STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF TRX16, A THIOREDOXIN-LIKE PROTEIN AND ALTERING SUBSTRATE SPECIFICITY OF SPI1, A PROTEASE INHIBITOR

PANKAJ KUMAR GIRI

NATIONAL UNIVERSITY OF SINGAPORE

2011

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF TRX16, A THIOREDOXIN-LIKE PROTEIN AND ALTERING SUBSTRATE SPECIFICITY OF SPI1, A PROTEASE INHIBITOR

PANKAJ KUMAR GIRI

A THESIS SUBMITTED

FOR

THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOLOGICAL SCIENCES

NATIONAL UNIVERSITY OF SINGAPORE

2011

This thesis is dedicated to my inspiring parents
for their love, endless support
and encouragement

ACKNOWLEDGEMENT

No words can express the profound respect and gratitude I have for my supervisors Prof. K Swaminathan and Prof. J. Sivaraman. I thank them for their perpetual guidance, unceasing cooperation and constant encouragement, without which this dissertation would have remained but a dream. They have not only led me with utmost scholarliness, but also in full earnestness fostered my own initiative and creativity. They have patiently guided me throughout the course and helped me streamline my efforts effectively.

I express my heartfelt gratitude to Prof. Ding Jeak Ling, for her support and encouragements. My special thanks to Dr. Gautam Sethi, Department of Pharmacology, NUS and his postdoctoral fellow Dr. Shanmugam Muthu Kumaraswamy for helping in ex-vivo studies based on HeLa cells for my project. I would like to thank Dr. Fan Jing-Song, who helped me during my NMR data collection and structure solution for one of my project.

I would like to convey my special thanks to Dr. Ping Yuan, Assistant professor, Li Ka Shing Institute of Health Sciences, CUHK, Hong Kong for the opportunity she gave me to work in her lab and for her guidance throughout my "Global research excellence programme under the CNCOO Grant 2011". I would like to thank Gan Jingyi and LI Peng for their help and support during my stay in Hong Kong.

I would like to extend my thanks to all my colleagues and friends from SBL-4 and 5 for their full support and help. A special thanks to Lissa who helped all along my whole duration of PhD. I want to thank friends Abdollah (NTU), Girish, Jack, Kang Wee,

Karthik, Sang, Smarajit, Shifali, Toan (NTU), and Vamsi, who shared with me numerous experiences and advice.

I am grateful to my ***parents*** *and* ***family members,*** *whose constant inspiration, persistent support and encouragement brought me to where I am now. I offer this thesis as a humble tribute to all their love, affection and blessings, which they have showered on me.*

I thank NUS for giving me the opportunity to pursue my PhD with a research scholarship.

"From small beginnings come great things…

… The distance does not matter. It's only the first step that is difficult"

"Sometimes the journey is as exciting as the destination"

Pankaj Kumar Giri
November2011

TABLE OF CONTENTS

SUMMARY

The causative agents of most diseases like cancer and Alzheimer's are proteins. The function of a protein can be fully appreciated only when we have a complete knowledge of its 3-dimensional structure, as structure and function go hand in hand. Decades of effort using X-ray crystallography and NMR have produced thousands of protein and complex with binding partner structures and these structures provide a rich source of data for learning the principles of how proteins interact and for rational design and engineering of therapeutics. Here, we are particularly keen on figuring out how proteins are involved in gene regulation under stress by interacting with their partners and the use of a rational approach for protein design and engineering to change the substrate specificity of a protease inhibitor.

This PhD thesis consists of five chapters. Chapter I deals with the literature survey and general introduction about the thioredoxin (Trx) system (an antioxidant) and briefly covers the various strategies of structure based protein design and engineering to develop drugs against a specific protease inhibitor. Chapter II deals with the structural and functional characterization of thioredoxin like protein 16 from *Carcinoscorpius rotundicauda* (Cr-TRP16), a 16 kDa Trx-like protein that contains a WCPPC motif. We present the NMR structure of the reduced form of Cr-TRP16, along with its regulation of NF-κB activity. Unlike other 16 kDa Trx-like proteins, Cr-TRP16 contains an additional Cys residue (Cys15, at the N-terminus), through which it forms a homo-dimer. Moreover we have explored the molecular basis of Cr-TRP16 mediated activation of NF-κB in the HeLa cell and show that Cr-TRP16 exists as a dimer under an oxidized condition and only the dimeric form binds to NF-κB and enhances its DNA-binding activity by directly

reducing the cysteines in the DNA-binding motif of NF-κB. The C15S mutant of Cr-TRP16 is unable to dimerize and hence does not bind to NF-κB.

Based on our finding and combined with the literature, we propose a model on how Cr-TRP16 is likely to bind to NF-κB. These findings elucidate the molecular mechanism by which NF-κB activation is regulated by Cr-TRP16. Chapter III reports the expression and purification of human thioredoxin like protein-6 (TXNL-6), a homolog of Cr-TRP16 and protects retinal cells from apoptosis under stress and characterization of its interaction with NF-kB.

Chapter IV presents the structure based rational design of altered specificity of a protease inhibitor. Protease inhibitors play a decisive role in maintaining homeostasis and eliciting antimicrobial activities. Invertebrates like horseshoe crab have developed unique modalities with serine protease inhibitors to detect and respond to microbial and host proteases. Two isoforms of immunomodulatory two-domain Kazal-like serine protease inhibitors, CrSPI-1 and CrSPI-2, have been recently identified in the hepatopancreas of the horseshoe crab, *Carcinoscorpius rotundicauda*. Full length and domain 2 of CrSPI-1 display powerful inhibitory activities against subtilisin. However the structure and function of CrSPI-1 domain-1 remain unknown. Here, we report the crystal structure of CrSPI-1-D1, refined at 2.0 Å resolution. Despite the close structural homology of CrSPI-1-D1 to rhodniin-D1 (a known thrombin inhibitor), CrSPI-1-D1 does not inhibit thrombin. This prompted us to modify the selectivity of CrSPI-1-D1, specifically towards thrombin. Here, we illustrate the use of the structural information of CrSPI-1-D1 to modify this domain into a potent thrombin inhibitor with IC_{50} of 26.3 nM. In addition, these studies demonstrate that besides the rigid conformation of the reactive site loop of

the inhibitor, the sequence is the most important determinant of the specificity of the inhibitor. This study will lead to significant applications to modify a multi-domain inhibitor protein to target several proteases. Chapter V provides the overall conclusion and future directions of these projects.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AA	Amino acids
A600/595	Absorbance at 600/595 nm
ATP	2'-deoxyadenosine 5'-triphosphate
bp	Base pair
BSA	Bovine serum albumin
cDNA	Complementary deoxyribonucleic acid
CD	Circular Dichroism
Cr	*Carcinoscropius rotundicauda*
Cr-SPI-1	*Carcinoscorpius rotundicauda* serine protease
Cr-TRP16	*Carcinoscropius rotundicauda* TRX1 (16 kDa)
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
DLS	Dynamic Light Scattering
E. coli	*Escherichia Coli*
EDTA	Ethylenediaminetetraacetic acid
EMSA	Electrophoretic mobility shift assay
Fig	Figure
FPLC	Fast Protein Liquid Chromatography
β-Gal	β−galactosidase
GSH	Glutathione
GST	Glutathione S-transferase
h	Hour
HSQC	Heteronuclear Single Quantum Correlation
IKK	Inhibitory κB kinase
iNOS	Inducible nitric oxide synthase
ITC	Isothermal Titration Calorimetry
Ka	Association constant
kb	Kilobase
Kd	Dissociation constant
kDa	Kilodalton
LB	Luria Bertani

LUC	Luciferase
mg	Milligram
min	Minute
ml	Milliliter
mM	Millimolar
MALDI-TOF	Matrix-assisted laser desorption ionization-Time of flight
mRNA	Messenger ribonucleic acid
NEB	New England Biolabs
NF-κB	Nuclear factor-κB
ng	Nanogram
NIK	NF-κB inducing kinase
Ni-NTA	Nickel nitrilo-triacetic acid
NLS	Nuclear localization signal
nM	Nanomolar
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Enhancement Spectroscopy
OD	Optical density
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	polyethylene glycol
ppm	Parts per million
RdCVF	Rod-derived cone viability factor
RF	Radio Frequency
rmsd	Root mean square deviation
RNA	Ribonucleic acid
RNase	Ribonuclease
ROS	Reactive oxygen species
rpm	Revolutions per minute
RT-PCR	Reverse transcriptase-polymerase chain reaction
SDS	Sodium dodecyl sulphate

SDS-PAGE	Sodium dodecyl sulfate - polyacrylamide gel electrophoresis
s	Second
T	Thymine
T1	Longitudinal relaxation
T2	Transverse relaxation
TD	Transactivation domain
TLR	Toll-like receptor
TOCSY	Total Correlation Spectroscopy
TNFα	Tumor necrosis factor α
Trx	Thioredoxin
TrxR	Thioredoxin reductase
TRNOE	Transferred Nuclear Overhauser Effect
TXNL-6	Thioredoxin-like 6 (Human 24 kDa TRX)
U	Unit
UV	Ultraviolet
V	Volt
v/v	Volume : volume ratio
w/v	Weight : volume ratio

LIST OF PUBLICATIONS

Giri P.K., Song FJ, Shanmugam MK,Ding JL,Sethi G, Swaminathan K, Sivaraman J (2011) (submitted) NMR structure of *Carcinoscorpius rotundicauda* thioredoxin-related protein 16 and its role in regulating transcription factor NF-kB activity.JBC

Giri P.K., Tang X, Thangamani S, Shenoy RT, Ding JL, Swaminathan K, Sivaraman J (2010) Modifying the substrate specificity of Carcinoscorpius rotundicauda serine protease inhibitor domain 1 to target thrombin. ***PLoS One*** **5: e15258**

CHAPTER-I: **GENERAL INTRODUCTION**

"Research in its highest expression,
is an open-minded expression inquiry; for truth,
to be found and revealed unreservedly,
for the information, instruction, advantage and welfare of ALL"
..... Williams J Gies

Aerobic life depends upon controlled combustion for energy supply. Controlled combustion is catalyzed and regulated by metabolic machinery that can be damaged by uncontrolled oxidative reactions that are associated with energy production. Due to the extreme threat of such uncontrolled oxidation, aerobic life evolved a complex set of antioxidant systems to control these reactions and repair or replace the damaged machinery. At the same time, enzyme systems evolved to produce reactive species for biological signaling, biosynthetic reactions, chemical defense, and detoxification functions. The presence of both toxic and beneficial consequences of reactive species precludes a simple definition of oxidative stress.

The following sections will present a general introduction about the thioredoxin system (an antioxidant) in the first part, while the second part briefly covers the various strategies of structure based protein design and engineering to develop drugs against specific protease inhibitors.

1.1 REACTIVE OXYGEN SPECIES (ROS)

Reactive oxygen species (ROS) is a collective term that describes the chemical species that are formed upon incomplete reduction of oxygen and includes the superoxide anion (O_{2^-}), hydrogen peroxide (H_2O_2) and the hydroxyl radical (HO). ROS are thought to mediate the toxicity of oxygen because of their greater chemical reactivity with regard to oxygen. Reactive oxygen species are highly reactive due to the presence of unpaired valence shell electrons. ROS are formed as a natural byproduct of the normal metabolism of oxygen and have important roles in cell signaling and homeostasis (Flohé et al., 1997; Novo and Parola, 2008; Quinn et al., 2002). Fig. 1.1 illustrates the mechanisms for the generation of ROS in living cells.

At the cellular level, ROS may act as second messengers in various signal transduction and elicit a wide spectrum of responses ranging from proliferation to growth or differentiation arrest, to senescence, to cell death by activating numerous major signaling pathways including phosphoinositide 3-kinase (PI-3K), NF-κB, phospholipase C-γ1 (PLC- γ1), p53, CREB, HSF and mitogen-activated protein kinases [MAPKs, which may classify into: extracellular signal-regulated kinases (ERKs), c-Jun N-terminal (JNK), p38 MAPK]. The magnitude and duration of the stress, as well as the cell type involved, are important factors in determining which pathways are activated and the particular outcome reflects the balance between these pathways (Martindale and Holbrook, 2002).

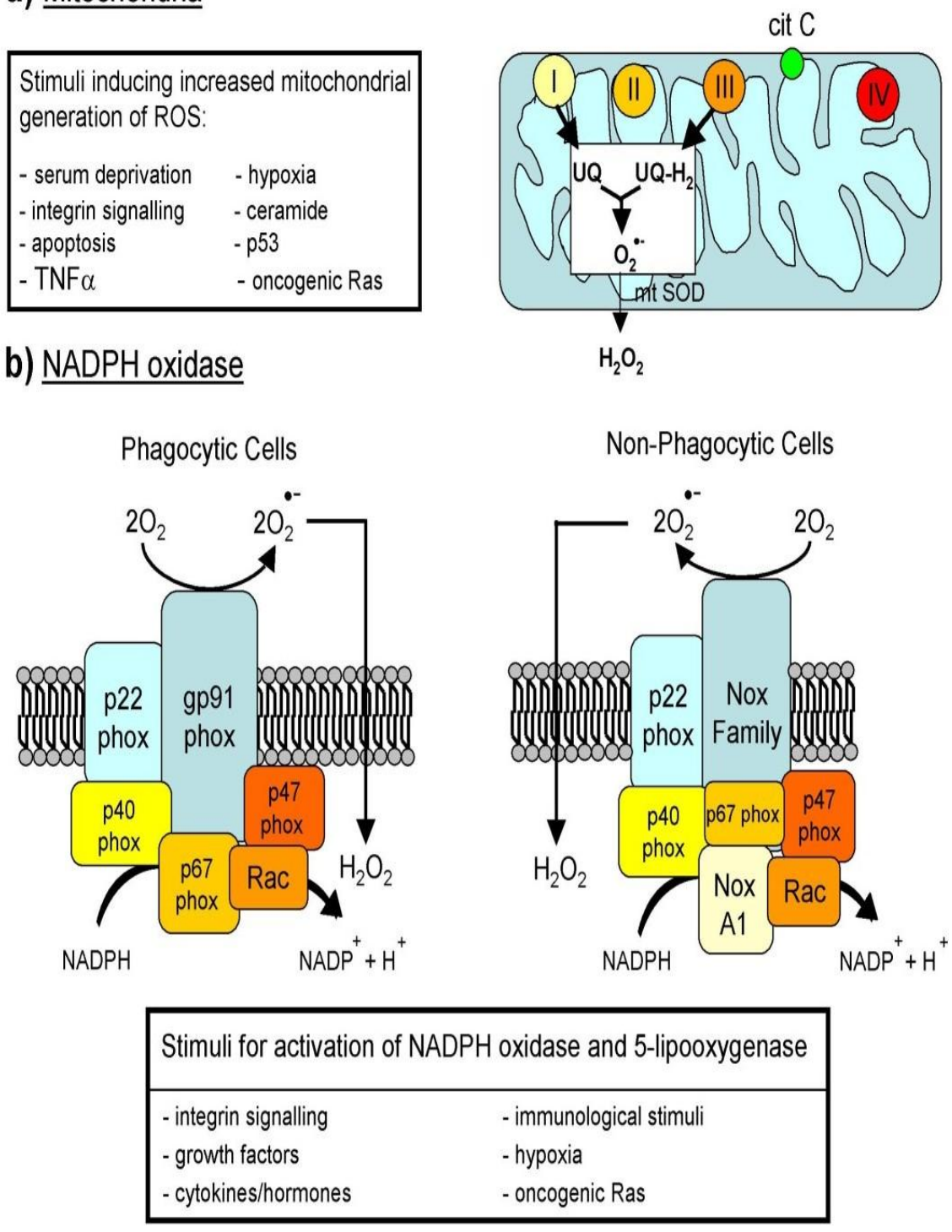


Figure I.1: Cellular sources of ROS in living cells. Adopted from(Novo and Parola, 2008).

The continuous efflux of ROS from endogenous and exogenous sources results in continuous and accumulative oxidative damage to cellular components (Comporti, 1989) and alters many cellular functions (Gracy et al., 1999). Among the biological targets most vulnerable to oxidative damage are proteinaceous enzymes (Davies et al., 1987; Levine and Stadtman, 2001), lipidic membranes (Davies et al., 1987), and DNA (Beckman and Ames, 1997; Chang et al., 2007) (Fig. 1.2).

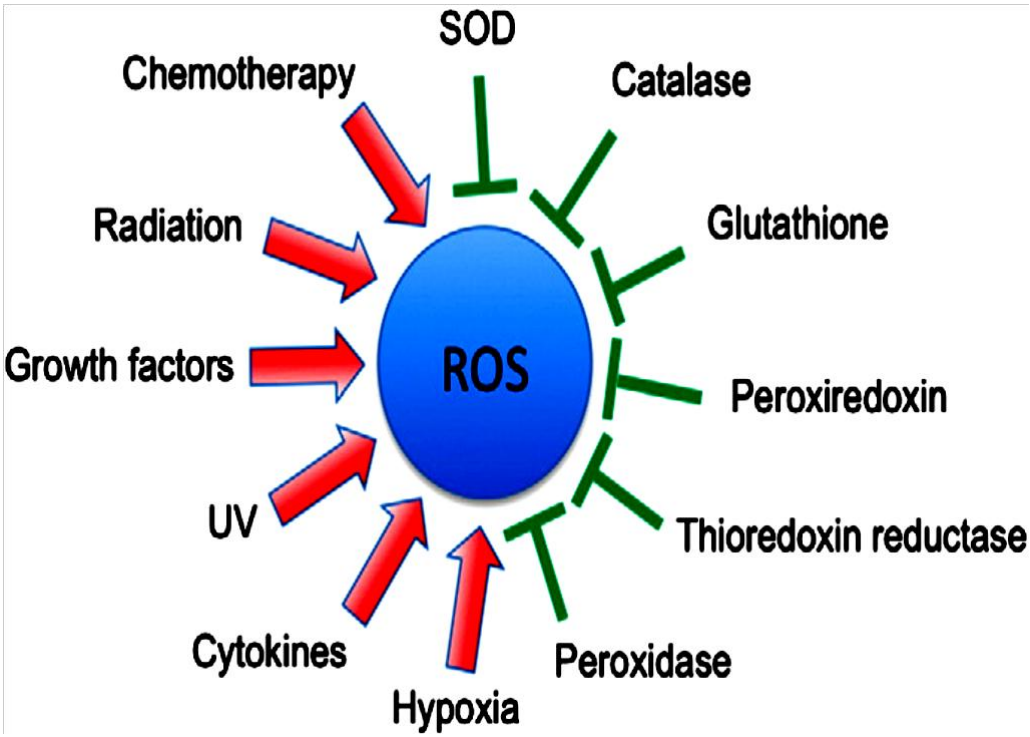


Figure I.2: Schematic representation of various activators and inhibitors of reactive oxygen species production. Adopted from (Reuter et al., 2010).

Numerous pathologies and disease states serve as sources for the continuous production of ROS (Baud and Ardaillou, 1986; Halliwell, 1994; Kawanishi et al., 2002; Levy, 1996; Venditti et al., 2002). More than 200 clinical disorders have been described in the literature in which ROS were important for the initiation stage of a disease or

produced during its course (Table 1.1). ROS may be important initiators and mediators in many types of cancer (Brown and Bicknell, 2001; Gracy et al., 1999; Nyska et al., 2002), heart diseases, endothelial dysfunction (Aikawa et al., 2001; Farré and Casado, 2001; Laroux et al., 2001), atherosclerosis and other cardiovascular disorders, inflammation and chronic inflammation (Laroux et al., 2001; Latha and Babu, 2001), burns (Latha and Babu, 2001), intestinal tract diseases (Blau et al., 2000), brain degenerative impairments (Giasson et al., 2002), diabetes (Opara, 2002), eye diseases (Goldstein et al., 1996), and ischemic and post ischemic e.g., damage to skin, heart, brain, kidney, liver, and intestinal tract pathologies (Sasaki and Joh, 2007).

Table I.1: A partial list of diseases that have been linked to reactive oxygen species.

Disease	**Reference**
Acute respiratory distress syndrome	(Wilson et al., 2001)
Aging	(Dugan and Quick, 2005)
Acute respiratory distress syndrome	(Hensley et al., 1996; Multhaup et al., 1997)
Alzheimer	(Nisha and Deshwal, 2011)
Atherosclerosis	(Lau et al., 2008)
Cancer	(Touyz, 2004)
Cardiovascular disease	(Muhammad et al., 2009)
Diabetes	(Di Virgilio, 2004)
Inflammation	(Moulton, 1996)
Inflammatory joint disease	(Bolaños et al., 2009)
Neurological disease	(Atabek et al., 2004)
Obesity	(Tabner et al., 2001)
Parkinson	(Kamp et al., 1992)
Pulmonary fibrosis	(Gelderman et al., 2007)
Rheumatoid arthritis	(Haurani and Pagano, 2007)
Vascular disease	(Wilson et al., 2001)

In several normal conditions ROS are produced and play a role in the pathogenesis of the physiological condition (Fig.1.3). These are exemplified during the aging process, where ROS production significantly increases as a result of impaired mitochondrial function and in the early stages of embryonic development (Lee and Wei, 2001). Other pathological disorders, which are associated with impaired metal metabolism, such as hemochromatosis (Eaton and Qian, 2002), Wilson disease (Rotilio et al., 2000), and thalassemia (Meral et al., 2000), in which iron is deposited in many organs, are known to increase significantly the concentration of ROS.

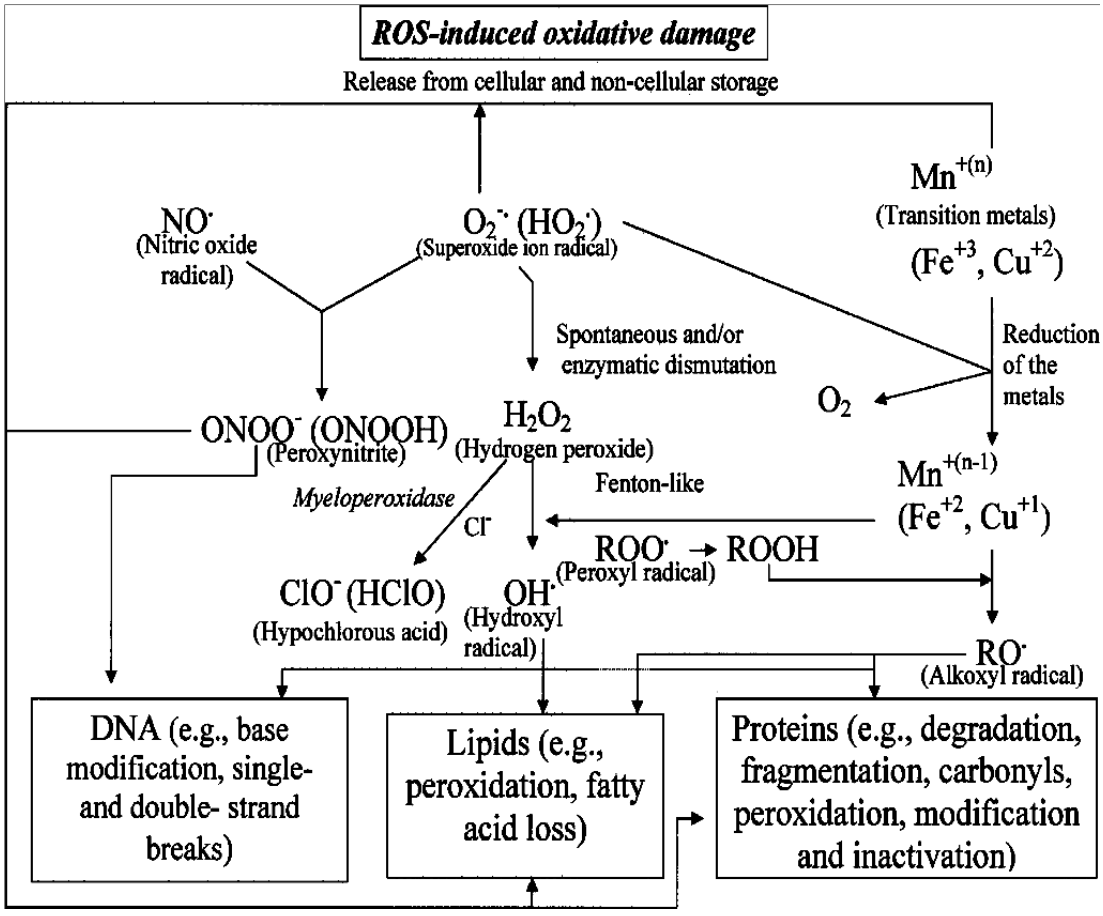


Figure I.3: Reactive oxygen species (ROS)-induced oxidative damage. Adopted from (Opara, 2002).

Reactive oxygen species (ROS) are known to be mediators of intracellular signaling pathways. However the excessive production of ROS may be detrimental to the cell as a result of the increased oxidative stress and loss of cell function. Hence, well tuned, balanced and responsive antioxidant systems are vital for proper regulation of the redox status of the cell. The biological system/cells are normally able to defend themselves against the oxidative stress/ROS induced damage by regulating the cellular reduction/oxidation (redox) status through the use of several antioxidant systems. Under pathologic conditions, the cells develop both enzymatic and non-enzymatic defense systems to reduce the concentration of these ROS (Stadtman, 1992).The enzyme superoxide dismutase (SOD) deals with the superoxide anions, while the hydrogen peroxide so formed from the SOD reaction is detoxified by the catalase (Loshchagin et al., 2002) and also by glutathione and thioredoxin dependent peroxidases (Flohé et al., 1997; Hofmann et al., 2002). The thioredoxin (Trx) system, and the glutathione (GSH) system are the two main ubiquitously expressed thiol-reducing antioxidant systems (Nordberg and Arnér, 2001). The Trx system (Trx, Trx reductase and NADPH) plays a crucial role in reducing oxidized cysteine groups on proteins (Fig.1.4).

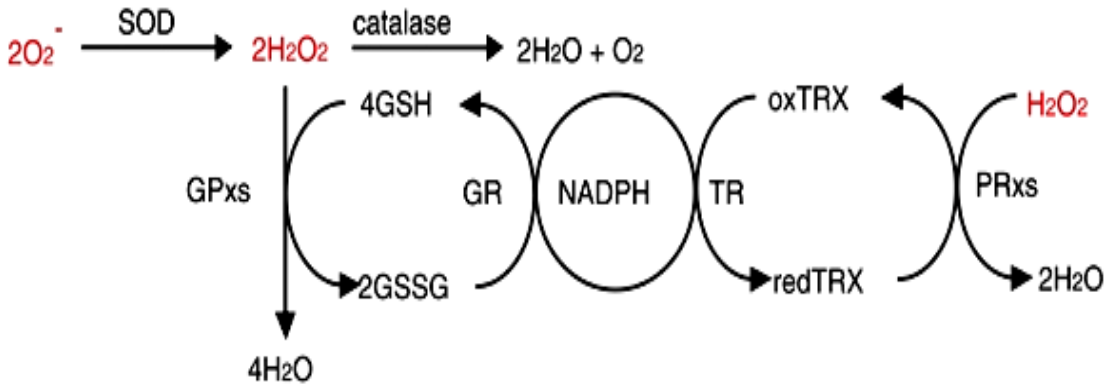


Figure I.4: O2- is converted into H2O2 by superoxide dismutases (SODs).H_2O_2 is then eliminated by catalase, glutathione peroxidases (GPxs), and peroxiredoxins (PRxs). During elimination of H_2O_2, reduced glutathione (GSH) is converted to disulfide form (GSSG) by GPxs, and then GSSG is recycled to GSH by glutathione reductase (GR). However, PRxs also catalyze H_2O_2 into H_2O by using reduced thioredoxin (Trx). Oxidized Trx is then recycled back to redTrx by thioredoxin reductase (TrxR). NADPH is essential for both recycling reactions. ROS are indicated using red characters (Nakano et al., 2005)

1.2 THIOREDOXIN SYSTEM

The thioredoxin system, comprising with thioredoxin (Trx), thioredoxin reductase (TrxR) and NADPH , was discovered by Reichard and coworkers in 1964 as a hydrogen donor for the enzymatic synthesis of cytidine deoxyribonucleoside diphosphate by ribonucleotide reductase from *Escherichia coli* (Laurent et al., 1964b) . It is the cells' major protein disulfide reductase, potentially being the physiological equivalent of a reducing agent like dithiothreitol (Holmgren, 1989b).

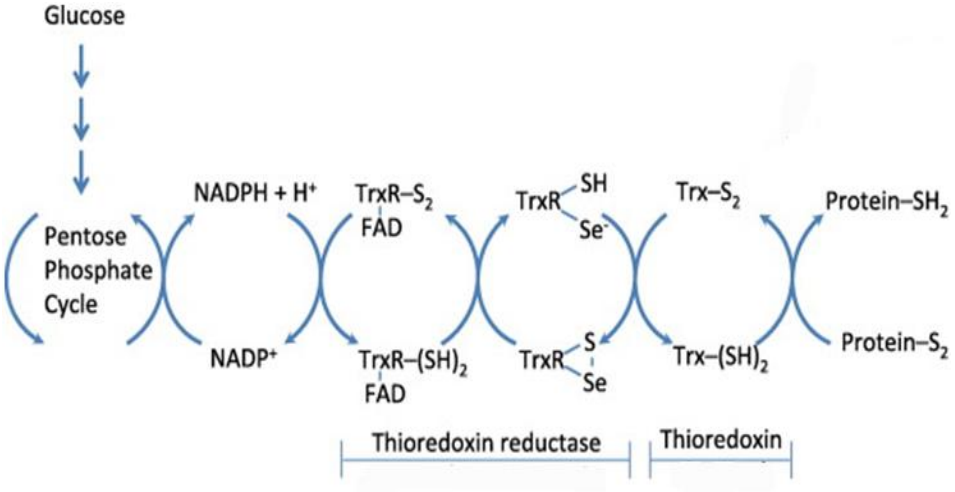


Figure I.5: Redox reactions catalyzed by a mammalian Trx system comprising thioredoxin reductase (TrxR), thioredoxin (Trx) and NADPH

(Holmgren and Lu, 2010). The electron source of the Trx system is NADPH, which is largely produced from the pentose phosphate pathway. The oxidized thioredoxin (Trx-S2) is reduced by NADPH and the selenoenzyme TrxR. Electrons are transferred from NADPH to FAD, then to the N-terminal redox active disulfide in one subunit of TrxR and finally to the C-terminal active site Gly–Cys–Sec–Gly of the other subunit (Elias S.J, 2009). Reduced thioredoxin [Trx-(SH)2] catalyzes disulfide bond reduction in many proteins. Upon oxidative stress, Trx can be secreted into plasma or cleaved into Trx80 lacking the C-terminal 20 or 24 amino acid residues (Pekkari et al., 2001).

Thioredoxin (Trx), which functions as a general protein-disulfide reductase, is commonly known to be a small ubiquitous protein of 12 kDa. It is evolutionarily conserved from prokaryotes to eukaryotes, plants, and animals (Holmgren, 1989b). The redox activity of Trx has been reported to reside in a conserved active site, Cys-Gly-Pro-Cys (CGPC), in which the two Cys residues undergo reversible oxidation, converting its dithiol group to a disulfide bond (Fig. 1.5) (Powis and Montfort, 2001).

Trx is maintained in its active reduced form by the thioredoxin reductase (TrxR), a selenocysteine-containing protein that uses the reducing power of NADPH (Powis and Montfort, 2001). The three-dimensional structure of Trx is conserved throughout evolution and consists of four or five central β-strands, surrounded by three or four α-helices (Fig. 1.6). The active site is located in a protrusion of the protein between the β2-strand and the α2-helix. Both the conserved active site sequence and the three-dimensional structure of Trx are the hallmarks of this superfamily (Martin, 1995b).

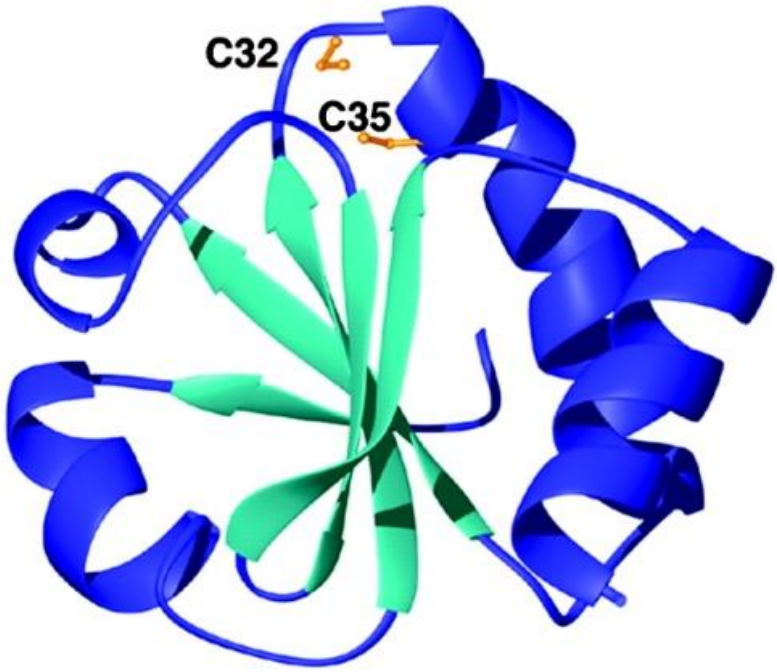


Figure I.6: The three-dimensional structure of E. coli thioredoxin(Kumar et al., 2004). The active site redox active disulfide cysteines (C32 and C35), shown in yellow, are located in a small cleft in the main body of the molecule.

1.2.1 Phylogenetic analysis of Thioredoxin (Trx)

Trx was first identified as an electron donor for ribonucleotide reductase for DNA synthesis in *Escherichia coli* (Laurent et al., 1964a).Trx, with a molecular weight of 12 kDa, is conserved throughout in all prokaryotic and eukaryotic species (Eklund et al., 1991b). The first characterized human Trx (Trx1) is a 12 kDa protein with a catalytic sequence of CGPC. Although bacterial 12 kDa Trx only contains two cysteine residues (at its catalytic site), the human 12 kDa Trx contains three other cysteine residues. The C-terminal Cys73 is involved in dimerization, and may convey unique biological properties to mammalian Trx (Holmgren, 1985).

A second slightly larger Trx-2 is a 166-amino acid protein with a molecular weight of 18 kDa, containing a conserved Trx catalytic site. It has been identified in the mitochondria of pig's heart (Spyrou *et al*, 1997). The 60-amino acid N-terminal extension of Trx-2 exhibits characteristics consistent with a mitochondrial translocation signal, and the mitochondrial localization of Trx-2 was confirmed by Western blotting (Miranda-Vizuete *et al*, 2000). A 32 kDa thioredoxin-like cytosolic protein was first cloned from a human testis cDNA library (Lee *et al*, 1998). The 289 amino acid protein has an N-terminal Trx domain of 105 amino acids, a conserved Trx active site (CGPC), and a high degree of homology to human Trx. It is ubiquitously expressed in the human testis. In 1997, a type of Trx called nucleoredoxin (NRX) with a WCPPC catalytic site was cloned from mice. Interestingly, this 435-amino acid protein is localized to the nucleus (Kurooka *et al*, 1997). Recently, a family of 16 kDa Trx has been identified from the nematodes and protozoa of the family Trypanosomatidae with an active site of WCPPC (Kunchithapautham et al., 2003).

The 14 kDa human Trx (TRP14), with an active motif of WCPDC, exhibits markedly different substrate specificity compared to the 12 kDa Trxs. Although TRP14 could reduce small disulfide-containing peptides, it did not reduce the disulfides of the known human Trx1 substrates ribonucleotide reductase and peroxiredoxin (Jeong *et al*, 2004b). Although the enzymes of the thioredoxin superfamily do not have a high level of sequence similarity, they nevertheless share a marked degree of structural similarity, and all having a common sequence CXXC in the active site.

The sequence alignment of these homologous proteins reveals that the catalytic WCPPC motif is largely conserved, indicating the potential importance of this motif in

the Trx function (Fig. 1.7). This highly conserved region is present at the end of an α-helix in all the enzymes belonging to the super family.

Table I.2: Homology (in percentage*) among thioredoxins from different species

	E. coli (108)	*C. nephridii* (105)	*C. nephridii* C2 (107)	*Anabaena* 7119 (106)	*A. nidulans* (107)	*Anabaena* 7120 (110)	*R. sphaeroides* (105)	*Rs. rubrum* (104)	Chlorobium (108)	Chromatium (107)	Spinach m (114)	Spinach f (113)	Chicken (104)	Rabbit (104)	Mouse (105)	Calf thymus (104)	Human (104)
SPECIES																	
E. coli	100	50	48	48	47	43	47	56	44	67	46	30	27	27	26	27	27
C. nephridii		100	51	50	52	29	56	52	45	52	50	30	30	32	30	32	30
C. nephridii C2			100	44	43	35	50	40	43	50	41	30	29	33	31	35	34
Anabaena 7119				100	84	39	52	51	51	52	57	30	26	29	26	29	30
A. nidulans R2					100	42	50	50	47	52	54	31	27	28	27	30	31
Anabaena 7120						100	35	37	30	39	32	23	21	21	21	22	22
R. sphaeroides							100	54	46	52	43	31	31	32	30	32	31
Rs. rubrum								100	40	52	43	31	31	28	26	28	30
Chlorobium									100	44	41	29	28	29	30	30	30
Chromatium										100	46	31	26	25	24	25	26
Spinach m											100	29	25	27	26	26	26
Spinach f												100	29	33	32	33	35
Chicken													100	68	69	69	68
Rabbit														100	79	88	87
Mouse															100	84	83
Calf thymus																100	90
Human																	100

* Only positive matches are scored. No penalties were assigned to gaps or loops. Percentages were calculated based on the number of residues (in parentheses) in the thioredoxin listed in the top line. Adopted from (Lillig and Holmgren, 2006).

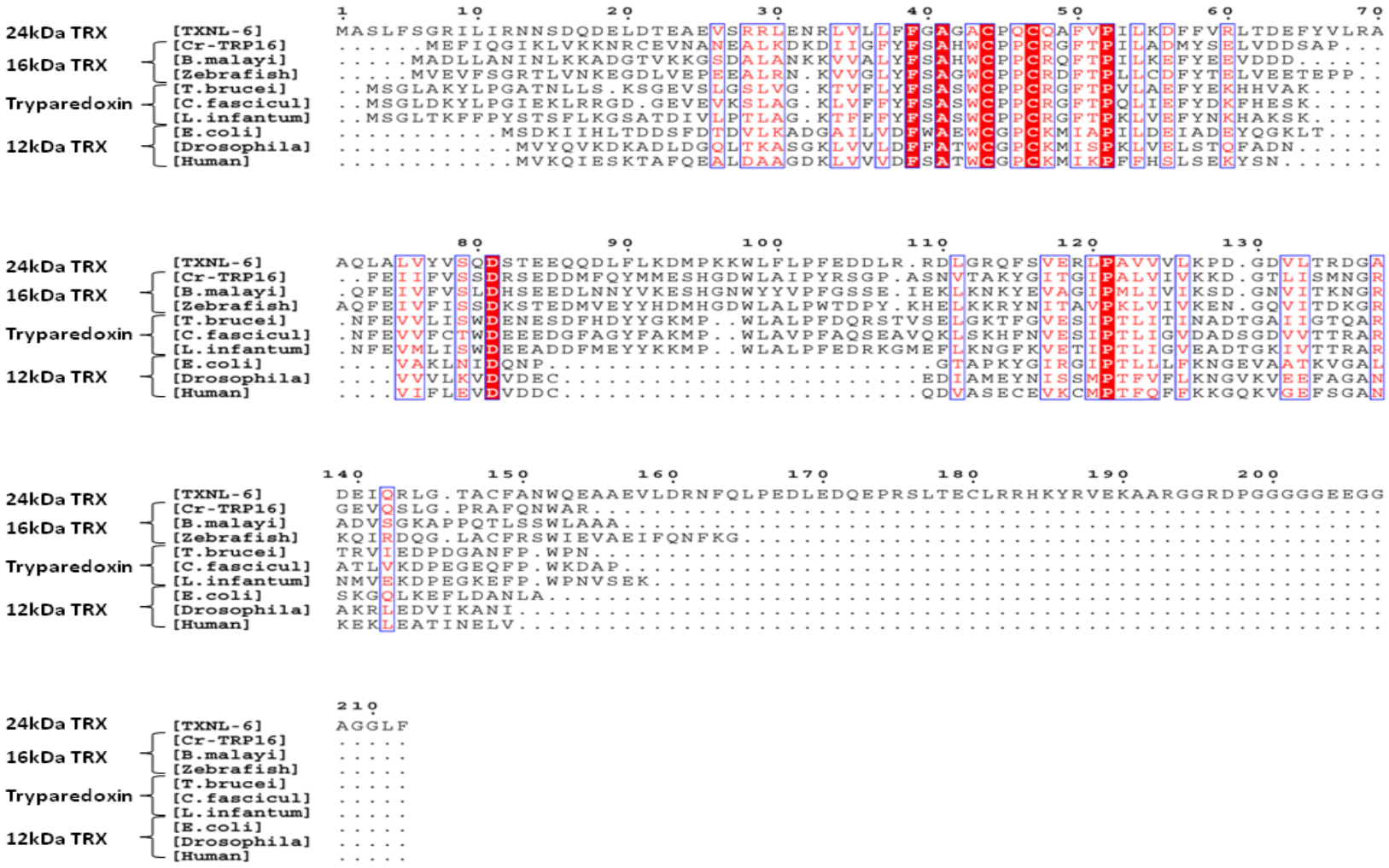


Figure I.7: Amino acid sequence comparison among thioredoxins from different species. Alignments were done by Clustal-X. The active sites are demarcated by a dashed line. This figure was created by using the program ESPript (Gouet et al., 1999)

Despite notable differences in the molecular mass and amino acid sequence of the catalytic site, those forms of Trxs (Table 1.2) appear to be functionally similar to the classical 12 kDa Trxs. Therefore, it appears that there is considerable flexibility in the two residues between the conserved Cys residues in the active site, and that the different catalytic sequences might confer diverse enzymatic activity and substrate specificity (Kunchithapautham et al., 2003). This observation supports the versatility of the Trx molecule and reflects its role in anti-oxidative protection of the host.

1.2.2 Biological roles of thioredoxin system

Thioredoxins are proteins that act as antioxidants by facilitating the reduction of other proteins by cysteine thiol-disulfide exchange (Holmgren, 1989a; Nordberg and Arnér, 2001). Trx proteins function as a redox sensor and transducer that impart information on the cellular redox status to proteins that do not possess their own redox-sensitive residues. For instance, the reduced form of Trx1, a mammalian cytosolic isoform, binds to and thereby inhibits the activity of apoptosis signal-regulating kinase 1 (ASK1), an upstream activator of the c-Jun N-terminal kinase (JNK) and p38 mitogen-activated protein kinase (MAPK) signaling pathways (Liu et al., 2000). Their alterations have been implicated in cardiovascular diseases (Aviram, 2000), diabetes (Davi et al., 2005), hepatic and renal diseases (Seki et al., 2002), Alzheimer's disease (Nunomura et al., 2006), Parkinson's disease (Wood-Kaczmar et al., 2006) and rheumatoid arthritis (Hitchon and El-Gabalawy, 2004).

Some human cancers show greatly increased Trx expression (Fujii et al., 1991; Gasdaska et al., 1994), indicating a potential role of Trx in tumorigenesis. Its interactions with different proteins of the metabolic and signaling pathways make Trx an attractive

target for therapeutic interventions. As illustrated in Fig.1.8, there are numerous systems with thiol-dependent redox mechanisms, which are related to important pathological states and human diseases.

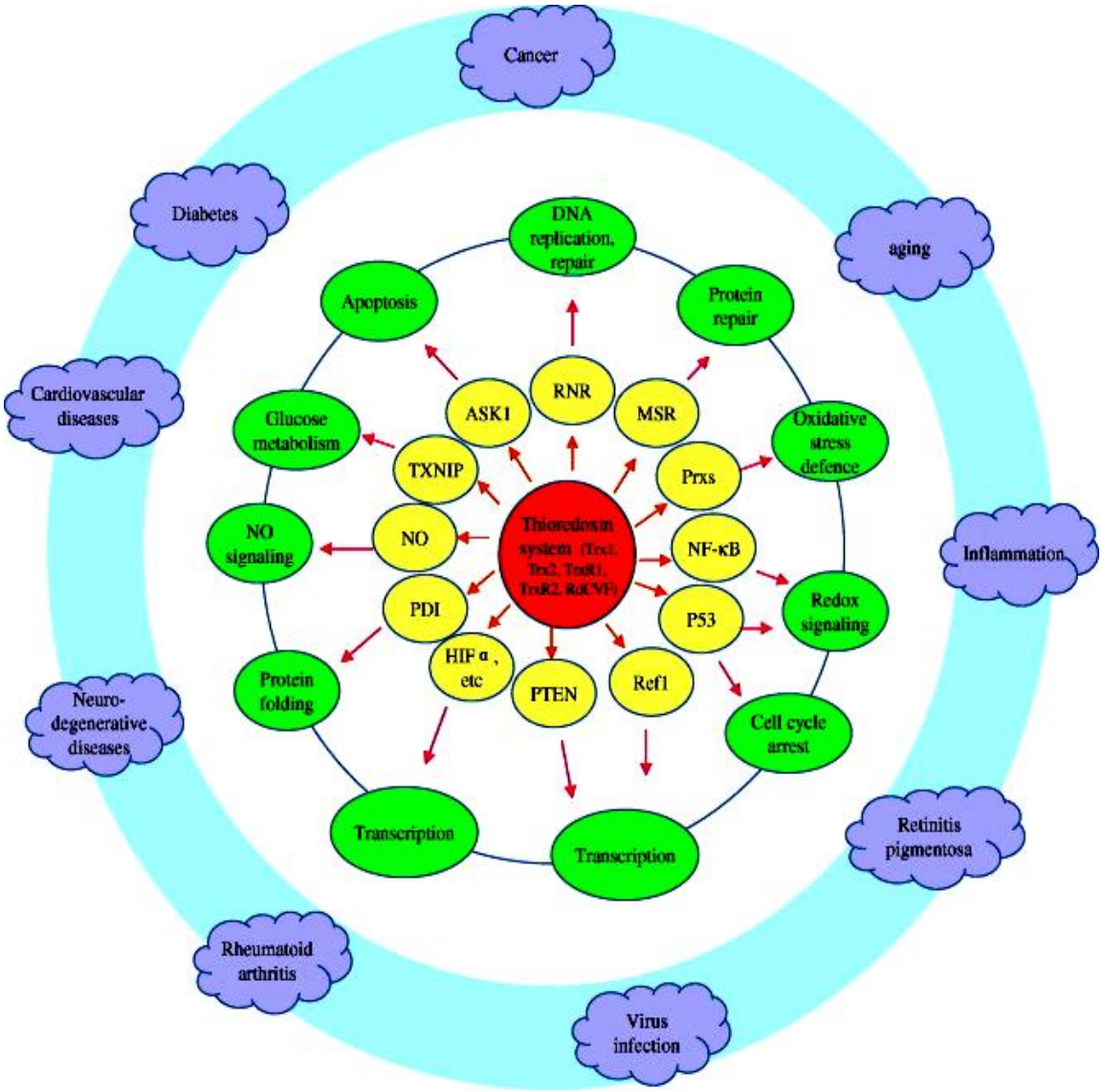


Figure I.8: Biological roles of the thioredoxin system. Adopted from (Holmgren and Lu, 2010).

TrxR1 and Trx1 in cytosol and nucleus, TrxR2 and Trx2 in mitochondria play critical roles in biochemical mechanisms. Trx1 reduces ribonucleotide reductase (RNR),

which is essential for DNA synthesis. Trx provides the electrons to methionine sulfoxide reductase (MSR), and Trx-dependent peroxidases (peroxiredoxins, Prxs) to repair of methionine sulfoxide residues in proteins or to protect against oxidative stress via removing hydrogen peroxide and peroxynitrite, respectively.

The Trx system operates in cellular redox signaling by controlling the activity of many transcription factors such as NF-jB, p53, Ref-1, HIFa, PTEN, AP-1, and glucocorticoid receptor, and so on (Lillig and Holmgren, 2006). Trx-(SH)2 can bind to and inactivate apoptosis signal-regulating kinase (ASK1) and regulate ASK1 dependent apoptosis (Saitoh et al., 1998). Thioredoxin interacting protein (TXNIP) binds to Trx-(SH)2 and regulates Trx activity (Yoshioka et al., 2004). TrxR can reduce protein disulfide isomerase (PDI),a critical player for disulfide bond formation (Lillig and Holmgren, 2006). Trx-(SH)2 affects the activity of some key proteins, such as caspases, via control of protein S-nitrosylation and denitrosylation.

The expression of Trx system proteins has been found to be changed in many diseases, including cancer, diabetes, cardiovascular and neurodegenerative diseases or rheumatoid arthritis (Lillig and Holmgren, 2006). Under the conditions of aging, inflammation and virus infection Trx levels are also changed. A Trx-like protein, rod-derived cone viability factor (TXNL-6) has been shown to be an essential factor to prevent cone loss, which induces retinitis pigmentosa (Fridlich et al., 2009; Wang et al., 2008)

1.2.3 Carcinoscorpius rotundicauda thioredoxin related protein 16

In addition to Trx, there are several Trx-like proteins, which possess the thioredoxin fold and function and exist in different sizes (12-32 kDa). Recently, a 16 kDa protein that contains a WCPPC motif has been isolated and characterized from the horse-shoe crab species *Carcinoscorpius rotundicauda*, and named as *Carcinoscorpius rotundicauda* Trx-like protein 16 (Cr-TRP16) (Wang et al., 2007). Notably, 16 kDa Trx-like proteins are functionally similar to the 12 kDa Trxs and contain two Cys residues at the active site. Unlike other 16 kDa Trx like proteins, Cr-TRP16 contains an additional Cys residue (Cys15, at the N-terminus) but it lacks the extra C-terminal Cys residue (Fig. 1.9), which is found in mammalian Trxs (Weichsel et al., 1996b).

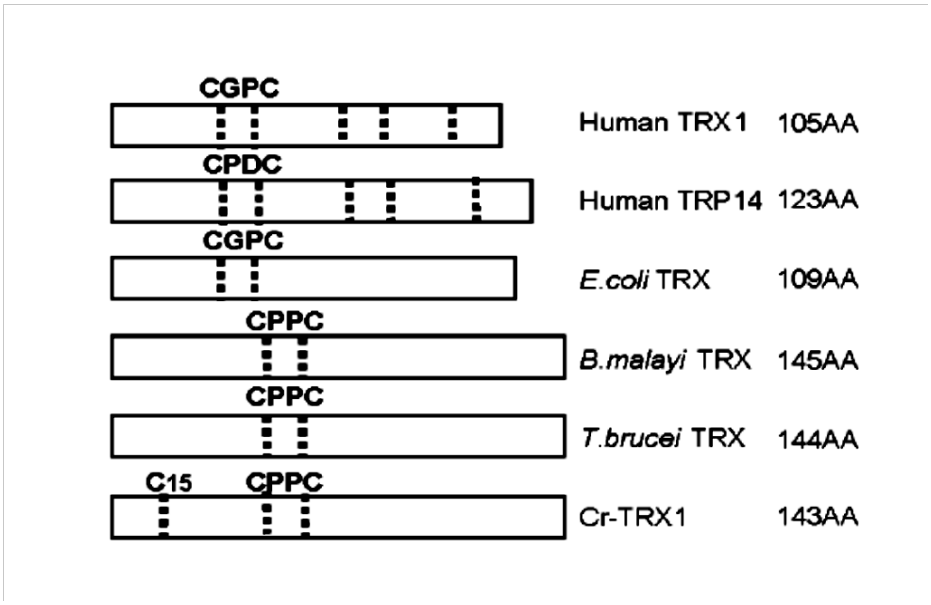


Figure I.9: Comparison of CXXC motif, numbers and positions of cysteine residues in various Trxs.The Cys residues are indicated by dashed lines. The sequences of catalytic motif are indicated as well. Adopted from (Wang et al., 2007)

Interestingly, like other 16 kDa thioredoxins, Cr-TRP16 also lacks the highly conserved Asp26, which is present in the *E. coli* Trx, and has been shown to play a crucial role for catalytic activity. Based on human Trx1 studies, it has been shown that mammalian 12 kDa Trxs have three conserved cysteine residues at positions 62, 69 and 73, besides the two conserved cysteines in the active site motif. Those Cys residues may impart unique biological properties to mammalian 12 kDa Trxs (Holmgren, 1989b). Furthermore, crystal structure also reveals that human Trx1 can form a dimer via Cys73, and the active site residues are buried in the dimer interface (Weichsel et al., 1996a). Surprisingly, although most of the 16 kDa Trxs, like the bacterial 12 kDa Trx, do not contain an extra Cys residue besides the ones in the active site, Cr-TRP16 possesses an extra Cys residue at the N-terminus (Wang et al., 2007).

Interestingly, while Cr-TRP16 lacks the C-terminal extra Cys residue, it is possible it can form a dimer via the N-terminal Cys15. However, at this stage, it is still unclear if such dimer also exists under physiological conditions, and therefore, the role of the extra N-terminal Cys in dimer formation needs further examination. Further mutational and structural studies would be useful to define the function of the extra N-terminal Cys15 residue in Cr-TRP16.

1.2.4 The influence of Cr-TRP16 in NF-κB signaling pathway

For a subtle balance between oxidants and antioxidants which is crucial for homeostasis, there is emerging evidence that the highly conserved intracellular redox state of Trx regulates signal transduction and gene expression (Sen and Packer, 1996). NF-kB is one of the redox-regulated proteins.

It is a member of the Rel family of transcription factors and exists in the cytoplasm in complex with its inhibitor protein, IκB (Verma, 2004). A wide variety of stimuli, including TNF-α, phorbol ester, bacterial lipopolysaccharide, and virus infection, can activate NF-κB (Menon et al., 1995).

NF-κB activation involves site-specific phosphorylation of IκB-α which results in the dissociation of the complex, unmasking of the NF-kB nuclear localization signal and nuclear entry of NF-κB to bind to its cognate DNA (Zabel and Baeuerle, 1990). Thus, for the stimuli that potently and rapidly modulate the nuclear activity of NF-κB, the IκB-α may represent a critical activation target (Verma, 2004). Several studies have suggested that Trx is a specifically potent antioxidant for NF-κB activation (Schenk et al., 1994). Similar to the human Trx1, Cr-TRP16 also up regulates the TNFα-induced NF-κB activation.

Cr-TRP16 exists as a dimer in the oxidized state. However, the underlying mechanism of dimerization and the regulation of NF-κB by Cr-TRP16 are so far unknown. For example, what is the physiological role of the Trx-dimer? Is Cys15 involved in Cr-TRP16 dimer formation? How is NF-κB regulated under oxidized and reduced cellular environments? It is possible that the oxidant-induced dimerization facilitates sensing cellular oxidative stress. Dimerization might remove Trx from the redox cycle, which is catalyzed by TrxR, since the dimer is not a suitable substrate for TrxR. In addition, dimerization of secreted Trx in the relatively oxidizing extracellular environment could be a way of limiting the growth stimulating effects of Trx (Ren et al., 1993). Thus, characterization of Trxs and Trx-like proteins from different organisms will

enhance the understanding of redox controls, which is another important oxidoreductase (Holmgren, 1989b).

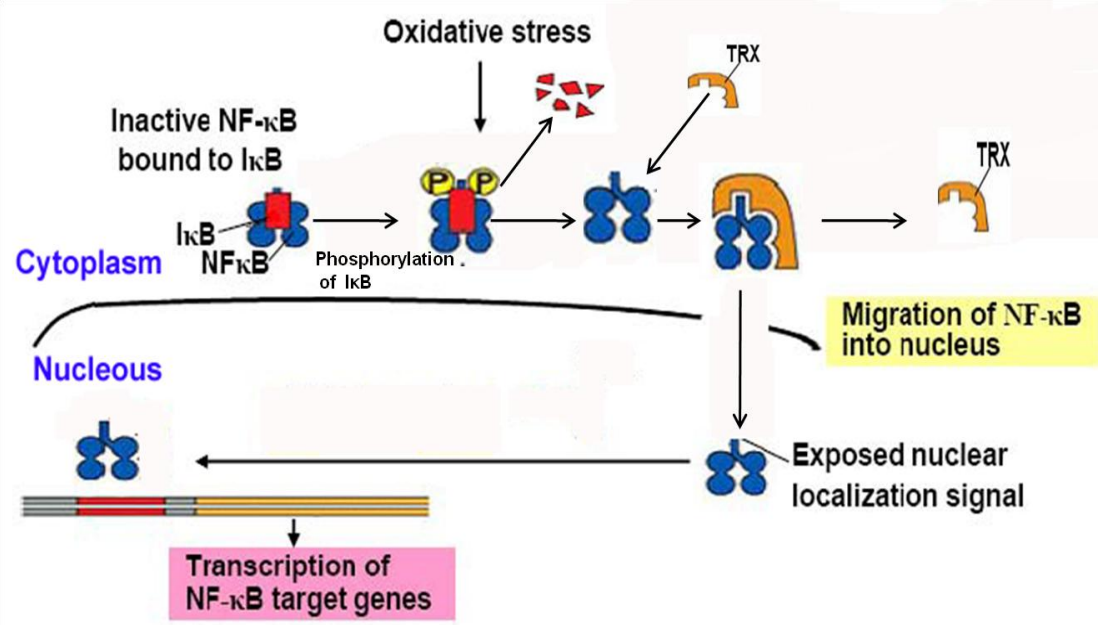


Figure I.10: Activation of NF-κB signaling pathway involves Trx.Oxidative stress causes activation signals to phosphorylate IκB which results in the dissociation of IκB from NF-κB and ubiquitination of IκB. Activated NF-κB translocates into the nucleus, and Trx enhances NF-κB DNA binding, resulting in the transcription of NF-κB dependent genes. Adapted with modification from Alberts *et al* (2002).

In order to understand the structure and functional implications of Cr-TRP-16 the NMR structure has been determined and subsequent biophysical studies and ex vivo studies identified the crucial regulatory role of Cr-TRP16 in NF-kB activity (Fig.1.10). In chapter 2 of this thesis these details are reported. In the next part of the general introduction chapter, the various strategies for structure based protein design and engineering to develop drugs against a specific protease inhibitor.

1.3 PROTEIN DESIGN AND ENGINEERING

Proteins are essential to life and most of them are enzymes, which serve as catalysts for nearly all the chemical reactions that define cellular metabolism. Their enormous rate accelerations and exact selectivities also make them extremely valuable outside the cell. Enzymes are being used increasingly in research, industry and medicine. Particularly, protein engineering, an integration of structural biology, chemistry, molecular biology and genetics, has much to offer for the modern day chemical and pharmaceutical industries.

Protein engineering is undergoing the most profound and exciting transformation in its history(Georgiou and DeWitt, 1999). It promises unprecedented expansion in the scope and applications of modified or improved enzymes with desired physical and catalytic properties. Two complementary strategies are currently available:

1. Rational redesign (Chen, 1999; Chen et al., 1995; Chen et al., 1996) and
2. Directed evolution(Arnold, 1997; Stemmer, 1994).

Although both approaches have met with great success, each has limitations. These techniques are not mutually exclusive (Fig. 1.11). The following section discusses some of these strategies in more detail.

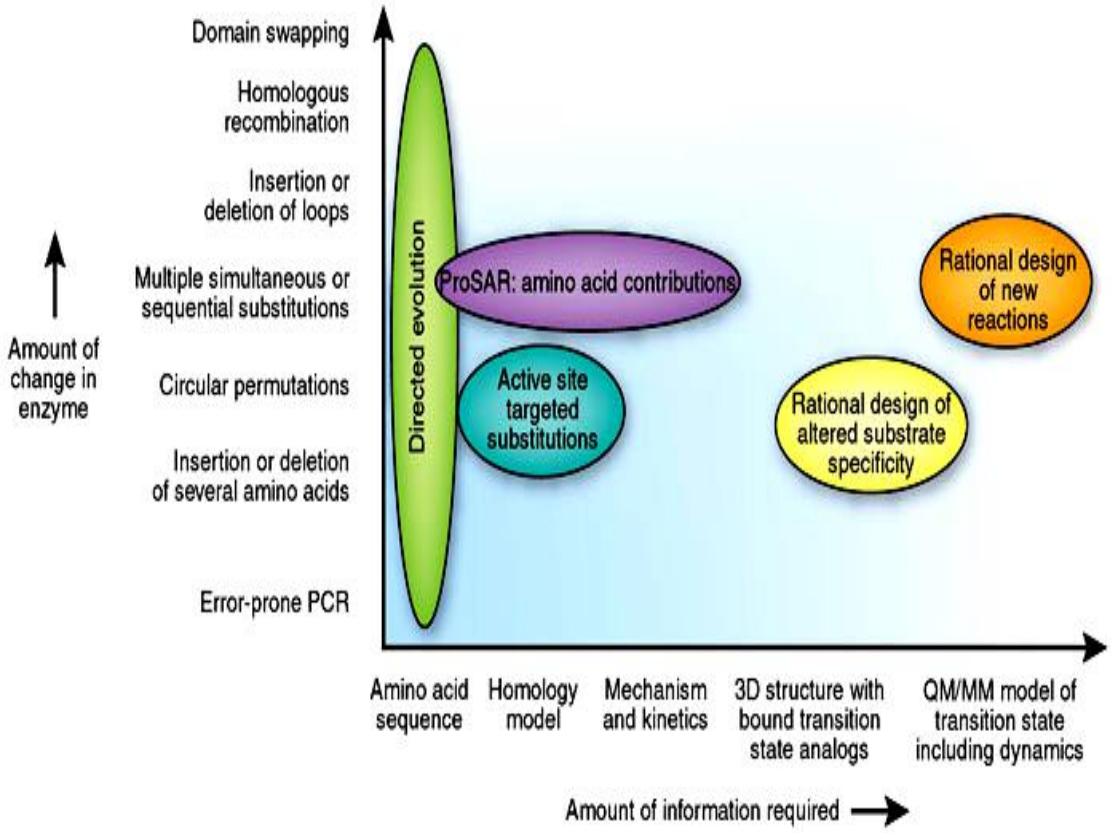


Figure I.11: Various strategies for protein design and engineering.Protein engineering methods differ widely based on the degree that the enzyme is changed and the amount of information available for rational design. ProSAR, or protein structure–activity relationship, is described in the text. 3D, three-dimensional; QM/MM, quantum mechanics or molecular mechanics(Kazlauskas and Bornscheuer, 2009).

1.3.1 Directed evolution strategies

In directed evolution, random mutagenesis is applied to a protein, and a selection regime is used to pick out variants that have the desired qualities. Further rounds of mutation and selection are then applied. This method mimics natural evolution and generally produces superior results to rational design (Nyska et al., 2002; Stemmer, 1994). An additional technique known as DNA shuffling mixes and matches pieces of

successful variants in order to produce better results. This process mimics the recombination that occurs naturally during sexual reproduction. The great advantage of directed evolution is that it requires no prior structural knowledge of a protein, nor is it necessary to be able to predict what effect a given mutation will have.

Indeed, the results of directed evolution experiments are often surprising in that desired changes are often caused by mutations that were not expected to have that effect. The drawback is that they require high-throughput, which is not feasible for all proteins. Large amounts of recombinant DNA must be mutated and the products screened for desired qualities. The sheer number of variants often requires expensive robotic equipment to automate the process. Furthermore, not all desired activities can be easily screened for.

1.3.2 Rational redesign strategies

Rational design means making an educated guess about which amino acids in an enzyme to alter, and then making the changes using targeted, or 'site-directed' mutagenesis of the corresponding gene. This approach works best when a high-resolution, three-dimensional structure of the enzyme has already been obtained using X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy. Particularly helpful are the structures in which the enzyme is 'caught in the act' with its substrate, a substrate-like inhibitor or product. These structures help to identify the residues that are at the business centre of the enzyme - its 'active site'. Even in the absence of a structure, it is still possible to create a computer model of a protein based upon the structure of the enzymes that have similar sequences; a large database that correlates sequences with structures is readily available.

1.3.3 Rational design and engineering of therapeutic proteins

An increasing number of engineered protein therapeutics are currently being developed, tested in clinical trials and marketed for use. The controlled manipulation of the physical, chemical and biological properties of proteins, enabled by structure-based simulation, is now being used to refine established rational engineering approaches and to advance new strategies. These methods provide clear, hypothesis driven routes to solve problems that plague many proteins and to create novel mechanisms of action.

A variety of strategies have emerged for modulating protein properties, such as efficacy, stability, specificity, immunogenicity and pharmacokinetics (PK). Mechanisms for altering these properties include manipulation of primary structure, incorporation of chemical and post-translation modifications and utilization of fusion partners. The most common route to optimization is site-directed mutagenesis, which is often performed in a brute force or trial-and-error manner. A smaller number of examples exists whereby semi-rational application of diversity methods, such as phage display, has been used to optimize a therapeutic candidate. Important recent developments are the creation and successful application of rational protein design methods and the determination of an increasing number of high-resolution protein structures. The controlled modification of specific biophysical properties of proteins can potentially impact a variety of therapeutic features (Table1.3).

Table I.3: The biophysical properties of proteins that can be optimized to obtain desired therapeutic outcomes

	Enable discovery	Mechanism of action	Pharmaco-kinetics	Immuno-genicity	Route of administration	Cost of goods	Shelf life
Stability	×				×	×	×
Solubility	×		×	×	×	×	×
Receptor binding affinity and specificity		×	×				
MHC binding affinity			×	×			
Oligomerization state		×	×		×		
Chemical modifications		×	×	×	×		×
Posttranslational modifications		×	×	×			×
Sequence diversity							
Conformational state		×					

[a] Abbreviation: MHC, major histocompatibility complex. Adapted from (Marshall et al., 2003).

An important subset of rational engineering methods consists of the approaches that utilize high-resolution, 3D structure information. The most sophisticated of these methods offers an extraordinary level of control over protein sequence and structure, a mechanism to explore sequence combinations that extends far beyond natural diversity, and the ability to couple multiple constraints algorithmically for simultaneous optimization of several protein properties. Furthermore, proven hypotheses can be reapplied to additional protein systems, thus saving discovery cost and time (Voigt et al., 2001; Voigt et al., 2002).

Table I.4: Some examples of protein engineering.

S.N.	Example	Method	Reference
1	Increased rate and extent of biodesulfurization of diesel by modification of dibenzothiophene mono-oxygenase	RACHITT	Coco *et al.* 2001
2	Generation of a subtilisin with a half-life at 65°C that is 50 times greater than wild type by recombining segments from five different subtilisin variants	StEP	Zhao *et al.* 1998
3	Conversion of a galactosidase into a fucosidase	Shuffling	Zhang *et al.* 1997
4	Enhanced activity of amylosucrase	Random mutagenesis plus shuffling	Van der Veen *et al.* 2004
5	Generation of novel DNA polymerases from a combination of rat DNA polymerase beta and African swine fever virus DNA polymerase X	SCOPE	O'Maille *et al.* 2002
6	Generation of novel β-lactamase by recombining two genes with 40% amino acid identity and 49% nucleotide sequence identity	SISDC	Hiraga & Arnold 2003

Adapted from (Marshall et al., 2003)

The exquisite specificity of biological therapeutics for their clinical targets has led to their continued development and application as medicines, despite competition from small molecule drugs. Several engineered protein therapeutics are currently being marketed (Table 1.4 and 1.5), and the annual sales of protein therapeutics are projected to exceed US$59 billion in 2010, which is twice the revenue generated in 2001

(http://www.pharmafile.com/Pharmafocus/ Features/feature.asp?fID=281). For well-validated targets, naturally occurring protein interaction partners constitute preselected 'lead' compounds with high affinity and specificity. However, because natural proteins are not evolved for utilization as drugs, lead optimization is frequently beneficial for development of a protein therapeutic. Modifications can influence the mechanism of action, side effects and efficacy, and satisfy practical constraints such as production costs, intellectual property and dosing frequency.

Table I.5: Engineered protein therapeutics on the market[a]

Name	Family	Indication	Modification	Property
Proleukin® (aldesleukin)	IL-2	Cancer	Mutated free cysteine	Decreased aggregation; improved bioavailability
Betaseron® (interferon beta-1b)	IFN-β	Multiple sclerosis	Mutated free cysteine	Decreased aggregation
Humalog® (insulin lispro)	Insulin	Diabetes	Monomer not hexamer	Fast acting
NovoLog® (insulin aspart)	Insulin	Diabetes	Monomer not hexamer	Fast acting
Lantus® (insulin glargine)	Insulin	Diabetes	Precipitates in dermis	Sustained release
Enbrel® (etanercept)	TNF receptor	Rheumatoid arthritis	Fc-Fusion	Longer serum half-life; increased avidity
Ontak® (denileukin diftitox)	Diptheria toxin-IL-2	Cancer	Fusion	Targets cancer cells
PEG-Intron® (peginterferon alfa-2b)	IFN-α	Hepatitis	PEGylation	Increased serum half-life; weaker receptor binding
PEGasys® (peginterferon alfa-2a)	IFN-α	Hepatitis	PEGylation	Increased serum half-life; weaker receptor binding
Neulasta™ (pegfilgrastim)	G-CSF	Leukopenia	PEGylation	Increased serum half-life
Oncaspar® (pegaspargase)	Asparaginase	Cancer	PEGylation	Decreased immunogenicity; increased serum half-life
Aranesp® (darbepoetin alfa)	Epo	Anemia	Additional glycosylation sites	Increased serum half-life; weaker receptor binding
Somavert® (pegvisomant)	Growth hormone	Acromegaly	PEGylation; binding site mutations	Novel mode of action; increased serum half-life

Adapted from (Marshall et al., 2003).

1.3.4 Structure-based design of altered specificity

For almost two decades, drug discovery and optimization efforts have utilized protein crystallographic methods to aid compound design. The target-based approach to drug discovery enables protein targets to be generally classified by both functional and structural characteristics. Drug design, also sometimes referred to as rational drug design or structure based drug design, is the inventive process of finding new medications based on the knowledge of the biological target(Gohlke et al., 2000).

Using a refined X-ray structure as a starting point for modeling the interaction between a protein and its ligands, several designed mutations that alter the substrate or cofactor specificity of an enzyme have been successfully made and tested during the past years e.g. substrate specificity of ∝-lytic protease has been changed (Bone et al., 1989; Conley et al., 2011; Huang et al., 2011; Huang et al., 2010; Kuwada et al., 2011; Lippow et al., 2010; Pan et al., 2010; Pleiss, 2011; Ren et al., 2011; Vouilleme et al., 2010). Improvement of thermo stability of the truncated fibrobacter succinogenes 1,3-1,4-β-d-glucanase (Huang et al., 2011), development of an artificial, tri-pedal, protein-based motor to move unidirectionally along a linear track (Kuwada et al., 2011), application of protein design in metabolic engineering(Pan et al., 2010; Pleiss, 2011), and innovative two-helix small protein scaffold as a novel and robust platform for imaging and therapy applications (Ren et al., 2011).

Most of these studies have relied on pairs of existing homologous proteins with different binding specificities examination of the differences in their amino acid sequences (in the context of a high-resolution structure of the protein) has provided the basis for the mutant designs.

The power of structure-based rational design is most impressive when it is used to generate novel mechanisms of action. In some cases, increasing the binding affinity for a target protein can produce an increase in biological activity. In other cases, it is possible to reduce undesired biological activities by decreasing the affinity for non target molecules. Many of the reported changes in substrate specificity have arisen fortuitously in the course of experiments designed to probe the basis of catalytic activity. Here below are a few examples for such strategy for the development of novel functions.

1.3.4.1 Engineering the substrate specificity in the cAMP-dependent protein kinase

The nucleotide activator specificity of cAMP-dependent protein kinase has been altered by a single amino acid change designed on the basis of predicted structural differences between homologous enzymes (Shabb et al., 1990). The cAMP kinase, a key regulatory mammalian enzyme, strongly prefers cAMP over cGMP as an activator, whereas the enzyme cGMP kinase shows the reversed specificity. Both proteins are homologous to the cAMP-binding catabolite gene activator protein, a bacterial protein whose structure has been determined crystallographically. Although most binding site residues are conserved, one amino acid appears to play a critical role in distinguishing between ligands: Ala334 in the cAMP-dependent enzyme is predicted to make no significant contacts with the nucleotide. The Ala334Thr substitution was made for the cAMP kinase, and its effects on specificity monitored. Whereas the wild-type protein productively binds cAMP with 200-fold higher affinity than cGMP, the engineered mutant actually has a 16-fold higher affinity for cGMP than for cAMP (Shabb et al., 1990). This example demonstrates the power of using the information contained in a

protein structure to relate known functional differences between homologous enzymes to their amino acid differences.

1.3.4.2 Engineering substrate specificity in alcohol dehydrogenase I

In many cases in which substrate specificity has been altered by directed mutagenesis, the engineered enzymes show much broader specificity than their wild-type precursors. In some instances, this broadened specificity has been anticipated (hoped for) and engineered into the design. Independent work on *Saccharomyces cerevisiae* alcohol dehydrogenase I (Creaser et al., 1990) and *Bacillus stearothermophilus* L-lactate dehydrogenase (Wilks et al., 1990), using homology modeled structures of the proteins to design mutations, has yielded engineered enzymes, which efficiently use a wide range of substrates to catalyze the same basic reaction.

Past success of protein design and engineering strategies provide not only invaluable lessons but are also a useful resource. Similar to these examples of protein engineering, our objective was to ultimately apply the structure based drug design approach on one of the most important class of enzymes i.e. proteases/protease inhibitors. We hope to understand sufficiently the relationship between structure and function so that we can rationally alter enzyme specificity at will. Although attaining such a state of enlightenment still eludes us, the success of recent structure-based experiments is quite encouraging. Specific targeting of protease function is achieved through tight regulation, narrow distribution and limited substrate recognition.

1.3.4.3 Engineering the substrate specificity in proteases inhibitors

Proteases have an important role in many signaling pathways, and represent potential drug targets for diseases ranging from cardiovascular disorders to cancer, as well as for combating many parasites and viruses (Adams and Kauffman, 2004). Although inhibitors of well-established protease targets such as angiotensin-converting enzyme and HIV protease have shown substantial therapeutic success, developing drugs for new protease targets has proved to be challenging in recent years (Turk, 2006b).

The protein engineering of proteases and protease inhibition is pursued for several reasons. The application of both site-directed and random mutagenesis technologies to the study of a specific protease or protease inhibitor has the potential to provide a wealth of structural, mechanistic and functional information, regarding that specific protein (Watzke et al., 2008). The questions that could not be even asked 10 years ago can now be answered by these techniques and, from the answers, even more challenging questions are developed. It is a common feature of protein engineering studies that each experiment, whether the outcome is positive or negative, has the potential to provide additional insight into protein structure-function relationships. Thus, for example, information regarding the enhanced stability of a protease obtained through protein engineering can be used to understand the general principles underlying protein stability.

Natural inhibitors of proteases are often characterized by broad selectivity and they can inhibit multiple related proteases. Protein engineering efforts using natural inhibitors as a starting point have yielded selective inhibitors. As an example, the classic thrombin inhibitor hirudin, initially isolated from the medicinal leech, led the way to selective thrombin antagonists.

Another example is ecotin, a natural *Escherichia coli* protein that can inhibit multiple trypsin-fold proteases and has been engineered, using a library display approach, to create a potent and selective inhibitor of plasma kallikrein (Stoop and Craik, 2003). The Kunitz domain, present in many natural protease inhibitors, has also been a focus of protein engineering efforts to generate specific protease inhibitors[86]. In fact, a Kunitz domain-based inhibitor of plasma kallikrein, ecallantide (Kalbitor; Dyax) has recently been approved for the treatment of hereditary angioedema (Levy and O'Donnell, 2006).

The macromolecule approach of developing protease inhibitors has the distinct advantage over the small-molecule approach in that the former can target larger surfaces on the protease and may yield more selective inhibitors. Structural studies of a protein-based inhibitor in complex with its specific protease could provide insights for the rational design of small-molecule inhibitors that mimic the binding interactions in terms of overall chemical architecture.

Serine proteases are among the first enzymes to be studied extensively (Neurath, 1985). Interest in this family has been maintained in part by an increasing recognition of their involvement in a host of physiological processes. In addition to the biological role played by digestive enzymes such as trypsin, serine proteases also function broadly as regulators through the proteolytic activation of precursor proteins (Neurath, 1984; Van de Venet al., 1993). Examples of this regulation include the processing of trypsinogen by enteropeptidase to produce active trypsin (Jing et al., 1999) and the cascades of zymogen activation that control blood clotting (Davie et al., 1991).On the basis of the known structures of the serine proteases, changes have been engineered into serine protease inhibitors (SPIs) in order to modify their specificities (Table.1.6).

Table I.6: Examples of successful strategies applied for the design and development of serine protease inhibitors

Disease	**Protease (class)**	**Drug name (trade name; company)**	**Drug type**	**Refs**
Diabetes	DPP4 (serine)	Sitagliptin (Januvia; Merck)	Non-peptidic	(Mattos et al., 2006)
Coagulation	Thrombin (serine)	Desirudin (Revasc/Iprivask; Novartis)	Peptidic	(Matheson and Goa, 2000)
Coagulation	Factor Xa (serine)	Rivaroxaban (Xarelto; Bayer)	Non-peptidic	(Drag and Salvesen, 2010)

Adopted with modification from (Drag and Salvesen, 2010).

The evolutionary mechanisms of SPIs serve to increase their variety and expand their functions, thus helping to meet the demands of the repertoire of endogenous and exogenous SPs an organism encounters. More importantly, the structural changes of a protease inhibitor in complex with its target protease can provide useful information on the interaction between the two proteins, thus allowing the development of analogs of that inhibitor with increased affinity towards the protease to achieve greater inhibition capacity. This motivated us to modify the selectivity of a serine protease inhibitor from *Carcinoscorpius rotundicauda* (CrSPI-1-D1) to specifically target thrombin. The details of this part of thesis have been mentioned in Chapter-III.

1.4 OBJECTIVES

There are two target proteins, Cr-TRP16 and Cr-SPI-1-D1, in this thesis. Both proteins are from *Carcinoscorpius rotundicauda (Cr).* The first part of the thesis (Chapter II) deals with the structural and functional characterization of a thioredoxin like protein 16 , whereas the second part of the thesis (Chapter 4) **be careful. Some place you are writing Chapter II (Roman) and in other places you write Chapter 4 (Arabic); be consistent and use only Arabic everywhere. represents the best example of a structure based protein engineering approach to modify the substrate specificity of a serine protease inhibitor (SERPIN). Thus the specific objective of this thesis is two-fold:

1. Structural and functional characterization of Cr-TRP16 (Chapter 2):
    a) To solve the NMR structure of Cr-TRP16
    b) To characterize the role of Cys15 in dimer formation of Cr-TRP16
    c) To delineate the mechanism of interaction between Cr-TRP16 and NF-kB
    d) To solve the X-ray structure of TXNL-6 (Chapter-3)
2. Structure based protein engineering approach to modify the substrate specificity of a SERPIN (Chapter 4):
        a) To solve the X-ray structure of Cr-SPI-1-D1
        b) To evaluate how rational structure based protein engineering strategy can be used to modify substrate specificity

CHAPTER-II: NMR STRUCTURE OF *carcinoscorpius rotundicauda* THIOREDOXIN-RELATED PROTEIN 16 AND ITS ROLE IN REGULATING NF-KB ACTIVITY

2.1 INTRODUCTION

Reactive oxygen species (ROS) are natural byproducts of several oxidative metabolic pathways and are known to mediate intracellular signaling pathways (Valko et al., 2007). However, excessive production of ROS is detrimental to the cell, due to increased oxidative stress (Marnett, 2000). Thioredoxin (Trx) plays a key role in maintaining an intracellular redox state, which is essential for protecting cells from oxidative damage (Nordberg and Arner, 2001). Trx is a 12 kDa ubiquitous and evolutionarily conserved protein in almost all organisms and functions as a disulfide reductase (Arnér and Holmgren, 2000; Holmgren, 1995). A conserved Cys-X-X-Cys sequence at the active site of Trx is involved in either reducing disulfide bonds or oxidizing sulfhydryls, both intra and extracellularly (Powis and Montfort, 2001b). The two cysteines are maintained in a reduced active form by thioredoxin reductase (TrxR), a selenocysteine-containing protein that uses the reducing power of NADPH (Powis and Montfort, 2001b). The disulfide reductase activity and reduction potential of the cysteines depend largely on the identity of the two XX residues (Carvalho et al., 2006).

Trx proteins also function as a redox sensor and transducer that impart information on the cellular redox status to proteins that do not possess their own redox-sensitive residues. For instance, the reduced form of Trx1, a mammalian cytosolic isoform, binds to and thereby inhibits the activity of apoptosis signal-regulating kinase 1 (ASK1), an upstream activator of the c-Jun N-terminal kinase (JNK) and p38 mitogen-activated protein kinase (MAPK) signaling pathways (Liu et al., 2000). Their alterations

have been implicated in cardiovascular diseases (Aviram, 2000), diabetes (Davi et al., 2005), hepatic and renal diseases (Seki et al., 2002), Alzheimer's disease (Nunomura et al., 2006), Parkinson's disease (Wood-Kaczmar et al., 2006) and rheumatoid arthritis (Hitchon and El-Gabalawy, 2004). Some human cancers show greatly increased Trx expression (Fujii et al., 1991; Gasdaska et al., 1994), indicating a potential role of Trx in tumorigenesis. Its interactions with different proteins of the metabolic and signaling pathways make Trx an attractive target for therapeutic interventions.

There are several Trx-like proteins, which possess the thioredoxin fold and function, and exist in different sizes (12-32 kDa). Recently, a 16 kDa protein that contains a WCPPC motif has been isolated and characterized from a species of the horseshoe crab, *Carcinoscorpius rotundicauda*, and named as Trx-related protein 16 (Cr-TRP16) (Wang et al., 2007). Notably, 16 kDa Trx-like proteins are functionally similar to the 12 kDa Trxs and contain two Cys residues at the active site. Unlike other 16 kDa Trx like proteins, Cr-TRP16 contains an additional Cys residue (Cys15, at the N-terminus) but it lacks the extra C-terminal Cys residue, which is found in mammalian Trxs (Weichsel et al., 1996b).

A subtle balance between oxidants and antioxidants in the cell is crucial for homeostasis. There is emerging evidence that the highly conserved intracellular redox state regulates the mechanisms of signal transduction and gene expression (Sen and Packer, 1996). NF-κB is one of the redox-regulated proteins. It is a member of the Rel family of transcription factors and exists in the cytoplasm in complex with its inhibitor protein, IκB (Verma, 2004). A wide variety of stimuli, including TNF-α, phorbol ester, bacterial lipopolysaccharide, and virus infection, can activate NF-κB (Menon et al.,

1995). NF-κB activation involves site-specific phosphorylation of IκB-α which results in the dissociation of the complex, unmasking of the NF-kB nuclear localization signal and nuclear entry of NF-κB to bind to its cognate DNA (Zabel and Baeuerle, 1990). Thus, for the stimuli that potently and rapidly modulate the nuclear activity of NF-κB, the IκB-α may represent a critical activation target (Verma, 2004). Several studies have suggested that Trx is a specifically potent antioxidant for NF-κB activation (Schenk et al., 1994). Similar to human Trx1, Cr-TRP16 also up-regulates the TNFα-induced NF-κB activation.

Cr-TRP16 exists as a dimer in the oxidized state. However, the underlying mechanism of dimerization and the regulation of NF-κB by Cr-TRP16 are so far unknown, for example, what is the physiological role of the Trx-dimer? Is Cys15 involved in Cr-TRP16 dimer formation? How is NF-κB regulated under oxidized and reduced cellular environments? It is possible that the oxidant-induced dimerization facilitates the sensing of cellular oxidative stress. Dimerization might remove Trx from the redox cycle, which is catalyzed by TrxR since the dimer is not a suitable substrate for TrxR. In addition, dimerization of secreted Trx in the relatively oxidizing extracellular environment could be a way of limiting the growth stimulating effects of Trx (Ren et al., 1993). Thus, characterization of Trxs and Trx-like proteins from different organisms will enhance the understanding of redox controls, which is another important oxidoreductase (Holmgren, 1989b).

In order to understand the NF-κB regulation by Cr-TRP16 and to address the above questions, we had undertaken the structural and functional characterization of Cr-TRP16, the first Trx fold structure from an arthropod. In this chapter we report the NMR structure of the reduced form of wild-type Cr-TRP16. Moreover the role of the Cys15

residue in dimerization and NF-κB regulation was studied by site-directed mutagenesis, biophysical and immunoprecipitation techniques.

2.2 EXPERIMENTAL PROCEDURES

2.2.1 Cloning

The full length Cr-TRP16 gene (encoding residues Met1-Arg143) was PCR amplified using a forward (CAGCATATGATGGAATTTATCCAAGGAAT) and a reverse (CCTCTCGAG TCTTGCCCAGTT) primer with an *Nde*I restriction site at the 5′ end and an *Xho*I site at the 3′ end (underlined), respectively and inserted into the pET-22b (+) vector (Invitrogen), which was previously linearized by the corresponding restriction enzymes. For mammalian expression, the Cr-TRP16 gene was PCR amplified with the following primers: Forward: GGAGGATCCATGGAATTTATCCAAGGA and Reverse:ATTCTCGAGTCATTTGTCGTCATCGTCCTTATAGTCTCTTGCCCAGTTC TGGAA that contained *BamH*I and *Xho*I restriction sites (underlined), respectively. The reverse primer introduced a FLAG tag at the C-terminus of the Cr-TRP16 protein. The insert was ligated into the previously linearized expression vector pcDNA3.1/V5-His/lacZ (Invitrogen).

2.2.2 Protein expression and purification

The pET22b:Cr-TRP16 construct was transformed into the bacterial strain *E. coli* BL21 (DE3) for expression. Optimal expression of the Cr-TRP16 protein was achieved by induction with 0.3 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) of 1 liter culture at 20 ºC. Cells were disrupted by a French press and the supernatant was collected after centrifuging at 10,000 g for 1 h at 4 ºC. His-tagged Cr-TRP16 proteins were purified

in two steps using Ni-NTA (Qiagen) affinity chromatography followed by a Superdex-75 gel filtration column on Akta Express (GE Healthcare). The buffer was exchanged to a solution containing 10 mM Tris (pH 7.0), 100 mM NaCl, 5 mM dithiothreitol, 5% glycerol, 1 mM EDTA and finally the protein was concentrated to 18 mg/ml. . For NMR study, ^{15}N-labeled and ^{13}C, ^{15}N-double labeled samples were prepared using M9 minimal medium with $^{15}NH_4Cl$ and ^{13}C-glucose as a sole source of nitrogen and carbon, respectively. Weakly aligned ^{15}N-labeled sample was prepared by the addition of 8 mg ml^{-1} of filamentous phage Pf1 (from ASLA biotech company) for residue dipolar couplings (RDCs) measurement.

2.2.3 NMR experiments and structure determination

All NMR experiments were carried out at 25 °C on a Bruker Avance 800 MHz spectrometer equipped with a TXI cryogenic probe. ^{1}H, ^{13}C and ^{15}N resonance assignments were performed by measuring the 3D HNCACB, 3D CBCACONH (Bax and Grzesiek, 1993) and 3D CCH-TOCSY (Fesik et al., 1990) spectra. Inter-proton distance restraints for structural calculation were obtained from 3D ^{13}C-edited NOESY-HSQC, 3D ^{15}N-edited NOESY-HSQC and 2D NOESY spectra using a 100 ms mixing time. $^{1}D_{NH}$ RDCs were measured using the IPAP method (27). The RDC values were obtained by subtracting the reference value in isotropic solution.Two- and three-dimensional NMR spectra were processed using the NMRPipe program (Delaglio et al., 1995) and data analysis was performed with the help of the Sparky program(Kneller and Kuntz, 1993) .The structure was calculated using the Xplor-NIH 2.24 software package (Schwieters et al., 2003). For the final Cr-TRP16 three-dimensional structure calculation, a total of of 1731 NOE distance restraints, 76 hydrogen bonds and 124 dihedral angle restraints that

were predicted by the TALOS program were (Cornilescu et al., 1999b) used. RDCs were used in the final cooling stage. At each stage, 100 structures were calculated using 30,000 steps of simulated annealing, and a final ensemble of 20 lowest energy structures was selected for figure preparation..

2.2.4 Site-directed mutagenesis

The NMR structure of Cr-TRP16 showed that Cys15 is located on the surface and it has been hypothesized to be involved in dimerization. This residue was mutated to serine in both the pET22b and pcDNA3.1/V5-His/lacZ constructs by the use of the inverse PCR based mutagenesis protocol (Ochman et al., 1988). The C15S mutant protein was expressed following the steps that were described for the wild type protein.

2.2.5 Analytical Ultra Centrifugation (AUC)

The dimeric state of affinity and gel filtration, followed by reverse phase HPLC, purified WT and C15S mutant Cr-TRP16 proteins was investigated by monitoring their sedimentation properties in sedimentation velocity experiments. 500 µl of samples at OD_{280nm} of 0.8 in 10 mM Tris-HCl (pH 7.0), 100 mM NaCl and 5% glycerol were used. Experiments were carried out in the absence of DTT in the buffer. Sedimentation velocity profiles were collected by monitoring absorbance at 280 nm. The samples were centrifuged at 40,000 rpm at 20 °C in a Beckman Optima XL-I centrifuge fitted with a six-hole AN-60 rotor and double-sector aluminum centerpieces and equipped with absorbance optics. The scans were analyzed using the Sedfit program (Brown and Schuck, 2006).

2.2.6 Western blotting

The human cervical epithelial carcinoma (HeLa) cell line was obtained from the American Type Culture Collection (Manassass, VA). Cells were cultured in the Rosewell Park Memorial Institute (RPMI 1640) medium containing 1X antibiotic-antimycotic solution with 10% FBS. The HeLa cells that were transfected with the Cr-Trp16 or C15S Cr-Trp16 plasmids were lysed in lysis buffer [20 mM Tris (pH 7.4), 250 mM NaCl, 2 mM EDTA (pH 8.0), 0.1% Triton X-100, 0.01 mg/ ml aprotinin, 0.005 mg/ ml leupeptin, 1 mM PMSF, and 4 mM NaVO4]. Lysates were then spun on a microfuge at 14,000 rpm for 10 min to remove insoluble material and resolved on 10% SDS gel. Nuclear protein extracts were prepared using a nuclear extraction kit (Active Motif, CA) according to the manufacturer's instructions. Protein concentration was determined with Bradford reagent and 30 µg of nuclear or cytoplasmic extract was taken for immunoblotting. After electrophoresis, proteins were electro transferred to a nitrocellulose membrane, blocked with 5% nonfat milk, and probed with the antibodies of interest overnight at 4 °C. The blot was washed, exposed to horseradish peroxidase (HRP) conjugated secondary antibodies for 1 h, and finally examined by chemluminescence (ECL; Amersham Pharmacia Biotech).

2.2.7 NF-κB DNA binding assay

To determine NF-κB activation, we performed DNA binding assay using the TransAM NF-κB kit (Active Motif, CA), according to the manufacturer's instructions and as previously described (Renard et al., 2001). Briefly, HeLa cells were transfected with the Cr-Trp16 or C15S Cr-Trp16 plasmids for 30 h and then stimulated with TNF-α (1 nM) for 6 h. Nuclear protein extracts were prepared using the nuclear extraction kit

(Active Motif, CA), according to the manufacturer's instructions. Protein concentration was determined with the Bradford reagent and 20 µg of nuclear extract was added to the plates precoated with an NF-κB consensus oligonucleotide sequence and incubated at room temperature for 1 h with continuous shaking at 100 rpm. The plates were washed and incubated with a p65 primary antibody for 1 h at room temperature. After the final incubation with a secondary antibody, the wells were then washed with washing buffer; color was developed by the addition of substrate solution and read after 10 min at 450 nm against a blank, which was read at a reference wavelength of 655 nm.

2.2.8 NF-κB dependent luciferase reporter assay

The NF-κB dependent reporter gene expression assay was performed as described previously (Mukhopadhyay et al., 2001). Briefly, about 10^5 per well HeLa cells were plated in ninety six-well plates in RPMI buffer containing 10% FBS. The NF-κB-responsive elements, linked to a luciferase reporter gene plasmid, were transfected with the wild type CR-TRP16 or C15S Cr-TRP16 plasmids. Transfection was performed using Lipofectamine (Invitrogen), according to the manufacturer's protocols. After 30 h of transfection, HeLa cells were treated with TNF-α (1 nM) for 6 h and then washed and lysed in luciferase lysis buffer (Promega). Luciferase activity was measured by a Tecan (Durham, NC) plate reader, which is equipped with a function to measure luminescence by using a luciferase assay kit (Promega) and was normalized to β-galactosidase activity. All luciferase experiments were performed in triplicate.

2.3 RESULTS

2.3.1 Purification of recombinant Cr-TRP16

For the expression of pET22b, the Cr-TRP16 construct was transformed into the bacterial strain *E. coli* BL21 (DE3). Optimal expression of the Cr-TRP16 protein was achieved by induction of 1 liter culture with 0.25 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 20 °C. Cells were disrupted by a French press and the supernatant was collected after centrifuging at 10,000 g for 1 h at 4 °C. The His-tagged Cr-TRP16 protein was purified in two steps using Ni-NTA (Qiagen) affinity chromatography, followed by a Superdex 75 gel filtration column on Akta Express (GE Healthcare) (Fig. 2.1). The buffer was exchanged to a solution containing 10 mM Tris (pH 7.4), 150 mM NaCl, 5 mM dithiothreitol, 5% glycerol, 1 mM EDTA and finally the protein was concentrated to 15 mg/ml. The quality of the protein was verified using native PAGE and dynamic light scattering (Dyanpro solutions) instrument. Native PAGE and DLS results (Fig. 2.2) show that the Cr-TRP16 protein was homogeneous.

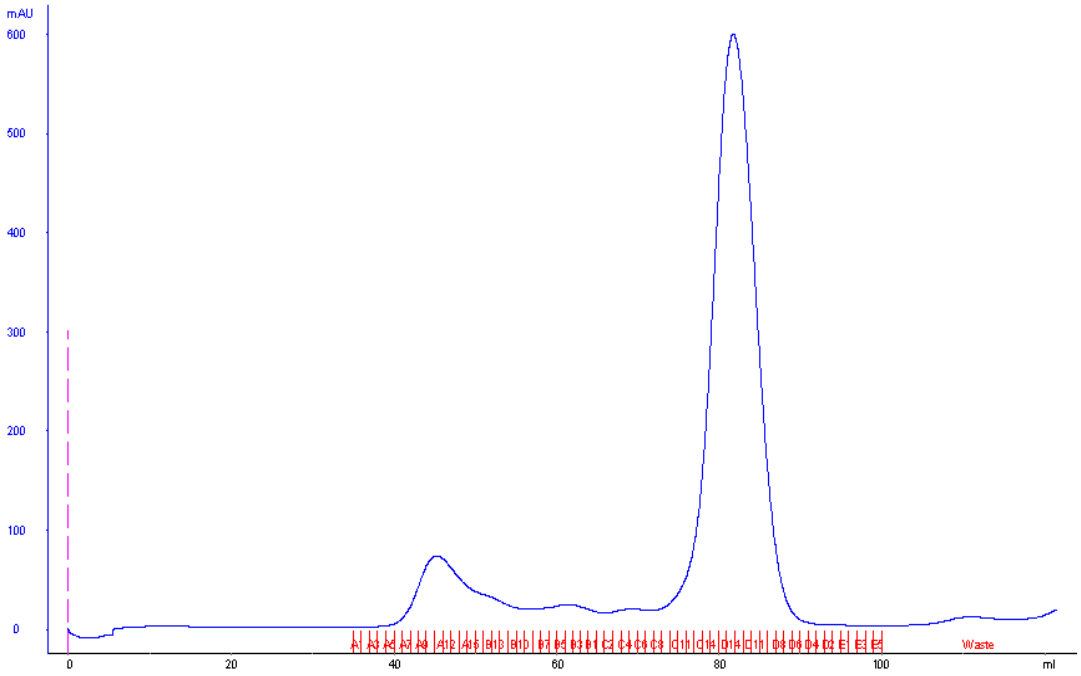


Figure II.1: FPLC profile of Cr-TRP16.After affinity purification the Cr-TRP16 was passed through Superdex-75 column and eluted as single peak, which corresponds to a molecular weight of monomeric Cr-TRP16.

Msr#	Time(s)	Temp(C)	Count Rate	Ampl	Diff Coeff	Radius(nm)	Polyd(nm)	PolydIndx	MW(KDa)	%Mass	Baseline	Sos Error
1	10.0	20.0	91967	0.618	973.	2.20	0.325	0.02	21.3	100.0	1.000	7.45
2	20.0	20.0	91851	0.624	987.	2.17	0.546	0.06	20.6	100.0	1.000	11.6
3	30.0	20.0	92475	0.622	968.	2.21	0.254	0.01	21.6	100.0	1.000	11.3
4	40.0	20.0	92037	0.605	977.	2.19	0.602	0.08	21.1	100.0	1.000	9.08
5	50.0	20.0	92399	0.607	978.	2.19	0.551	0.06	21.0	100.0	1.000	7.19
6	60.0	20.0	92349	0.597	938.	2.28	0.402	0.03	23.2	100.0	1.000	8.14
7	70.0	20.0	93473	0.610	975.	2.20	0.650	0.09	21.2	100.0	1.000	8.19
8*	80.0	20.0	120708	0.453	936.	2.29	0.759	0.11	23.3	100.0	1.038	13.4
9	90.0	20.0	96969	0.527	929.	2.30	0.686	0.09	23.7	100.0	1.003	11.2
10	100.0	20.0	93993	0.579	961.	2.23	0.411	0.03	21.9	100.0	1.000	7.19
11	110.0	20.0	93157	0.603	986.	2.17	0.375	0.03	20.7	100.0	1.000	6.39
12	120.0	20.0	92307	0.636	971.	2.20	0.101	0.00	21.4	100.0	1.000	10.9
13	130.0	20.0	92502	0.639	971.	2.20	0.521	0.06	21.4	100.0	1.000	10.8
14	140.0	20.0	92991	0.645	953.	2.25	0.735	0.11	22.4	100.0	1.000	9.99
15	150.0	20.0	92695	0.643	961.	2.23	0.394	0.03	21.9	100.0	1.000	6.40
16	160.0	20.0	95363	0.635	959.	2.23	0.758	0.12	22.0	100.0	1.000	11.4
17	170.0	20.0	95281	0.645	952.	2.25	0.446	0.04	22.4	100.0	1.000	10.6
Aves:												
Mono		20.0	93238	0.615	965.	2.22	0.485	0.05	21.7	100.0	1.000	9.24
Bi-1		0.0	0	0.000	0.000	0.000	----	----	0.000	0.0	0.000	0.000
Bi-2				0.000	0.000	0.000			0.000	0.0		

Figure II.2: Dynamic light scattering (DLS) profile of Cr-TRP16. DLS data of Cr-TRP16 measured for 15 mg/ml. The hydrodynamic radius and poly dispersity index of the Cr-TRP16 protein shows monodisperse homogeneous sample.

2.3.2 Overall structure

The structure of full length Cr-TRP16 (143 aa) was solved by NMR and refined to a final RMSD of 0.76±0.13 Å (back bone) for 20 best structures (Fig. 2.3 A). Cr-TRP16 consists of seven β-strands in two β-sheets (Fig.2.3 B), similar to other Trxs. The first β-sheet accommodates both parallel and anti-parallel strands (β7↑β6↓β3↑β4↑β5↑). The residues of strands β3 and β4 are mostly buried and conserved in thioredoxins. These residues are mostly hydrophobic and are part of a hydrophobic cluster, formed by Phe31, Tyr32, Phe33, Phe35, Trp37, Phe44, Tyr52, Phe62, Phe66, Phe77 Tyr79, Trp87, Tyr92 and Tyr104. The second β-sheet, formed as a β-hairpin (β1↑β2↓), is solvent accessible

and carries five charged residues: Lys11, Lys12, Arg14, Glu16, and Glu21. This segment is highly variable among thioredoxin like proteins. The longest helix, α1 (residues 43-57), is located on the surface of the protein (Fig. 2.4).

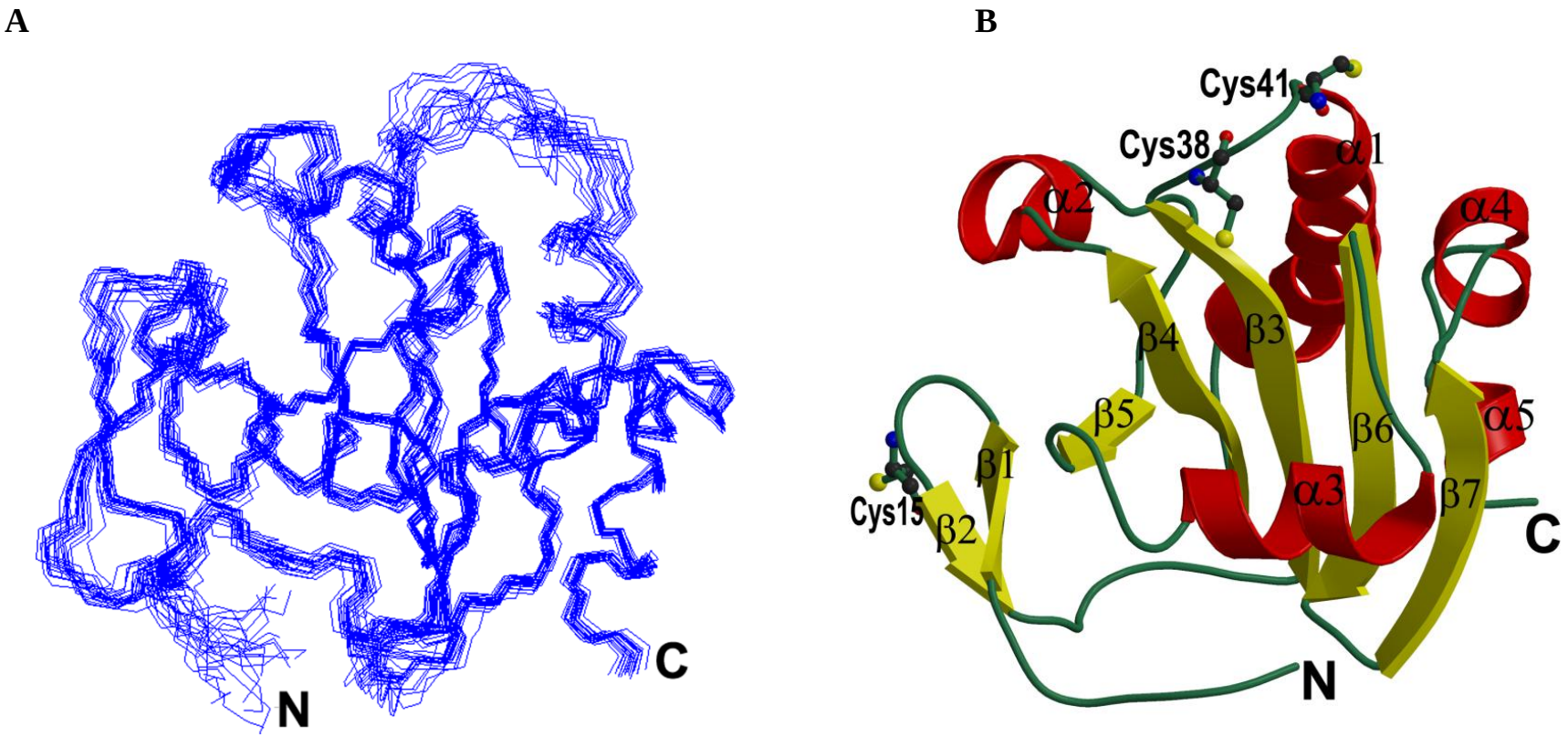


Figure II.3: Structure of Cr-TRP16. *(A)* The best 20 backbone structures of reduced WT Cr-TRP16 thioredoxin after simulated annealing refinement. *(B)* The ribbon diagram of the Cr-TRP16 molecule. Cr-TRP16 consists of seven β-strands in two β-sheets, similar to other Trxs. The first β-sheet (β1β2) is an anti-parallel β-hairpin and the second β-sheet contains both parallel and anti-parallel strands (β7↑β6↓β3↑β4↑β5↑). There are four α-helices (α1, α2, α3 and α4). The active site cysteines (Cys38 and Cys41) and the N-terminal Cys15 residue that promotes dimerization are drawn as ball-and-sticks. This figure and Fig. 2.3 A were prepared using the program Molscript (Kraulis, 1991).

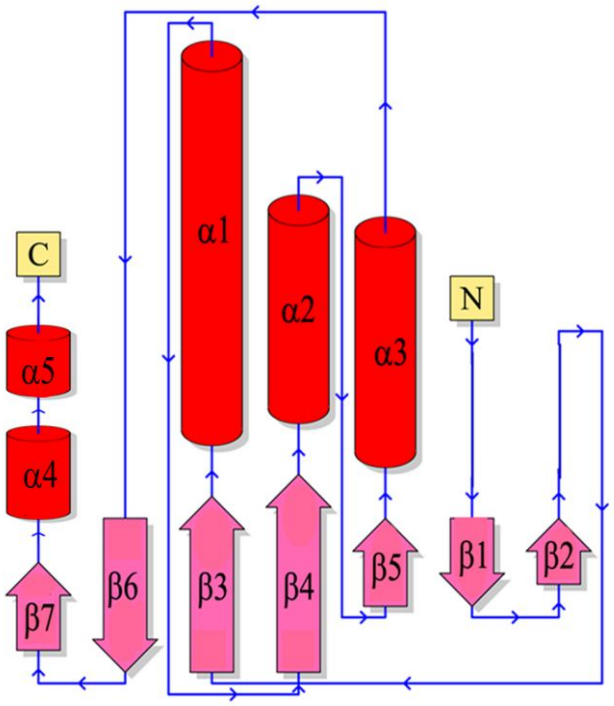


Figure II.4: The topology diagram of Cr-TRP16.

2.3.3 Sequence and structural homology

Proteins with the Trx fold and conserved active site-sequence are classified under the Trx superfamily (Matsuo et al., 2002). Comparison of Trxs and Trx-like protein sequences (Fig. 2.3 A) show that the Trx core domain is well conserved with the central five β-strands, surrounded by four α-helices (Eklund et al., 1991a). Cr-TRP16 has an overall sequence identity of 18% with the 12 kDa Trxs from mammalian, insect, and bacterial species. Most invariant residues are clustered around the active site. The 12 kDa Trxs are characterized by a conserved WCGPC motif in the active site whereas Cr-TRP16 has a WCPPC motif.

Table II.1: NMR data and structure determination details for reduced Cr-TRP16

Parameters	
All NOE distance restraints[a]	1728
Intra-residue	717
Inter-residue	
Sequential (|i–j| = 1)	435
Medium-range (1 <|i–j|≤5)	227
Long-range (|i–j|≥5)	349
Hydrogen bonds restraints	76
Dihedral angle restraints(φ, ψ)[b]	124
Residual dipolar coupling restraints	54
Deviations from idealized covalent geometry[c]	
RMSD of bond lengths (Å)	0.0021±0.00004
RMSD of bond angles (°)	0.3634±0.0086
RMSD of improper angles (°)	0.3329±0.008
Deviations from experimental restraints	
RMSD of distance restraints (Å)	0.0231±0.0005
RMSD of dihedral angle restraints (°)	0.2170±0.057
Ramachandran plot analysis (%)[d]	
Residues in allowed regions	95.5%
Residues in generously allowed regions	3.0%
Residues in disallowed regions	1.5%
Average RMSD from mean structure (Å)[e]	
All residues (2-141)	0.72±0.15
Regular 2° structure region	0.47±0.10

[a] All NMR experiments were performed on a Bruker AVANCE 800 spectrometer equipped with a TXI cryoprobe at 25 °C. 1H, 13C and 15N resonance assignments were assigned by measuring the 3D HNCACB, 3D CBCACONH (Bax and Grzesiek, 1993) and 3D CCH-TOCSY (Fesik et al., 1990) spectra. Interproton distance restraints for structural calculation were obtained from 3D 13C-edited NOESY-HSQC(Görler et al., 1999), 3D 15N-edited NOESY-HSQC(Görler et al., 1999) and 2D NOESY (Haase et al., 2011) spectra using a 100 ms mixing time. All data were processed with NMRPipe

(Delaglio et al., 1995) and analyzed with SPARKY(Kneller and Kuntz, 1993). The distance restraints were obtained by classifying the NOE cross peaks into three categories: strong (1.8–2.9 Å), medium (1.8–3.5 Å), and weak (1.8–5.0 Å). Solution structures of Cr-TRP16 was calculated using the torsion angle dynamics simulated annealing method within XPLOR-NIH(Schwieters et al., 2003).

[b] Dihedral angles of backbone φ and ψ were predicted by TALOS (Cornilescu et al., 1999a) using the chemical shifts of Cα, Cβ, Hα, N, and HN.

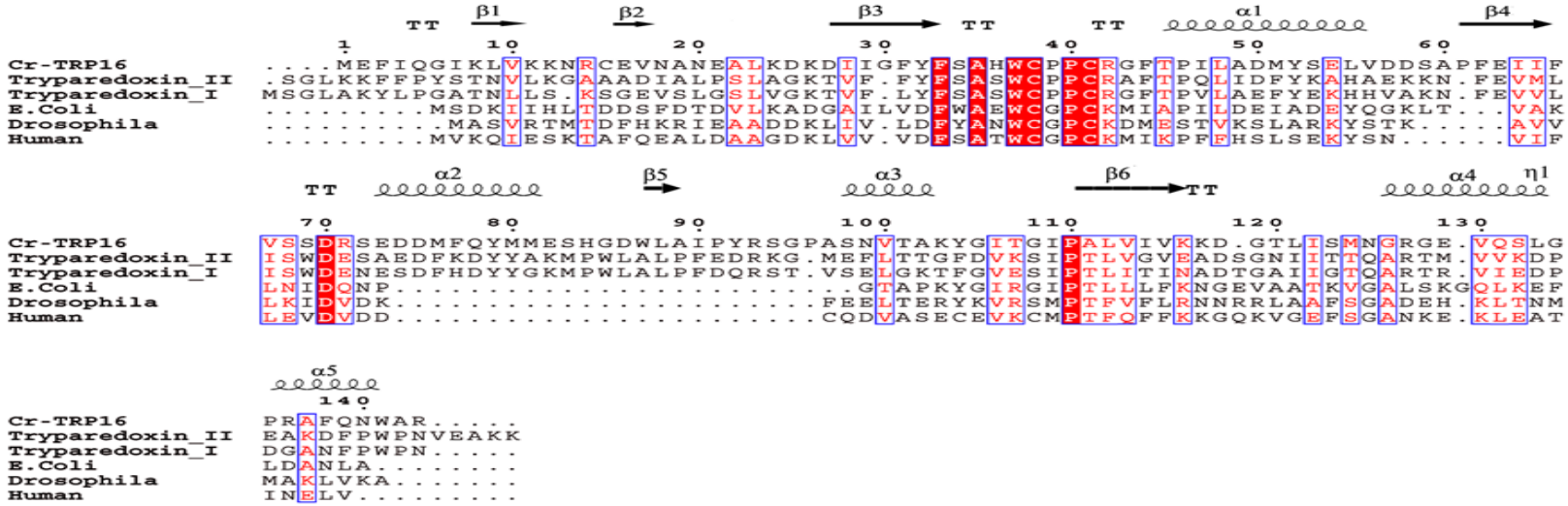
[c] Twenty lowest-energy conformers with no NOE violations greater than 0.5 Å and no torsion angle violations greater than 5° were selected from 100 conformers to represent the NMR ensembles.

[d] Calculated with PROCHECK-NMR (Laskowski et al., 1996).

[e] Calculated with MOLMOL(Koradi et al., 1996); averages are over backbone atoms.

So far, no Trx structure is reported from arthropods. A search for topologically similar domains in the Protein Data Bank (PDB) using the DALI program (Holm and Sander, 1998) reveals that the structural features of Cr-TRP16 resemble a typical Trx (Martin, 1995a)**.** The closest match is observed between the Trx-related protein *tryparedoxin II* (PDB code 1FG4), a 16 kDa Trx like protein from *Crithidia fasciculate*, with 26% sequence identity and an RMSD of 2.9 Å for 126 Cα atoms. The structure-based sequence alignment reveals that most of the structurally invariant residues are located at the N-terminal loop region between β2 and β3 (Fig. 2.5 A).

A



B

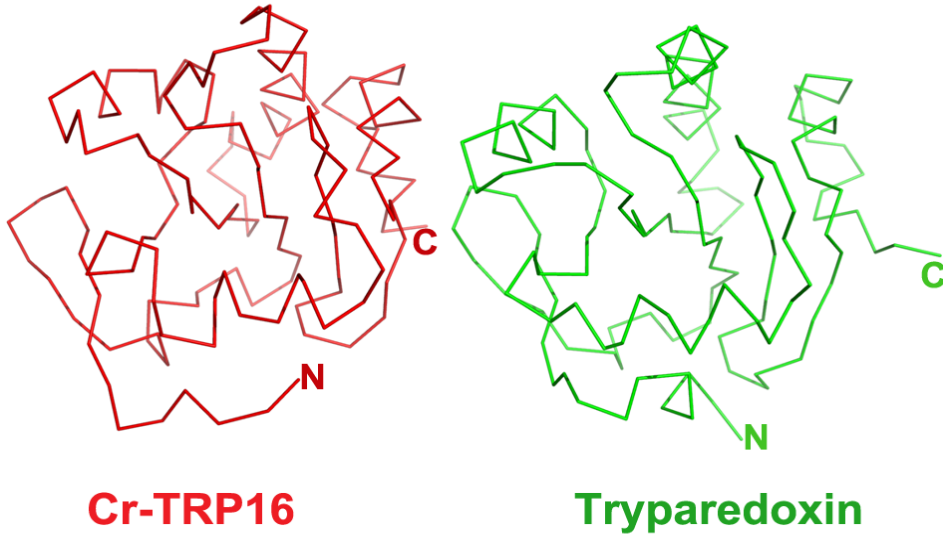


Figure II.5: Comparison of Cr-TRP16 with Trypaerdoxin. *(A)* Structure based sequence comparison between Cr-TRP16 and other Thioredoxins. Alignment was done by Clustal W. Identical and conserved residues are highlighted in red and outlined in blue, respectively. This figure was created by using the program ESPript (Gouet et al., 1999). (B*)* Superimposition of Cr-TRP16 (red) and Trypaerdoxin II (green) (PDB code 1FG4) from *Crithidia fasciculate*. The alignment was performed using the program COOT (Emsley and Cowtan, 2004). The RMSD between is 2.9 Å for 126 Cα atoms.

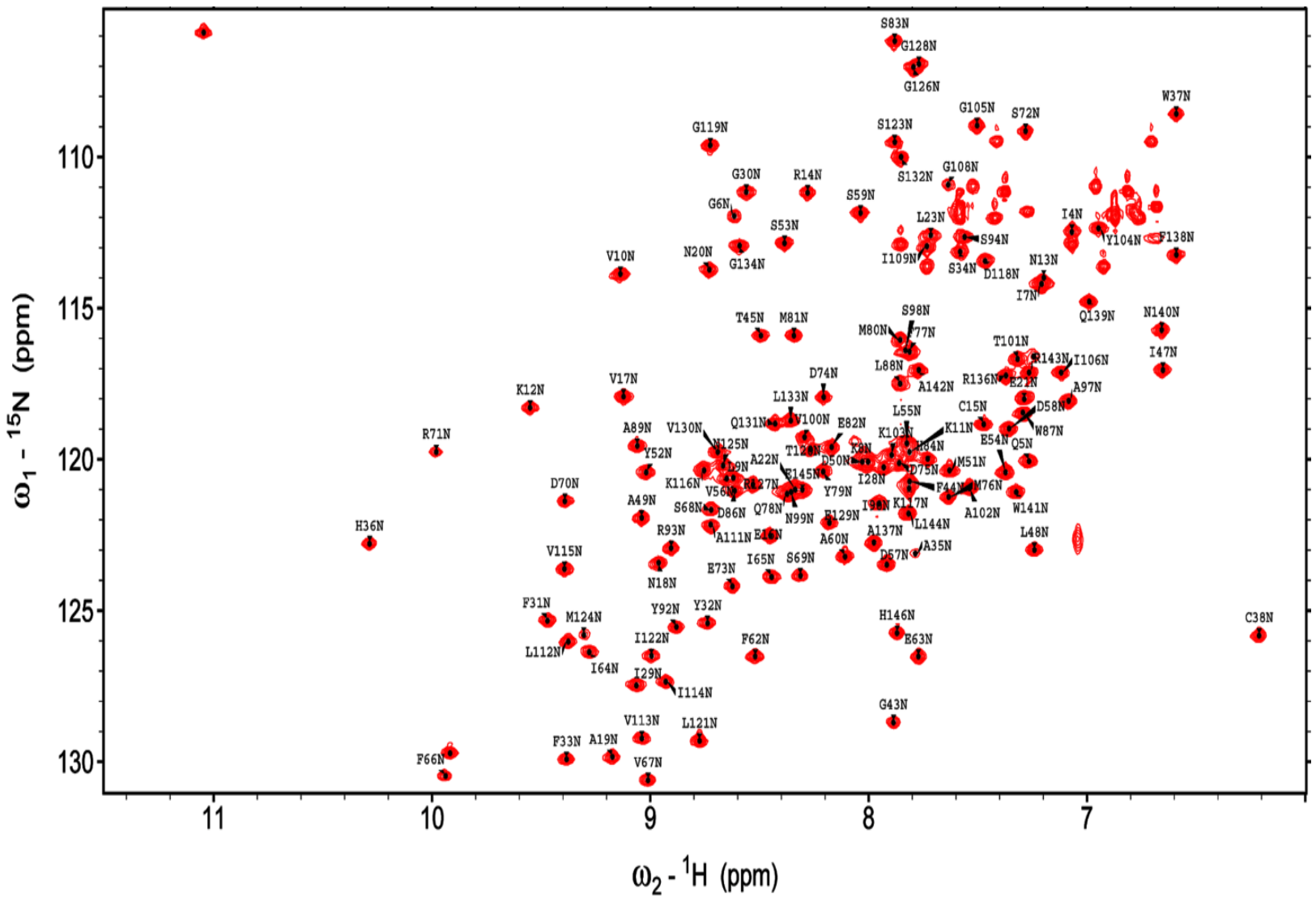
2.3.4 Dimerization of Cr-TRP16

Human 12 kDa Trx1 forms a homodimer through an intermolecular disulfide bond *via* Cys73 (Weichsel et al., 1996a). Interestingly, Cr-TRP16 can also form a stable dimer. Despite low sequence identity (18%) between Cr-TRP16 and hTrx1, Cr-TRP16 could also enhance TNFα-induced NF-κB transcriptional activation by increasing the NF-κB DNA-binding activity. In addition to the two conserved Cys residues (Cys38 and Cys41) at the active site, Cr-TRP16 has an additional Cys residue at position 15, which is involved in the formation of an intermolecular disulphide bond. The previous report (Wang et al., 2007), based on mass spectrometry, has shown that only the two conserved Cys residues at the active site are redox active, similar to human Trx1.

2.3.5 ^{1}H- ^{15}N-HSQC NMR spectroscopy

In order to characterize the mode of dimerization and to identify the critical residues involved in dimer formation, the ^{1}H-^{15}N-HSQC NMR spectra of reduced and oxidized wild type Cr-TRP16 were compared. Sequential assignments were completed by using the standard triple-resonance methods. The assigned spectra of reduced and the overlay ^{1}H-^{15}N-HSQCNMR spectra of reduced and oxidized Cr-TRP16 are shown in Figs. 2.6 A and 2.6 B, respectively. 2D NMR experiments show that the majority of the amides are unaffected by dimerization. Fig. 2.6 B shows a total of 26 amide signal chemical shifts have changed. Nearly all of these residues are located in the surface exposed region around the N-terminus Cys15 and the second β-sheet (β1β2). This suggests that this part of the molecule is involved in dimerization.

A


ω1 - 15N (ppm)
ω2 - 1H (ppm)
110
115
120
125
130
11
10
9
8
7
S83N
G128N
G126N
W37N
G105N
S72N
G119N
S123N
S132N
G108N
G30N
R14N
S59N
G6N
L23N
I4N
S53N
S94N
Y104N
F138N
V10N
N20N
G134N
I109N
S34N
D118N
N13N
I7N
T45N
M81N
S98N
M80N
F77N
Q139N
N140N
T101N
I47N
L88N
R143N
I106N
K12N
V17N
D74N
A142N
R136N
E21N
A97N
L133N
L55N
C15N
D58N
R71N
A89N
V130N
Q131N
V100N
E82N
K103N
K11N
W87N
E54N
Q5N
Y52N
T126N
K8N
H84N
A22N
D50N
M51N
D9N
E145N
D75N
I28N
D70N
K116N
R127N
F44N
M76N
V56N
Y79N
A49N
S68N
D86N
Q78N
N99N
E129N
I98N
K117N
A102N
W141N
H36N
R93N
A111N
E16N
L144N
A137N
A60N
A35N
L48N
V115N
I65N
S69N
D57N
E73N
N18N
F31N
Y92N
Y32N
H146N
C38N
M124N
I122N
F62N
E63N
L112N
I64N
I29N
I114N
G43N
V113N
L121N
F33N
A19N
F66N
V67N

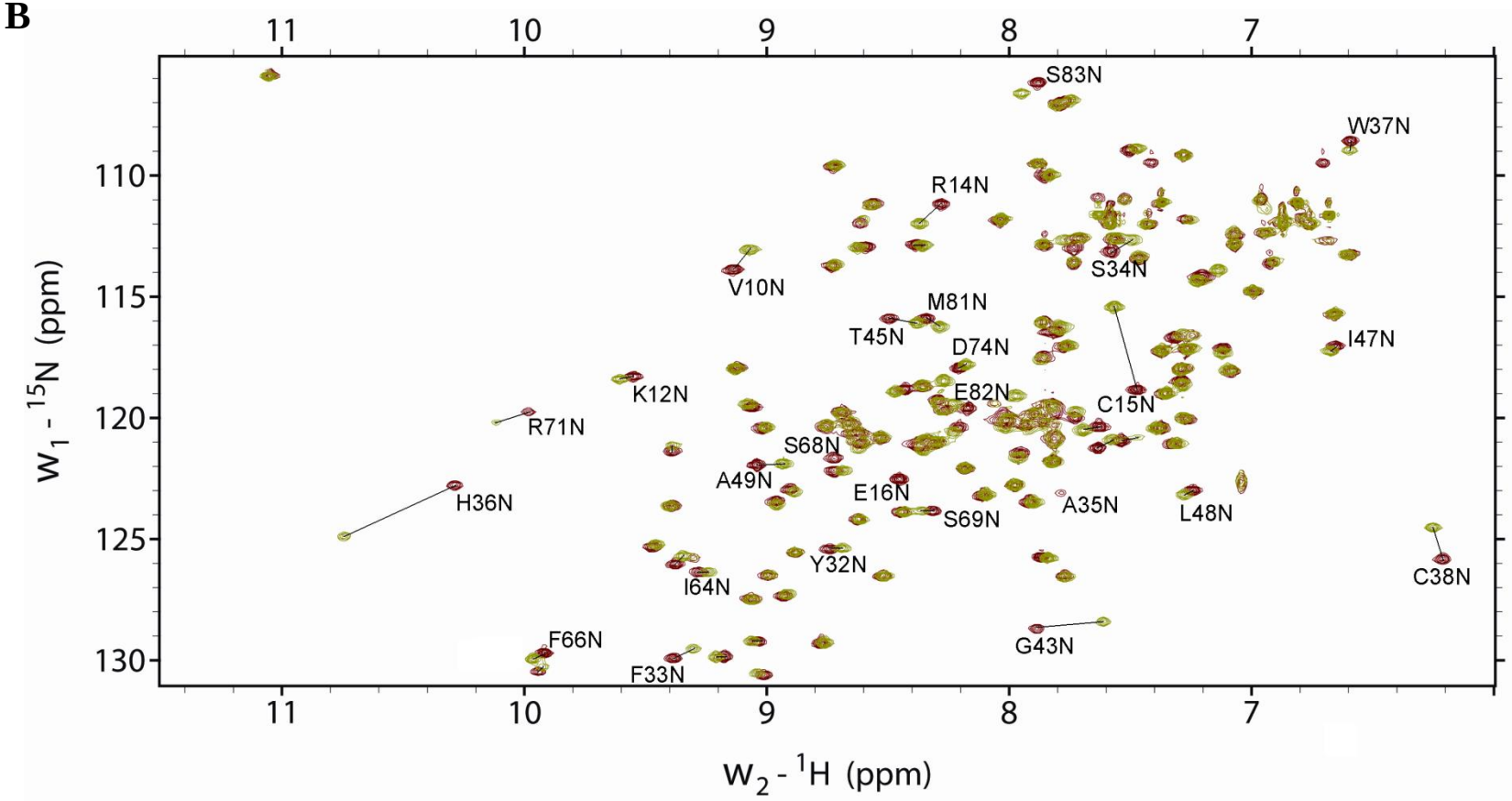
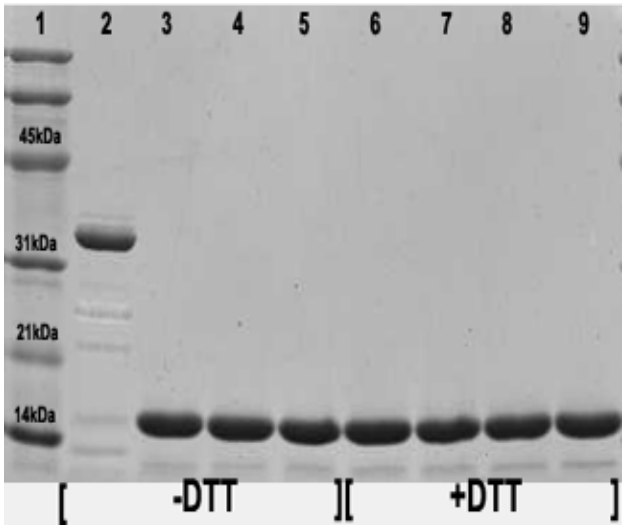


Figure II.6: Superposition of 1H-15N HSQC spectra of oxidized and reduced wild type Cr-TRP16 1H-15N HSQC spectra of CrTRP16 in its (A) reduced and (B) overlay of oxidized and reduced forms. Backbone amide resonances are labeled according to sequential assignments. Lines indicate shifts displayed by amino acid residues participating in the interactions leading to dimerization. The spectra were acquired at 303 K in 10 mM Tris-HCL buffer (pH 7.0).

2.3.6 Analytical ultracentrifugation (AUC)

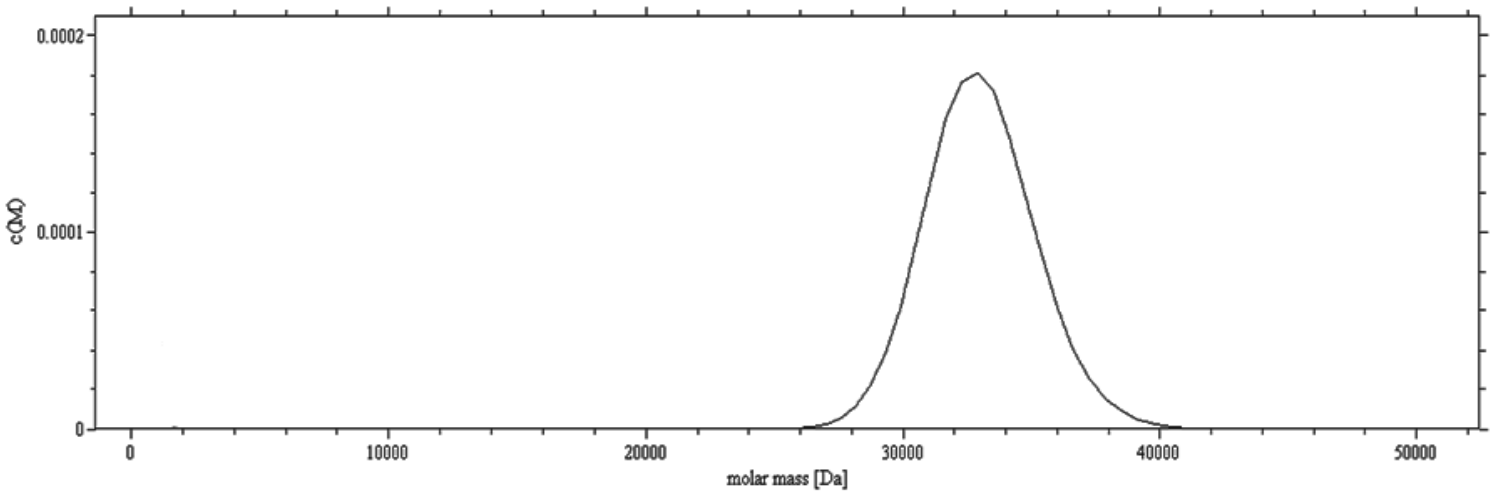
The Cys15Ser (C15S) mutant was generated to elucidate the role of the additional Cys at position 15 (Cys15) in dimer formation and its physiological significance in the NF-κB function. Dynamic light scattering (DLS) measurements on reduced and oxidized C15S Cr-TRP16 show an apparent molecular weight corresponding to only a monomer. Non-reducing and SDS PAGE for C15S Cr-TRP16 show a single band at an apparent molecular weight of 17 kDa (Fig. 2.7 B), whereas non-reducing PAGE for the wild type protein shows an apparent dimer (34 kDa) band. In the AUC analysis of the oxidized wild type Cr-TRP16 protein (0.8 OD at 280 nm), there is only one peak corresponding to an apparent molecular weight of a dimeric Cr-TRP16 (34 kDa, Fig. 2.7Ai) while the reduced Cr-TRP16 protein shows a peak at 17 kDa, corresponding to monomeric Cr-TRP16 (Fig. 2.7Aii). On the other hand, we observe only one peak with an apparent molecular weight of 17 kDa for both the oxidized and reduced (Fig. 2.7B) C15S mutant of Cr-TRP16. These results are in agreement with our hypothesis that Cys15 is responsible for dimerization.

A

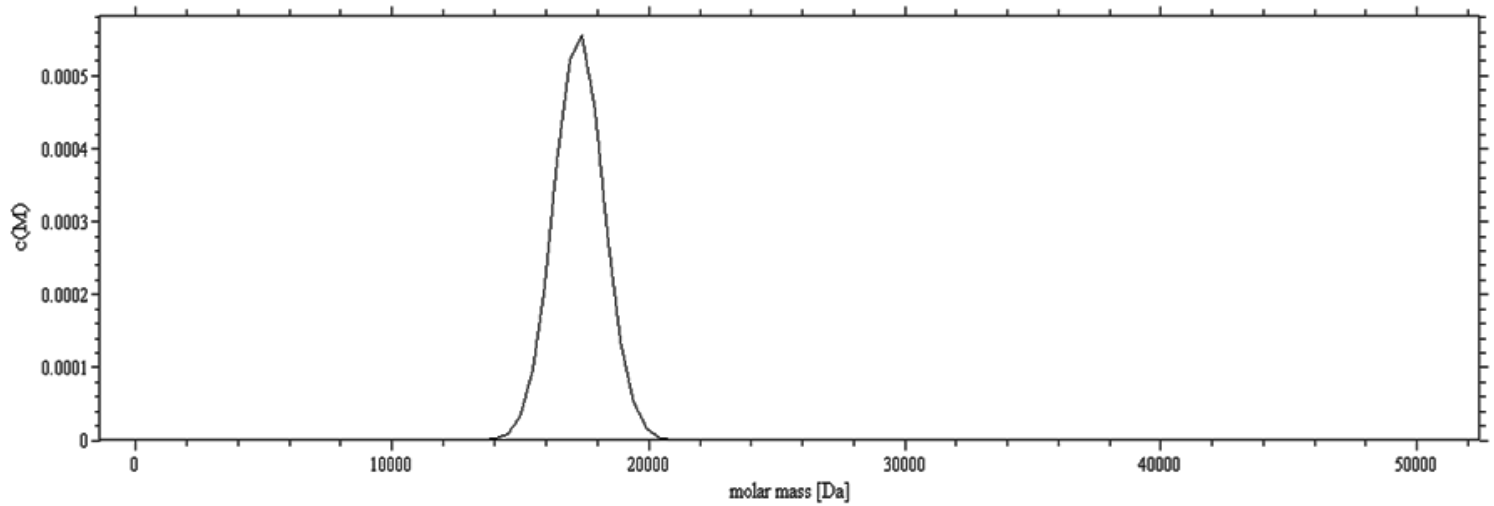


B

i


0.0002
0.0001
0
c(M)
0
10000
20000
30000
40000
50000
molar mass [Da]

ii


0.0005
0.0004
0.0003
0.0002
0.0001
0
c(M)
0
10000
20000
30000
40000
50000
molar mass [Da]

C

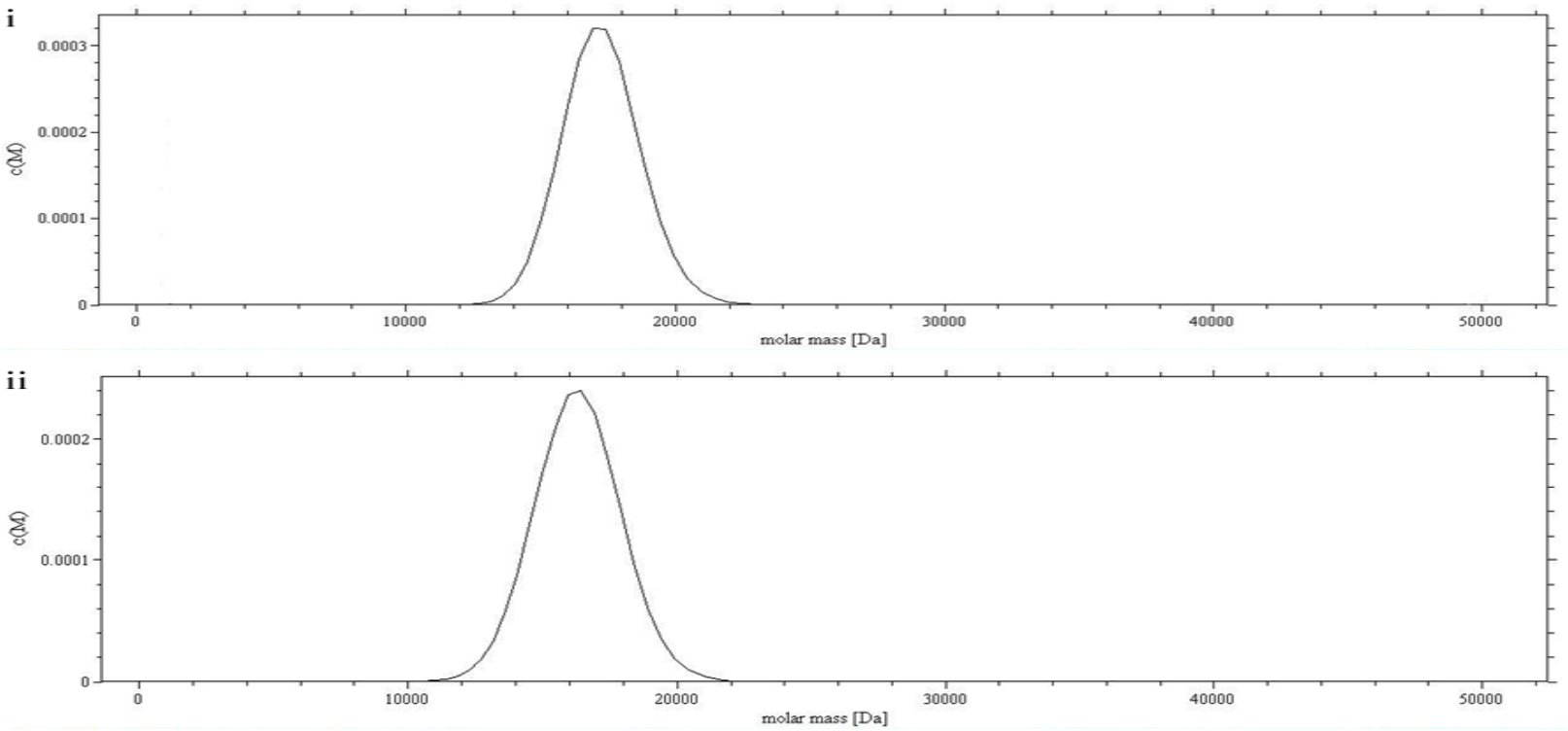


Figure II.7: Study of the dimerization of Cr-TRP16 by sedimentation velocity analysis.The experiments were conducted in TRIS buffer (Tris-HCl 10 mM pH 7.0, 100 mM NaCl, and 5% glycerol) at a rotor speed of 40,000 rpm, and a rotor temperature of 20 ºC. The experiments were carried out in the absence of DTT in the buffer. The sedimentation velocity profiles were collected by monitoring the absorbance at 280 nm. (A) SDS-PAGE analysis of Cr-TRP16 dimerization. . FPLC purified Cr-TRP16 was subjected to reversed-phase HPLC by using a C_8 column. The separated proteins were analyzed by SDS-PAGE and coomassie blue staining under nonreducing (–DTT) and reducing (+DTT) conditions, as indicated. Lanes 2 and 3 represent oxidized and reduced wild type Cr-TRP16, respectively, under a non-reducing condition. Similarly lanes 4 and 5 represent C15S mutant Cr-TRP16 under a non-reducing condition. Lanes 5 to 8 represent similar loading patterns under a reducing (+DTT) SDS-PAGE condition as indicated earlier. A molecular marker is shown at Lane 1 of the gel. This analysis shows that only wild type Cr-TRP16 can form a dimer. (B) Molecular weight profile of (i) oxidized and (ii) reduced wild type Cr-TRP16. (C) Molecular weight profile of oxidized C15S Cr-TRP16 mutant. The scans were analyzed using Sedfit program (Brown and Schuck, 2006).

2.3.7 Cr-TRP16 increases TNF-α induced nuclear translocation of p65 and p50

Biophysical and mutational studies have confirmed that Cr-TRP16 exists as a dimer and Cys15 is responsible for its dimerization. However, it is unclear whether Cr-TRP16 exists as a dimer under physiological conditions. In order to address this, we expressed wild type and C15S mutant Cr-TRP16 proteins, and stimulated the cells with TNF-α, one of the most potent inducers of NF-κB activation (Sethi and Tergaonkar, 2009). Western blot analysis with anti-p50 antibody revealed minimal change in the NF-κB p50 protein, suggesting that the overexpression of Cr-TRP16 does not change the expression of p50 (Fig. 5A). Similarly, it does not show any significant effect on the TNF-α induced IκBα degradation (Fig. 5A). Subsequently, we investigated the effect of Cr-TRP16 overexpression on nuclear translocation of NF- κB. Western blot analysis with nuclear extract using anti-p50, anti-p65 and anti-phosphorylated p65 antibodies revealed that the overexpression of wild type Cr-TRP16 induced the nuclear translocation of the p50-p65 heterodimer and increased the level of phosphorylated p65 while no appreciable increase was observed in C15S Cr-TRP16 transfected cells (Figs.2.5B and C).

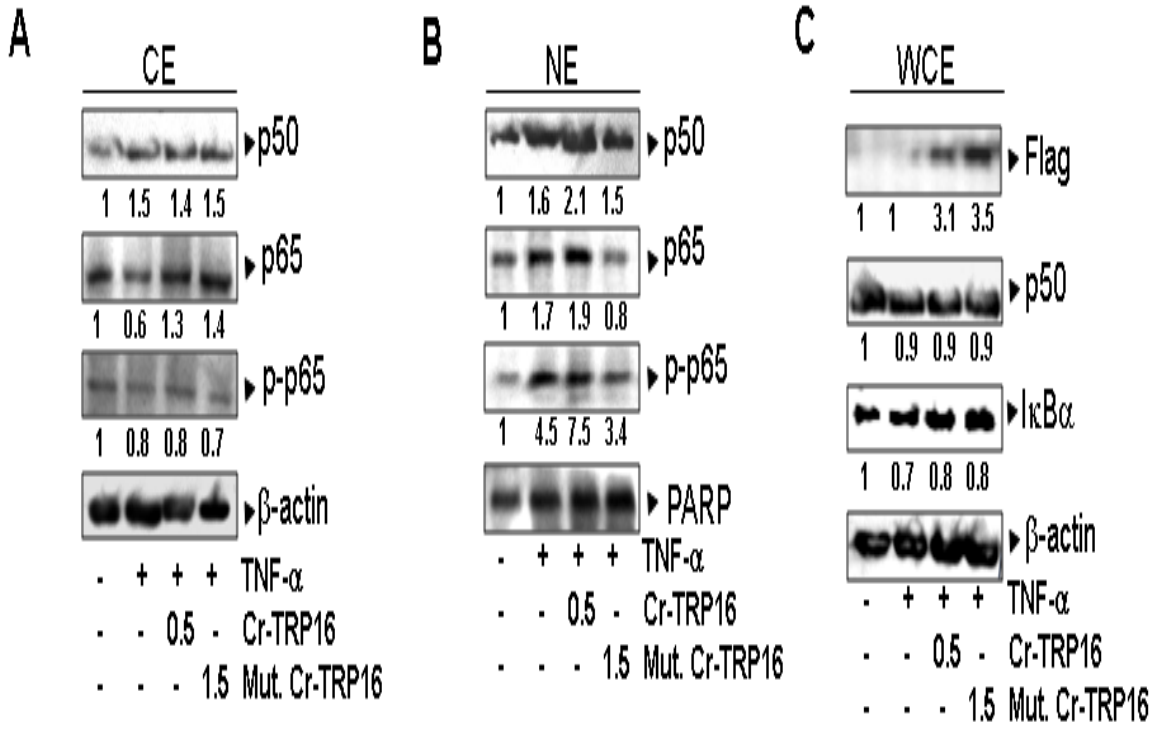


Figure II.8: Effect of Cr-TRP16 on the expression and subcellular localization of NF-κB. (A) Effects of Cr-TRP16 expression on NF-κB p50, p65 and IκBα expression. HeLa cells were transfected with Cr-TRP16 as described in materials and methods. Whole cell lysates were subjected to Western blot analysis. The FLAG-tagged Cr-TRP16 was detected with anti-FLAG antibody. (B) and (C) HeLa were transfected with the plasmids Cr-TR16 or C15S Cr-TRP16 for 30 h and then stimulated with TNF-α (1 nM) for 6 h. Nuclear and cytoplasmic protein extracts were prepared using nuclear extraction kit (Active motif, CA) according to the manufacturer's instructions. Protein concentration was determined with Bradford reagent. The nuclear and cytoplasmic extracts were subjected to Western blot analysis. The level of actin served as loading control.

2.3.8 Cr-TRP16 augments TNF-α induced NF-κB DNA binding activity

Next we investigated the effect of wild type and the C15S mutant Cr-TRP16 on NF-κB DNA binding activity. Nuclear extracts were prepared and tested for NF-κB activity using an ELISA based DNA binding assay. Fig. 2.9A represents the effect of Cr-TRP16 on constitutive NF-κB DNA binding activity. This result demonstrates that Cr-TRP16 does not activate NF-κB on its own but enhances TNF-α-induced NF-κB activation. As shown in Fig. 6B, wild type Cr-TRP16 increases TNF-α induced NF-κB DNA binding activity while there is no significant change observed in the C15S Cr-TRP16 transfected cells. These results clearly indicate that over expression of Cr-TRP16 significantly enhances TNF-α induced NF-κB activation in HeLa cells (Fig. 2.9).

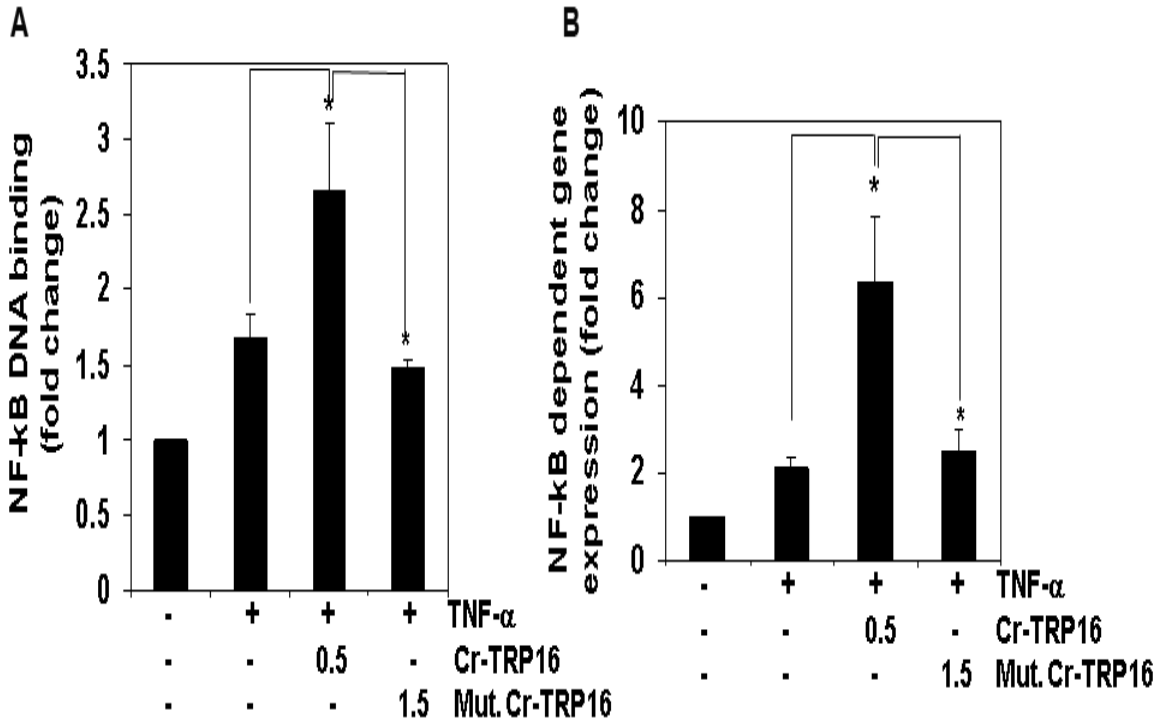


Figure II.9: TNFα induced NF-κB DNA-binding activity. **(A) TNF-α induced NF-κB DNA-binding activity.** HeLa cells were transfected with indicated amounts of wild type and mutant (C15S) Cr-TRP16 for 30 h and then stimulated with TNF-α (1 nM) for 6 h. Nuclear extracts from HeLa cells stimulated with TNF-α were assayed for NF-κB p65 activation. Nuclear protein extracts were prepared using a nuclear extraction kit (Active motif, CA) according to the manufacturer's instructions and assayed for NF-κB p65 activation using the TransAM NFκB p65 kit. **(B) Effect of over expression of Cr-TRP16 on the NF-κB activity.** HeLa cells were transfected with indicated amounts of wild type and C15S Cr-TRP16 along with NF-κB regulate luciferase reporter construct and co-transfected with β-galactosidase vector. Luciferase activity was measured using luciferase assay kit (Promega). Relative light units were measured by Tecan (Durham, NC) plate reader with function to measure luminescence and the data was normalized to β-galactosidase activity. All luciferase experiments were done in triplicate and repeated three or more times.

2.3.9 Cr-TRP16 augments TNF-α induced NF-κB-dependent reporter gene expression

Although we observe that Cr-TRP16 increases NF-κB activation in the NF-κB DNA binding assay, DNA binding alone does not always correlate with NF-κB-dependent gene transcription, indicating there could be some additional regulatory steps (Ahn et al., 2007). To determine the effects of Cr-TRP16 on TNF-α induced NF-κB dependent reporter gene expression, we transiently transfected HeLa cells with Cr-TRP16 or C15S Cr-TRP16 along with NF-κB to regulate a luciferase reporter construct and co-

transfected with a β-galactosidase vector. A 2-fold increase in luciferase activity was observed after stimulation with TNF-α alone (Fig. 2.9 C) while the wild type Cr-TRP16 transfected HeLa cells show a 6-fold increase in TNF-α mediated NF-κB dependent luciferase expression. Interestingly, no significant effect was observed in the cells transfected with C15S Cr-TRP16. These results show that Cr-TRP16 can potentiate TNF-α inducible NF-κB-dependent reporter gene expression.

Overall, our results clearly demonstrate that the Cys15 residue is crucial for the dimerization of Cr-TRP16 and only the dimeric form plays a pivotal role under physiological conditions in the regulation of NF-κB activation.

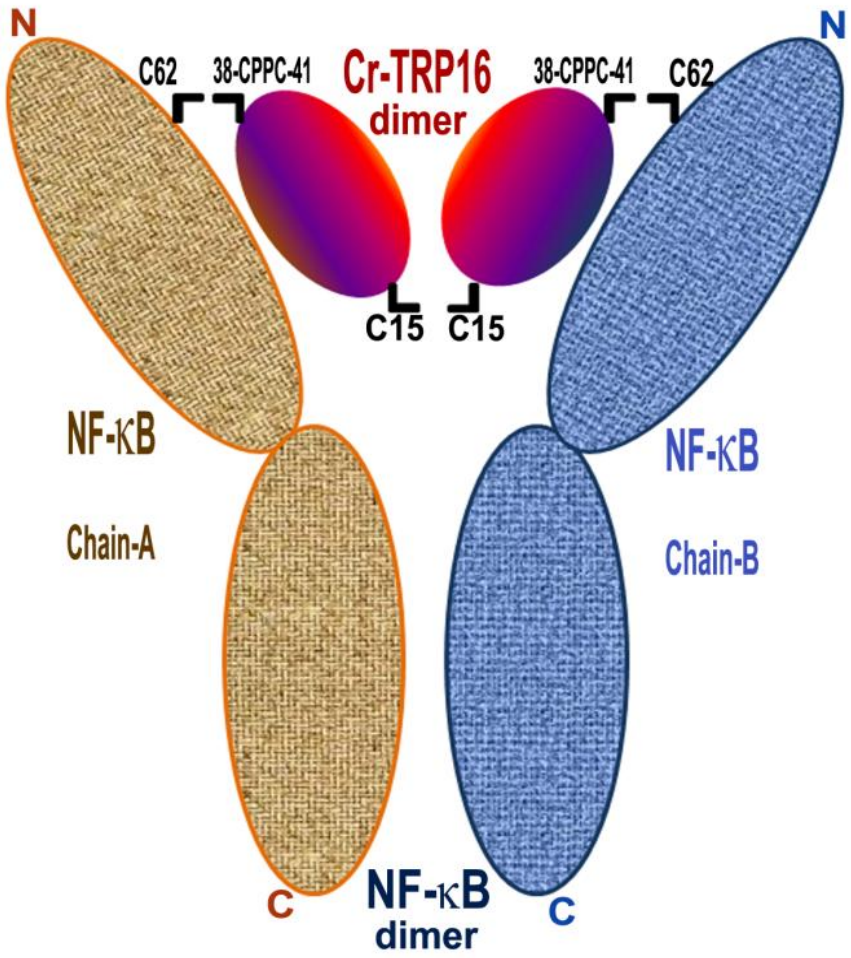


Figure II.10: Model for the interaction of Cr-TRP16 dimer with NF-κB dimer.Based on the previously published human NF-κB p50 homodimer:DNA complex and human Trx: NF-κB p50 peptide complex structures and our structural, biophysical and biochemical results, we propose that Cr-TRP16 forms a homodimer by making a disulfide link through Cys15 and is likely to occupy the location where DNA binds to NF-κB. The active site motif WCPPC of Cr-TRP16 (through Cys38 and Cys41) is likely to interacts with the redox active Cys 62 of NF-kB dimer. For the ease of residue numbering, a p50 homodimer is assumed for NF-κB and the C-terminal cysteines of NF-κB which promote dimerization of NF-κB are not shown for simplicity.

2.4 DISCUSSION

Sequence alignment reveals that the catalytic WCPPC motif is evolutionarily conserved and indicates the potential importance of this motif even in 16 kDa Trx-like proteins. Although the active site of the 16 kDa Trx-like proteins in mammals has undergone marked changes, several other regions remain well conserved, when compared to 12 kDa Trxs. This observation suggests that the members of the 16 kDa Trx-like protein family have evolutionarily diverged from the 12 kDa Trxs at an early stage from *C.elegans* to human (Wang et al., 2007). Several multi-domain Trx-like proteins have evolved, retaining the Trx fold (Atkinson and Babbitt, 2009). Although the Trx superfamily has received much attention, only a few structural and functional studies have been undertaken to analyze the multi-domain Trx proteins. Cr-TRP16 has a fold similar to that of tryparedoxin, a 16 kDa Trx-like protein from the parasitic protozoa *Crithidia fasciculata* (*Trypanosoma*) (Montemartini et al., 2000). In particular, the active site motif, CPPC, is homologous to that of the corresponding motif in the protozoan Trx-like protein (Hofmann et al., 2001). However, functionally the Cr-TRP16 is quite different from the tryparedoxin. In contrast, despite notable differences in the amino acid sequence of the catalytic site, Cr-TRP16 appears to be functionally similar to human Trx (Qin et al., 1995) and regulates NF-κB activity (Wang et al., 2007).

While most of the 16 kDa Trx-like proteins do not have an extra cysteine, Cr-TRP16 possesses an additional Cys residue at the N-terminus. This prompted us to investigate the functional significance of this Cys and hypothesize that Cr-TRP16 could form a homo dimer and regulate NF-κB, like hTrx1. Mutation, AUC, non-reducing SDS and 2D-NMR experiments confirm that only wild type Cr-TRP16 can form a dimer under

an oxidized condition and not the C15S mutant. However, it was unclear, whether the dimer or monomer is the active form *in vivo*. In order to clarify the physiological role of the N-terminal Cys15 residue of Cr-TRP16, we overexpressed the wild type and C15S mutant proteins in HeLa cells, stimulated the cells with TNF-α and measured NF-κB activation. We found that wild type Cr-TRP16 could increase the TNF-α induced NF-κB DNA binding activity but not the C15S mutant-transfected cells (Fig. 2.9A&B).

Moreover we observed a similar pattern of results in TNF-α induced NF-κB dependent luciferase expression in the cells that were transfected with the wild type and C15S mutant proteins (Fig. 2.9C). Even though the C15S mutant contains intact redox active Cys38 and Cys41 residues, it could not regulate the NF-κB activity, suggesting that Cr-TRP16 exists as a dimer under the physiological condition and Cys15 is the crucial residue for the dimer formation. These results are consistent with our biophysical studies and present evidence that the dimer of Cr-TRP16 is the naturally occurring form.

Studies by Matthews *et al.* (1992) and Meyer *et al*. (1993) show that the DNA-binding activity of NF-κB follows redox regulation through its modulatory Cys residues. Furthermore, the NF-κB components, p50 and p65, contain well-conserved cysteine residues in their DNA-binding loops (Ghosh et al., 1995; Muller et al., 1995). Activation of the DNA-binding properties of human NF-κB requires that the cysteine at position 62 of the p50 subunit be reduced. In the p50 homodimer:DNA complex structure (Ghosh et al., 1995; Muller et al., 1995), the p50 homodimer binds to DNA in a channel, whose one end is closed by the C-terminal dimerization domain and the front is closed by two loops that contain Cys62, one from each subunit. If a disulphide is formed between the two Cys62 residues in uncomplexed NF-κB, then DNA cannot bind to the p50 homodimer.

Similarly, if a disulphide bridge is formed in the p50 homodimer:DNA complex, then DNA cannot be released (Qin et al., 1995). This implies that the activation of NF-κB DNA binding activity begins with IκB degradation, reduction of NF-κB by Trx followed by translocation of NF-κB to the nucleus (Flohé et al., 1997; Powis and Montfort, 2001a) for effective DNA binding and transactivation (Flohé et al., 1997). Based on this argument and our structure-function analyses, we propose a model on how Cr-TRP16 could bind to the p50 homodimer or p50/p65 heterodimer *in vivo* (Fig. 2.10). Thioredoxin and glutathione reductase are responsible for maintaining a reduced environment in the cell. In a glutathione reductase free system, Trx and Trx-like proteins are solely responsible for maintaining a reduced environment. Several diseases, including cancer, autoimmune diseases, and infectious diseases, are mediated through the effects of Trx and TrxR and unbalanced NF-kB activation.

The structure based functional studies clearly demonstrated that (1) Cr-TRP16 regulate NF-kB similar to 12 kDa human Trx, (2) C15 is the key residue for dimerization, and (3) only the dimeric Cr-TRP16 regulates Nf-kB activity. This study provides insights into the interactions of Trxs and Trx-like proteins with other metabolic pathways and their physiological relevance and will eventually lead to promising therapeutic development to modulate the action of Trx and NF-κB. In the next chapter, we are going to discuss the expression and purification of human thioredoxin like protein-6 (TXNL-6), a homolog of Cr-TRP16 and protect retinal cells from apoptosis under stress and characterization of its interaction with NF-kB.

CHAPTER-III: CHARACTERIZATION OF HUMAN THIOREDOXIN LIKE PROTEIN-6 (TXNL-6)

3.1 INTRODUCTION

Photoreceptor cell death by apoptosis is responsible for many eye diseases including age-related macular degeneration which is estimated to affect more than 8 million individuals in the United States alone (Donovan and Cotter, 2002). It has largely been accepted that reactive oxidative species (ROS) and a series of caspase and calpain enzymes plays a key role in both the initiation and execution of apoptosis (Gómez-Vicente et al., 2005). Therefore, the regulation of the cellular redox state and apoptotic process may serve as potential therapeutic strategies applicable to retinitis pigmentosa and other human retinal diseases (Sanvicens et al., 2004).

It has been shown that nerve growth factor-dependent activation of NF-κB contributes to the survival of sympathetic neurons (Maggirwar et al., 1998) and the activation of NF-κB is necessary for the survival of S-type neuroblastoma (Bian et al., 2002). Rod-derived cone viability factor, (RdCVF) a truncated mouse thioredoxin is specifically expressed by rod photoreceptor cells and prevents the apoptosis of cone cells. However the protective mechanism of RdCVF and the implications of its human homologue, thioredoxin-like 6 (TXNL6), on the apoptosis of retinal cells remain unknown. Recently, Leveillard et al. (Léveillard et al., 2004) showed that RdCVF, could slow down cone degeneration in both the chick and mouse models. However, the protective mechanism of RdCVF against apoptosis of retinal cells is still unknown.

In vertebrates, there are two families of Trxs with high homology to the 16 kDa Cr-TRP16; one is 16 kDa and other is 24 kDa (TRX6). However invertebrates only

possess the 16 kDa Trxs and bacteria are devoid of this homolog. This observation suggests that the 16 kDa TRXs have evolutionarily diverged from the 12 kDa Trxs at an early stage and the 16 kDa Trxs probably underwent gene duplication and divergence in vertebrates and gave rise to the 24 kDa TXNL-6 (Zhang, 2003). Interestingly, unlike the invariable WCPPC catalytic motif in the 16 kDa TRXs of invertebrates, the active site of the 24 kDa TXNL-6 of vertebrates has undergone marked changes (Fig. 1). For human TXNL-6, even the most conserved Cys residues have been replaced with Ser residues. It is therefore reasonable to postulate that the duplication and further evolution of the gene encoding in vertebrates have probably relaxed the selection pressures and accelerated the evolution of novel catalytic motifs and functions (Force *et al*, 1999; Zhang, 2003). Therefore, it is interesting to investigate if the 24 kDa TXNL-6 in mammals still conserve and exhibit the basic enzymatic functions of the invertebrate 16 kDa counterparts. Simultaneously, it would be pertinent to explore the evolutionary, structure and functional similarities /differences between the Cr-TRP16 and human TXNL-6.

It has been postulated that TXNL-6 reduces the Cys residue in the DNA binding motif of NF-κb to enhance its binding to the κB promoter site (Wang et al., 2008). However still a few questions remain unanswered. Why does TXNL-6 not appear to affect the accumulation of ROS in contrast to the universally expressed human Trx1, which prevents photoreceptor cell damage by scavenging singlet oxygen and hydroxyl radicals? How does TXNL-6 interact with NF-κB? In order to address these questions we sought to determine the crystal structure of TXNL-6 and its complex with NF-κBp50 and subsequent functional studies to explore and understand the mechanism of interaction

between them to protect the photoreceptor cells from apoptosis. Here, we report the expression, purification and characterization of TXNL-6 and NF-kB.

3.2 RESULTS AND DISCUSSION

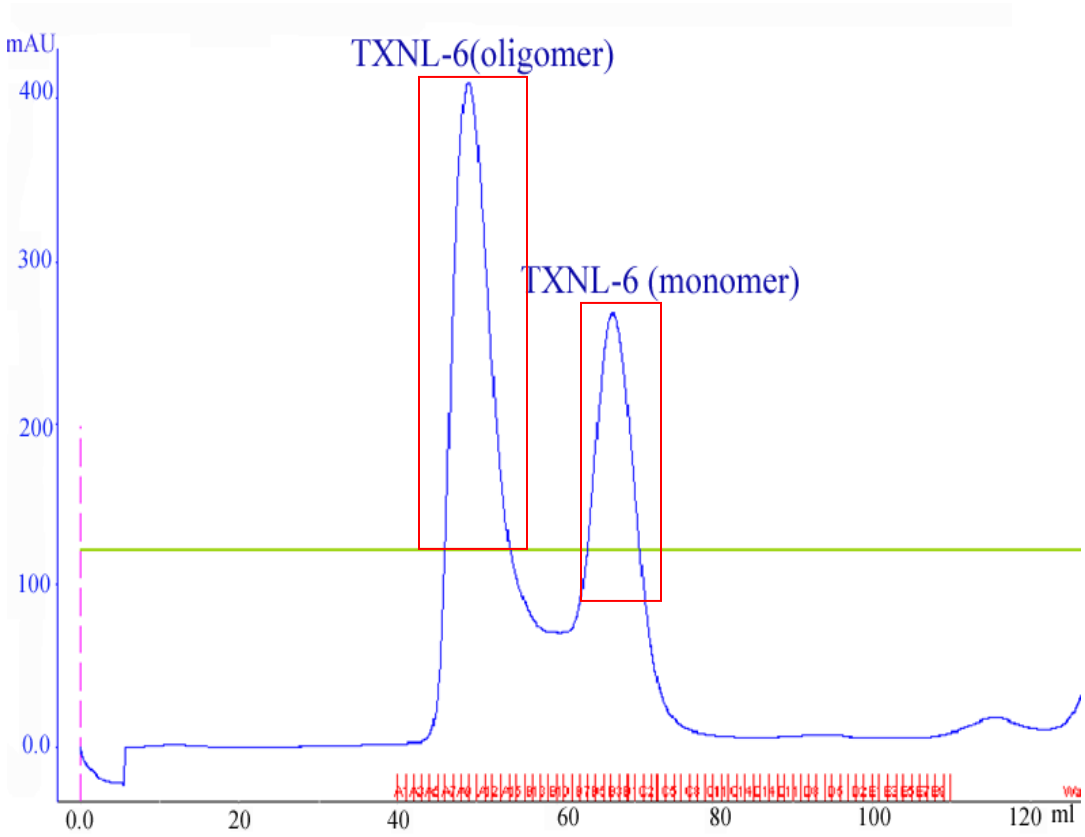
3.2.1 Cloning

The full length TXNL-6 gene (encoding residues Met1-Phe212) was PCR amplified using a forward (CAG GGA TCC ATG GCC TCC CTG TTC TCT) and a reverse (CCT CTC GAG TCA GAA CAG CCC CCC GG) primer with a *BamH*-I restriction site at the 5′ end and an *Xho*I site at the 3′ end (underlined), respectively. The insert was ligated into the previously linearized pET-22b (+) expression vector (Invitrogen).

3.2.2 Protein expression and purification

The pET22b: TXNL-6 construct was transformed into the bacterial strain *E. coli* BL21 (DE3) for expression. Optimal expression of the TXNL-6 protein was achieved by induction with 0.3 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) of 1 liter culture at 18 ºC. Cells were disrupted by a French press and the supernatant was collected after centrifuging at 10,000 g for 1 h at 4 ºC. His-tagged TXNL-6 proteins were purified in two steps using Ni-NTA (Qiagen) affinity chromatography followed by a Superdex 75 gel filtration column on Akta Express (GE Healthcare). The buffer was exchanged to a solution containing 10 mM Tris (pH 7.4), 150 mM NaCl, 5 mM dithiothreitol, 5% glycerol, 1 mM EDTA and finally the protein was concentrated to 18 mg/ml.

A



B

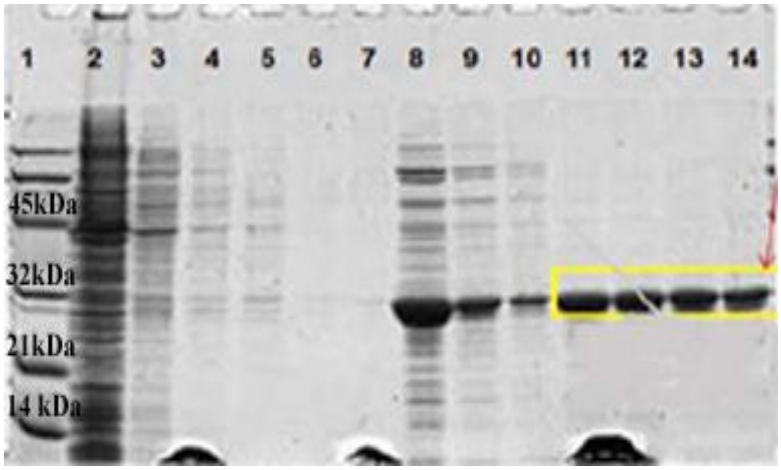


Figure III.1: Gel filtration profile of purified TXNL-6. A.) after affinity purification the TXNL-6 was passed through Superdex-75 column and eluted as two peaks. First peak corresponds to void volume and the second

peak is the monomeric TXNL-6 B.) Samples from different stages of purification were collected, separated by 12% SDS-PAGE and visualized by Coomassie staining. Eluted TXNL-6 migrated in the gel with an apparent molecular weight of 24 kDa. Lanes: 1. Low molecular weight ladder; 2. Flow through after binding Ni-NTA beads; 3-7. Column washes; 8-10. First peak of gel filtration elution; 11-14. Second peak of gel filtration elution.

The quality of protein was verified using Native-PAGE and Dynamic light scattering (Dyan pro solutions) instrument. Native-PAGE and DLS results showed that TXNL-6 proteins were non-homogeneous.

3.2.2 Mutagenesis of surface exposed charged residues of TXNL-6

To enhance crystallization probability, we have carried out the mutagenesis of the surface exposed charged residues of cluster 1,2 and 3, as predicted by the SERp server[http://www.doe-mbi.ucla.edu/Services/SER](Goldschmidt et al., 2007). Due to the presence surface exposed charged residues, the protein could form non-specific intermolecular interaction which can cause aggregation, and prevent crystallization. Lysine → Alanine mutation has been used not only to overcome the protein aggregation but also to improve crystallization kinetics. We obtained monomeric TXNL-6 in compression to the oligomer after mutagenesis. Protein quality was improved drastically after mutagenesis (Fig.3.2)

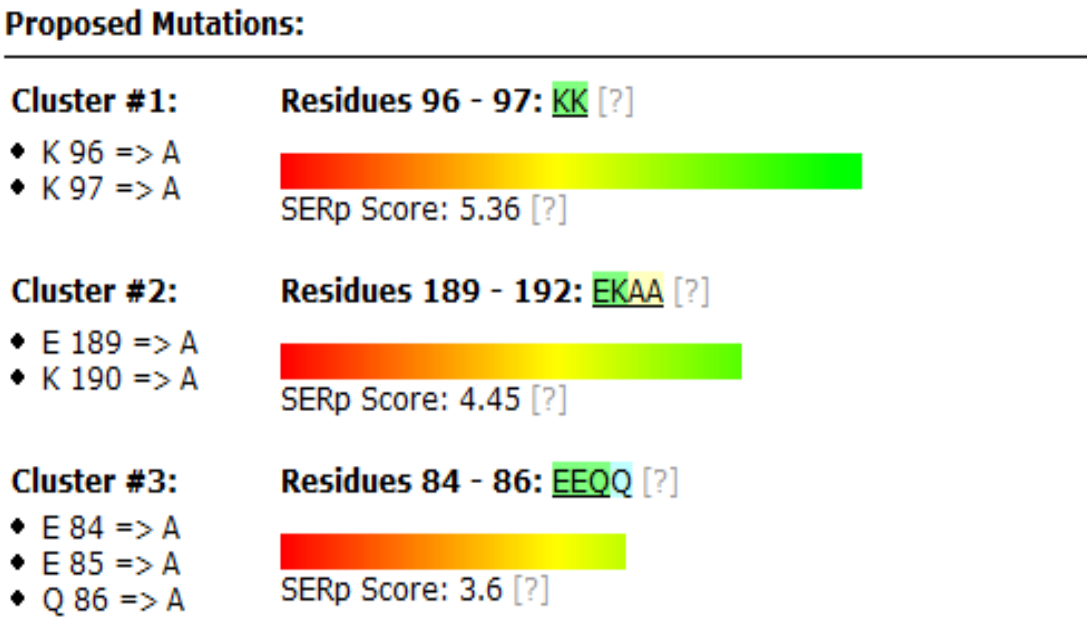


Figure III.2: SERp predicted surface exposed charged residues clusters.This summary shows 3 cluster candidate(s) that contain residues suitable for mutation designed to confer enhanced crystallizability. The cluster candidates are presented in the order of predicted success (best candidates with higher scores are shown first). Ideal mutation candidates are non-conserved, high entropy residues that lie in surface-exposed, entropy-rich regions of the protein. Good candidates for mutation typically have clusters scores of 3.0 and above.

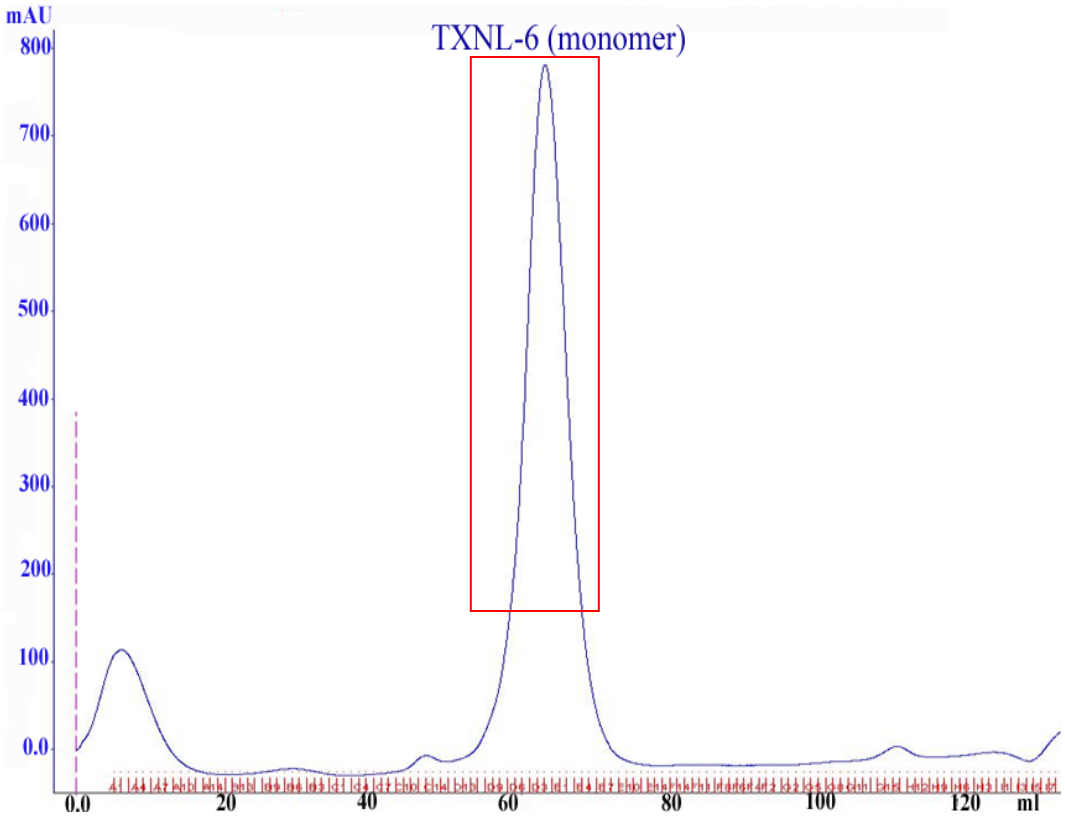


Figure III.3: FPLC profile of TXNL-6 after surface exposed mutagenesis.After affinity purification the TXNL-6 was passed through Superdex-75 column and eluted as single peak, which corresponds to a molecular weight of monomeric TXNL-6.

3.2.3 In vitro interaction between TXNL-6 and NF-kB

It has been shown that TXNL-6 forms an intermediate with various target molecules through disulfide linkage(Wang et al., 2008). To explore further the association between TXNL-6 and NF-kB p50, and to trap the physical association between both proteins, we used cross linking reagents such as diamide, which convert the free sulfhydryls to disulfides by cysteine oxidation.

Mst#	Time(s)	Temp(C)	Count Rate	Ampl	Diff Coeff	Radius(nm)	Polyd(nm)	Polydindx	MW(KDa)	%Mass	Baseline	Sos Error
1	10.0	20.0	609982	0.718	844.	2.53	0.235	0.01	28.5	100.0	1.000	2.23
2	20.0	20.0	617581	0.734	813.	2.63	0.403	0.02	31.9	100.0	1.000	4.08
3	30.0	20.0	607844	0.711	822.	2.60	0.854	0.11	30.8	100.0	1.000	2.70
4	40.0	20.0	616197	0.696	814.	2.63	0.719	0.07	31.8	100.0	1.000	2.43
5	50.0	20.0	611813	0.684	807.	2.65	0.738	0.08	32.6	100.0	1.000	2.81
6	60.0	20.0	608375	0.684	812.	2.64	0.303	0.01	32.1	100.0	1.000	1.92
7	70.0	20.0	610266	0.679	807.	2.65	0.319	0.01	32.7	100.0	1.000	2.43
8	80.0	20.0	596414	0.715	805.	2.66	0.572	0.05	32.8	100.0	1.000	1.38
9	90.0	20.0	610789	0.705	786.	2.72	0.773	0.08	35.3	100.0	1.000	3.40
10	100.0	20.0	601436	0.725	803.	2.66	0.667	0.06	33.1	100.0	1.000	2.78
11	110.0	20.0	598256	0.738	798.	2.68	0.306	0.01	33.7	100.0	1.000	2.19
12	120.0	20.0	609825	0.728	798.	2.68	0.332	0.02	33.7	100.0	1.000	2.34
13	130.0	20.0	603402	0.730	801.	2.67	0.201	0.01	33.3	100.0	1.000	2.46
14	140.0	20.0	594229	0.734	800.	2.68	0.381	0.02	33.5	100.0	1.000	1.43
15	150.0	20.0	604238	0.728	801.	2.67	0.474	0.03	33.3	100.0	1.000	2.47
16	160.0	20.0	615900	0.700	796.	2.69	0.346	0.02	34.0	100.0	1.001	2.73
17	170.0	20.0	609138	0.688	794.	2.70	0.540	0.04	34.3	100.0	1.000	3.73
Aves:												
Mono		20.0	607393	0.712	806.	2.66	0.480	0.04	32.8	100.0	1.000	2.56
Bi-1		0.0	0	0.000	0.000	0.000	----	----	0.000	0.0	0.000	0.000
Bi-2				0.000	0.000	0.000			0.000	0.0		

Figure III.4: Dynamic light scattering (DLS) profile of mutated TXNL-6.DLS data of TXNL-6 measured for 10 mg/ml. The hydrodynamic radius and poly dispersity index of the TXNL-6 protein shows monodisperse homogeneous sample.

3.2.3.1 Cloning of NF-kBp50 protein

The NF-kB gene (encoding residues 1-435, 1-366, 1-244 and 43-244) were PCR amplified using a forward and a reverse primer with a *BamH*-I restriction site at the 5′ end and an *Xho*I site at the 3′ end , respectively. The insert was ligated into the previously linearized expression vector pGEX-6P-1 (GE healthcare).

3.2.3.2 NF-kBp50 protein expression and purification

The expression vector pGEX-6P-1: NF-kB (43-244) was transformed into *E. coli* BL21 and induced with 0.2 mM IPTG at 18 ˚C overnight. The bacteria pellets were harvested, sonicated and lysed in 50mM Tris-HCl pH 8.0, 100 mM NaCl, 2 mM

dithiothreitol containing a protease inhibitors cocktail (Sigma). After binding to GST sepharose 4B beads and washing away all unbound proteins, the NF-kB protein was cleaved with thrombin and eluted from the column. The eluted protein was further purified on a FPLC Hiload 16/60 Superdex75 gel filtration column using an AKTA FPLC UPC-900 system (GE Healthcare) (Fig.3.5). Dynamic light-scattering experiments were performed on the purified NF-kB samples using a DynaPro (Protein Solutions) dynamic light-scattering instrument.

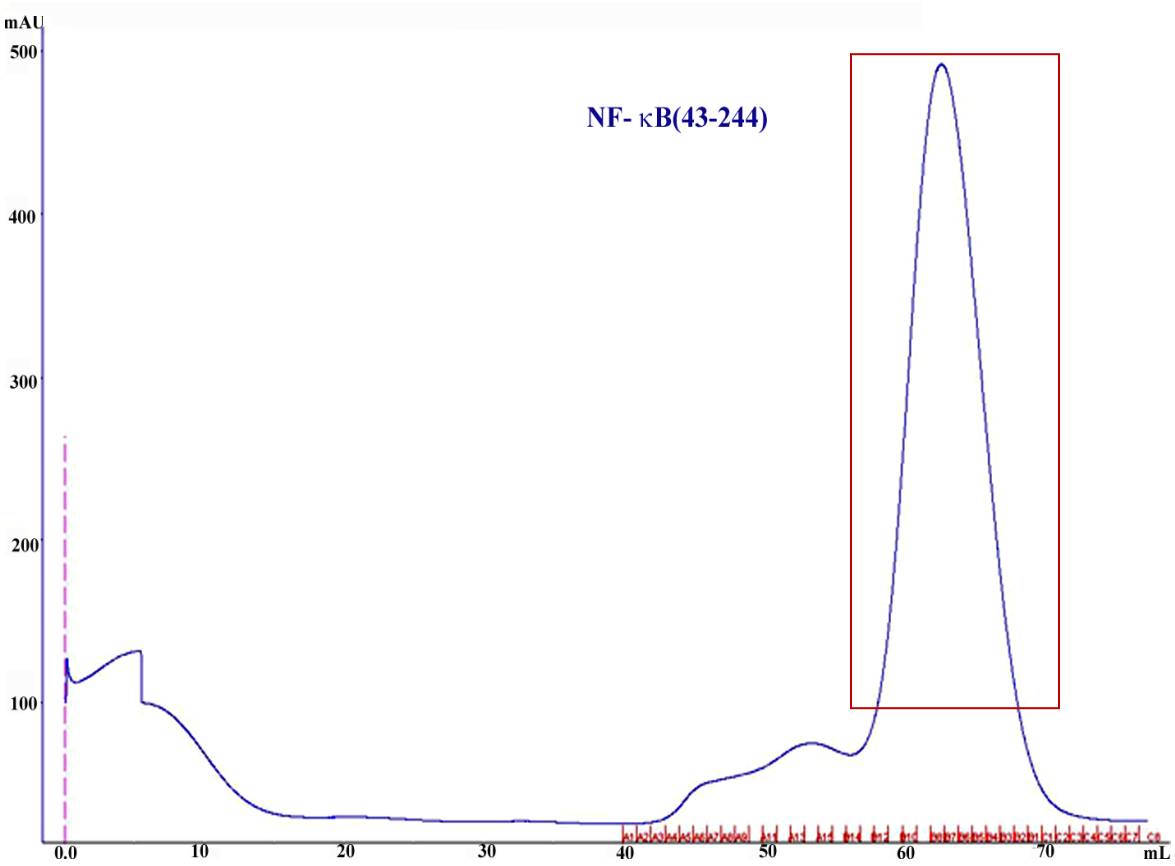


Figure III.5: The FPLC profile of NF-κB (43-244).After affinity purification NF-κB (43-244) was passed through Superdex-75 column and eluted as a single peak, which corresponds to the molecular weight of monomeric NF-κB (43-244).

3.2.3.3 Complex formation of NF-kB p50(43-244) with TXNL-6

To form the NF-kB p50 (43-244) / TXNL-6 complex, the purified 6XHis tagged TXNL-6 protein was allowed to interact with GST: NF-kB in the presence of diamide, which would cross link the two proteins by the conversion of free sulfhydryls to disulfides by cysteine oxidation(Wang et al., 2008). After extensive washing, the proteins were eluted from the resin by using 30 mM reduced glutathione. Subsequently, we passed the presumable complex through Ni-NTA resin to retain only the TXNL-6 / NF-kB hetero-dimer. After extensive washing, the cross-linked heterodimer was eluted by 300 mM imidazole and further purified using a FPLC Hiload 16/60 Superdex200 gel filtration column using an AKTA FPLC UPC-900 system (GE Healthcare). Non-reducing SDS-PAGE analysis shows only the hetrodimer of TXNL-6 and NF-kBp50 (43-244), (Fig. 3.6) and the result was further confirmed by peptide mass finger printing (Fig. 3.7).

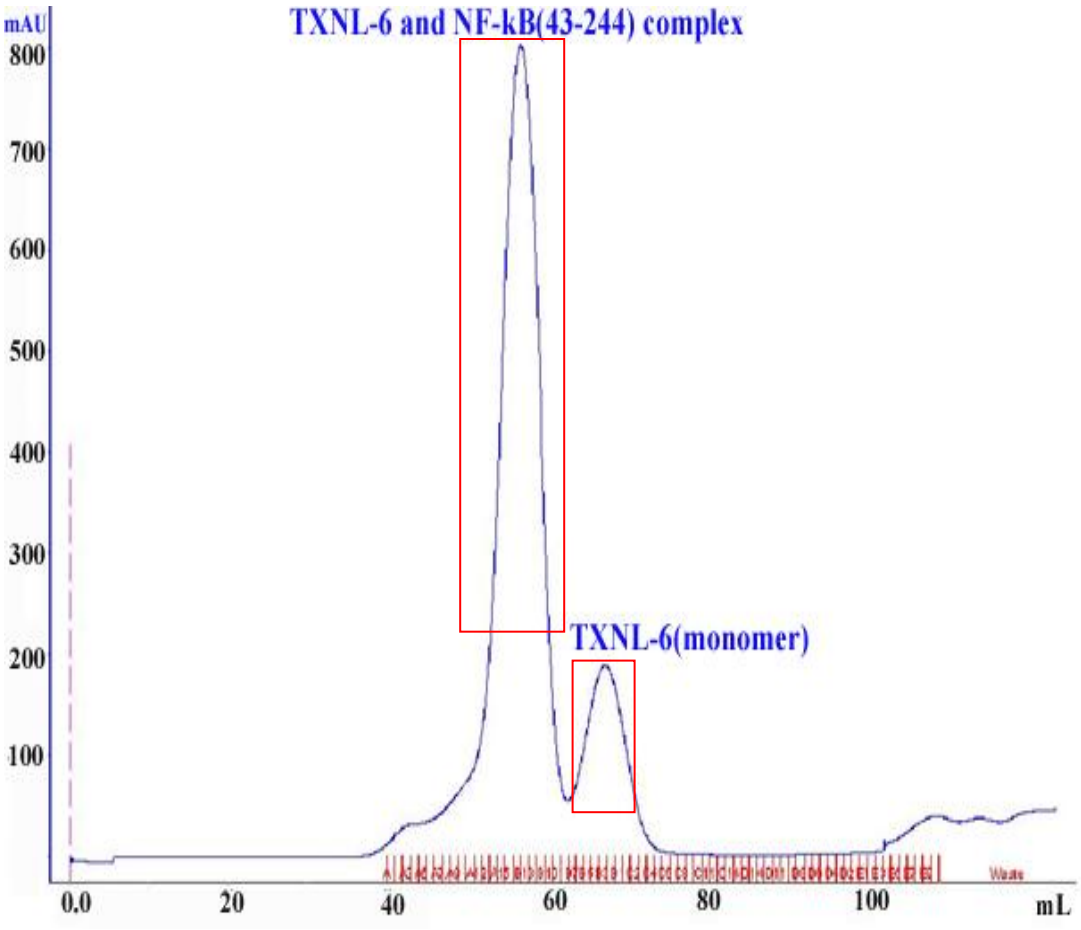


Figure III.6: FPLC profile of TXNL6 and NF-κB p50 complex protein. After affinity purification of TXNL6 and NF-κB p50 complex were passed through Superdex-200 column. There were two peaks in the elution profile. The complex was eluted as first peak in the gel filtration column where as the second peak corresponds to unbound monomeric TXNL-6.

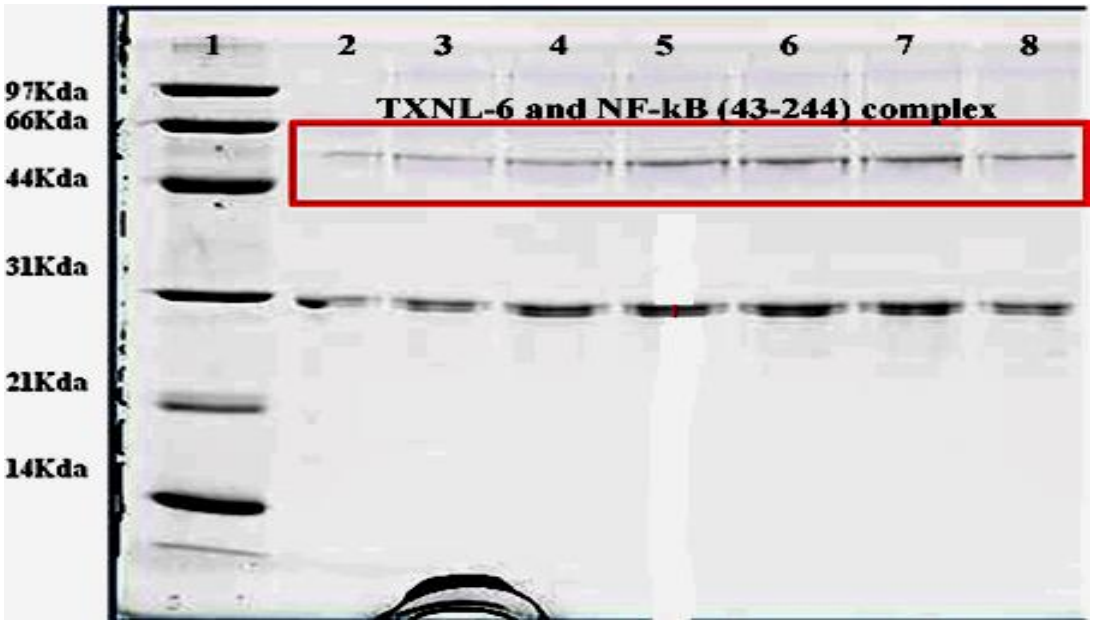


Figure III.7: In vitro interaction between TXNL6 and NF-κB p50 subunit (non-reducing SDS-PAGE).GST- NF-κB(43-244), bound to glutathione sepharose beads, was incubated with human TXNL-6.The mixture was kept at 4 °C with 20 mM diamide for overnight .After elution from the GST-column, it was further passed through Ni-NTA beads to eliminate free NF-kB band. The TXNL6 and NF-κB p50 complex was finally purified by using a Superdex-200 column followed by Ni-NTA affinity purification. Samples from gel-filtration were collected, analyzed by 12% non-reducing SDS-PAGE and visualized by Coomassie staining. The eluted TXNL6 and NF-κB p50 complex migrated in the gel with an apparent molecular weight of 47 kDa (lane2-8). A molecular marker is shown at lane 1 of the gel.

Mascot Search Results

User :
Email :
Search title : SampleSetID: 942, AnalysisID: 1467, MaldiWellID: 51199, SpectrumID: 143785, Path=\Michelle2009\090319\A3-B24
Database : NCBInr 090202 (7777454 sequences; 2682183781 residues)
Timestamp : 19 Mar 2009 at 10:00:05 GMT
Significant hits: gi|189178 nuclear factor kappa-B DNA binding subunit
gi|19923987 nucleoredoxin-like 1 [Homo sapiens]

Probability Based Mowse Score

Ions score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 52 indicate identity or extensive homology (p<0.05).
Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits.

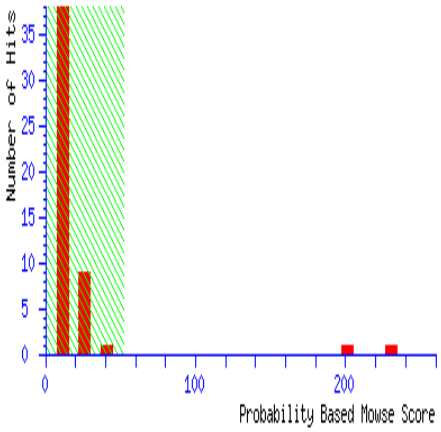


Figure III.8: Identification of TXNL6 and NF-κB p50 elution peak by peptide mass fingerprint. The peptide mass fingerprint (PMF) of peak 1 from gel-filtration (Fig. 3.7) result revealed the identity of the TXNL6 and NF-κB p50 heterodimer.

3.2.4 Crystallization of TXNL-6 and its complex with NF-kB-p50

Crystallization screens for TXNL-6 and its complex with NF-kB were setup at 4 and 25 °C by the hanging drop vapor diffusion method using Hampton Research crystallization screens, JCSG+ (Qiagen),Nextal (Qiagen) and JB crystallization screens (Jena Biosciences) with drops containing equal volumes (1 µl each) of reservoir and protein solution (12 mg/ml) against 0.5 ml of reservoir .Presently drops are under observation. Once, we will have crystals, structure and subsequent functional studies will be performed.

In the next chapter, we are going to discuss the structure based rational design of altered specificity of a serine protease inhibitor.

CHAPTER-IV: MODIFYING THE SUBSTRATE SECIFICITY OF *Carcinoscorpius rotundicauda* SERINE PROTEASE INHIBITOR DOMAIMN 1 TO TARGET THROMBIN

4.1 INTRODUCTION

The innate immune system is the first line of inducible host defense against various pathogens and their products (Hoebe et al., 2004). Secreted proteases serve important roles in pathogen virulence. Several families of protease inhibitors from the host play an important role in innate immunity by inactivating and clearing the proteases from the pathogens. Horseshoe crab hemocytes contain granules filled with several serine protease zymogens. During mechanical injury or pathogen invasion, the granules are released into the extracellular milieu by exocytosis, and precursor forms of clotting enzymes are activated by a serine protease cascade triggered by bacterial endotoxin. This pathogen induced cascade is regulated by three serpins, also known as *Limulus* intracellular coagulation inhibitors (LICI-1, LICI-2 and LICI-3) (Armstrong, 2001; Ding et al., 1993; Gettins, 2002; Muta and Iwanaga, 1996). Protease inhibitors, thus plays multiple roles by maintaining homeostasis and eliciting innate immunity (Hoffmann et al., 1999). This defense system is essential for the survival and perpetuation of all multicellular organisms (Hoffmann et al., 1999; Salzet, 2001).

The Kazal family is one amongst 18 families of serine protease inhibitors, and is mainly divided into two groups: the classical and the non-classical inhibitors. Non-classical Kazal inhibitors (Hemmi et al., 2005) consist of one to seven repeated domains, with each domain constituting 50-60 amino acid residues. Regardless of whether a domain is functionally active, it contains a reactive site loop (RSL) exposed at the

surface. The serine protease inhibitor functions as a substrate analogue, but the resulting enzyme-inhibitor complex is very stable (Kanost, 1999).

We recently reported a two-domain non-classical Kazal serine protease inhibitor from the hepatopancreas of *Carcinoscorpious rotundicauda* (CrSPI) with a possible dual function of inactivating pathogen protease (subtilisin) and host protease (furins). The full length and domain 2 of CrSPI-1 have been shown to contain full inhibitory activities against subtilisin. However, the function of the domain 1 of CrSPI (hereafter referred to as CrSPI-1-D1) is not yet characterized (Jiang et al., 2009). Analysis of the CrSPI-1-D1 sequence shows that it is significantly homologous to that of rhodniin-D1 from *Rhodnius prolixus*, which is a thrombin inhibitor (van de Locht et al., 1995). A number of endogenous thrombin inhibitors are available, and the most potent one is hirudin from the medicinal leech, *Hirudo medicinalis* (Rydel et al., 1991).

In spite of several studies on serine protease inhibitors, CrSPIs are relatively new and potent (Jiang et al., 2009). There are several unexplored potentials and unanswered questions about CrSPIs, for example, what is the structural homology of the CrSPI domains, among themselves and other SPIs? What is the variance of target specificity and inhibition? In order to address these questions we have undertaken the structural and functional studies on CrSPI-1-D1.

In this chapter, we report the crystal structure of CrSPI-1-D1 refined up to 2.0Å resolution. Despite the close structural homology of CrSPI-1-D1 to rhodniin-D1, the native CrSPI-1-D1 does not inhibit thrombin. This motivated us to modify the selectivity of the CrSPI-1-D1 to specifically target thrombin. We show that sequential mutations in the RSL region of CrSPI-1-D1 generated a potent and specific thrombin inhibitor. The

full length CrSPI-1 with this modified role of CrSPI-1-D1 as a thrombin modulator, might play a central role in regulating not only hemostasis but also inflammation, and may provide a close link between these processes and how they might co-evolve in the biological system. Furthermore, the possibilities to further develop this D1 mutant into a shorter yet active anti-thrombin holds potentials for biomedical applications as a coagulation modulator(Drag and Salvesen, 2010; Gurm and Bhatt, 2005; Markwardt, 2002; Pfau, 2003; Steinmetzer et al., 2001; Turk, 2006a; Vacca, 2000; Weitz and Crowther, 2002).

4.2 EXPERIMENTAL PROCEDURES

4.2.1 Plasmid and strain construction

The CrSPI-1-D1 (encoding Cys1-Glu40) was PCR amplified using forward CTACTGGATCCTGTCCTCAT and reverse GCAGAGTTCGAATTCCTAGCAAGT TTC CCA primers that were designed to introduce a *Bam* H1 site to the 5` end and an *Eco* R1 site to the 3` end. Such PCR fragments were then digested with *Bam* H1 and *Eco* R1, and ligated into pET-M vector, which were previously linearized by compatible restriction enzymes, and transformed into *Escherichia coli*, BL 21.

4.2.2 Expression and Purification

Optimal expression of the CrSPI-1-D1 was conducted by induction with 0.5mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) of 1 liter culture at 25 ºC. The cells were then disrupted by French Press and the supernatant were collected after centrifuging at 10,000x g for 1 h at 4 ºC. His-tagged CrSPI-1-D1 proteins were purified in two steps using Ni-NTA (Qiagen) affinity chromatography followed by a Superdex 75 gel filtration

column on the Akta Express (GE Healthcare). The buffer was exchanged to a solution containing 20 mM Tris (pH-8.5), 150 mM NaCl, 5 mM dithiothreitol and finally concentrated up to 10 mg/ml.

4.2.3 Crystallization and structure determination

Initial crystallization conditions were screened at 25 ^{o}C in the hanging drop vapor diffusion technique using Hampton Research crystallization screens and JB crystallization screens (Jena Biosciences) with drops containing equal volumes (1 µl) of reservoir and protein solution of 10 mg/ml against 0.5 ml of reservoir. Small rod-shaped crystals were formed within 2–3 days. Further optimization by equilibrating 1 µl CrSPI-1-D1 protein solution of 15 mg/ml and 1 µl reservoir solution (0.4 M mono ammonium dihydrogen sulphate, 0.1 M Tris-HCl pH 8.5) using hanging drop vapor diffusion technique at 20 ºC led to best diffraction-quality crystals. The crystals diffracted up to 2.0Å and belonged to space group $P2_1$ with solvent content is approximately of 35% (Vm = 1.9Å^3/kDa).

Crystals were cryo-protected in the reservoir solution supplemented with 25–30% glycerol, and flash cooled at 100 K. The diffraction data were obtained using a CCD detector (Platinum 135) mounted on a Bruker Microstar Ultra rotating anode generator (Bruker AXS, Madison, WI). All datasets were processed with HKL2000(Otwinowski and Minor, 1997). The structures were solved by molecular replacement with PHASER(McCoy et al., 2007). Subsequently the models were manually built by using COOT (Emsley and Cowtan, 2004), followed by refinement using CNS(Brunger et al., 1998). The data collection and refinement statistics are provided in Table 4.1.

4.2.4 Site-directed mutagenesis

Based on the rhodniin-thrombin complex structure (PDB code 1tbq), residues Thr4, Tyr5, Lys6 and Pro7 of CrSPI-1-D1 were mutated to Ala, Leu, His and Arg respectively, the corresponding residues 8-11 of rhodniin that are crucial for interaction with thrombin (Table 4.2). We used inverse PCR based mutagenesis (Ochman et al., 1988) to generate all mutants. In total, we generated 15 mutants (single to tetra). All mutant inhibitor proteins were expressed *E. coli* (BL21DE3) using optimized expression conditions and purified by His-tag based affinity and size exclusion column chromatography. The gel filtration purified CrSPI1-D1 was further purified by reverse phase chromatography through an analytical Jupiter C18 column on a Smart workstation (GE Healthcare) using a gradient (15-40% over 60 min) of buffer B (80% acetonitrile in 0.1% TFA). The molecular masses of the RP-HPLC purified mutants were determined by ESI-MS on a Perkin-Elmer Sciex API III triple-stage quadrupole instrument equipped with an ionspray interface.

4.2.5 CD spectroscopy

Far-UV CD spectra (260–190 nm) of CrSPI-1-D1 dissolved in 20 mM Tris-HCl buffer (pH 7.4) at a 30 μM protein concentration were collected using a Jasco J-810 spectropolarimeter (Easton, MD). All measurements were carried out at room temperature using 0.1-cm path length cuvettes with a scan speed of 50 nm/min, a resolution of 0.2 nm, and a bandwidth of 2 nm.

4.2.6 Stability verification of CrSPI-1-D1 mutants against serine proteases

20 µL of 1mg/ml CrSPI-1-D1 mutants were incubated with fixed quantity 1 µL of 1mg/ml of different serine proteases such as thrombin, trypsin, chymotrypsin, elastase and subtilisin at 37 ºC for 30 min. Reaction was stopped by heating the sample with 5X SDS loading dye at 100 ºC. One-dimensional (1-D) SDS PAGE was carried out following a standard protocol.

4.2.7 Inhibition of Thrombin Amidolytic Activity

The buffer used in all functional assays was 20 mM Tris-HCl, pH 7.4. For all thrombin amidolytic activity assay, we used S2238 (H-D-Phenylalanyl-L-pipecolyl-Larginine-p-nitroaniline dihydrochloride), which is a chromogenic substrate for thrombin from Chromogenix (Milano, Italy). To measure the inhibition activity of different CrSPI-1-D1 proteins on thrombin activity, we performed all reactions in 96-wells microtiter plates. For each inhibition assay, 50 µl of 4.5 nM α-human thrombin was pre-incubated for 30 min at 37 ºC with increasing amounts (10 to 70 nM) 50 µl of CrSPI-1-D1 in a total reaction volume of 200 µl, prior to adding 100 µl of S2238. The rate of formation of colored product, p-nitroaniline, was read using an enzyme-linked immunosorbent assay plate reader at 405 nm for 10 min. Appropriate negative controls without the thrombin was assayed simultaneously. Percentage inhibition was calculated by taking the rate of increase in absorbance in the absence of inhibitor as 0%. A decrease in absorbance indicated the inhibitory effect of CrSPI-1-D1 on thrombin activity.

4.2.8 Isothermal Titration Calorimetry (ITC)

The ITC experiments were carried out using VP-ITC calorimeter (Microcal, LLC) at 20 °C using 300 μM of the protein in the sample cell and 40 μM of α-human thrombin in the injector. All samples were thoroughly degassed and then centrifuged to get rid of precipitates. Volumes of 10 μl per injection were used for the different experiments. For every experiment, the heat of dilution for each ligand was measured and subtracted from the calorimetric titration experimental 30 runs for the protein. Consecutive injections were separated by at least 4 min to allow the peak to return to the baseline. The ITC data was analyzed using a single site fitting model using Origin 7.0 (OriginLab Corp.) software.

4.3 RESULTS

4.3.1 Overall structure

The structure of CrSPI-1-D1 was solved by molecular replacement method and refined to a final R-factor of 0.21 (R_{free}=0.25) at 2.0Å resolution. The model has been refined with good stereo chemical parameters (Table 4.1). There are two CrSPI-1-D1 monomers in the asymmetric unit. The structure of CrSPI-1-D1 mostly consists of loops with a two-strand (Val8-Gly10 and Gly13-Tyr16) β-sheet and a two-turn α-helix. In addition, a single turn α-helix (Trp33-Cys36) is present at the C-terminal. A disulphide bond is located between Cys1 and Cys20 to help maintain the rigidity of the RSL. The carboxyl terminus is linked to the N-terminal through a second disulphide bridge, Cys9-Cys36 (Fig. 4.1).

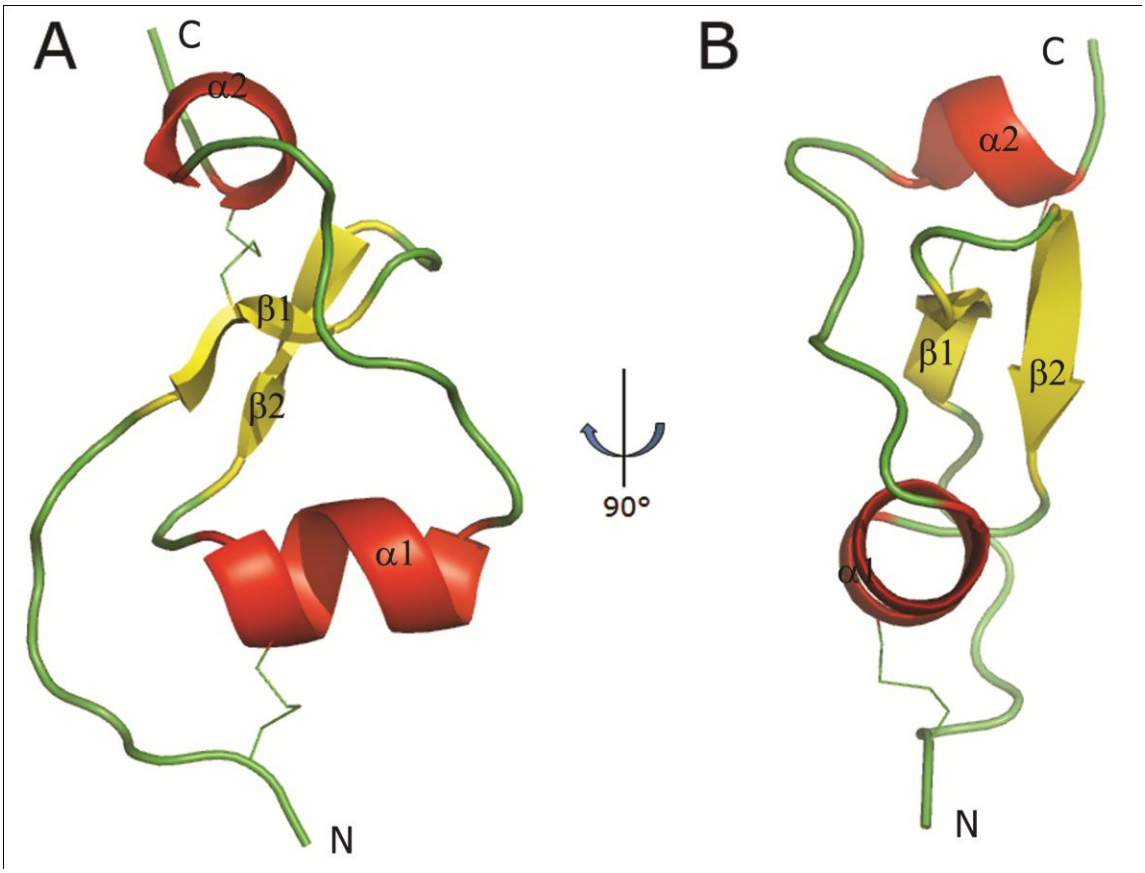


Figure IV.1: Structure of CrSPI-1-D1.*(A)* Ribbon diagram of the CrSPI-1-D1. *(B)* 90˚ rotated side view. α-Helix, β-strands and random coils are depicted in red, yellow and green, respectively. The disulfide bridges are shown in green. The secondary structures, N- and C-termini, are labeled. This figure and the following figures of this manuscript were prepared using the program PyMoL(DeLano and Lam, 2005) .

4.3.2 Structural comparison

A search for topologically similar domains within the PDB database using the DALI program (Holm and Sander, 1998) revealed that the structural features of CrSPI-1-D1 resemble the typical non-classical Kazal type inhibitor (Hemmi et al., 2005). The highest structural similarity is observed between hirudin, the leech-

derived tryptase inhibitor from *H. medicinalis* and CrSPI-1-D1, yielding an rmsd of 1.9 Å for 36 Cα atoms (pdb code 1ldt). This is followed by a thrombin protease inhibitor, rhodniin domain 1 (rhodniin-D1) from *Rhodnius prolixus,* which yielded an rmsd of 2.0 Å for 36 Cα atoms (pdb code 1tbq).

In addition to the structural homology, the CrSPI-1-D1 and rhodniin-D1 display 42% sequence identity while only 35% sequence identity was observed with hirudin. The structure-based sequence alignment revealed that most of the structurally invariant residues are located at the carboxy terminus, including the RSL, β1, β2 and α1 of CrSPI-1-D1 (Figure 4.2). These observed features provided a clue that CrSPI-1-D1 might specifically target thrombin after modifications of a few residues in the RSL, and this prompted us to change the specificity of CrSPI-1-D1 to target thrombin.

Table IV.1: Data collection and refinement statistics of CrSPI-1-D1

Experiment	
Cell parameters (Å, °)	a= 25.5, b= 37.2, c= 36.5, α=90, β= 99.8, γ=90
Space group	$P2_1$
Data collection [a]	
Resolution range (Å)	50.0-2.0 (2.07-2.00)
Observed Reflections	28,930
Unique Reflections	4,535
Redundancy	6.4 (3.2)
Completeness (%)	97.3 (80.7)
Overall ($I/\sigma I$)	31.3
R_{sym} [b]	0.046 (0.061)
Refinement	
Resolution range (Å)	20.0-2.0
Number of Reflections used	4432
R factor[c]/ R_{free} [d] (%)	21.47/25.56
RMSD bond lengths (Å)	0.009
RMSD bond angles (°)	1.660
Average B-factors ($Å^2$)	18.046
Main chain (# atoms)	16.045 (312)
Side chain (# atoms)	18.394 (264)
Water (# atoms)	27.559 (56)
Ramachandran Plot	
Most favored region (%)	93.4
Additional allowed regions (%)	4.9
Generously allowed regions (%)	1.6
Disallowed regions (%)	0.0

a. Numbers in parentheses refer to the highest resolution shell.

b. $R_{sym} = \Sigma|I_i - \langle I_i \rangle| / \Sigma| I_i|$

c. $Rf_{actor} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$

d. Rfree equals the R factor against 5.9% of the data removed prior to refinement.

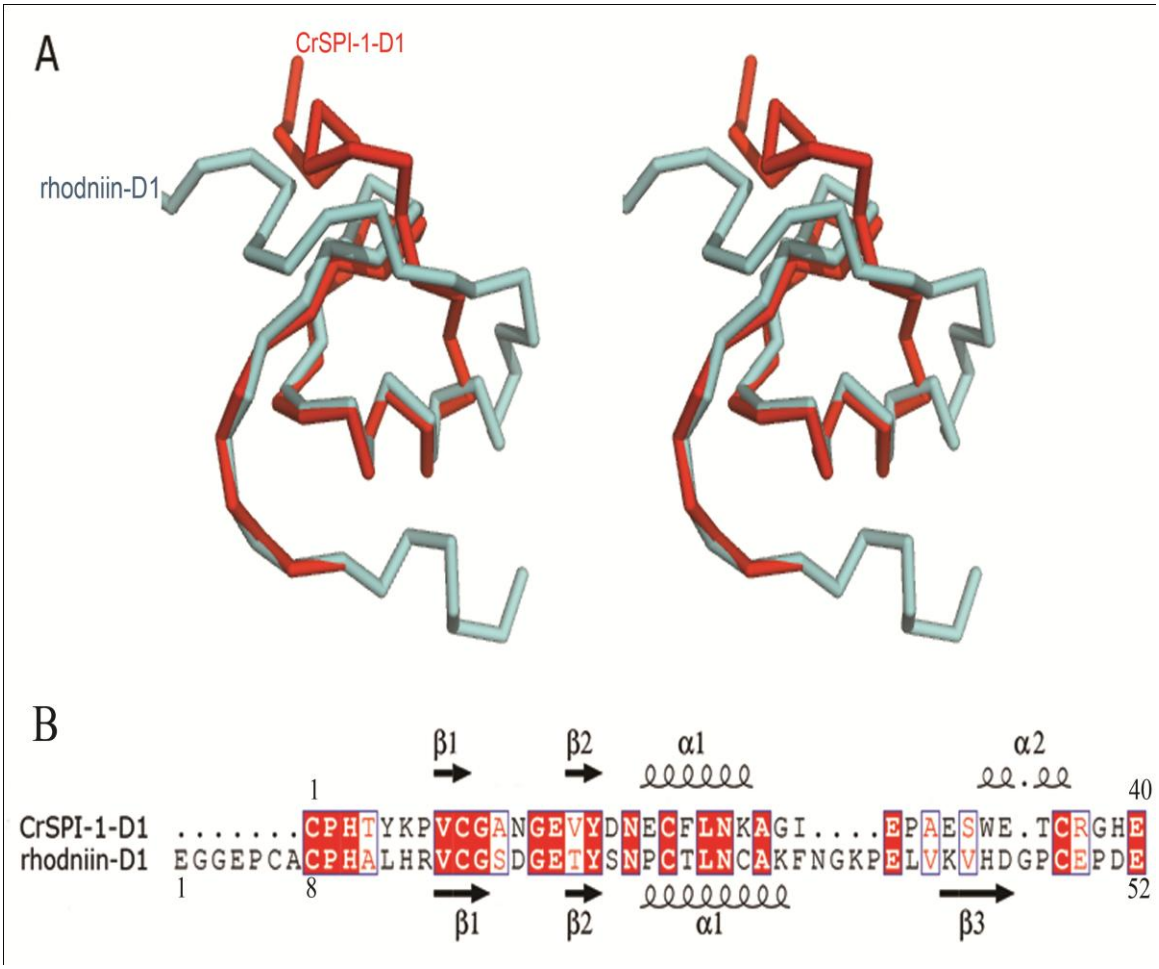


Figure IV.2: Comparison of CrSPI-1-D1 with rhodniin-D1. *(A)* Stereo Cα superposition of CrSPI-1-D1 (red) and rhodniin-D1 (cyan). The RMSD between CrSPI-1-D1 and rhodniin-D1 is 2.0 Å for 36 Cα atoms. *(B)* Structure based sequence alignment between CriSPI-1-D1 and rhodniin-D1. This alignment was performed using the program COOT (Emsley and Cowtan, 2004). The secondary structural elements for CrSPI-1-D1 and rhodniin-D1 are shown at the top and the bottom, respectively. The conserved residues are highlighted in red boxes outlined in blue. This figure was created by using the program ESPript (Gouet et al., 1999).

4.3.3 The reactive-site loop

Although the sequence of the reactive-site loop (RSL) is different in several families of serine protease inhibitors, the conformation of the RSL is similar (Jiang et al., 2009; van de Locht et al., 1995). Like other Kazal-type inhibitors, the disulfide bonds formed by cysteine residues at the P3 and P5' positions (Cys1 and Cys9 in CrSPI-1-D1) hold the RSL in a relatively rigid conformation. Besides, there are several internal hydrogen bonds (<3.2 Å) which help maintain the rigidity of the RSL in the CrSPI-1-D1. Figure 3c shows selected hydrogen bonding contacts between RSL and CrSPI-1-D1. Notably, strong intra-molecular H-bonds (< 3.0 Å) were observed between the carbonyl oxygen of Pro2 (P2 position) and amide nitrogen of Thr4 (P1' position); Asn18 and Phe21, ND2 of Asn18 interacts with the main chain carbonyl atoms of Pro2 and Thr4 at the P2 and P1' positions of the RSL, respectively (Table 4.2).

Table IV.2: Interaction involved for rigidity of reactive site loop of CrSPI-1-D1

Interaction		Distance (Å)
CrSPI-1-D1	O (Pro2,P1′ position)-ND2 (Asn18)	2.94
	O (Thr4,P2′ position)-ND2 (Asn18)	2.85
	OD1 (Asn18)- N(Phe21)	2.86
	O(Asn18)-N(Phe21)	3.14
	O (Pro2,P1′ position)-N (Thr4,P2 position)	3.57
rhodniin-D1	O (Pro9,P1′ position)-ND2 (Asn25)	3.50
	O (Ala 11,P2 position)-ND2 (Asn25)	2.80
	OD1 (Asn18)- N(Thr28)	3.17
	O(Asn18)-N(Thr28)	3.13
	O (Pro9,P1′ position)-N (Ala 11,P2 position)	3.25
OMTYK3	O (Thr 17,P1′ position)-ND2 (Asn33)	2.84
	O (Glu 19,P2 position)-ND2 (Asn33)	3.00
	OD1 (Asn33)- N(Asn 36)	3.00
	O(Asn18)-N(Asn 36)	3.20
	O (Thr17,P1′ position)-N (Glu19,P2 position)	3.75

Similar interactions were observed in rhodniin-D1 and other protease inhibitors such as the turkey ovomucoid third domain, OMTKY3, although there are different amino acids in those positions (Maynes et al., 2005). In addition to the S-S bonds, these hydrogen bonds are essential to maintain the rigidity of the RSL during the inhibition of the cognate enzyme. Although a similar rigid conformation is found in these inhibitors, they recognize the substrates differently. This clearly shows that in addition to the rigid conformation, the sequence of the RSL dictates the selectivity towards a particular protease. Thus, we have mutated the RSL side chains of CrSPI-1-D1 to specifically target thrombin.

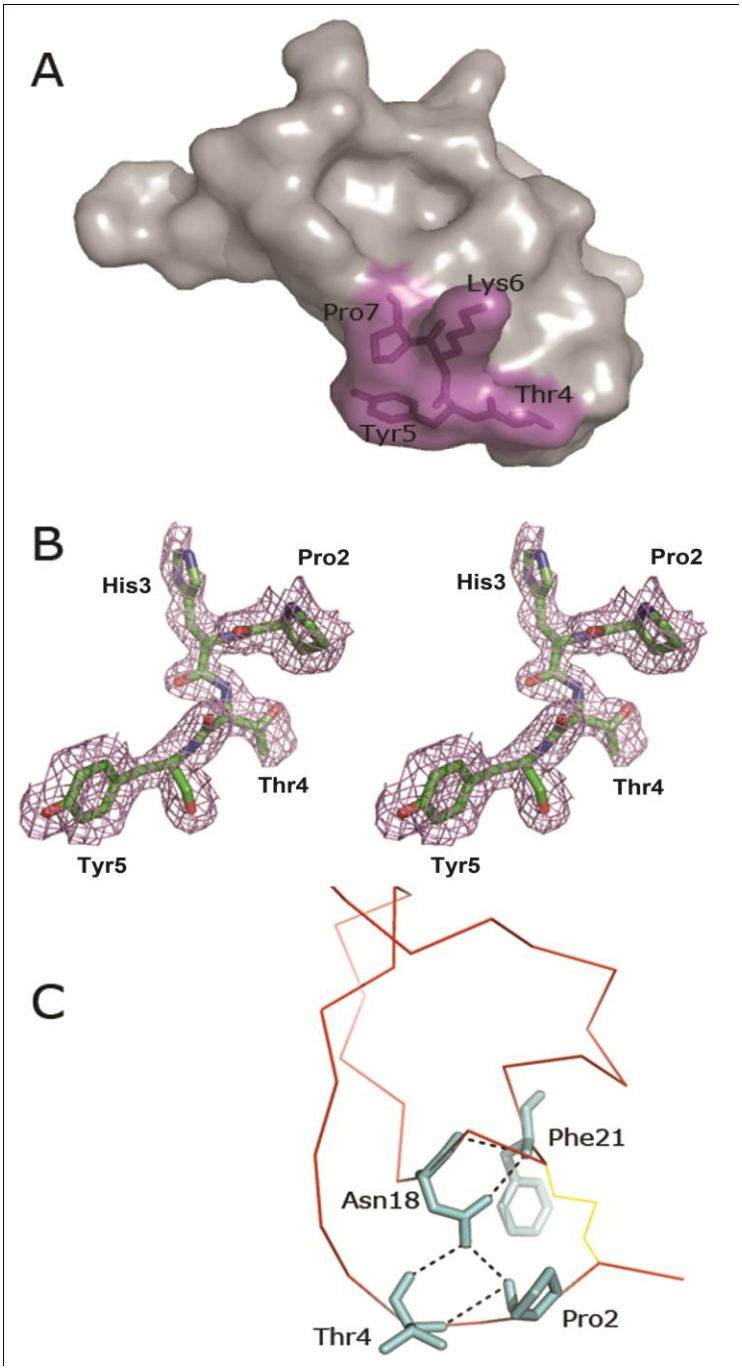


Figure IV.3: The reactive-site loop (RSL). *(A)* The site-directed mutation on CrSPI-1-D1. A transparent surface representation of the CrSPI-1-D1 is shown with all mutated residues in stick representation in magenta. *(B)* Stereo view of the electron density map. Simulated annealing *Fo-Fc* omit map of the N-terminal region of CrSPI-1-D1 showing the key residues in reactive-site loop. All residues shown in this figure as well as all atoms

within 2 Å of these residues were omitted prior to refinement and map calculation. The map was contoured at a level of 2.0 σ. *(C)* A close view shows the interactions aid in maintaining the rigidity of the RSL. Cα of the CrSPI-1-D1 is shown in red. The disulfide bonds are shown in yellow, stick line while the hydrogen bonds are shown in black, dotted lines.

4.3.4 Mutations to change the specificity

Following the structure determination of CrSPI-1-D1, the next main objective was to elucidate the inhibitory efficiency of this domain. Our previous studies showed that full length as well as domain 2 of CrSPI-1 is a specific inhibitor of subtilisin, however the specificity of domain-1 is not yet established (Jiang et al., 2009). An analysis of P3 to P4` residues of the RSLs of various substrates like binding serine protease inhibitors such as for subtilisin, thrombin, trypsin, chymotrypsin and furin was performed to identify the minimum side chains of CrSPI-1-D1 to be mutated to alter the selectivity (Table 4.3). The closest similarities were observed with RSLs of rhodniin-domain-1. P3, P2 and P1 of CrSPI-1-D1 and rhodniin-D1 are similar, but P1`, P2`, P3` and P4` were different. Complex crystal structure of rhodniin and thrombin showed that the N terminal domain of rhodniin interacts with the active-site cleft region of thrombin (PDB 1tbq). In addition to the interactions of Pro9, His10 and Alall, the side chain of Leu12 occupies the S2' site of thrombin. His13 mediates a hydrogen bond and stacks with aromatic residues in S3'. Arg14 at P4' allows charge compensation of Glu39 from thrombin. The clustering of the positively charged inhibitor residues at P3' and P4' might be particularly beneficial for thrombin binding (Lombardi et al., 1999). Based on these observations we have introduced mutations in the RSL region of CrSPI-1-D1, which was previously of uncharacterized function, to specifically target thrombin.

Table IV.3: Reactive site loop regions from P3 to P4' position of selected serine protease inhibitors

Inhibitors	P3	P2	P1	P1'	P2'	P3'	P4'	Inhibitor against	[ref.]
CrSPI-1 domain-1	Cys1	Pro2	His3	Thr4	Tyr5	Lys6	Pro7	Uncharac--terized	(Jiang et al., 2009)
CrSPI-1 domain-2	Cys46	Thr47	Glu48	Glu49	Tyr50	Asp51	Pro52	Subtilisin	(Jiang et al., 2009)
rhodniin domain-1	Cys8	Pro9	His10	Ala11	Leu12	His13	Arg14	Thrombin	(van de Locht et al.,
rhodniin domain-2	Asp61	Gly62	Asp63	Glu64	Tyr65	Lys66	Pro67	Thrombin	(van de Locht et al.,
LDTI (Leech Derived Tryptase Inhibitor)	Cys6	Pro7	Lys8	Ile9	Leu10	Lys11	Pro11	Trypsin	(Stubbs et al., 1997)
Tomato inhibitor II domain-1	Cys3	Thr4	Arg5	Glu6	Cys7	Gly8	Asn9	Subtilisin	(Barrette-Ng et al.,
Tomato inhibitor II dommain-2	Cys60	Thr61	Phe62	Asn63	Cys64	Asp65	Pro66	Subtilisin	(Barrette-Ng et al.,
OMTKY3	Cys16	Thr17	Leu18	Glu19	Tyr20	Arg21	Pro22	Subtilisin	(Maynes et al., 2005)
Spn4A (Furin inhibitor from Drosophila)	Arg371	Lys372	Arg373	Ala374	Ile375	Met376	Ser377	Furin	(Richer et al., 2004)
Human PI8 (Furin Inhibitor)	Asn337	Ser338	Arg339	Cys340	Ser341	Arg342	Met343	Furin	(Leblond et al., 2006)

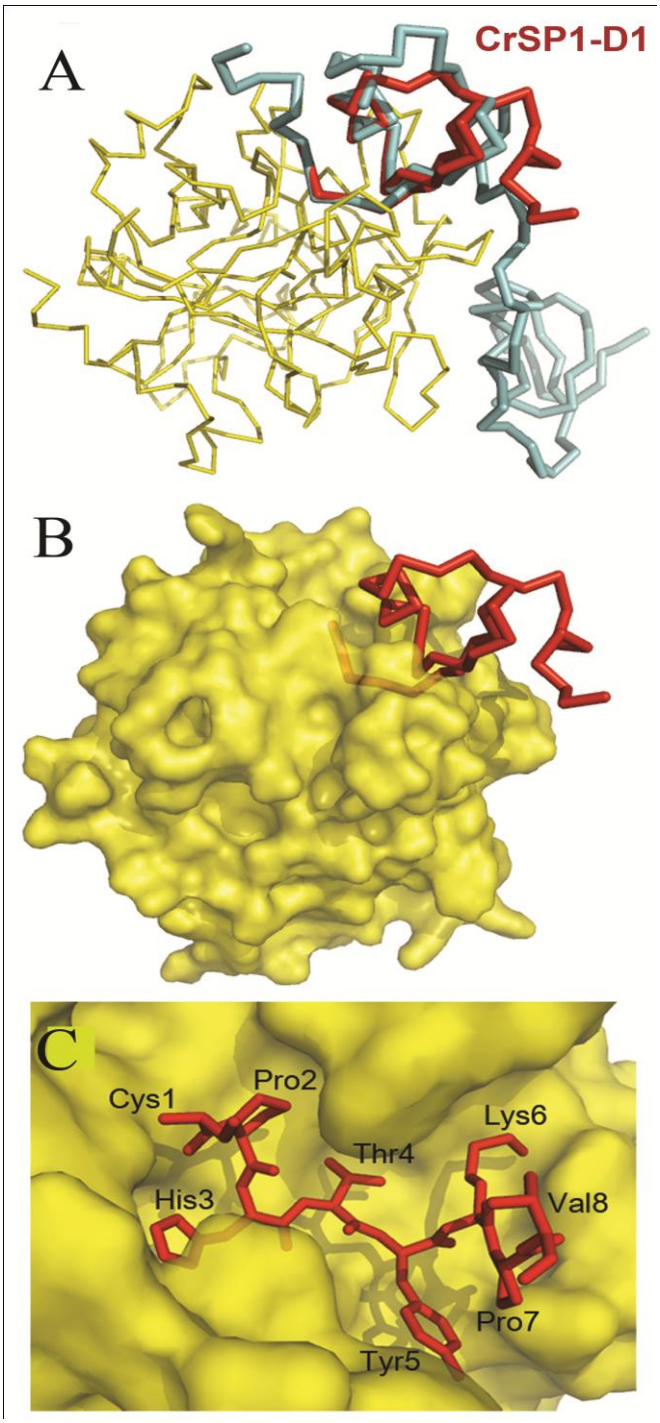


Figure IV.4.5: Modeling complex of CrSPI-1-D1 with thrombin. ***(A)*** Complex structure of rhodniin with thrombin (pdb code 1tbq). Rhodniin and thrombin are shown in cyan and yellow, respectively. CrSPI-1-D1 (red) was superimposed on rhodniin-D1. ***(B)*** Thrombin is shown as surface representation in this model complex. ***(C)*** Close up view of the RSL of CrSPI-1-D1 in the active site of thrombin.

Our approach was to mimic the P1`, P2`, P3` and P4` residues (Thr4, Tyr5, Lys6 and Pro7) of CrSPI-1-D1 to rhodniin-D1 (Ala11, Leu12, His13 and Arg14); to sequentially mutate and evaluate the implication of these four residues towards thrombin inhibition. In addition to the tetra mutant, we have tried all possible single, double and triple mutants. A total of 15 mutants (Table 4.4) have been generated and their thrombin inhibition was studied. All the mutants were expressed in bacteria and purified as wild type CrSPI-1-D1 (Fig. 4.5) .The CD spectrum was recorded on all 15 mutants of CrSPI-1-D1, which indicated that these mutants share the same α/β structure as the wild type CrSPI-1-D1 (Fig. 4.6). Furthermore, the ESI-MS spectrum showed their expected molecular mass (Fig.4.9).

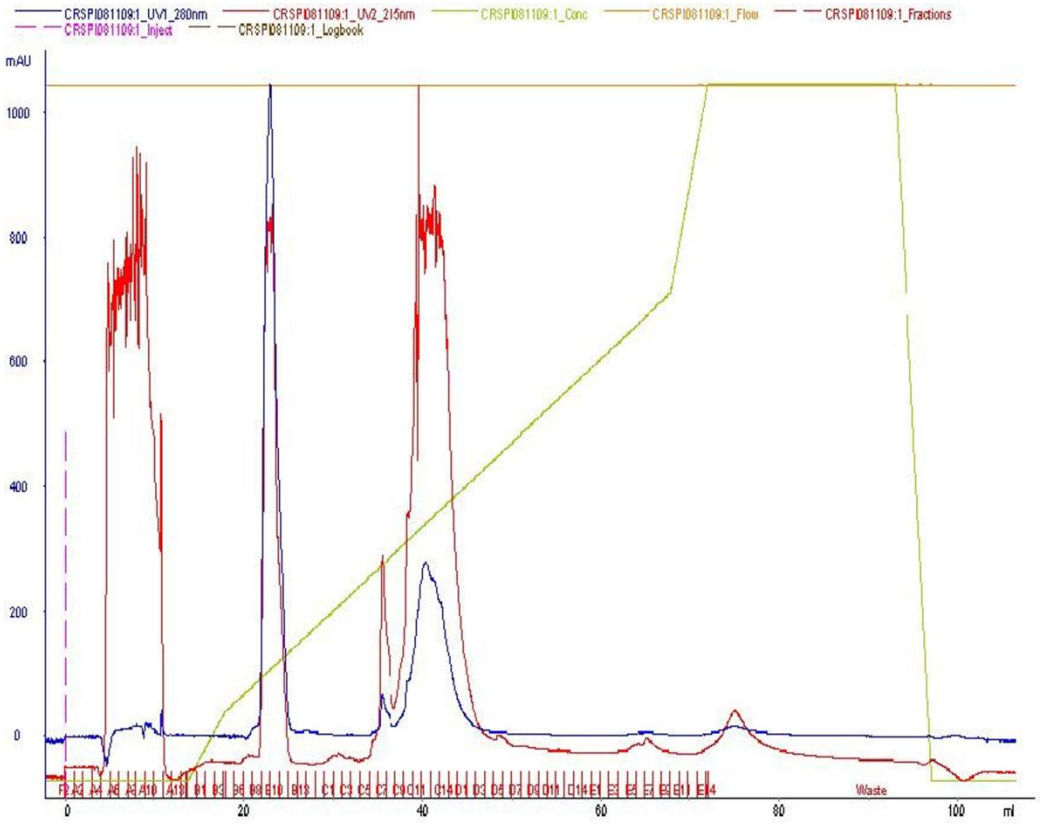


Figure IV.5: Reverse Phase-HPLC profile of CrSPI-1-D1.Gel filtration purified CrSPI-1-D1 was loaded onto an analytical Jupiter C18 analytical

column on SMART Workstation (GE-healthcare) using a gradient (15 - 40% over 60 min) of buffer B (80% ACN in 0.1% TFA. Figure shows the elution of protein monitored at 215 nm. The peak (indicated with the arrow) contains a single homogenous CrSPI-1-D1 taken for kinetics studies.

We have verified the stability of the CrSPI-1-D1 mutants as a possible inhibitor against different serine proteases such as thrombin, trypsin, chymotrypsin, elastase and subtilisin. Notably only the tetra mutant is stable against thrombin, whereas other serine proteases degrade the modified CrSPI-1-D1, which seemed to act more as substrate rather than an inhibitor (Fig. 4.7). It suggests that CrSPI-1-D1 mutant is thrombin specific. In the following section we describe the inhibition studies of CrSPI-1-D1 mutants with thrombin.

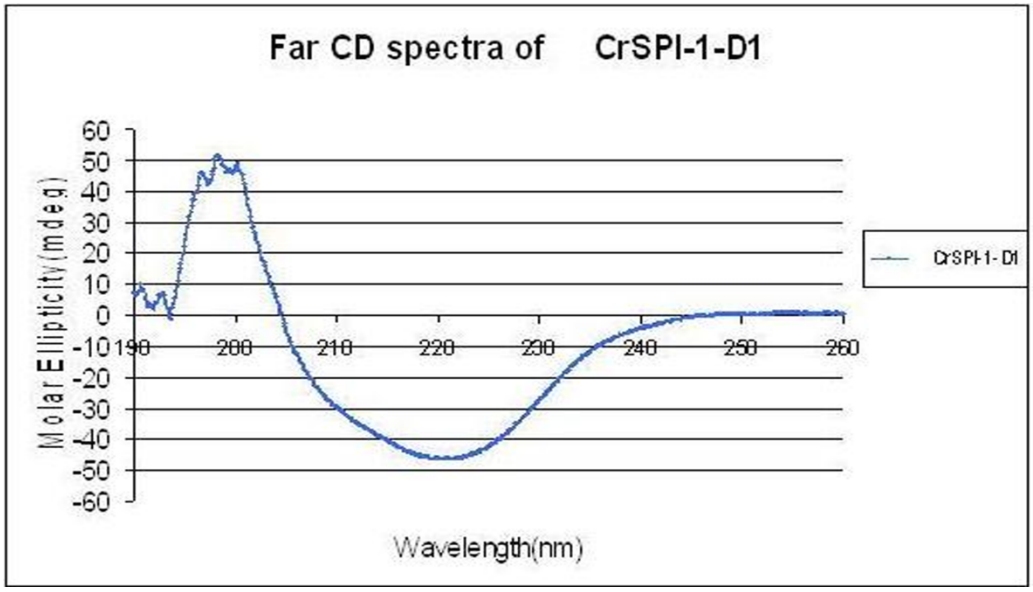


Figure IV.6: CD spectroscopy profile of reverse phase HPLC purified CrSPI-1-D1.Far-UV CD spectra (260–190 nm) of CrSPI-1-D1 dissolved in 20 mM Tris-HCl buffer (pH 7.4) at a 30 μM protein concentration were

collected using a Jasco J-810 spectropolarimeter (Easton, MD). All measurements were carried out at room temperature using 0.1-cm path length cuvettes with a scan speed of 50 nm/min, a resolution of 0.2 nm, and a bandwidth of 2 nm. The CD spectrum of the tetra mutant of CrSPI-1-D1 indicated that it assumed an α/β structure.

Table IV.4: IC_{50} and dissociation constant (Kd) for the inhibition of thrombin by various variants of CrSPI-1-D1

S.No.	**Mutants**	**IC_{50}**	**Kd**
1	T4A	ND	ND
2	Y5K	ND	ND
3	K6H	ND	ND
4	P7R	ND	ND
5	T4A,Y5K	ND	ND
6	T4A,K6H	ND	ND
7	T4A,P7R	ND	ND
8	Y5K,K6H	ND	ND
9	Y5K,P7R	ND	ND
10	K6H,P7R	ND	ND
11	T4A,Y5K,K6H	ND	ND
12	T4A,Y5K,P7R	ND	ND
13	T4A,K6H,P7R	ND	ND
14	Y5K,K6H,P7R	ND	ND
15	T4A,Y5K,K6H,P7R	26.3nM	4μM

ND- not detected

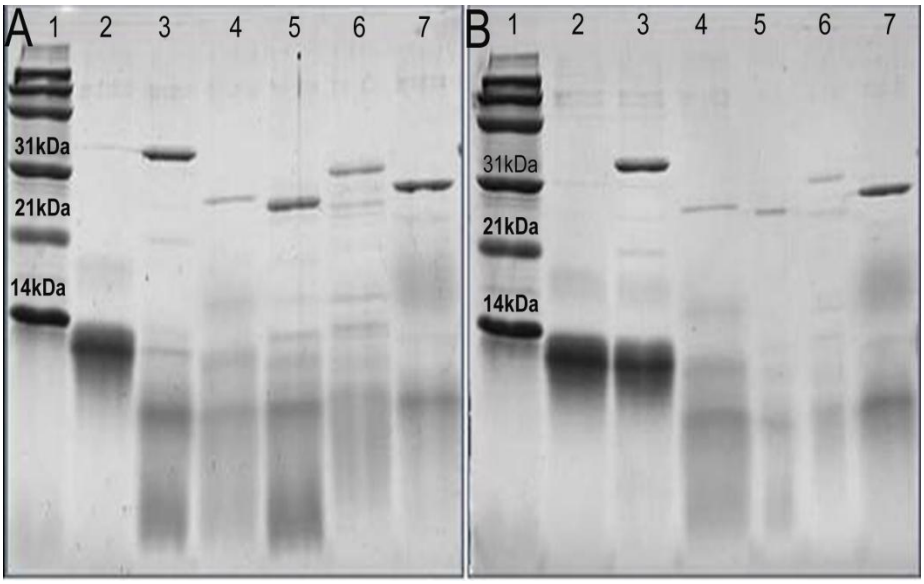


Figure IV.7: The specificity of CrSPI-1-D1 tetra mutant for thrombin ascertained by comparison with other proteases.SDS-PAGE analysis for the interaction of CrSPI-1-D1 wild type and tetra mutant with different proteases. *A)* Lane 1 protein marker ; Lane 2 CrSPI-1-D1 alone and Lane 3-7 CrSPI-1-D1 wild type incubated with α-human thrombin, chymotrypsin, trypsin, elastase and subtilisin, respectively, for 37 ºC for 30 min. *B)* Lane 1 protein marker; Lane 2 T4A,Y5K,K6H,P7R CrSPI-1-D1 alone and Lane 3-7 T4A,Y5K,K6H,P7R CrSPI-1-D1 incubated with α-human thrombin, chymotrypsin, trypsin ,elastase and subtilisin, respectively, for 37 ºC for 30 min.

4.3.5 Thrombin inhibition assay

Previously it has been shown that hirudin has very high inhibitory activity against the α-human thrombin(Markwardt, 1994). We chose to study the properties of these CrSPI-1-D1 variants under a similar condition as hirudin:α-human thrombin complex. Out of all 15 CrSPI-1-D1 mutants, only tetra mutant (T4A, Y5K, K6H, P7R) showed the highest significant inhibition with α-human thrombin in a dose-dependent manner.

Fig. 4.10 shows the typical dose-response curves. Wild type CrSPI-1-D1 showed no inhibition, whereas the tetra mutant exhibited strong inhibition against thrombin. The dose response plot of the fractional velocity as a function of different concentrations of tetra mutant CrSPI-1-D1 showed that 26.3 nM of tetra mutant CrSPI-1-D1 was sufficient to inhibit 50% of 4.5 nM thrombin (Fig. 4.8). Since the IC_{50} value of 26.3 nM is within a factor of 10 of the concentration of thrombin and CrSPI-1-D1, it is ascertained that the mode of inhibition follows the typical kazal domain's mode of inhibition. Following the inhibition studies, we verified the binding affinities of these mutants of CrSPI-1-D1 with α-human thrombin using ITC experiments.

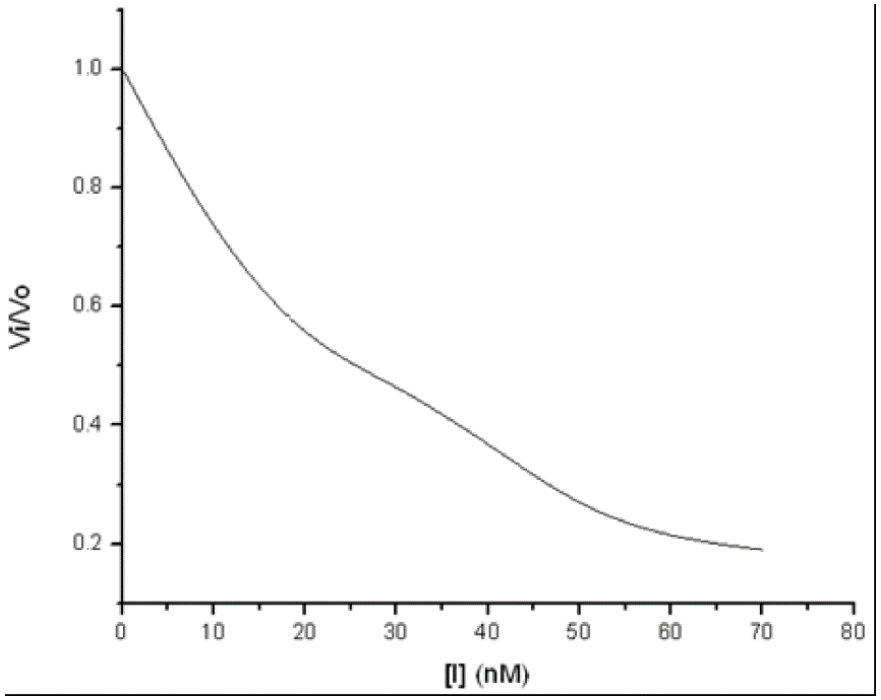


Figure IV.8: Determination of IC50 values based on dose response plots of fractional velocity as a function of different tetra mutant CrSPI-1-D1 concentration.OD_{405} was monitored based on hydrolysis of 0.1mM S2238 (H-D-Phenylalanyl-L-pipecolyl-Larginine-p-nitroaniline dihydrochloride)

and formation of colored product, p-nitroaniline, by 4.5 nM thrombin in a total reaction volume of 200µL. Vo and Vi are the initial velocities in the absence and presence of tetra mutant CrSPI-1-D1, respectively and were calculated from Beer-Lambert's Law. Hence, the tetra mutant CrSPI-1-D1 concentration predicted to block 50% of the activity of a fixed concentration of thrombin was obtained on a graphical plot of Vi/Vo versus inhibitor concentration.

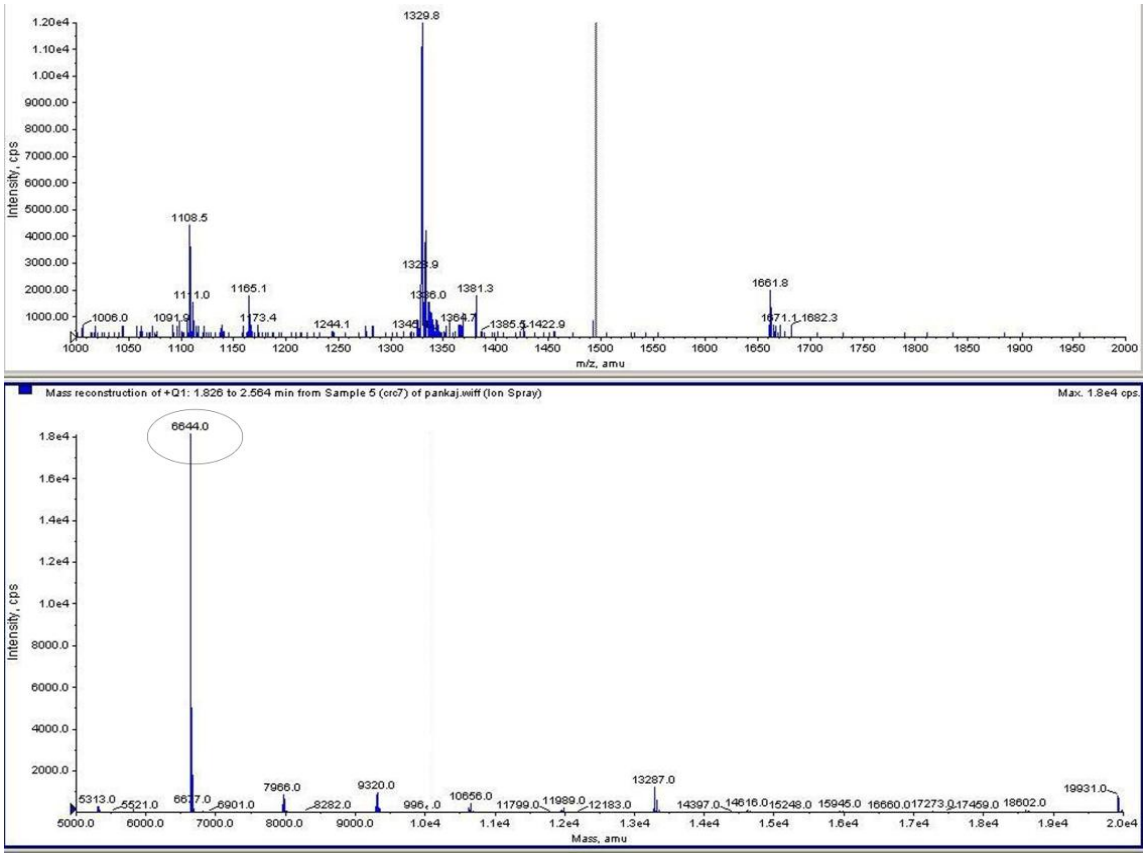


Figure IV.9: ESI/MS profile of reverse phase HPLC purified CrSPI-1-D1. The spectrum shows a series of multiply charged ions, corresponding to the correct molecular mass of 6644± 0.22 Da. The purity and mass of all mutant proteins of CrSPI-1-d1 were determined by electro spray ionization mass spectrometry using an API 300 liquid chromatography tandem mass spectrometry system (PerkinElmer Life Sciences Sciex, Selton, CT).

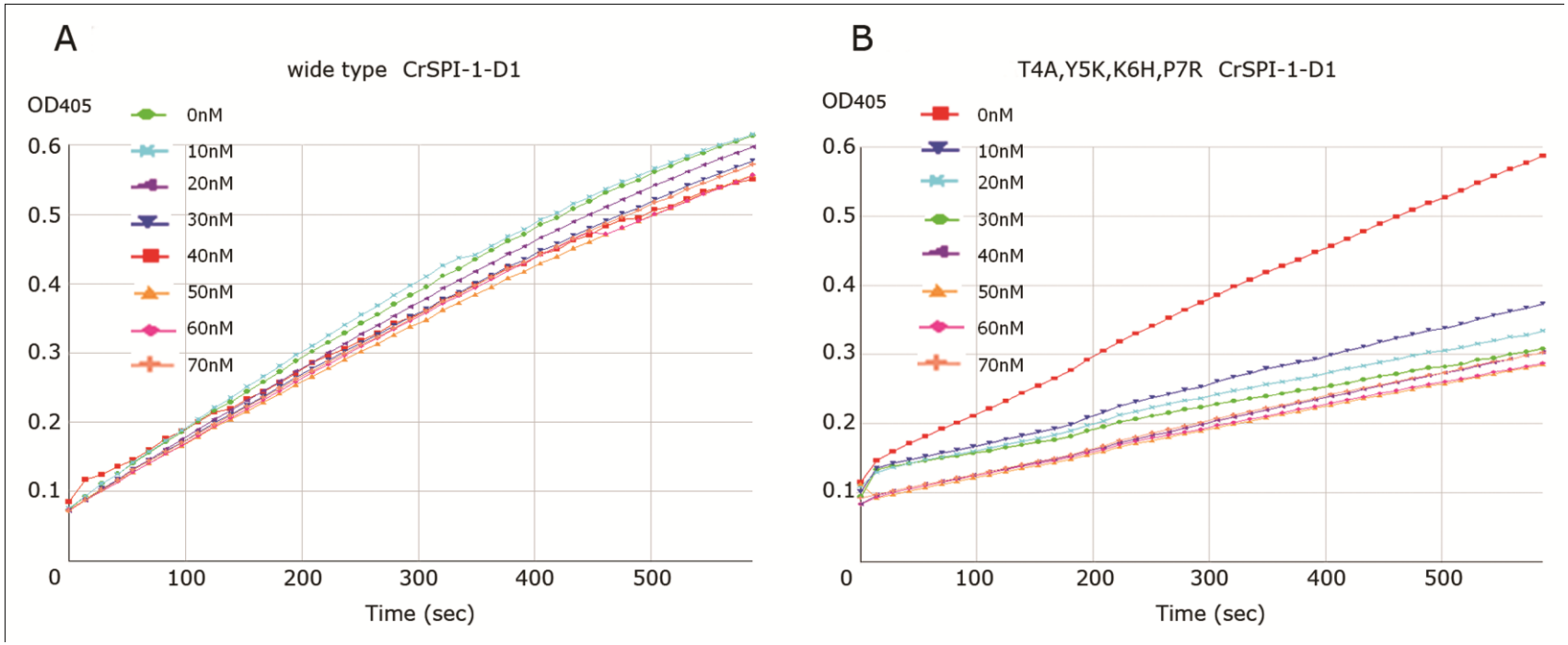


Figure IV.10: Concentration dependent Inhibition of α-human thrombin by CrSPI1-D1 and its mutant: *(A)* wild type CrSPI-1-D1 and *(B)* T4A,Y5K,K6H,P7R CrSPI-1-D1. Residual protease activity was measured as OD_{405} based on hydrolysis of 0.1 mM S2238 (H-D-Phenylalanyl-L-pipecolyl-Larginine-p-nitroaniline dihydrochloride) and formation of colored product, p-nitroaniline, by 4.5 nM thrombin in a total reaction volume of 200 µL. The concentration of wild type CrSPI-1-D1 and T4A, Y5K, K6H, P7R CrSPI-1-D1 are indicated next to the curve.

4.3.6 Isothermal Titration Calorimetry (ITC) studies

To verify the interactions between the CrSPI-1-D1 and thrombin, we have performed ITC experiments with wild type CrSPI-1-D1 and selected mutants against thrombin. The wild type CrSPI-1-D1 and the mutants which lacked thrombin inhibition did not show any binding with thrombin. Consistent with the results of thrombin inhibition assays, only the tetra mutant showed interactions with α-human thrombin with dissociation constant (Kd) of 4 μM (Fig. 4.11). The model used for the ITC analysis is a single site binding model assuming a stoichiometric ratio of 1:1 (CrSPI-1-D1: thrombin).

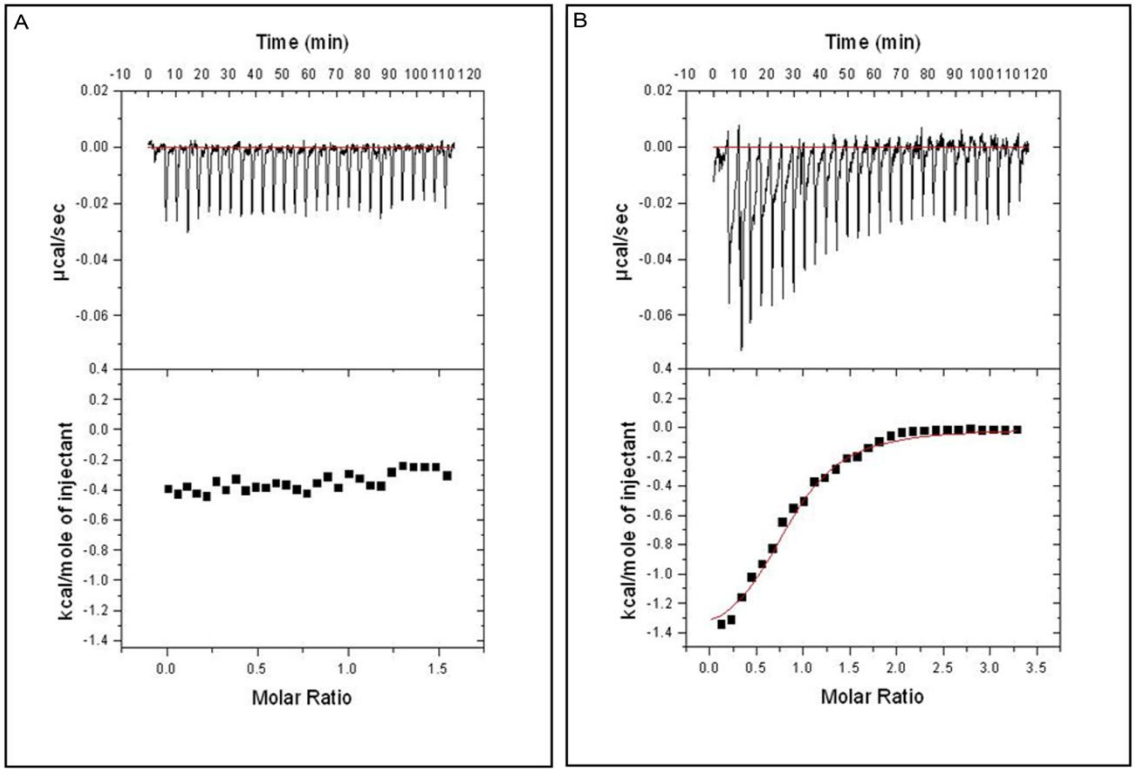


Figure IV.11: Isothermal Titration Calorimetry analysis. *(A)* α-human thrombin – wild type CrSPI-1-D1 titration. *(B)* α-human thrombin - T4A, Y5K, K6H, P7R CrSPI-1-D1 titration. The upper panels show the injection profile after baseline correction and the bottom panel shows the integration (heat release) for each injection.

4.3.7 IC_{50} and affinity

For enzyme inhibitors, the dissociation constant Kd is usually termed the inhibition constant Ki . It is important to be aware that data from enzyme inhibition is often reported not as Ki(or, equivalently, Kd), but instead as the IC50, the concentration of ligand that reduces enzyme activity by 50%. The difference between IC50 and Ki results from the fact, in an enzymatic binding assay of a competitive enzyme inhibitor, the inhibitor is not the only molecule trying to bind the enzyme's active site. It is competing with the substrate, so the concentration of ligand needed to reduce the enzymatic activity by 50% depends on the concentration of substrate and how tightly it binds the enzyme. IC_{50} is not a direct indicator of affinity although the two can be related at least for competitive agonists and antagonists by the Cheng-Prusoff equation.[4] In general, then, the IC50 is expected to be greater than Ki (Cheng-Prusoff equation). IC50 is not a direct indicator of affinity, nor is Kd/Ki a direct indicator of inhibition. However, the two can be related using the Cheng-Prusoff equation:

$$K_i = \frac{IC_{50}}{1 + \frac{[S]}{K_m}}$$

where K_i is the binding affinity of the inhibitor, IC_{50} is the functional strength of the inhibitor, [S] is substrate concentration and K_m is the concentration of substrate at which enzyme activity is at half maximal (but is frequently confused with substrate affinity for the enzyme, which it is not). Whereas the IC_{50} value for a compound may vary between experiments depending on experimental conditions, the K_i is an absolute value.

Binding of a protein with a drug-like molecule often leads to a small amount of heating or cooling of the solution; i.e., the reaction can be exothermic or endothermic. Isothermal titration calorimetry(ITC) measures binding through detection of these changes. An ITC assay tends to require a larger quantity of sample (protein and ligand) than an enzyme inhibition assay, since it does not have the same property of amplification (see above). However, ITC has the advantage of not requiring the development of a specialized substrate whose reaction can be detected; the natural heat release of the binding reaction is that that is needed. Note, however, that some binding

reactions neither take up nor release much heat, and these are not so easily monitored by ITC.

It is also worth reiterating that there are additional sources of variation among binding measurements for a given protein and ligand. These include the choice of enzyme substrate in enzyme IC50 assays, the pH, temperature, and ionic strength, and potentially the choice of cosolutes such as buffers and redox agents.

4.5 DISCUSSION

The *Carcinoscorpius rotundicauda* is an ancient invertebrate that has survived for several hundred million years, and thus termed a 'living fossil'. Being able to efficiently defend against the multitude of pathogens that thrive in its habitat and survive in this harsh environment, suggests that it possesses a very powerful innate immune defense system. Serine Protease Inhibitors (SPIs) serve important roles in immunity by inactivating and clearing the proteases from the invading pathogens, which use them as virulence factors. How did multidomain SPIs arise? The SPI domains are 'evolutionarily mobile' (Ikeo et al., 1995). In the process of evolution, domains from different families of SPIs could have been shuffled and fused in a single inhibitor, resulting in a multidomain inhibitor.

The evolutionary mechanisms of SPIs serve to increase their variety and expand their functions, thus helping to meet the demands of the repertoire of endogenous and exogenous SPs an organism encounters. Thus, knowing the structure of an inhibitor usually provides insights into its inhibitory functions. More importantly, the structural changes of a protease inhibitor in complex with its target protease can provide useful information on the interaction between the two proteins, thus allowing the development of analogs of that inhibitor with increased affinity towards the protease to achieve greater inhibition capacity. This motivated us to modify the selectivity of CrSPI-1-D1 to specifically target thrombin and here we show that selected mutation in the RSL region of CrSPI-1-D1 led to a potent and specific thrombin inhibitor.

We have determined the crystal structure of CrSPI-1-D1 refined up to 2.0Å resolution, from the horseshoe crab, *C. rotundicauda*. Although the native CrSPI-1-D1 itself is highly homologous to the thrombin inhibitor, rhodniin domain 1 (rhodniin-D1) from *R. prolixus*, native CrSPI-1-D1 does not inhibit thrombin. Therefore, our directed mutation of the RSL represents a structure-based drug design approach in the conversion of an uncharacterized CrSPI-1-D1 into a potent thrombin inhibitor with an IC_{50} of 26.3 nM. Furthermore, our studies revealed that besides the rigid conformation of the RSL, the sequence is most important to dictate the specificity of the inhibitor. This study adds an important implication to modifying a multidomain inhibitor protein. The CrSPI-1 has been shown to target two molecules of proteases. The modified domain D1 targets thrombin, whereas the wild type domain D2 targets subtilisin (Jiang et al., 2009). Moreover, this may lead to further development of the D1 mutant into a shorter active anti-thrombin inhibitor for therapeutic interventions.

.

CHAPTER-V: CONCLUSION AND FUTURE DIRECTION

5.1 CONCLUSIONS

5.1.1 Cr-TRP16 and its role in NF-kB signaling pathways

The first part of this thesis, reports the NMR structure of the reduced form of wild-type Cr-TRP16, along with functional studies. We have demonstrated the role of the C15 residue of in the dimer formation of Cr-TRP16 by site directed mutagenesis and biophysical methods such as Analytical Ultracentrifugation (AUC), NMR and gel-filtration. Moreover the *ex-vivo* study in HeLa cells for the wild type and C15S Cr-TRP16 were found consistent with the biophysical results and reveals the role of the additional N-terminal Cys residue in the regulation of NF-kB activity under oxidative stress. These studies shed light in understanding the interactions of Trxs and Trx-like proteins with other metabolic pathways and their physiological relevance and will eventually lead to developing promising therapeutic strategies to modulate the action of Trx and NF-κB.

5.1.2 Modifying the substrate specificity of a Cr inhibitor to target thrombin

The second part of this thesis reports the crystal structure of CrSPI-1-D1 from *C. rotundicauda*, refined at 2.0Å resolution. Although the structure of the native CrSPI-1-D1 itself is highly homologous to the thrombin inhibitor, rhodniin domain 1 (rhodniin-D1) from *R. prolixus*, native CrSPI-1-D1 does not inhibit thrombin. Therefore, the site directed mutations on the Reactive Site Loop (RSL) represents a structure-based drug design approach to transform uncharacterized CrSPI-1-D1 into a potent and specific inhibitor of thrombin with an IC_{50} of 26.3 nM. Moreover results presented in this part of

thesis reveal that in addition to the rigid conformation of the RSL, the sequence is most important to dictate the specificity of the inhibitor. This study adds an important implication to modify any multi-domain inhibitor protein to target any specific substrate.

5.2 FUTURE DIRECTION

Based on the experiments and results reported in this thesis continuation of these studies is proposed in the following sections. In vertebrates, there are two families of TRXs with high homology to the 16 kDa Cr-TRP16; one is 16 kDa and another one is 24 kDa (TXNL6). However only invertebrates possess the 16 kDa TRXs and bacteria are devoid of these homologues. This observation suggests that the 16 kDa TRXs have evolutionarily diverged from the 12 kDa TRXs at an early stage and the 16 kDa TRXs probably underwent gene duplication and divergence in the vertebrates and gave rise to the 24 kDa TXNL-6 (Zhang, 2003).

5.2.1 Structural insights into the mechanism of TXNL-6 / NF-κB complex in protection of human photoreceptor cells from photo oxidative damage

Interestingly, unlike the invariable WCPPC catalytic motif in Cr-TRP16, the active sites of the 24 kDa human TXNL-6 have undergone marked changes. Therefore, it will be interesting to investigate if human TXNL-6 still conserve and exhibit the basic structural and functional features like the Cr-TRP16 counterparts. Simultaneously, it would be pertinent to explore the structural and functional similarities /differences between the 16 kDa and 24 kDa TRXs.

5.2.2 Development of smaller and less immunogenic potent thrombin inhibitor

Given the central role of thromboembolism in cardiovascular disease, inhibition of thrombosis has become a key therapeutic strategy. Our structure directed protein engineering of the reactive site loop (RSL) leads to the conversion of an uncharacterized CrSPI-1-D1 into one of the smallest (40 residues) potent thrombin inhibitor. Moreover, this may lead to further development of the D1 variant into a shorter (10-12 residues) and less immunogenic active anti-thrombin inhibitor for therapeutic interventions.

BIBLIOGRAPHY

Adams, J., M. Kauffman, 2004. Development of the Proteasome Inhibitor Velcade™ (Bortezomib). Cancer Investigation 22, 304-311.

Ahn, K.S., G. Sethi, B.B. Aggarwal, 2007. Nuclear factor-kappa B: from clone to clinic. Curr Mol Med 7, 619-37.

Aikawa, R., T. Nagai, M. Tanaka, Y. Zou, T. Ishihara, H. Takano, H. Hasegawa, H. Akazawa, M. Mizukami, R. Nagai, I. Komuro, 2001. Reactive Oxygen Species in Mechanical Stress-Induced Cardiac Hypertrophy. Biochemical and Biophysical Research Communications 289, 901-907.

Armstrong, P.B., 2001. The contribution of proteinase inhibitors to immune defense. Trends Immunol 22, 47-52.

Arnér, E.S.J., A. Holmgren, 2000. Physiological functions of thioredoxin and thioredoxin reductase. European Journal of Biochemistry 267, 6102-6109.

Arnold, F.H., 1997. Design by directed evolution. FASEB Journal 11.

Atabek, M.E., H. Vatansev, I. Erkul, 2004. Oxidative Stress in Childhood Obesity. Journal of Pediatric Endocrinology and Metabolism 17, 1063-1068.

Atkinson, H.J., P.C. Babbitt, 2009. An Atlas of the Thioredoxin Fold Class Reveals the Complexity of Function-Enabling Adaptations. PLoS Comput Biol 5, e1000541.

Aviram, M., 2000. Review of human studies on oxidative damage and antioxidant protection related to cardiovascular diseases. Free Radical Research 33, S85-S97.

Barrette-Ng, I.H., K.K. Ng, M.M. Cherney, G. Pearce, U. Ghani, C.A. Ryan, M.N. James, 2003. Unbound form of tomato inhibitor-II reveals interdomain flexibility and conformational variability in the reactive site loops. J Biol Chem 278, 31391-400.

Baud, L., R. Ardaillou, 1986. Reactive oxygen species: production and role in the kidney. American Journal of Physiology - Renal Physiology 251, F765-F776.

Bax, A., S. Grzesiek, 1993. Methodological Advances in Protein Nmr. Accounts of Chemical Research 26, 131-138.

Beckman, K.B., B.N. Ames, 1997. Oxidative decay of DNA. J Biol Chem 272, 19633-6.

Bian, X., A.W. Opipari Jr, A.B. Ratanaproeksa, A.E. Boitano, P.C. Lucas, V.P. Castle, 2002. Constitutively active NFκB is required for the survival of S-type neuroblastoma. Journal of Biological Chemistry 277, 42144-42150.

Blau, S., R. Kohen, P. Bass, A. Rubinstein, 2000. Relation Between Colonic Inflammation Severity and Total Low-Molecular-Weight Antioxidant Profiles in Experimental Colitis. Digestive Diseases and Sciences 45, 1180-1187.

Bolaños, J.P., M.A. Moro, I. Lizasoain, A. Almeida, 2009. Mitochondria and reactive oxygen and nitrogen species in neurological disorders and stroke: Therapeutic implications. Advanced Drug Delivery Reviews 61, 1299-1315.

Bone, R., J.L. Silen, D.A. Agard, 1989. Structural plasticity broadens the specificity of an engineered protease. Nature 339, 191-195.

Brown, N., R. Bicknell, 2001. Hypoxia and oxidative stress in breast cancer: Oxidative stress - its effects on the growth, metastatic potential and response to therapy of breast cancer. Breast Cancer Res 3, 323 - 327.

Brown, P.H., P. Schuck, 2006. Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation. Biophys J 90, 4651-61.

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

Carvalho, A.P., P.A. Fernandes, M.J. Ramos, 2006. Similarities and differences in the thioredoxin superfamily. Prog Biophys Mol Biol 91, 229-48.

Chang, C.M., C.C. Yu, H.T. Lu, Y.F. Chou, R.F. Huang, 2007. Folate deprivation promotes mitochondrial oxidative decay: DNA large deletions, cytochrome c oxidase dysfunction, membrane depolarization and superoxide overproduction in rat liver. Br J Nutr 97, 855-63.

Chen, R., 1999. A general strategy for enzyme engineering. Trends in Biotechnology 17, 344-345.

Chen, R., A. Greer, A.M. Dean, 1995. A highly active decarboxylating dehydrogenase with rationally inverted coenzyme specificity. Proceedings of the National Academy of Sciences of the United States of America 92, 11666-11670.

Chen, R., A. Greer, A.M. Dean, 1996. Redesigning secondary structure to invert coenzyme specificity in isopropylmalate dehydrogenase. Proceedings of the National Academy of Sciences of the United States of America 93, 12171-12176.

Comporti, M., 1989. Three models of free radical-induced cell injury. Chem Biol Interact 72, 1-56.

Conley, G.P., M. Viswanathan, Y. Hou, D.L. Rank, A.P. Lindberg, S.M. Cramer, R.C. Ladner, A.E. Nixon, J. Chen, 2011. Evaluation of protein engineering and process optimization approaches to enhance antibody drug manufacturability. Biotechnology and Bioengineering 108, 2634-2644.

Cornilescu, G., F. Delaglio, A. Bax, 1999a. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. Journal of Biomolecular NMR 13, 289-302.

Cornilescu, G., F. Delaglio, A. Bax, 1999b. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR 13, 289-302.

Creaser, E.H., C. Murali, K.A. Britt, 1990. Protein engineering of alcohol dehydrogenases; effects of amino acid changes at positions 93 and 48 of yeast ADH1. Protein Engineering 3, 523-526.

Davi, G., A. Falco, C. Patrono, 2005. Lipid peroxidation in diabetes mellitus. Antioxid Redox Signal 7, 256-68.

Davie, E.W., K. Fujikawa, W. Kisiel, 1991. The coagulation cascade: initiation, maintenance, and regulation. Biochemistry 30, 10363-10370.

Davies, K.J., S.W. Lin, R.E. Pacifici, 1987. Protein damage and degradation by oxygen radicals. IV. Degradation of denatured protein. J Biol Chem 262, 9914-20.

Delaglio, F., S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6, 277-93.

DeLano, W.L., J.W. Lam, 2005. PyMOL: A communications tool for computational models. Abstracts of Papers of the American Chemical Society 230, U1371-U1372.

Di Virgilio, F., 2004. New pathways for reactive oxygen species generation in inflammation and potential novel pharmacological targets. Current Pharmaceutical Design 10, 1647-1652.

Ding, J.L., M.A. Navas, 3rd, B. Ho, 1993. Two forms of factor C from the amoebocytes of Carcinoscorpius rotundicauda: purification and characterisation. Biochim Biophys Acta 1202, 149-56.

Donovan, M., T.G. Cotter, 2002. Caspase-independent photoreceptor apoptosis in vivo and diffrential expression of apoptotic protease activating factor-1 and caspase-3 during retinal development. Cell Death and Differentiation 9, 1220-1231.

Drag, M., G.S. Salvesen, 2010. Emerging principles in protease-based drug discovery. Nat Rev Drug Discov 9, 690-701.

Dugan, L.L., K.L. Quick, 2005. Reactive Oxygen Species and Aging: Evolving Questions. Sci. Aging Knowl. Environ. 2005, pe20-.

Eaton, J.W., M. Qian, 2002. Molecular bases of cellular iron toxicity. Free Radical Biology and Medicine 32, 833-840.

Eklund, H., F.K. Gleason, A. Holmgren, 1991a. Structural and functional relations among thioredoxins of different species. Proteins: Structure, Function, and Genetics 11, 13-28.

Eklund, H., F.K. Gleason, A. Holmgren, 1991b. Structural and functional relations among thioredoxins of different species. Proteins: Structure, Function, and Bioinformatics 11, 13-28.

Elias S.J, A., 2009. Focus on mammalian thioredoxin reductases — Important selenoproteins with versatile functions. Biochimica et Biophysica Acta (BBA) - General Subjects 1790, 495-526.

Emsley, P., K. Cowtan, 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60, 2126-32.

Farré, A.L., S. Casado, 2001. Heart Failure, Redox Alterations, and Endothelial Dysfunction. Hypertension 38, 1400-1405.

Fesik, S.W., H.L. Eaton, E.T. Olejniczak, E.R.P. Zuiderweg, L.P. McIntosh, F.W. Dahlquist, 1990. 2D and 3D NMR spectroscopy employing carbon-13/carbon-13 magnetization transfer by isotropic mixing. Spin system identification in large proteins. Journal of the American Chemical Society 112, 886-888.

Flohé, L., R. Brigelius-Flohé, C. Saliou, M.G. Traber, L. Packer, 1997. Redox Regulation of NF-kappa B Activation. Free Radical Biology and Medicine 22, 1115-1126.

Fridlich, R., F. Delalande, C. Jaillard, J. Lu, L. Poidevin, T. Cronin, L. Perrocheau, G. Millet-Puel, M.-L. Niepon, O. Poch, A. Holmgren, A. Van Dorsselaer, J.-A. Sahel, T. Léveillard, 2009. The Thioredoxin-like Protein Rod-derived Cone Viability Factor (RdCVFL) Interacts with TAU and Inhibits Its Phosphorylation in the Retina. Molecular & Cellular Proteomics 8, 1206-1218.

Fujii, S., Y. Nanbu, H. Nonogaki, I. Konishi, T. Mori, H. Masutani, J. Yodoi, 1991. Coexpression of adult T-cell leukemia-derived factor, a human thioredoxin homologue, and human papillomavirus DNA in neoplastic cervical squamous epithelium. Cancer 68, 1583-91.

Gasdaska, P.Y., J.E. Oblong, I.A. Cotgreave, G. Powis, 1994. The predicted amino acid sequence of human thioredoxin is identical to that of the autocrine growth factor

human adult T-cell derived factor (ADF): thioredoxin mRNA is elevated in some human tumors. Biochim Biophys Acta 1218, 292-6.
Gelderman, K.A., M. Hultqvist, L.M. Olsson, K. Bauer, A. Pizzolla, P. Olofsson, R. Holmdahl, 2007. Rheumatoid arthritis: The role of reactive oxygen species in disease development and therapeutic strategies. Antioxidants and Redox Signaling 9, 1541-1567.
Georgiou, G., N. DeWitt, 1999. Enzyme beauty. Nat Biotech 17, 1161-1162.
Gettins, P.G., 2002. Serpin structure, mechanism, and function. Chem Rev 102, 4751-804.
Ghosh, G., G.V. Duyne, S. Ghosh, P.B. Sigler, 1995. Structure of NF-[kappa]B p50 homodimer bound to a [kappa]B site. Nature 373, 303-310.
Giasson, B.I., H. Ischiropoulos, V.M.Y. Lee, J.Q. Trojanowski, 2002. The relationship between oxidative/nitrative stress and pathological inclusions in Alzheimer's and Parkinson's diseases. Free Radical Biology and Medicine 32, 1264-1275.
Gohlke, H., M. Hendlich, G. Klebe, 2000. Knowledge-based scoring function to predict protein-ligand interactions. Journal of Molecular Biology 295, 337-356.
Goldschmidt, L., D.R. Cooper, Z.S. Derewenda, D. Eisenberg, 2007. Toward rational protein crystallization: A Web server for the design of crystallizable protein variants. Protein Science 16, 1569-1576.
Goldstein, I.M., P. Ostwald, S. Roth, 1996. Nitric Oxide: A Review of Its Role in Retinal Function and Disease. Vision Research 36, 2979-2994.
Gómez-Vicente, V., M. Donovan, T.G. Cotter, 2005. Multiple death pathways in retina-derived 661W cells following growth factor deprivation: Crosstalk between caspases and calpains. Cell Death and Differentiation 12, 796-804.
Görler, A., W. Gronwald, K.-P. Neidig, H.R. Kalbitzer, 1999. Computer Assisted Assignment of13C or15N Edited 3D-NOESY-HSQC Spectra Using Back Calculated and Experimental Spectra. Journal of Magnetic Resonance 137, 39-45.
Gouet, P., E. Courcelle, D.I. Stuart, F. Metoz, 1999. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 15, 305-8.
Gracy, R.W., J.M. Talent, Y. Kong, C.C. Conrad, 1999. Reactive oxygen species: the unavoidable environmental insult? Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis 428, 17-22.
Gurm, H.S., D.L. Bhatt, 2005. Thrombin, an ideal target for pharmacological inhibition: a review of direct thrombin inhibitors. Am Heart J 149, S43-53.
Haase, M., M. Anlauf, M. Schott, S. Schinner, E. Kaminsky, W.A. Scherbaum, H.S. Willenberg, 2011. Erratum to: A new mutation in the menin gene causes the multiple endocrine neoplasia type 1 syndrome with adrenocortical carcinoma. Endocrine.
Halliwell, B., 1994. Free radicals, antioxidants, and human disease: curiosity, cause, or consequence? The Lancet 344, 721-724.
Haurani, M.J., P.J. Pagano, 2007. Adventitial fibroblast reactive oxygen species as autacrine and paracrine mediators of remodeling: Bellwether for vascular disease? Cardiovascular Research 75, 679-689.
Hemmi, H., T. Kumazaki, K. Yoshizawa-Kumagaye, Y. Nishiuchi, T. Yoshida, T. Ohkubo, Y. Kobayashi, 2005. Structural and functional study of an Anemonia

elastase inhibitor, a "nonclassical" Kazal-type inhibitor from Anemonia sulcata. Biochemistry 44, 9626-36.

Hensley, K., D.A. Butterfieldld, N. Hall, P. Cole, R. Subramaniam, R. Mark, M.P. Mattson, W.R. Markesbery, M.E. Harris, M. Aksenov, M. Aksenova, J.F. Wu, J.M. Carney, 1996. Reactive Oxygen Species as Causal Agents in the Neurotoxicity of the Alzheimer's Disease-Associated Amyloid Beta Peptidea. Annals of the New York Academy of Sciences 786, 120-134.

Hitchon, C.A., H.S. El-Gabalawy, 2004. Oxidation in rheumatoid arthritis. Arthritis Res Ther 6, 265-78.

Hoebe, K., E. Janssen, B. Beutler, 2004. The interface between innate and adaptive immunity. Nat Immunol 5, 971-4.

Hoffmann, J.A., F.C. Kafatos, C.A. Janeway, R.A. Ezekowitz, 1999. Phylogenetic perspectives in innate immunity. Science 284, 1313-8.

Hofmann, B., H.J. Hecht, L. Flohé, 2002. Peroxiredoxins. Biological Chemistry 383, 347-364.

Hofmann, B., H. Budde, K. Bruns, S.A. Guerrero, H.M. Kalisz, U. Menge, M. Montemartini, E. Nogoceke, P. Steinert, J.B. Wissing, L. Flohe, H.J. Hecht, 2001. Structures of tryparedoxins revealing interaction with trypanothione. Biol Chem 382, 459-71.

Holm, L., C. Sander, 1998. Touring protein fold space with Dali/FSSP. Nucleic Acids Res 26, 316-9.

Holmgren, A., 1989a. Thioredoxin and glutaredoxin systems. Journal of Biological Chemistry 264, 13963-13966.

Holmgren, A., 1989b. Thioredoxin and glutaredoxin systems. J Biol Chem 264, 13963-6.

Holmgren, A., 1995. Thioredoxin structure and mechanism: conformational changes on oxidation of the active-site sulfhydryls to a disulfide. Structure 3, 239-243.

Holmgren, A., J. Lu, 2010. Thioredoxin and thioredoxin reductase: Current research with special reference to human disease. Biochemical and Biophysical Research Communications 396, 120-124.

Huang, J.W., Y.S. Cheng, T.P. Ko, C.Y. Lin, H.L. Lai, C.C. Chen, Y. Ma, Y. Zheng, C.H. Huang, P. Zou, J.R. Liu, R.T. Guo, 2011. Rational design to improve thermostability and specific activity of the truncated Fibrobacter succinogenes 1,3-1,4-β-d-glucanase. Applied Microbiology and Biotechnology, 1-11.

Huang, X., Y. Yin, J. Liu, 2010. Design of Artificial Selenoenzymes Based on Macromolecular Scaffolds. Macromolecular Bioscience 10, 1385-1396.

Ikeo, K., K. Takahashi, T. Gojobori, 1995. Different evolutionary histories of kringle and protease domains in serine proteases: a typical example of domain evolution. J Mol Evol 40, 331-6.

Jiang, N., S. Thangamani, C.F. Chor, S.Y. Wang, I. Winarsih, R.J. Du, J. Sivaraman, B. Ho, J.L. Ding, 2009. A novel serine protease inhibitor acts as an immunomodulatory switch while maintaining homeostasis. J Innate Immun 1, 465-79.

Jing, H., K.J. Macon, D. Moore, L.J. DeLucas, J.E. Volanakis, S.V.L. Narayana, 1999. Structural basis of profactor D activation: from a highly flexible zymogen to a novel self-inhibited serine protease, complement factor D. EMBO J 18, 804-814.

Kamp, D.W., P. Graceffa, W.A. Pryor, S.A. Weitzman, 1992. The role of free radicals in asbestos-induced diseases. Free Radical Biology and Medicine 12, 293-315.
Kanost, M.R., 1999. Serine proteinase inhibitors in arthropod immunity. Dev Comp Immunol 23, 291-301.
Kawanishi, S., Y. Hiraku, M. Murata, S. Oikawa, 2002. The role of metals in site-specific DNA damage with reference to carcinogenesis. Free Radical Biology and Medicine 32, 822-832.
Kazlauskas, R.J., U.T. Bornscheuer, 2009. Finding better protein engineering strategies. Nat Chem Biol 5, 526-529.
Kneller, D.G., I.D. Kuntz, 1993. Ucsf Sparky - an Nmr Display, Annotation and Assignment Tool. Journal of Cellular Biochemistry, 254-254.
Koradi, R., M. Billeter, K. Wüthrich, 1996. MOLMOL: A program for display and analysis of macromolecular structures. Journal of Molecular Graphics 14, 51-55.
Kraulis, P., 1991. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. Journal of Applied Crystallography 24, 946-950.
Kumar, J.K., S. Tabor, C.C. Richardson, 2004. Proteomic analysis of thioredoxin-targeted proteins in Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America 101, 3759-3764.
Kunchithapautham, K., B. Padmavathi, R.B. Narayanan, P. Kaliraj, A.L. Scott, 2003. Thioredoxin from Brugia malayi: Defining a 16-Kilodalton Class of Thioredoxins from Nematodes. Infect. Immun. 71, 4119-4126.
Kuwada, N.J., M.J. Zuckermann, E.H.C. Bromley, R.B. Sessions, P.M.G. Curmi, N.R. Forde, D.N. Woolfson, H. Linke, 2011. Tuning the performance of an artificial protein motor. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics 84.
Laroux, F.S., K.P. Pavlick, I.N. Hines, S. Kawachi, H. Harada, S. Bharwani, J.M. Hoffman, M.B. Grisham, 2001. Role of nitric oxide in inflammation. Acta Physiologica Scandinavica 173, 113-118.
Laskowski, R.A., J.A. Rullmannn, M.W. MacArthur, R. Kaptein, J.M. Thornton, 1996. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 8, 477-86.
Latha, B., M. Babu, 2001. The involvement of free radicals in burn injury: a review. Burns : journal of the International Society for Burn Injuries 27, 309-17.
Lau, A.T.Y., Y. Wang, J.-F. Chiu, 2008. Reactive oxygen species: Current knowledge and applications in cancer research and therapeutic. Journal of Cellular Biochemistry 104, 657-667.
Laurent, T.C., E.C. Moore, P. Reichard, 1964a. Enzymatic Synthesis of Deoxyribonucleotides. Journal of Biological Chemistry 239, 3436-3444.
Laurent, T.C., E.C. Moore, P. Reichard, 1964b. Enzymatic Synthesis of Deoxyribonucleotides .4. Isolation + Characterization of Thioredoxin Hydrogen Donor from Escherichia Coli B. Journal of Biological Chemistry 239, 3436-&.
Leblond, J., M.H. Laprise, S. Gaudreau, F. Grondin, W. Kisiel, C.M. Dubois, 2006. The serpin proteinase inhibitor 8: an endogenous furin inhibitor released from human platelets. Thromb Haemost 95, 243-52.
Lee, H.-C., Y.-H. Wei, 2001. Mitochondrial alterations, cellular response to oxidative stress and defective degradation of proteins in aging. Biogerontology 2, 231-244.

Léveillard, T., S. Mohand-Saïd, O. Lorentz, D. Hicks, A.C. Fintz, E. Clérin, M. Simonutti, V. Forster, N. Cavusoglu, F. Chalmel, P. Dollé, O. Poch, G. Lambrou, J.A. Sahel, 2004. Identification and characterization of rod-derived cone viability factor. Nature Genetics 36, 755-759.

Levine, R.L., E.R. Stadtman, 2001. Oxidative modification of proteins during aging. Experimental Gerontology 36, 1495-1502.

Levy, J.H., 1996. The Human Inflammatory Response. Journal of Cardiovascular Pharmacology 27, 31-37.

Levy, J.H., P.S. O'Donnell, 2006. The therapeutic potential of a kallikrein inhibitor for treating hereditary angioedema. Expert Opinion on Investigational Drugs 15, 1077-1090.

Lillig, C.H., A. Holmgren, 2006. Thioredoxin and Related Molecules–From Biology to Health and Disease. Antioxidants & Redox Signaling 9, 25-47.

Lippow, S.M., T.S. Moon, S. Basu, S.H. Yoon, X. Li, B.A. Chapman, K. Robison, D. Lipovšek, K.L.J. Prather, 2010. Engineering enzyme specificity using computational design of a defined-sequence library. Chemistry and Biology 17, 1306-1315.

Liu, H., H. Nishitoh, H. Ichijo, J.M. Kyriakis, 2000. Activation of apoptosis signal-regulating kinase 1 (ASK1) by tumor necrosis factor receptor-associated factor 2 requires prior dissociation of the ASK1 inhibitor thioredoxin. Molecular and Cellular Biology 20, 2198-2208.

Lombardi, A., G. De Simone, S. Galdiero, N. Staiano, F. Nastri, V. Pavone, 1999. From natural to synthetic multisite thrombin inhibitors. Biopolymers 51, 19-39.

Loshchagin, O.V., R.I. Kovalenko, A.D. Nozdrachev, I.N. Yanvareva, B.I. Krivoruchko, 2002. Possible Role of Catalase in Adaptation to Diving of Semi-Aquatic Rodents <i>Ondatra zibethica</i&gt. Journal of Evolutionary Biochemistry and Physiology 38, 90-95.

Maggirwar, S.B., P.D. Sarmiere, S. Dewhurst, R.S. Freeman, 1998. Nerve growth factor-dependent activation of NF-κB contributes to survival of sympathetic neurons. Journal of Neuroscience 18, 10356-10365.

Markwardt, F., 1994. The development of hirudin as an antithrombotic drug. Thromb Res 74, 1-23.

Markwardt, F., 2002. Historical perspective of the development of thrombin inhibitors. Pathophysiol Haemost Thromb 32 Suppl 3, 15-22.

Marnett, L.J., 2000. Oxyradicals and DNA damage. Carcinogenesis 21, 361-70.

Marshall, S.A., G.A. Lazar, A.J. Chirino, J.R. Desjarlais, 2003. Rational design and engineering of therapeutic proteins. Drug Discovery Today 8, 212-221.

Martin, J.L., 1995a. THIOREDOXIN - A FOLD FOR ALL REASONS. Structure 3, 245-250.

Martin, J.L., 1995b. Thioredoxin a fold for all reasons. Structure (London, England : 1993) 3, 245-250.

Martindale, J.L., N.J. Holbrook, 2002. Cellular response to oxidative stress: Signaling for suicide and survival*. Journal of Cellular Physiology 192, 1-15.

Matheson, A.J., K.L. Goa, 2000. Desirudin: A Review of its Use in the Management of Thrombotic Disorders. Drugs 60, 679-700.

Matsuo, Y., K. Hirota, H. Nakamura, J. Yodoi, 2002. Redox regulation by thioredoxin and its related molecules. Drug News & Perspectives 15, 575-580.
Mattos, C., C.R. Bellamacina, E. Peisach, A. Pereira, D. Vitkup, G.A. Petsko, D. Ringe, 2006. Multiple Solvent Crystal Structures: Probing Binding Sites, Plasticity and Hydration. Journal of Molecular Biology 357, 1471-1482.
Maynes, J.T., M.M. Cherney, M.A. Qasim, M. Laskowski, Jr., M.N. James, 2005. Structure of the subtilisin Carlsberg-OMTKY3 complex reveals two different ovomucoid conformations. Acta Crystallogr D Biol Crystallogr 61, 580-8.
McCoy, A.J., R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, 2007. Phaser crystallographic software. J Appl Crystallogr 40, 658-674.
Menon, S.D., G.R. Guy, Y.H. Tan, 1995. Involvement of a putative protein-tyrosine phosphatase and I kappa B-alpha serine phosphorylation in nuclear factor kappa B activation by tumor necrosis factor. J Biol Chem 270, 18881-7.
Meral, A., P. Tuncel, uuml, G. rmen, E. r, Ouml, R. zbek, zt, E. rk, nay, 2000. LIPID PEROXIDATION AND ANTIOXIDANT STATUS IN beta-THALASSEMIA. Pediatric Hematology and Oncology 17, 687-693.
Montemartini, M., P. Steinert, M. Singh, L. Flohe, H.M. Kalisz, 2000. Tryparedoxin II from Crithidia fasciculata. Biofactors 11, 65-6.
Moulton, P.J., 1996. Inflammatory joint disease: The role of cytokines, cyclooxygenases and reactive oxygen species. British Journal of Biomedical Science 53, 317-324.
Muhammad, S., A. Bierhaus, M. Schwaninger, 2009. Reactive oxygen species in diabetes-induced vascular damage, stroke, and Alzheimer's disease. Journal of Alzheimer's Disease 16, 775-785.
Mukhopadhyay, A., C. Bueso-Ramos, D. Chatterjee, P. Pantazis, B.B. Aggarwal, 2001. Curcumin downregulates cell survival mechanisms in human prostate cancer cell lines. Oncogene 20, 7597-609.
Muller, C.W., F.A. Rey, M. Sodeoka, G.L. Verdine, S.C. Harrison, 1995. Structure of the NF-kappa B p50 homodimer bound to DNA. Nature 373, 311-7.
Multhaup, G., T. Ruppert, A. Schlicksupp, L. Hesse, D. Beher, C.L. Masters, K. Beyreuther, 1997. Reactive oxygen species and Alzheimer's disease. Biochemical Pharmacology 54, 533-539.
Muta, T., S. Iwanaga, 1996. The role of hemolymph coagulation in innate immunity. Curr Opin Immunol 8, 41-7.
Nakano, H., A. Nakajima, S. Sakon-Komazawa, J.H. Piao, X. Xue, K. Okumura, 2005. Reactive oxygen species mediate crosstalk between NF-[kappa]B and JNK. Cell Death Differ 13, 730-737.
Nisha, K., R.K. Deshwal, 2011. Antioxidants and their protective action against DNA damage. International Journal of Pharmacy and Pharmaceutical Sciences 3, 28-32.
Nordberg, J., E.S. Arner, 2001. Reactive oxygen species, antioxidants, and the mammalian thioredoxin system. Free Radic Biol Med 31, 1287-312.
Nordberg, J., E.S.J. Arnér, 2001. Reactive oxygen species, antioxidants, and the mammalian thioredoxin system. Free Radical Biology and Medicine 31, 1287-1312.
Novo, E., M. Parola, 2008. Redox mechanisms in hepatic chronic wound healing and fibrogenesis. Fibrogenesis & Tissue Repair 1, 5.

Nunomura, A., R.J. Castellani, X. Zhu, P.I. Moreira, G. Perry, M.A. Smith, 2006. Involvement of oxidative stress in Alzheimer disease. J Neuropathol Exp Neurol 65, 631-41.

Nyska, A., A. Dayan, R.R. Maronpot, 2002. New Tools in Therapeutic Research—Prostatic Cancer and Models. Toxicologic Pathology 30, 283-287.

Ochman, H., A.S. Gerber, D.L. Hartl, 1988. Genetic applications of an inverse polymerase chain reaction. Genetics 120, 621-3.

Opara, E.C., 2002. Oxidative stress, micronutrients, diabetes mellitus and its complications. The Journal of the Royal Society for the Promotion of Health 122, 28-34.

Otwinowski, Z., W. Minor, 1997. Processing of X-ray diffraction data collected in oscillation mode. Macromolecular Crystallography, Pt A 276, 307-326.

Pan, S.Y., S. Pan, Z.L. Yu, D.L. Ma, S.B. Chen, W.F. Fong, Y.F. Han, K.M. Ko, 2010. New perspectives on innovative drug discovery: An overview. Journal of Pharmacy and Pharmaceutical Sciences 13, 450-471.

Pekkari, K., J. Avila-Cariño, Å. Bengtsson, R. Gurunath, A. Scheynius, A. Holmgren, 2001. Truncated thioredoxin (Trx80) induces production of interleukin-12 and enhances CD14 expression in human monocytes. Blood 97, 3184-3190.

Pfau, R., 2003. Structure-based design of thrombin inhibitors. Curr Opin Drug Discov Devel 6, 437-50.

Pleiss, J., 2011. Protein design in metabolic engineering and synthetic biology. Current Opinion in Biotechnology 22, 611-617.

Powis, G., W.R. Montfort, 2001a. Properties and biological activities of thioredoxins. Annu Rev Pharmacol Toxicol 41, 261-95.

Powis, G., W.R. Montfort, 2001b. Properties and biological activities of thioredoxins. Annu Rev Biophys Biomol Struct 30, 421-55.

Qin, J., G.M. Clore, W.M.P. Kennedy, J.R. Huth, A.M. Gronenborn, 1995. Solution structure of human thioredoxin in a mixed disulfide intermediate complex with its target peptide from the transcription factor NFºB. Structure (London, England : 1993) 3, 289-297.

Quinn, M.T., S.D. Swain, T.T. Rohn, 2002. Neutrophil priming in host defense: Role of oxidants as priming agents. Antioxidants & Redox Signaling 4, 69-83.

Ren, G., J.M. Webster, Z. Liu, R. Zhang, Z. Miao, H. Liu, S.S. Gambhir, F.A. Syud, Z. Cheng, 2011. In vivo targeting of HER2-positive tumor using 2-helix affibody molecules. Amino Acids, 1-9.

Ren, X., M. Bjornstedt, B. Shen, M.L. Ericson, A. Holmgren, 1993. Mutagenesis of structural half-cystine residues in human thioredoxin and effects on the regulation of activity by selenodiglutathione. Biochemistry 32, 9701-8.

Renard, P., I. Ernest, A. Houbion, M. Art, H. Le Calvez, M. Raes, J. Remacle, 2001. Development of a sensitive multi-well colorimetric assay for active NFkappaB. Nucleic Acids Res 29, E21.

Reuter, S., S.C. Gupta, M.M. Chaturvedi, B.B. Aggarwal, 2010. Oxidative stress, inflammation, and cancer: How are they linked? Free Radical Biology and Medicine 49, 1603-1616.

Richer, M.J., C.A. Keays, J. Waterhouse, J. Minhas, C. Hashimoto, F. Jean, 2004. The Spn4 gene of Drosophila encodes a potent furin-directed secretory pathway serpin. Proc Natl Acad Sci U S A 101, 10560-5.

Rotilio, G., M.T. Carr, L. Rossi, M.R. Ciriolo, 2000. Copper-Dependent Oxidative Stress and Neurodegeneration. IUBMB Life 50, 309-314.

Rydel, T.J., A. Tulinsky, W. Bode, R. Huber, 1991. Refined structure of the hirudin-thrombin complex. J Mol Biol 221, 583-601.

Saitoh, M., H. Nishitoh, M. Fujii, K. Takeda, K. Tobiume, Y. Sawada, M. Kawabata, K. Miyazono, H. Ichijo, 1998. Mammalian thioredoxin is a direct inhibitor of apoptosis signal-regulating kinase (ASK) 1. EMBO J 17, 2596-2606.

Salzet, M., 2001. Vertebrate innate immunity resembles a mosaic of invertebrate immune responses. Trends Immunol 22, 285-8.

Sanvicens, N., V. Gómez-Vicente, I. Masip, A. Messeguer, T.G. Cotter, 2004. Oxidative stress-induced apoptosis in retinal photoreceptor cells is mediated by calpains and caspases and blocked by the oxygen radical scavenger CR-6. Journal of Biological Chemistry 279, 39268-39278.

Sasaki, M., T. Joh, 2007. Oxidative Stress in Digestive DiseaseGuest Editor: Yuji NaitoOxidative Stress and Ischemia-Reperfusion Injury in Gastrointestinal Tract and Antioxidant, Protective Agents. Journal of Clinical Biochemistry and Nutrition 40, 1-12.

Schenk, H., M. Klein, W. Erdbrugger, W. Droge, K. Schulze-Osthoff, 1994. Distinct effects of thioredoxin and antioxidants on the activation of transcription factors NF-kappa B and AP-1. Proc Natl Acad Sci U S A 91, 1672-6.

Schwieters, C.D., J.J. Kuszewski, N. Tjandra, G.M. Clore, 2003. The Xplor-NIH NMR molecular structure determination package. J Magn Reson 160, 65-73.

Seki, S., T. Kitada, T. Yamada, H. Sakaguchi, K. Nakatani, K. Wakasa, 2002. In situ detection of lipid peroxidation and oxidative DNA damage in non-alcoholic fatty liver diseases. Journal of Hepatology 37, 56-62.

Sen, C.K., L. Packer, 1996. Antioxidant and redox regulation of gene transcription. FASEB J 10, 709-20.

Sethi, G., V. Tergaonkar, 2009. Potential pharmacological control of the NF-kappaB pathway. Trends Pharmacol Sci 30, 313-21.

Shabb, J.B., L. Ng, J.D. Corbin, 1990. One amino acid change produces a high affinity cGMP-binding site in cAMP-dependent protein kinase. Journal of Biological Chemistry 265, 16031-4.

Stadtman, E., 1992. Protein oxidation and aging. Science 257, 1220-1224.

Steinmetzer, T., J. Hauptmann, J. Sturzebecher, 2001. Advances in the development of thrombin inhibitors. Expert Opin Investig Drugs 10, 845-64.

Stemmer, W.P.C., 1994. DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution. Proceedings of the National Academy of Sciences of the United States of America 91, 10747-10751.

Stoop, A.A., C.S. Craik, 2003. Engineering of a macromolecular scaffold to develop specific protease inhibitors. Nat Biotech 21, 1063-1068.

Stubbs, M.T., R. Morenweiser, J. Sturzebecher, M. Bauer, W. Bode, R. Huber, G.P. Piechottka, G. Matschiner, C.P. Sommerhoff, H. Fritz, E.A. Auerswald, 1997. The three-dimensional structure of recombinant leech-derived tryptase inhibitor in

complex with trypsin. Implications for the structure of human mast cell tryptase and its inhibition. J Biol Chem 272, 19931-7.

Tabner, B.J., S. Turnbull, O. El-Agnaf, D. Allsop, 2001. Production of reactive oxygen species from aggregating proteins implicated in Alzheimer's disease, Parkinson's disease and other neurodegenerative diseases. Current topics in medicinal chemistry 1, 507-517.

Touyz, R.M., 2004. Reactive oxygen species and angiotensin II signaling in vascular cells: implications in cardiovascular disease. Brazilian Journal of Medical and Biological Research 37, 1263-1273.

Turk, B., 2006a. Targeting proteases: successes, failures and future prospects. Nat Rev Drug Discov 5, 785-99.

Turk, B., 2006b. Targeting proteases: successes, failures and future prospects. Nat Rev Drug Discov 5, 785-799.

Vacca, J.P., 2000. New advances in the discovery of thrombin and factor Xa inhibitors. Curr Opin Chem Biol 4, 394-400.

Valko, M., D. Leibfritz, J. Moncol, M.T. Cronin, M. Mazur, J. Telser, 2007. Free radicals and antioxidants in normal physiological functions and human disease. Int J Biochem Cell Biol 39, 44-84.

van de Locht, A., D. Lamba, M. Bauer, R. Huber, T. Friedrich, B. Kroger, W. Hoffken, W. Bode, 1995. Two heads are better than one: crystal structure of the insect derived double domain Kazal inhibitor rhodniin in complex with thrombin. EMBO J 14, 5149-57.

Venditti, P., I.R. Costagliola, S.D. Meo, 2002. H₂O₂ Production and Response to Stress Conditions by Mitochondrial Fractions from Rat Liver. Journal of Bioenergetics and Biomembranes 34, 115-125.

Verma, I.M., 2004. Nuclear factor (NF)-kappaB proteins: therapeutic targets. Ann Rheum Dis 63 Suppl 2, ii57-ii61.

Voigt, C.A., S.L. Mayo, F.H. Arnold, Z.G. Wang, 2001. Computationally focusing the directed evolution of proteins. J Cell Biochem Suppl Suppl 37, 58-63.

Voigt, C.A., C. Martinez, Z.-G. Wang, S.L. Mayo, F.H. Arnold, 2002. Protein building blocks preserved by recombination. Nat Struct Mol Biol 9, 553-558.

Vouilleme, L., P.R. Cushing, R. Volkmer, D.R. Madden, P. Boisguerin, 2010. Engineering peptide inhibitors to overcome PDZ binding promiscuity. Angewandte Chemie - International Edition 49, 9912-9916.

Wang, X.W., Y.C. Liou, B. Ho, J.L. Ding, 2007. An evolutionarily conserved 16-kDa thioredoxin-related protein is an antioxidant which regulates the NF-kappaB signaling pathway. Free Radic Biol Med 42, 247-59.

Wang, X.W., B.Z. Tan, M. Sun, B. Ho, J.L. Ding, 2008. Thioredoxin-like 6 protects retinal cell line from photooxidative damage by upregulating NF-κB activity. Free Radical Biology and Medicine 45, 336-344.

Watzke, A., G. Kosec, M. Kindermann, V. Jeske, H.-P. Nestler, V. Turk, B. Turk, K.U. Wendt, 2008. Selective Activity-Based Probes for Cysteine Cathepsins. Angewandte Chemie International Edition 47, 406-409.

Weichsel, A., J.R. Gasdaska, G. Powis, W.R. Montfort, 1996a. Crystal structures of reduced, oxidized, and mutated human thioredoxins: evidence for a regulatory homodimer. Structure 4, 735-51.

Weichsel, A., J.R. Gasdaska, G. Powis, W.R. Montfort, 1996b. Crystal structures of reduced, oxidized, and mutated human thioredoxins: evidence for a regulatory homodimer. Structure (London, England : 1993) 4, 735-751.

Weitz, J.I., M. Crowther, 2002. Direct thrombin inhibitors. Thromb Res 106, V275-84.

Wilks, H.M., D.J. Halsall, T. Atkinson, W.N. Chia, A.R. Clarke, J.J. Holbrook, 1990. Designs for a broad substrate specificity keto acid dehydrogenase. Biochemistry 29, 8587-8591.

Wilson, J.N., J.D. Pierce, R.L. Clancy, 2001. Reactive oxygen species in acute respiratory distress syndrome. Heart and Lung: Journal of Acute and Critical Care 30, 370-375.

Wood-Kaczmar, A., S. Gandhi, N.W. Wood, 2006. Understanding the molecular causes of Parkinson's disease. Trends Mol Med 12, 521-8.

Yoshioka, J., P.C. Schulze, M. Cupesi, J.D. Sylvan, C. MacGillivray, J. Gannon, H. Huang, R.T. Lee, 2004. Thioredoxin-Interacting Protein Controls Cardiac Hypertrophy Through Regulation of Thioredoxin Activity. Circulation 109, 2581-2586.

Zabel, U., P.A. Baeuerle, 1990. Purified human I[kappa]B can rapidly dissociate the complex of the NF-[kappa]B transcription factor with its cognate DNA. Cell 61, 255-265.

Zhang, J., 2003. Evolution by gene duplication: an update. Trends in ecology & evolution (Personal edition) 18, 292-298.