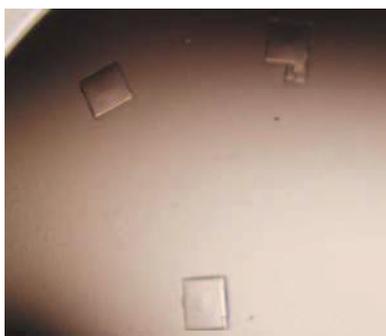
Brown spider envenomation results in dermonecrosis, intravascular coagulation, haemolysis and renal failure, mainly owing to the action of sphingomyelinases D (SMases D), which catalyze the hydrolysis of sphingomyelin to produce ceramide 1-phosphate and choline or the hydrolysis of lysophosphatidylcholine to produce lysophosphatidic acid. Here, the heterologous expression, purification, crystallization and preliminary X-ray diffraction analysis of LgRec1, a novel SMase D from *Loxosceles gaucho* venom, are reported. The crystals belonged to space group $P2_12_12$, with unit-cell parameters $a = 52.98$, $b = 62.27$, $c = 84.84$ Å and diffracted to a maximum resolution of 2.6 Å.

1. Introduction

Sphingomyelinase D activity has only been observed in the pathogenic bacteria *Corynebacterium pseudotuberculosis*, *C. ulcernas*, *Archanobacterium haemolyticum* and *Vibrio damsel* (McNamara *et al.*, 1995) and in spider venoms of the genus *Loxosceles*. Except for the enzyme from *V. damsel*, the bacterial and spider venom SMases (sphingomyelinases) possess similar molecular masses (30–35 kDa) and display significant sequence homology in the N-terminal regions (Tambourgi *et al.*, 1998; van Meeteren *et al.*, 2004). Interestingly, infection by *C. pseudotuberculosis* or envenomation by *Loxosceles* results in similar pathologies (Binford *et al.*, 2005), suggesting a possible evolutionary relationship (Bernheimer *et al.*, 1985).

Spiders from the genus *Loxosceles* are widely distributed (Platnick, 2009) and envenomation results in severe dermonecrotic lesions. Sphingomyelinases D (SMases D; EC 3.1.4.41), also referred to as phospholipases D (PLDs), which are the principal causative agents of lesions, are the major components of the venoms of many spiders belonging to the genus *Loxosceles* (Gomes *et al.*, 2011; Gremski *et al.*, 2014). These enzymes catalyze the hydrolysis of sphingomyelin, which results in the concomitant production of ceramide 1-phosphate and choline or the hydrolysis of lysophosphatidylcholine, generating the lipid mediator lysophosphatidic acid (van Meeteren *et al.*, 2004). Recently, the conversion of lysophosphatidylcholine and sphingomyelin to cyclic phosphates has been reported (Lajoie *et al.*, 2013). The physiological symptoms of envenomation include dermonecrosis, intravascular coagulation, haemolysis and renal failure (Tambourgi *et al.*, 1998; van Meeteren *et al.*, 2004).

Based on their structural properties, spider venom SMases D have been divided into two classes (Murakami *et al.*, 2006): class I enzymes possess a single disulfide bridge (Cys53–Cys57), whereas the class II members possess an additional disulfide bridge that stabilizes the flexible loop which partially occludes the catalytic site (Murakami *et al.*, 2005; de Giuseppe *et al.*, 2011; Ullah *et al.*, 2011).

We have previously determined the structures of enzymes belonging to both classes I (Murakami *et al.*, 2005, 2006) and II (de Giuseppe *et al.*, 2011; Ullah *et al.*, 2011) and the present structural study of SMase D from *L. gaucho* (LgRec1) was carried out to correlate structural differences with the observed variation in toxicity and enzymatic specificity.
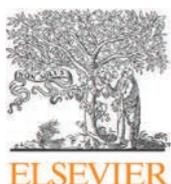
Crystal structure of mature 2S albumin from *Moringa oleifera* seeds

Anwar Ullah [a, **], Ricardo Barros Mariutti [a], Rehana Masood [a], Icaro Putinhon Caruso [a],
Gustavo Henrique Gravatim Costa [b], Cristhyane Millena de Freita [b],
Camila Ramos Santos [c], Leticia Maria Zanphorlin [d], Márcia Justino Rossini Mutton [b],
Mario Tyago Murakami [c], Raghuvir Krishnaswamy Arni [a, *]

[a] Multiuser Center for Biomolecular Innovation, Department of Physics, IBILCE/UNESP, São Jose do Rio Preto, SP, Brazil
[b] Departamento de Tecnologia, FCAV/UNESP, Campus de Jaboticabal, Brazil
[c] Brazilian Biosciences National Laboratory (LNBio), National Center for Research in Energy and Materials, Campinas, SP, 13083-970, Brazil
[d] Brazilian Bioethanol Science and Technology Laboratory (CTBE), Brazilian Center for Research in Energy and Materials (CNPEM), Campinas, SP, Brazil

ABSTRACT

2S albumins, the seed storage proteins, are the primary sources of carbon and nitrogen and are involved in plant defense. The mature form of *Moringa oleifera (M. oleifera)*, a chitin binding protein isoform 3-1 (m*Mo*-CBP$_3$-1) a thermostable antifungal, antibacterial, flocculating 2S albumin is widely used for the treatment of water and is potentially interesting for the development of both antifungal drugs and transgenic crops. The crystal structure of m*Mo*-CBP$_3$-1 determined at 1.7 Å resolution demonstrated that it is comprised of two proteolytically processed α-helical chains, stabilized by four disulfide bridges that is stable, resistant to pH changes and has a melting temperature (T$_M$) of approximately 98 °C. The surface arginines and the polyglutamine motif are the key structural factors for the observed flocculating, antibacterial and antifungal activities. This represents the first crystal structure of a 2S albumin and the model of the pro-protein indicates the structural changes that occur upon formation of m*Mo*-CBP$_3$-1 and determines the structural motif and charge distribution patterns for the diverse observed activities.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Moringa oleifera (*M. olifera*) a hardy, drought resistant tree, native to the foothills of the Himalayas is widely cultivated in subtropical and tropical regions since the bark, leaves, seeds and roots are rich in vitamins and mineral salts and are widely used in herbal medicine [1], as a dietary supplement [2], for the extraction of edible oil [3], as natural flocculants for water purification [4,5] and more recently, as flocculants in the industrial production of alcohol from sugar cane [6].

A number of cationic peptides and low molecular weight basic proteins have been isolated from the crude extract of *M. oleifera* seeds [7,8]. One of these, the chitin binding protein (*Mo*-CBP$_3$), is a 14 kDa thermostable protein has been the focus of a disproportionately large number of studies since it is widely used for water

purification and also possess antigungal and antibacterial activities [9].

Mo-CBP$_3$ belongs to the 2S albumin family of proteins that are synthesized as precursors, which are then proteolytically cleaved to form the mature protein. Four isoforms of this precursor have been isolated from *M. oleifera* seeds namely *Mo*-CBP$_3$-1, *Mo*-CBP$_3$-2, *Mo*-CBP$_3$-3 and *Mo*-CBP$_3$-4, which differ from each other by only a few amino acid residues [9]. The *Mo*-CBP$_3$-1 precursor (p*Mo*-CBP$_3$-1) consists of 163 amino acids including the N-terminal signal peptide and the linker peptides. Solution structures (NMR) of several 2S albumin precursors have been deposited with the Protein Data Bank (http://www.rcsb.org) and the only crystal structure currently available is that of the sweet protein, Mabinlin II (*Capparis masaikai* Levl., PDB ID: 2DS2) [10].

The crystal structure of mature *Mo*-CBP$_3$-1 (m*Mo*-CBP$_3$-1), a typical 2S seed storage albumin determined at 1.70 Å resolution provides a model for understanding the diversity of the structures of this large family of albumins, serves as a model for the structures of the precursors and indicates the structural basis for the observed flocculating activities.

---

* Corresponding author.
** Corresponding author.
E-mail addresses: anwar.ms90@yahoo.com (A. Ullah), arni@sjrp.unesp.br (R.K. Arni).

Three-Dimensional Structures and Mechanisms of Snake Venom Serine Proteinases, Metalloproteinases, and Phospholipase A$_2$s

M. A. Coronado[a], F. R. de Moraes[a], A. Ullah[a], R. Masood[a], V. S. Santana[a], R. Mariutti[a], H. Brognaro[a], D. Georgieva[b], M. T. Murakami[c], C. Betzel[b] and R. K. Arni[a]*

[a]Department of Physics, IBILCE/UNESP, Multi User Center for Biomolecular Innovation, São Jose do Rio Preto-SP, Brazil
[b]Institute of Biochemistry and Molecular Biology, University of Hamburg, Laboratory of Structural Biology of Infection and Inflammation, c/o DESY, Hamburg, Germany
[c]National Laboratory for Biosciences, National Center for Research in Energy and Materials, Campinas, Brazil

Abstract

High-resolution crystal structures provide accurate information of the positions of the atoms which can be used to understand substrate specificity, secondary binding sites, and catalytic mechanisms. Detailed structural information and mechanisms of serine proteinases, metalloproteinases, and phospholipases A2s are presented.

Introduction

Structural information gleaned at the molecular and atomic levels, when correlated with biochemical and biophysical details capable of generating a coherent picture of the salient structural features and interactions that modulate and determine biomolecular recognition, specificity, and hydrolysis, provides us with very powerful tools to decipher, step-by-step, complex biological phenomena, thus permitting a profound understanding of the basic underlying role of molecular architectural diversity in efficiently performing essential, distinct, chemical reactions that are central to life.

The enormous number of protein and DNA sequences currently deposited in the data banks (e.g., http://www.expasy.org, http://www.uniprot.org) indicate that the one-dimensional representation of protein sequences generally contains a trace of the fingerprint of evolution, and often, only a faint residual of the ancestral protein is retained in the protein linear amino acid sequence. However, upon closer examination, the application of this concept exposes its fundamental limitations, and hence, its utility is strictly limited since (a) proteins that perform the same function or catalyze similar reactions often share only very low-sequence identity, (b) proteins that are about 20 % identical in their primary sequences may still catalyze distinct reactions and modulate different functions, and (c) point mutations in the active sites or cofactor-binding sites often produce proteins that catalyze different reactions or in extreme cases result in enzymatically inactive proteins.

A fundamental conceptual bridge linking the linear protein sequence and its primary biological function is encoded in the three-dimensional fold or, in other words, in the exact positions of the atoms of the protein in three dimensions. Central to this concept is the fact that the three-dimensional (3D) structure of a protein is more highly conserved during evolution (Bajaj and Blundell 1984; Finkelstein and Ptitsyn 1987) than the linear amino acid sequence of the protein, and consequently, the shape of a protein, the spatial distribution of its atoms, and the surface charge, solvent

---

*Email: arni@sjrp.unesp.br