Structural and functional characterization of protein(s) involved in mycobacterial DNA repair

Thesis submitted to

JAWAHARLAL NEHRU UNIVERSITY

New Delhi, India

In partial fulfillment for the award of the degree of

DOCTOR OF PHILOSOPHY

By

ANKITA SHUKLA

MOLECULAR AND STRUCTURAL BIOLOGY DIVISION

CSIR-CENTRAL DRUG RESEARCH INSTITUTE

Lucknow, India

2020

Dedicated to
My Late Grandfather
&
Parents

सी.एस.आई.आर.-केन्द्रीय औषधि अनुसंधान संस्थान, लखनऊ
(वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद्)
सेक्टर 10, जानकीपुरम विस्तार, सीतापुर रोड, लखनऊ – 226 031 (भारत)
CSIR - Central Drug Research Institute
(Council of Scientific & Industrial Research)
Sector 10, Janakipuram Extension, Sitapur Road, Lucknow - 226 031 (India)

This is to certify that the thesis entitled '**Structural and functional characterization of protein(s) involved in mycobacterial DNA repair**' being submitted to the Jawaharlal Nehru University, New Delhi in partial fulfillment of the requirements for the award of the degree of Doctor of Philosophy,comprises work done by **Ms. Ankita Shukla** under my guidance at the CSIR-Central Drug Research Institute, Lucknow. The work is original and has not been submitted, in part or full, for any other degree or diploma of any university.

R-Ravishankar 03/02/2020

Dr. Ravishankar Ramachandran
(Supervisor)
Senior Principal Scientist, HoD
Molecular and Structural Biology Division
CSIR-CDRI
Lucknow

Ankita Shukla

Ankita Shukla

Acknowledgement

Undertaking this PhD has been an important milestone and truly life-changing experience for me. It is my pleasure to express my sincere thanks and deepest gratitude to all those people who led their unconditional support as well as guidance and assistance during my long period of PhD research work.

Primarily, praises and thanks to the God, the Almighty, for His showers of blessings throughout my research work to complete it successfully. Then I express my sincere gratitude to my mentor Dr. R. Ravishankar for his inspiring guidance, enduring patience, constructive criticisms and constant encouragement. His constant support in every step of learning made me more confident. He always provided opportunities to learn new techniques and explore the growing area of structural biology. It was a great privilege to work and study under his guidance and he is mostly responsible for whatever I will achieve in future. I acknowledge Dr. J. V. Pratap for his influential advice and feedback throughout the journey. I would like to thank my DAC committee members Dr. J. K. Ghosh and Dr. Arunava Das Gupta for report evaluation and valuable suggestions. I also express thanks to Dr. Shakeel Ahmed, Dr. Saman Habib, Dr. Tejendar Thakur, Dr. Sohail Akhtar, Dr. M.I. Siddiqui, and Dr. Amogh Sahasrabuddhe. I would also express thanks to Dr. Ravishankar Ambapathi for carrying out NMR experiments, Dr. Mukesh pashupuleti for synthesizing the peptides used in the study and Dr. Y. K. Manju for carrying out in vivo knockout experiments.

I am also grateful to Dr. Ashish Ganguly, IMTECH, Chandigarh, for initially mentoring about the Small angle X-ray scattering experiments. I would also express my sincere thanks to Dr. Radha Chauhan, NCCS, Pune for helping me with the computational facilities for solving Cryo-EM data. I am also thankful to Dr. Ruchir kant for maintaining X-ray and SAXS facility. Mr. J.P Srivastava is acknowledged for assisting in official reimbursement paper work. Daily lab chore help from Santosh, Arvind bhaiya and Deepak bhaiya are also acknowledged.

I express my sincere thanks to my seniors; Dr. Sandeep Kumar Srivastava, Dr. Vandna, Dr. Taran Khanum whose studies on DNA repair laid the foundation of my research work by exploring deep in to the pathway. I also wish to mention my other seniors from X-tal group for their help and assistance; Dr. Sharan, Dr. Ashok, Dr. Aparna, Dr. Arjun, Dr. Pragati, Dr.

Sonal, Dr. Kiran, Dr. Amit Gaur and Dr. Abhishek Dey. My heartfelt thank is reserved for Nelam for supporting during the last phase of PhD. I also express my gratitude to my colleagues and friends Shikha, Sanjay, Vijay, Jyoti, Sadik, Anas, Anil, Pankaj and Dr. Ahmad for all the help and fun time we spent together. I also acknowledge my juniors who recently joined the X-ray crystallography group; Chandan, Farheen, Mohini, Surendra, Faiz, Rajlaxmi and Shalu.

As it is rightly said that journey becomes easy when travelled together, I owe my deep hearted thanks to Mohammad Afsar for valuable friendship. Together we explored the different aspects of the structural biology and DNA repair. We fought, we argued but at last, here we are accomplishing our goal. Thank you for supporting me through all thick and thin.

Nothing would have been possible without the constant support of my family. I thank my parents for encouraging me in all of my pursuits and inspiring me to follow my dreams. My late grandfather was also the biggest pillar of strength and I dedicate this journey to him, as he was the source of inspiration for me. I also thank my husband Himanshu for unfailingly support. Without his encouragement and understanding, it would have been not possible for me to successfully achieve this goal. My brother Abhishek and sister-in-law Ayushi are also acknowledged for their emotional support during the journey. I also thank my in laws for being constant supportive.

I thank Wellcome trust – DBT alliance for access to the ESRF, beamline and Dr. Deepak Nair towards SAXS data collection. Along with, I also thank Dr. Ravi Makde for allowing us to access to collect X-ray data at BL21, Indus2 RRCAT (Raja Ramanna Centre for Advanced Technology), Indore, and India. Financial assistance from UGC, New Delhi is highly acknowledged.

Lastly, I acknowledge all those people who knowingly and unknowingly helped me in the successful completion of PhD.

Ankita Shukla

CONTENTS

LIST OF ABBREVIATIONS

Å	Angstrom
aa	Amino acids
Ab	antibody
AMP	Adenosine monophosphate
AP	Apurinic/ Apyrimidinic
ATP	Adenosine triphosphate
BER	Base excision repair
BLAST	Basic Local Alignment Search Tool
bp	Base pair
BSA	Bovine Serum albumin
CBB	Coomassie brilliant blue
CFU	Colony forming unit
C-terminal	Carboxy terminal
CV	Column volume
DSB	Double Strand Break
dsDNA	Double stranded DNA
DTT	Dithiothreitol
EDTA	Ethylendiaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
EtBr	Ethidium bromide
FRET	Fluorescence resonance energy transfer
gDNA	Genomic DNA
GST	Glutathione S-transferase
HEPES	N-[2-hydroxyethyl] –piperazine ethanesulfonic acid
dIC_{50}	Half maximal inhibitory concentration
IDCL	Inter domain connector loop
IPTG	Isopropyl β-D-thiogalactopyranoside
ITC	Isothermal titration calorimetry
KD	Dissociation constant
KA	Association constant
kDa	Kilo Dalton

Ki	Inhibitory constant
Km	Michaelis constant
Kb	Kilo base pair(s)
LB	Luria-Bertani medium
LD_{50}	Lethal dose
Lp-BER	Long patch base excision repair
MCS	Multiple cloning site
MIC	Median inhibitory concentration
MMR	Mismatch repair
MMS	Methyl methanesulfonate
Mtb	Mycobacterium tuberculosis
NAD+/NADH	Nicotinamide adenine dinucleotide
ND	Not determined
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
Ni-NTA	Nickel-nitrilotriacetic acid
N-terminal	Amino-terminal
PAGE	Polyacrylamide gel electrophoresis
PBG	Peptide binding groove
PBS	Phosphate buffered saline
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PDB	Protein data bank
PIP	Protein interacting peptide
PMSF	Phenylmethyl sulfonyl fluoride
r.m.s.d	Root mean square deviation
RNase	Ribonuclease
RNS	Reactive nitrogen species
ROS	Reactive oxygen species
Rpm	Revolutions per min
rRNA	Ribosomal RNA
RT	Room temperature
SASBDB	Small Angle Scattering Biological Data Bank
SAXS	Small angle X-ray scattering

SDS	Sodium dodecyl sulfate
SP-BER	Short patch base excision repair
SPR	Surface plasmon resonance
ss-DNA	Single stranded DNA
TAE	Tris-acetate-EDTA
TB	Tuberculosis
TBE	Tris-borate-EDTA
TE	Tris-EDTA
THF	Tetrahydrofuran
TRIS	Tris [hydroxymethyl] aminomethane
UV	Ultraviolet
Vmax	Maximum reaction rate
WT	Wild type
β-ME	β-mercaptoethanol

PREFACE

Mycobacterium tuberculosis is the etiological agent of tuberculosis (TB). The dreaded disease accounts for ~25% of all cases known in the world in India. The pathogen is successful as it is able to evade the host immune system and persist for long years. This also involves protecting itself against the host-responses. In this context DNA repair is of paramount importance for the survival of the bug. In the absence of a mismatch repair (MMR) pathway in the mycobacterium, the non-canonical MMR by EndoMS/ NucS or DNA Base Excision Repair (BER) pathway assumes importance for countering host-inflicted oxidative DNA damaging responses. The BER pathway involves multiple players that detect specific types of damage and conclude with repair of the damage.

Our laboratory has been involved in studying the specific and delicate interplay of protein-protein interactions in mycobacterial pre-replicative BER and their implications to the repair outcome. Understanding the underlying molecular mechanisms is expected to suggest alternate therapeutic strategies. Indeed, components of the pathway like the NAD^{+}-dependent DNA ligase (MtbLigA), Sliding DNA ß-clamp, Class II AP-endonuclease, glycosylases and other interacting factors have been shown to be important for the survival of the pathogen and also critical players. In this context, our laboratory has identified several classes of inhibitors that can distinguish between human DNA Lig I and MtbLigA. We have also identified specific inhibitors that disrupt DNA ß-clamp –partner protein interactions, among others. We had also earlier reported the characterization of the DNA ß-clamp-XthA complex and showed that the mode of interactions is DNA substrate dependent.

We set out to identify, characterize and functionally understand the roles of other BER complexes (BERosomes), the underlying molecular mechanisms and inhibitors that could disrupt the complexes for further development as a therapeutic strategy. We have carried out the study using a combination of structural biology tools like X-ray crystallography, Small-Angle X-ray scattering (SAXS), Surface plasmon resonance, Isothermal calorimetry, specifically designed *in vitro* and *in vivo* assays, and computational biology approaches. The present work is focused on the structural characterization of NAD^{+}-dependent DNA ligase with the series of crystal and solution structures of adenylation reaction intermediates that explains the cofactor-induced serial-remodeling of the active site. This also involved identifying small-molecule fragment inhibitors and demonstrating their mode of interactions with LigA through co-crystal structures. We also report the structural characterization of an intriguing MtbLigA-MtbXthA complex. While

XthA is an initial player in the BER pathway, LigA is the terminal enzyme that completes the repair. Among other results, we have identified a novel fundamental interaction of the bacterial BER pathway that is apparently conserved. The studies suggest a coordinating function for XthA whereby it engages initially with LigA to prevent the undesirable consequences of futile cleavage and ligation cycles that might derail bacterial BER.

The thesis is divided in to seven sections that briefly described below-

Chapter 1 This chapter acquaints the reader with the different kinds of DNA damage and the repair mechanisms. The focus is on base excision repair pathway (BER) and the BERosomes formed at the damaged site. This chapter also summarizes the potential of BER pathway and protein-protein interactions at the damaged site as a target for new therapeutic treatment.

Chapter 2 Here, the various techniques and experimental approaches used to clone and purify recombinant proteins/ mutants reported in the present study are detailed. It also covers the various biochemical and biophysical assays used to characterize the respective protein-protein interactions. This chapter details the methods used to solve the crystal and solution structures of NAD^+-dependent DNA ligase wild type and mutants. It also covers the experimental approaches used to identify and evaluate specific inhibitors of MtbLigA and MtbXthA.

Chapter 3 reports the results regarding the mechanism of cofactor-induced serial remodeling of MtbLigA with the help of crystal and solution structures of various reaction intermediates. Also presented, is the mechanism of NMN release from the active site of MtbLigA in order to prevent the reaction reversal. The chapter also reports the details of conformational changes of MtbLigA in the presence of nicked DNA substrate with the help of SAXS studies and the link to the overall ligase reaction.

Chapter 4 reports the determination/ identification of the residues that play a crucial role in the relative movements of Ia and 1b subdomains of the MtbLigA adenylation domain and also the conserved residues of BRCT domain. Structure guided mutational analysis of residues from the adenylation and BRCT domains respectively has identified novel roles during DNA ligation reaction.

Chapter 5 details the structural characterization of the MtbLigA-MtbXthA complex. With the help of various techniques including SAXS, SPR, SEC, ITC and FRET we determined the motifs and residues on both MtbLigA and MtbXthA responsible for interaction. We also rationalized why the initial player (XthA) of the BER pathway interacts with the final

component (LigA) and suggest that this is a conserved BER interaction that is needed to prevent futile cleavage and ligation cycles.

Chapter 6 reports the evaluation of small molecule fragments from the commercially procured BioNet fragment library as potential and specific inhibitors of the MtbLigA. The chapter gives the details of the *in vitro* and *in vivo* antibacterial assays used to test their inhibition potency. The studies including co-crystallization of small molecule fragments with the adenylation (AdD) domain have led to the identification of small-molecule fragments that competitively bind to the NMN binding site.

Chapter 7 is the final chapter and reports the identification of small molecule fragments from the BioNet library and *Medicines for Malaria* initiative MMV pathogen box as potential and specific inhibitors of the MtbXthA. The chapter gives the details of the *in vitro* and *in vivo* antibacterial assays used to test their inhibition potency.

The X-ray crystallography data and Small Angle X-ray scattering data reported in the study have been submitted to the Protein Data Bank (PDB) and Small Angle Scattering Biological Data Bank (SASBDB) respectively and are listed below-

S.No	X-ray Structures	Database	Accession number
1.	Adenylation domain (AdD) domain of MtbLigA	RCSB Protein Data Bank	PDB ID- 6KJM
2.	AdD domain active site mutant- K123A	RCSB Protein Data Bank	PDB ID- 6KSD
3.	AdD domain active site mutant- K123A	RCSB Protein Data Bank	PDB ID- 6KSC
4.	AdD domain Ib subdomain mutant- H236Y	RCSB Protein Data Bank	PDB ID- 6KRH
5.	AdD domain Ia subdomain mutant- E22A	RCSB Protein Data Bank	PDB ID- 6KDU
6.	AdD domain co-crystal structure with small molecule fragment 9R0345	RCSB Protein Data Bank	PDB ID- 6KKV

S.No	Small angle X-ray Scattering (SAXS) Structures	Database	Accession number
1.	AdD domain SAXS structure with NAD^+	Small angle scattering biological data bank (SASBDB)	SASBDB ID-SASDEC6
2.	AdD domain SAXS structure with NMN	Small angle scattering biological data bank (SASBDB)	SASBDB ID-SASDEU5
3.	AdD domain SAXS structure with NMN and AMP	Small angle scattering biological data bank (SASBDB)	SASBDB ID-SASDG65
4.	*Apo*-MtbLigA SAXS structure	Small angle scattering biological data bank (SASBDB)	SASBDB ID-SASDDN9
5.	MtbLigA with nicked DNA substrate SAXS structure	Small angle scattering biological data bank (SASBDB)	SASBDB ID-SASDDV9
6.	MtbXthA SAXS structure	Small angle scattering biological data bank (SASBDB)	SASBDB ID-SASDDD8
7.	MtbXthA-MtbLigA complex SAXS structure	Small angle scattering biological data bank (SASBDB)	SASBDB ID-SASDEW3
8.	MtbXthA-MtbLigA complex with nicked DNA substrate SAXS structure	Small angle scattering biological data bank (SASBDB)	SASBDB ID-SASDEX3
9.	*Apo*-MtbLigA SAXS structure	Small angle scattering biological data bank (SASBDB)	SASBDB ID-SASDDN9
10.	MtbLigA with nicked DNA substrate SAXS structure	Small angle scattering biological data bank (SASBDB)	SASBDB ID-SASDDV9
11.	MtbXthA SAXS structure	Small angle scattering biological data bank (SASBDB)	SASBDB ID-SASDDD8
12.	MtbXthA-MtbLigA complex SAXS structure	Small angle scattering biological data bank (SASBDB)	SASBDB ID-SASDEW3
13.	MtbXthA-MtbLigA complex with nicked DNA substrate SAXS structure	Small angle scattering biological data bank (SASBDB)	SASBDB ID-SASDEX3

CHAPTER 1 | INTRODUCTION

1. Introduction

For all organisms, DNA repair is one of the most essential systems for maintaining the genomic integrity life. The several exogenous and endogenous stresses continuously attack the genome that leads to the accumulation of lesions throughout the DNA and thus ultimately results in the death if not repaired. The cells are equipped with inordinately well-conserved DNA repair pathways that are responsible for removing these deleterious lesions and ease their fallouts. Therefore, in nutshell the DNA repair pathways maintain the genomic stability. The importance of these pathways are highlighted by the recipients of 2015 Nobel prize in chemistry by Tomas Lindhal, Aziz Sancar and Paul Modrich for their originative work in determining DNA repair pathways.

1.1 Types of DNA Damage

1.1.1 Damage of endogenous origin

1.1.1.1 Mismatch, one base transition to another

There are four nitrogenous bases in DNA that may convert into one another or into rare bases with non-canonical base pairing. The hydrolysis of nitrogenous base leads to the base conversion. The most common hydrolytic damage is deamination, for example, deamination of cytosine creates uracil. Since uracil pairs with A more expeditiously than G, thereby leading to the formation of mismatch. In the next replication cycle, this will leads to substitution of a C:G with A:T pair. So just like this several deamination leads to the formation of some other base producing the overall mismatch in the genome. The final protein resulted may be partially or completely altered due to the substitution of amino acid which occurred because of conversion of one base to another. Alternatively, mismatches may insert a premature stop signal in mRNA, leading in production of truncated and/or unstable protein or no protein at all.

1.1.1.2 Loss of nitrogenous bases, hydrolytic damages

The distinctive example of hydrolytic loss of nitrogenous bases is DNA depurination. This a very commonly encountered type of DNA damage occurring spontaneously $5x10^3$ - $1x10^4$ times per genome per day in human cells (Lindahl 1993). Several *in vitro* experiments show that DNA depurination occurs much quicker than depyrimidination and at neutral pH and

37°C temperature, the rate of loss of pyrimidine bases from DNA is about 5% of the rate of loss of purine bases (Chatterjee and Walker 2017). At 37°C, the *E. coli* genome (~10^6 bp) loses on the average one base per cell generation, while at 85°C the genome of *Thermus thermophilus* (˜10^6 bp) may lose up to 300 bases per cell generation (Caglayan and Wilson 2015). Thus, the loss of nitrogenous bases in DNA is strongly dependent on the temperature, finally creating abasic sites in the DNA.

1.1.1.3 Oxidation of nitrogenous bases, oxidative stress

The purines (A, G) as well as pyrimidine (C, T) both are affected by the oxidation of nitrogenous bases. For example, purine base oxidation results in 8-oxopurines or alternatively in formamidopyrimidines (**Fig. 1.1**). The oxidation products of pyrimidine are generally thymine glycol, uracil glycol, 5-hydroxycytosine and 5-hydroxyuracil. The exposure of some environmental agents or endogenous lipid peroxidation may leads to the formation of ethenobases. Enhanced formation and persistency of 1,N(6)-ethenoadenine in DNA is not linked with more prominent susceptibility to carcinogenesis in alkylpurine-DNA-N-glycosylase knockout mice treated with vinyl carbamate (Barbin, Wang et al. 2003). The cellular metabolism or UV exposure or by some chemical exposure leads to the formation of reactive oxygen/nitrogen species (ROS, RNS) which in turn cause severe DNA damage including strand breaks, abasic sites, oxidized rings etc (Boiteux and Radicella 1999, Wallace 2002, Wallace 2013). Stand breaks immediately after base oxidations in DNA caused by the same damaging agent (most significantly, free radical species) can induce both types of DNA damage. Oxidative DNA damage accumulation (oxidative stress) is considered one of the deadliest DNA damage.

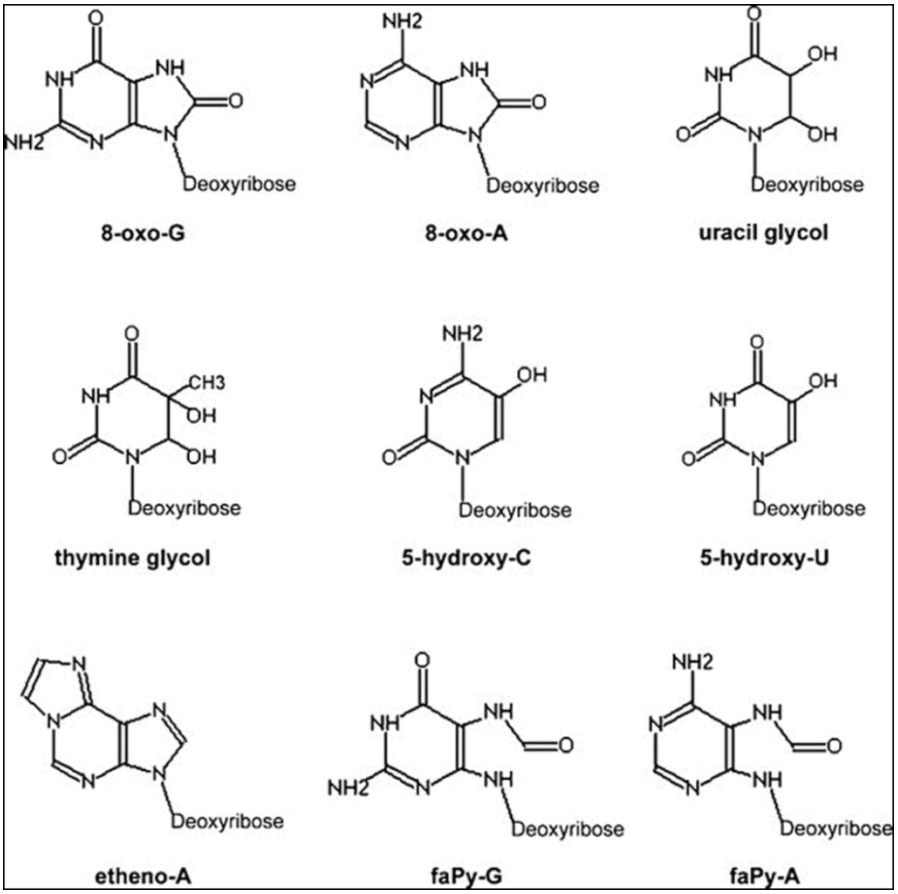


Figure1.1 Common type of oxidative DNA damage. Figure adapted from Chakarov. S et al, 2014.

1.1.1.4 DNA breaks of endogenous origin

SSBs (single strand breaks), formed by disruption of phosphodiester bond, are among the most common instances of DNA damage. Strand breaks also occur due to genotoxic stress resulting from normal cell metabolism. While double strand breaks, occur during replication, relaxation of supercoils etc. Double-strand breaks (DSBs) are more deleterious and are less common in living cells. When the phosphodiester backbone of both DNA strand is broken leads to the double-strand DNA breaks (DSB). Double-strand breaks in normal cells frequently result from more than one single-strand breaks, occurring in the neighborhood of each other, on opposite strand. The occurrence of a DSB may start or/and take part in recombination leading to either homologous recombination, Non-homologous recombination (NHEJ) depending on the presence of type of damage.

1.1.2 Damage of exogenous origin

1.1.2.1 Dimerization due to radiation

The dimer formation mostly outcome of exposure to electromagnetic radiation in the UV range (100-300 nm, UV-B and 300-400 nm, UV-A) while strong ionizing radiation produces DNA breaks and weak ionizing radiation leads to the production of reactive free species (oxidative damage) (Wallace 2002, Wallace 2013). Dimers of any type do not naturally subsist in DNA. Dimerization may occur between bases belonging to the same strand of DNA or between bases from different DNA strands. It produces two different types of dimers in DNA: 6-4 photoproducts and cyclobutane pyrimidine dimers.

1.1.2.2 Formation of bulky adducts

The polycyclic aromatic hydrocarbons (PAH) are among the ordinarily encountered and strong adduct-forming environmental agents. The benzopyrene (a compound of tobacco smoke) leads to the formation of adduct-forming agent in vivo i.e. benzopyrene-diol-epoxide. The other agents responsible for formation of adducts in DNA are trioxsalen (psoralen), formaldehyde (creates DNA-DNA as well as DNA-protein crosslinks), mitomycin C (a crosslinking agent), some of the rawer anticancer drugs, such as ellipticine (Stiborova, Sejbal et al. 2004), the dye ethidium bromide (an intercalating agent), and many others. The commonly used compounds in cleaning liquids, liquids used in polishing, disinfectants or some industrial solvents are also are adduct-forming agents. The bulky adducts containing regions of DNA needs more than one type of repair mechanism at a time since the process is complex. Nucleotide excision repair pathway and homologous recombination pathway is usually involve in this.

1.1.2.3 Alkylation damage

The damage by alkylation is usually in the form of methylation on selected bases, forming a variety of modified bases like 5-methylcytosine, 7-methylguanine, 1-methyladenine, etc. The alkylated bases forms non-canonical base pairing thus leading to the mismatch mutagenesis. The alkylated bases are also highly cytotoxic as they block the replication process. The most common alkylating agents are halolalkanes (chloroform, tetra/tri chloromethane), nitrosoureas, alkyl sulfonates, etc.

1.2 DNA repair

The continuous exposures to the DNA damaging agents are hazardous to the life. The exposure to oxidation, alkylation and hydrolysis leads to single strand break lesions and if these lesions are not repaired may lead to the permanent mutation. While agents like ionizing radiation, host defense system and antibiotic treatment etc. lead to DNA double strand break (DSB's). Thus, the several conserved pathways to repair or extenuate the lesions are created during the course of evolution. The process of recombination (homologous or non-homologous) generally repairs the DSB's while SSB's are amended by direct reversal or by the excision of the damaged component (**Fig. 1.2**).

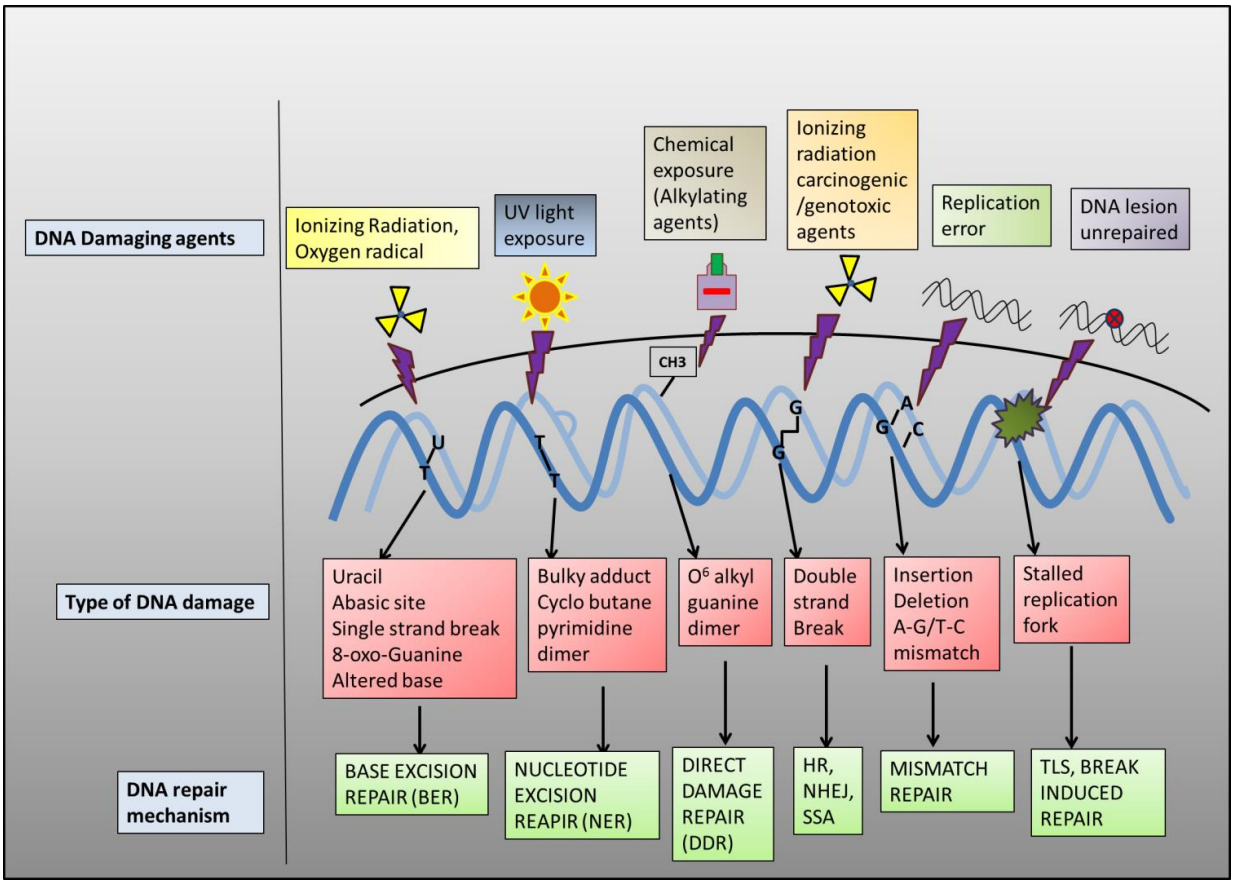


Figure 1.2 The DNA damages eliciting various DNA repair responses. The term HR is, homologous recombination; NHEJ is, Non-homologous end joining; SSA is, single strand annealing; and TLS is, translesion synthesis.

1.2.1 Double strand break repair

The most lethal damage is when both strands in the DNA double helix are damaged and may lead to the lasting effect on the genome or may cause genome rearrangement. These damages generally occur at the time of replication or by the exposure of ionizing radiation or

genotoxic chemicals. The DSB's are repaired by the two elementary pathways *viz.* homologous recombination (HR) and non-homologous end joining (NHEJ). A new mechanism of end joining is also reported i.e. alternate end joining (AEJ).

1.2.1.1 Homologous recombination (HR)

The intact DNA duplex homologous to DSB site act as template for the re-synthesis of broken strand(Cromie, Connelly et al. 2001). The 5'-3' exonuclease resect the DSB site containing DNA at one or both end. The exonuclease action generates the 3'-OH single strand extension that further intrudes on the intact sister DNA. This invading reaction is catalyzed by RecA enzyme (Rad251 in eukaryotes). The intruding strand act as a primer for DNA polymerase, which copies the information across the break (**Fig. 1.3**). In prokaryotes and lower eukaryotes, RecBCD or RecFOR complexes initiate the HR. The residual single strand nicks are further sealed by the action of DNA ligase (LigA in bacteria).

1.2.1.2 Non-homologous end joining (NHEJ)

In non-homologous joining, the strands are directly ligated without the need of homologous DNA template. The term was first coined by Moore and Harber (Moore and Haber 1996). The NHEJ process is error prone as it introduces substitution, deletion, duplication, etc. In bacteria, NHEJ is constituted by the two main enzymes Ku and LigD (Weller, Kysela et al. 2002, Bowater and Doherty 2006). Ku binds to the ds DNA in a sequence independent manner that then recruits the LigD to complete the overall DNA repair process (Zhu and Shuman 2010, de Vega 2013, de Ory, Carabana et al. 2019) (**Fig. 1.4**).

1.2.1.3 Alternate end joining repair (AEJ)

A new mechanism of end joining was reported in *E. coli*, *which* does not rely on the NHEJ repair. Since proteins Ku and LigD are, absent in *E. coli* so they definitely rely on some mechanism other than NHEJ for repairing DSBs. The AEJ leads to the resection of several bidirectional resections. The studies shows that RecBCD complex and LigA is required to accomplish the AEJ (Chayot, Montagne et al. 2010).

1.2.2 Directed reversal repair (DDR)

It is one of the simplified repair processes that correct the damage in single step without the template requirement. The type of damage decides the DDR and this does not break the

phosphodiester bond. For example, the enzyme MGMT (methyl guanine methylase transfer) reverses the guanine methylation (Mishina, Duguid et al. 2006).

1.2.3 Mismatch repair

Sometimes the incorrectly added nucleotide escapes the proofreading by replication machinery and gets logged permanently in to the new replicated strand thus leading to the deleterious lesion in the pathogen. The specialized pathway, which deals with this kind of deleterious situations, is known as mismatch repair process. It was identified that Muts detects the mismatched nucleotide (Junop, Obmolova et al. 2001)by recognizing the DNA distortion. This in turn recruits MutL which then recruits MutH (Spampinato and Modrich 2000, Lee, Chang et al. 2005) followed by the UvrD (Guarne, Ramon-Maiques et al. 2004) which unwinds the DNA from the cleaved point to the mismatched nucleotide. Exonuclease then cleaves creating a single stranded gap, which is then filled by DNA polymerase and finally ligated by DNA ligase. However, mismatch repair process had not been discovered until now in mycobacterium tuberculosis.

1.2.4 Nucleotide excision repair (NER)

Nucleotide excision repair pathway recognizes wide range of lesions comprising from bulky distorting one to the non- distorting lesions. This pathway is diverse in its action as it recognizes any modification that transforms the normal base in to the damaged bases. The versatility of the pathway works by recognizing the distortion in the double helix shape of DNA. The major enzymes UvrA, UvrB, UvrC and UvrD execute the whole chore (Crowley, Boubriak et al. 2006, Savery 2007). The UvrA-UvrB complex performs the scanning along the DNA with energy requiring from ATP. UvrA detects the DNA distortion and exits the site leaving behind the UvrB to make the ds DNA in to an ss-bubble around the distortion. Then the UvrC is recruited which performs the incision activity on both side of lesion (Verhoeven, van Kesteren et al. 2002). UvrD that helps in the unwinding during the process follows UvrC. At last DNA polymerase I and DNA ligase completes the repair and seal the gap (**Fig. 1.5**). The NER is further subdivided in to two sub pathways, global genome repair NER (GG-NER) which comes in action when error occurs in non-transcribed area and transcription coupled NER (TC-NER) which corrects the error in the genes actively involved in expression (Kitsera, Rodriguez-Alvarez et al. 2019).

1.2.5 Base excision repair (BER)

1.2.5.1 The core of BER

One of the conserved repair pathways responsible for correction of most non-bulky lesions that cannot be processed by the direct reversal process. The process begins with recognition of damaged base by DNA glycosylase (Barnes and Lindahl 2004) which catalyzes the N-glycosidic bond and cleaves the damaged base thereby creating Abasic (AP) site. The abasic sites are further processed by the two-sub pathways- short patch BER and long patch BER. The short patch BER repair the individual nucleotide damage and is initiated by AP lyase that cleaves the DNA backbone on the 3'- side of the abasic site while long patch BER requires the additional two proteins – PCNA and flap endonuclease (Fen-1). Polymerase delta and epsilon are required along with proliferating cell nuclear antigen (PCNA) in displacing multiple downstream bases and simultaneously synthesizing DNA. This displacement results in the generation of a single stranded 5' DNA flap substrate, which is, do not respond to ligation. The flap is further processed by flap endonuclease (Kim, Biade et al. 1998), thus facilitating ligation of DNA ends by DNA ligase. The single-nucleotide short-patch BER pathway is completed by the action of DNA ligase III while DNA ligase I completes the nick ligation in long-patch BER (**Fig. 1.6**). The inhibition of LP-BER occurred when PCNA was depleted from the cell extracts using polyclonal antibody suggesting it is too required for long patch repair (Frosina, Fortini et al. 1996).

The factors that determine the selection of either the short patch or long patch BER pathway in cell are still not properly understood, although it has been hypothesized that both the nature of the DNA glycosylase and the lesion may influence the selection. The other hypothesis, which explains the switch, is the relative ATP concentration near the abasic site. In addition, the types of lesion, cell cycle stage also determine the one of the two-sub pathway. Both SP-BER and LP-BER exist in humans but SP-BER is absent in yeast *S. cerivisae* because various mammalian short-patch proteins homologs are absent in it, including pol β, DNA ligase III, XRCC1, etc. It is also hypothesized that if 5'-dRP intermediate produced by AP endonuclease is efficiently removed by dRP lyase (dRPase) activity of pol β, the BER goes forward by short-patch mechanism. Moreover, if it is not efficiently removed then to avoid generating a nick that is futile to the action of DNA ligase the pathway proceeds by long patch (Klungland and Lindahl 1997).

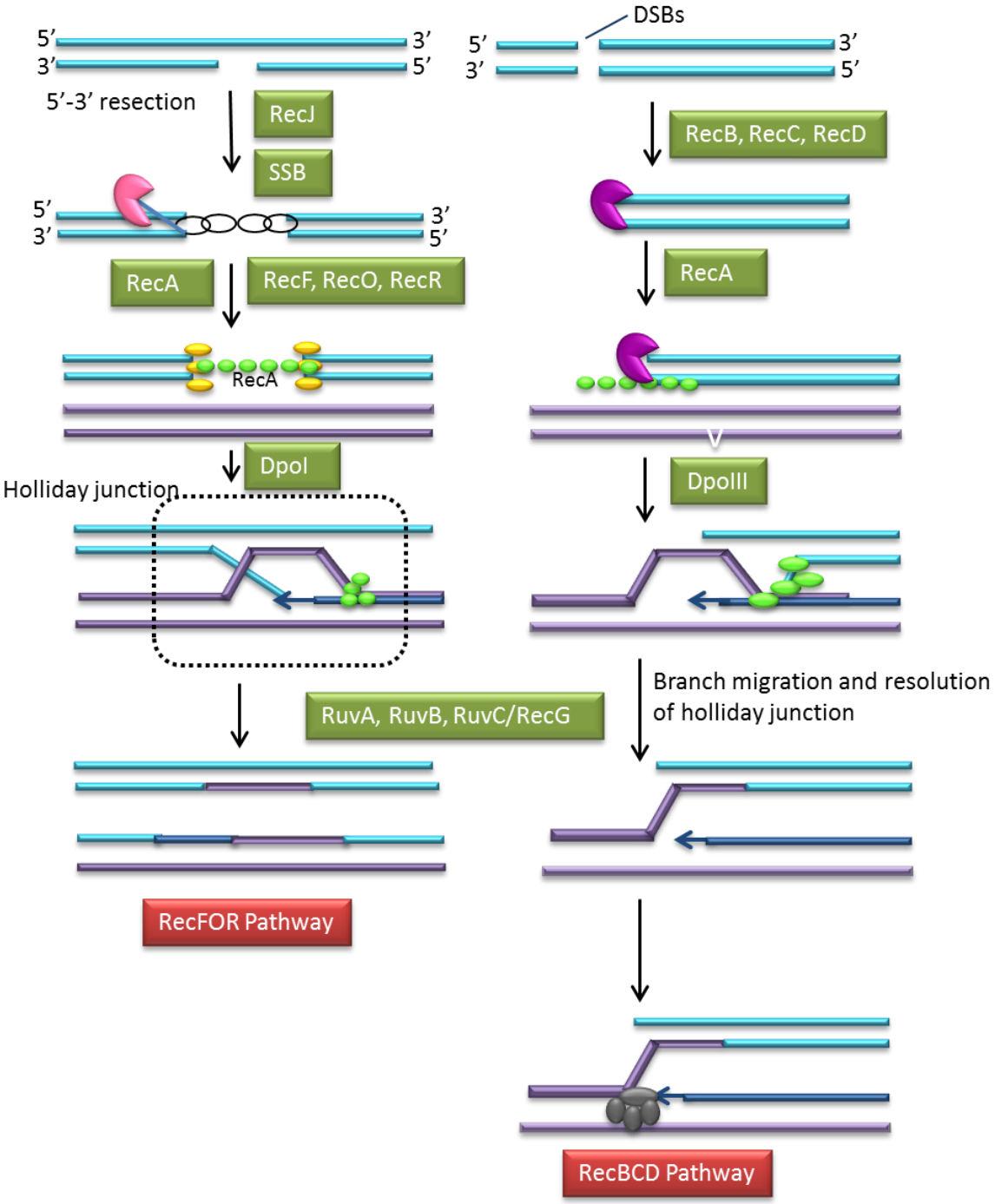


Figure 1.3 Schematic representation of homologous recombination pathway in *Mycobacterium tuberculosis.*

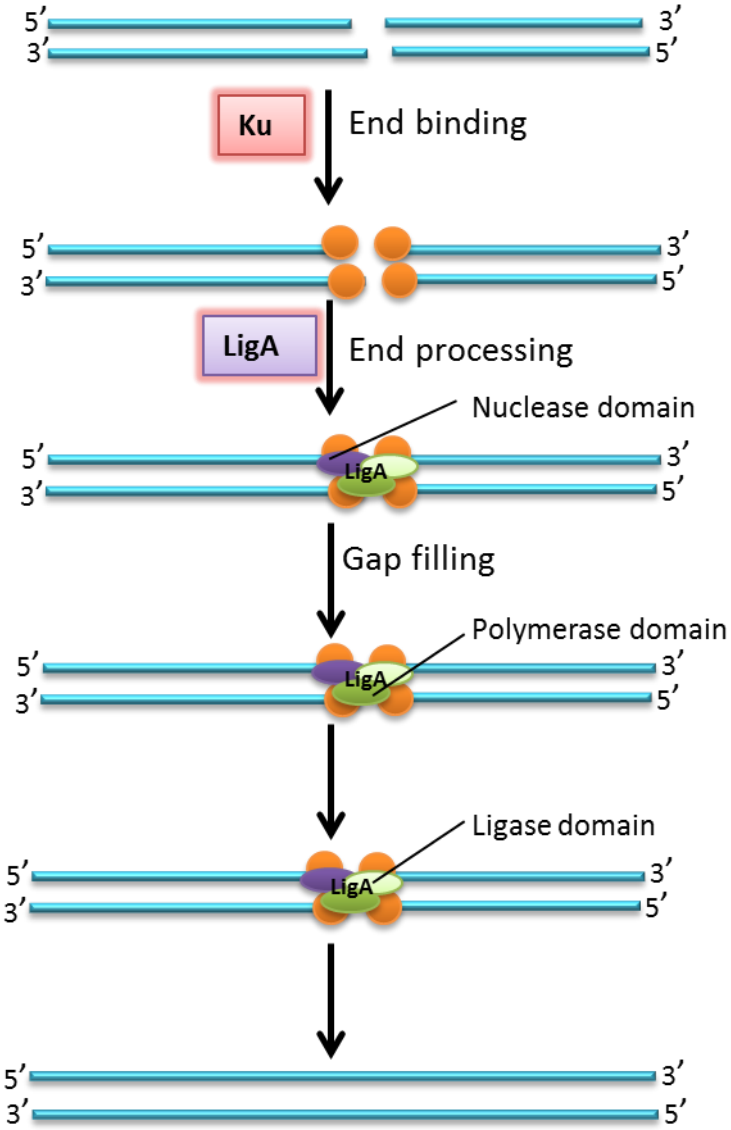


Figure 1.4 Schematic representation of Non-homologous recombination repair (NHEJ) pathway in *Mycobacterium tuberculosis.*

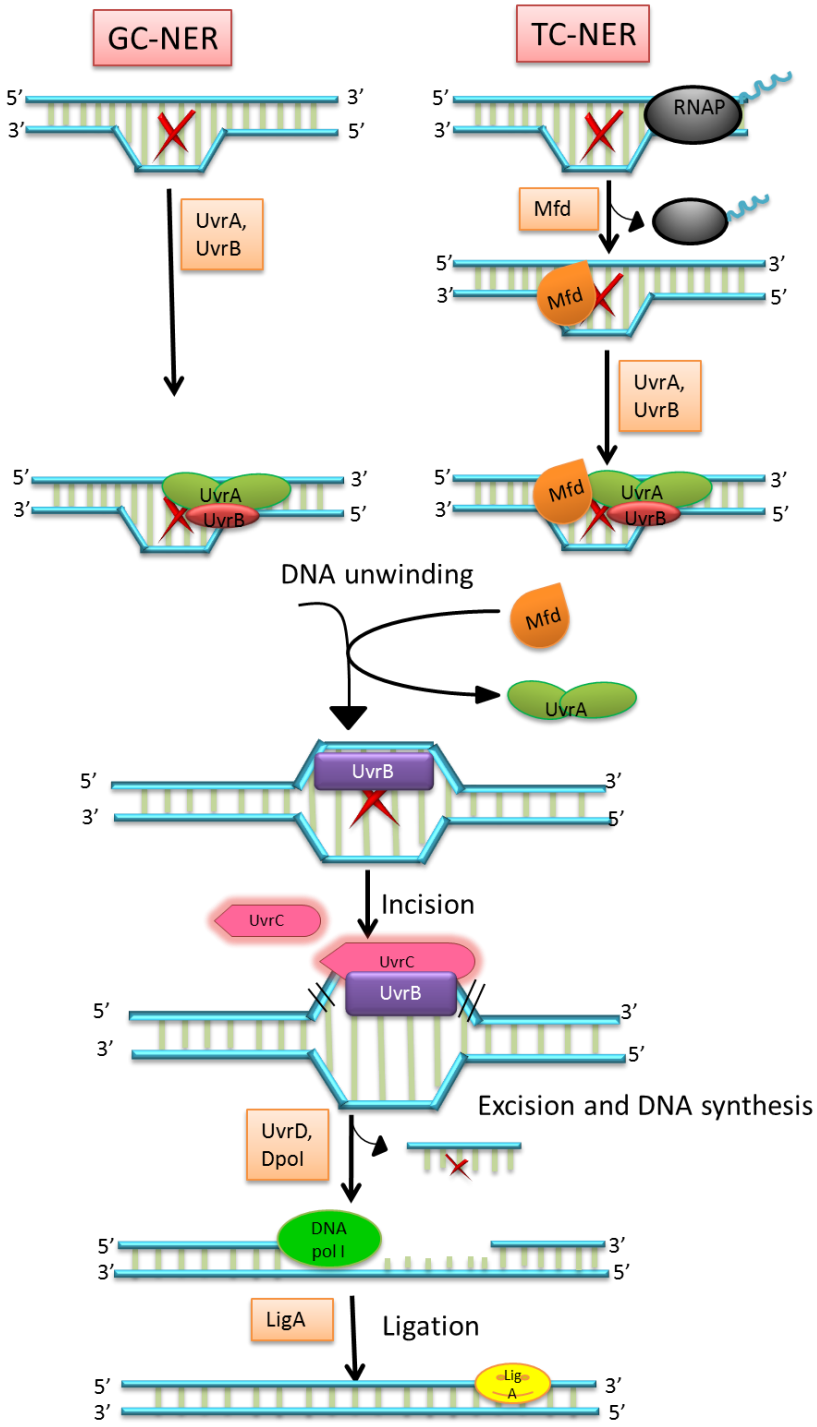


Figure 1.5 Schematic representation of Nucleotide excision repair (NER) pathway in *Mycobacterium tuberculosis.*

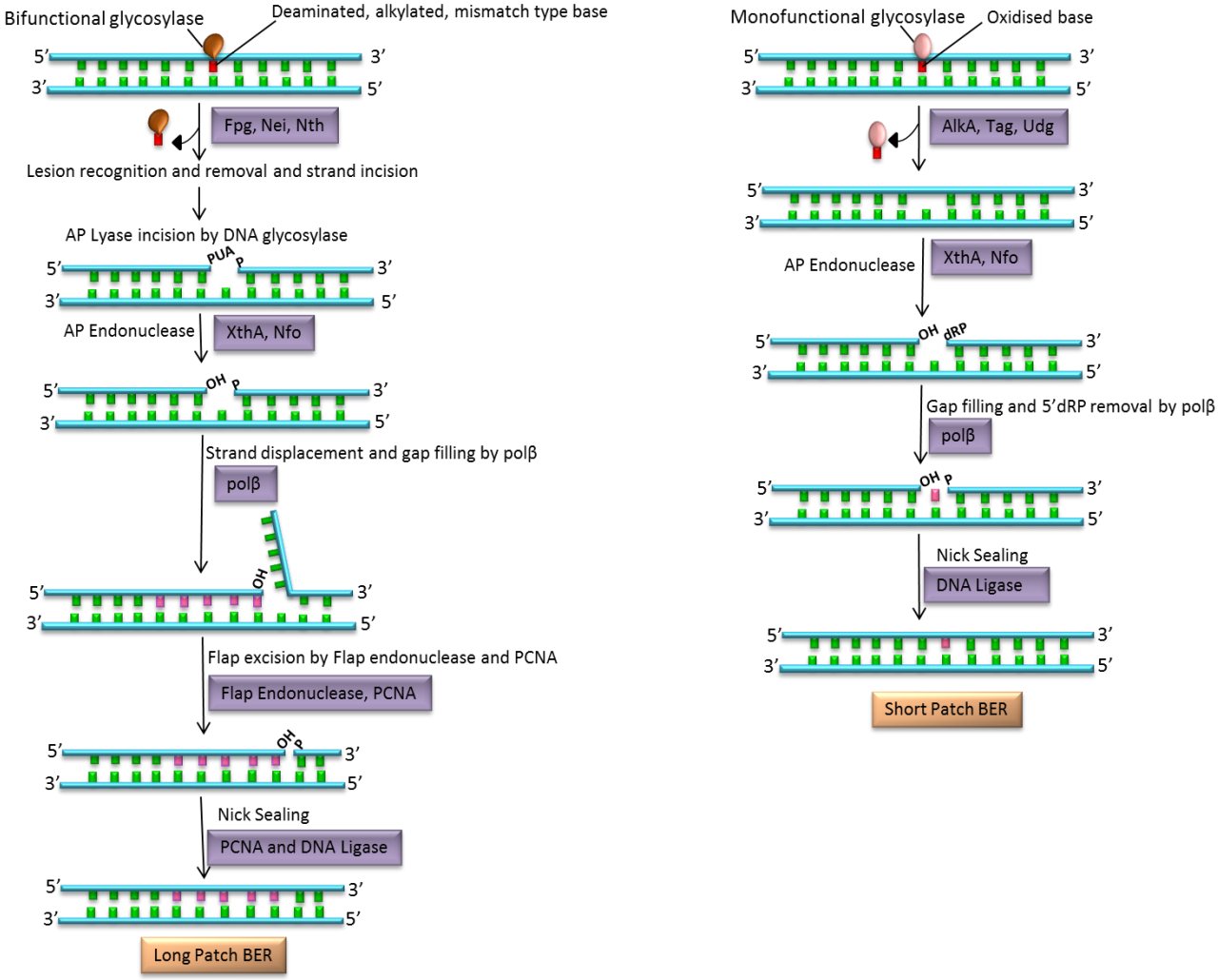


Figure 1.6 Schematic representation of Base excision repair (BER) pathway in *Mycobacterium tuberculosis.*

Table 1.1 List of proteins involved in Base excision and Nucleotide excision repair pathways of *Mycobacterium tuberculosis.*

Proteins involved in BER/NER pathway	Activity performed in BER/NER	Gene annotation	Crystal structure from *M. tb* (PDB ID)
Udg	Class 5 uracil DNA glycosylase	Rv1259	4WS4
Ung	Class 1 uracil DNA glycosylase	Rv2976c	3A7N
Fpg	Formamidopyrimidine DNA glycosylases	Rv2924	-
Fpg2	Formamidopyrimidine DNA glycosylases	Rv0944	-
Nei1	Endonuclease VIII	Rv2464c	-
Nei2	Endonuclease VIII	Rv3297	-
Nth	Endonuclease III	Rv3674	-
MutY	Adenine DNA glycosylase	Rv3589	-
TagA	Glycosylase	Rv1210	-
AlkA	MethylphosphotriesterDNA-protein-cysteine S -methyltransferase	Rv1317	-
Mpg	Glycosylase	Rv1688	-
Nfo	Endonuclease VI	Rv0670	-
XthA	Exonuclease III	Rv0427c	-
Dut	dUTPase and dCTPase activities	Rv2697c	1SM8,1SMC,2PY4,4GCY ISLH, 3LOJ ,1SIX
MuT 1,2,3 and 4	8-oxo(dGTP)/8oxo(GTP) Hydrolase,dGTP,dTTP hydrolase, NTPases	Rv2985,Rv1160, Rv0413, Rv3908	-
RdgB	NTPase for removal of hypoxanthine/xanthine triphosphates	Rv1341	-
MazG	Nucleoside triphosphate pyrophosphohydrolase	Rv1021	-
UvrA	Exonuclease/ATPase	Rv1638	3ZQJ
UvrB	Stimulates ATPase activity of UvrA	Rv1633	-
UvrC	Attaches to UvrA-B and incises the damaged DNA on both sides of the	Rv1420	-

	damaged site		
UvrD1/D2	3'-5'DNA helicase	Rv0949, Rv3198	-
Mfd	Involved in TC-NER	Rv1020	-
PolA	DNA polymerase	Rv1629	-
DinP	Y-family DNA polymerases V	Rv3056	-
DinX	Y-family DNA polymerases IV	Rv1537	-
DnaE2	Error-prone polymerase	Rv3370	-
LigA	NAD^+-dependent DNA ligase	Rv3014c	1ZAU,6KJM(adenylation domain),6KDU,6KRH,6KSD,6KSC (AdD domain mutants), 3SGI(BRCT deleted LigA), 6KKV(AdD domain with inhibitor)
LigB	ATP-dependent DNA ligase	Rv3062	-
LigC	ATP-dependent DNA ligase	Rv3731	-
LigD	ATP-dependent DNA ligase	Rv0938	1VSO(ligase domain) 2IRU(polymerase domain),2R9L(polymerase domain)

1.3 DNA ligase: the *sine qua non* of genome integrity

As the result of discontinous DNA synthesis on the lagging strand of the replication fork there comes about the break in the phosphodiester backbone. Also their occurs break during the repair and recombination process which are all finally sealed by DNA ligase. The first DNA joining was reported in 1967 in the extracts of *E. coli* cells and also the bacteriophage T4- infected *E. coli* cells and later was also reported in the extracts of mammalian cells (Soderhall and Lindahl 1976). Due to the essentiality of the lig gene, the genetic inactivation

causes pleotropic phenotypes which includes lethality as well as hypersensitivity to DNA damaging agents. The DNA ligases comes in the nucleotidyl transferase superfamily (Shuman 1996), which also includes RNA ligase and mRNA capping enzyme. The phosphodiester bond formation catalysis is summarised in three overall steps by utilization of cofactor (ATP or NAD^+) by DNA ligase. In the first step of ligation reaction, NAD^+ binds to Ia subdomain of adenylation domain through NMN part and rightly place the AMP into the active site (Sriskanda and Shuman 2002, Gajiwala and Pinko 2004). The transfer of AMP from active site lysine to the 5'PO_4 nick of the DNA occurs in second step. The oxygen of 5'PO4 of DNA attacks the phosphorous of ligase-adenylate and the lysine in active site is the leaving group. Finally in the last (third) step, the nick sealing is attained and AMP is freed from the active site later the DNA. There occurs slight variation in sequence alignment and domain organization of ATP and NAD^+-dependent DNA ligases. During the ligation reaction steps, the orientation around DNA of both ligase is somewhat dissimilar. The particularity of both ligases towards the cofactor is largely imparted by their respective binding site residues which constitute structural dynamics that clutches and fit the PP_i and nicotinamide mononucleotide as the departuring group (Hakansson, Doherty et al. 1997, Shuman and Lima 2004).

1.2.5.3 The domain architecture of NAD^+ dependent DNA Ligase

Highly modular structure in NAD^+-dependent DNA ligase consisting of unique arrangement of one each domain including adenylation domain, oligomer binding fold domain, helix-hairpin-helix motif domain and BRCT domain assists the distinct structural switches into protein throughout the reaction (**Fig. 1.7 A and B**). Nick sensing mechanism is an inherent property in DNA ligase by having strong affinity towards 3'OH and 5'PO_4 nick as compare to duplex DNA or gapped DNA, which depends on the covalent adenylation of 5'PO_4 of the nick (Sriskanda and Shuman 1998, Odell and Shuman 1999).

1. **Adenylation domain**

This domain is further composed of two subdomain viz. Ia and Ib. The crystal structure of the adenylation domain of *Efa*LigA complexed with NMN (PDB: 1TA8) and NAD^+ (PDB: 1TAE) demonstrated the flexibility of Ia subdomain around the Ib subdomain to complete the binding niche for NAD^+. The Small angle X-ray scattering of adenylation domain of MtbLigA as reported in the present study also shows the Ia subdomain rotation around Ib during the process of deadenyaltion and adenylation.

The adenylation domain (AdD) of MtbLigA is composed of mix α, β fold that belongs to the nucleotidyl transferase (NTase) superfamily including mRNA capping enzymes with the conservation of active site (Srivastava, Dube et al. 2005). The structural constitution of Ia subdomain has 2 alpha helices while Ib subdomain has 7 beta sheets surrounded by 9 helices. The residues in Ia subdomain E22, E26 and E87 in *M. tb* plays crucial role during the closed confirmation of AdD domain at time of adenylation of active site. The other residues of Ia subdomain which constitute the NMN binding pocket are Y32, I39, D41 and G72. The nicotinamide ring is base stacked between F44 and Y31 while I39 and D41 forms the hydrogen bonding with the amide N. The KXDG motif is located in the loop between the two antiparallel beta sheets in Ib subdomain that forms the active site. The phosphate group of the nucleotide is exposed outside while the adenine base of the cofactor is stacked in the hydrophobic site between the aromactic residues. The tyrosine is highly conserved in the other NAD^+-dependent DNA ligases (Y227 in *E. fa* and Y226 in *T. fi*) but in MtbLigA, tyrosine is replaced by histidine (H236). The other aromatic residue is V289 in *E. fa* and *T. fi* while V299 in *M. tb.* the mutational analysis of these residues in Ia subdomain (E22A, E26A, E87A) and Ib subdomain (H236Y, K123A and K123R) have shown to be detrimental to its activity as well as in survival. The mutational analysis in *E. coli* LigA (Zhu and Shuman 2010) these residues are crucial and are involved in interaction with the cofactor.

Our group has earlier reported the AdD domain crystal structure with AMP captured in novel syn and anti conformation (PBD : 1ZAU and PDB: 3SGI) (Srivastava, Dube et al. 2005) and where the switch in contact of AMP with active site lysine can be seen (**Fig. 1.8 A and B**). Here in this study we report the series of crystal structure which showcase the whole remodeling of active site with AMP as well as NMN bound in their respective niche. The conversion of AMP from *syn* to *anti* and then to *syn* before the entry of DNA is summarised through the variours crystal structures reported here. Overall, the crsytal structure and solution structure suggest that two subdomain of adenylation domain shows large relative displacemnt during the course of reaction and exist in 'open' and 'closed' state.

2. Oligomer binding domain (OB)

OB fold is derivative of a greek key motif found in proteins that binds with the single stranded or double stranded DNA or RNA. The biochemical studies of T7 ATP-dependent DNA ligase revealed that it binds with ds-DNA and also enhances the ligation activity of enzyme (Lee, Chang et al. 2000). The entomopoxvirus which is a NAD+-dependent DNA

ligase possess complete OB fold (Sriskanda and Shuman 2002). Moreover the structure guided mutational analysis of OB domain in E.co LigA and HuLig III had also implicated that this domain is involved in interaction with DNA (Mackey, Niedergang et al. 1999). Recently it has been observed that OB domain of *Prochlorococcus marinus* makes extensive contacts with the 5'- end of the nick and also with the phosphodiester backbone of the complement strand (Williamson and Leiros 2019).

3. Zn finger and Helix-hairpin-helix motif (Zn and HhH)

The four cysteine residues are conserved in the C-terminal of NAD^+-dependent DNA ligases, which binds with zinc and DNA. Among the ATP dependent DNA ligases, Human Lig III is reported to have Zn finger at N-terminal (**Fig. 1.9**). So far it is concluded that Zn finger helps in DNA binding and also act as structural support to HhH and BRCT domain (Jeon, Shin et al. 2004).

The HhH motif resides in the C-terminal of all the eubacterial NAD^+-dependent DNA ligases and the number of helices may vary from species to species. In *T. fi* it provides one of the two DNA binding sites. The EcoligA structure with DNA reveals the conserved interactions of different residues with DNA minor grooves (Nandakumar, Nair et al. 2007).

4. BRCA 1 like C-terminal domain (BRCT)

BRCT domains are present in the tumour suppressor protein BRCA1 (Callebaut and Mornon 1997), and the domain is identified in over 40 other proteins, specifying a superfamily that includes DNA ligase III-a and the essential human DNA repair protein XRCC1. It was reported that carboxy termini (C-terminal) of DNA ligase III-a and XRCC1 is responsible for their interaction. The C terminal was loacted close to the region containing the BRCT domain (Nash, Caldecott et al. 1997). The *T. fi* structure depicts the BRCT domain with four stranded parallel beta sheet flanked by three alpha helices. There is a lacuna of unanimity in the literature regarding the essentiality and purpose of the BRCT domain of NAD^+-dependent DNA ligases. The fact that a second EcoligA paralog (LigB) and two entomopoxvirus ligases lack BRCT domain but have NAD^+-dependent nick sealing activity (Sriskanda and Shuman 2002), suggests that this structural module is not a limiting necessity for NAD^+-dependent ligation. However our group had earlier reported the effects of deleting the BRCT domain of MtbLigA (Srivastava, Dube et al. 2007), which proves to be deleterious. Several other reports also suggest that deleting the domain may act deleterious to mild or minimal (Cho, Prindle et

al. 2003, Ward, Kim et al. 2006, Gilson, Greer et al. 2012), depending on which LigA enzyme is being investigated.

In this study we further focused on the BRCT domain and found it to be essential not only for the overall ligation activity but also for the protein-protein interaction. Its deletion proves to be lethal as observed in the growth of temperature sensitive *E. coli* GR501 strain complemented with MtbLigA. The *in vitro* assay also shows the abolish of ligation activity in the BRCT deleted mutant of MtbLigA. We further zoomed our studies on the residues involved in the protein-protein interaction based on the conservation. We selected few important residues mainly glycines which were highly conserved and also suggested to mediate the interaction. Their *in vitro* interaction analysis further clarify the exact role played by each residue as well as their complementation in temperature sensitive strain results in the delayed growth or abolish the complete growth.

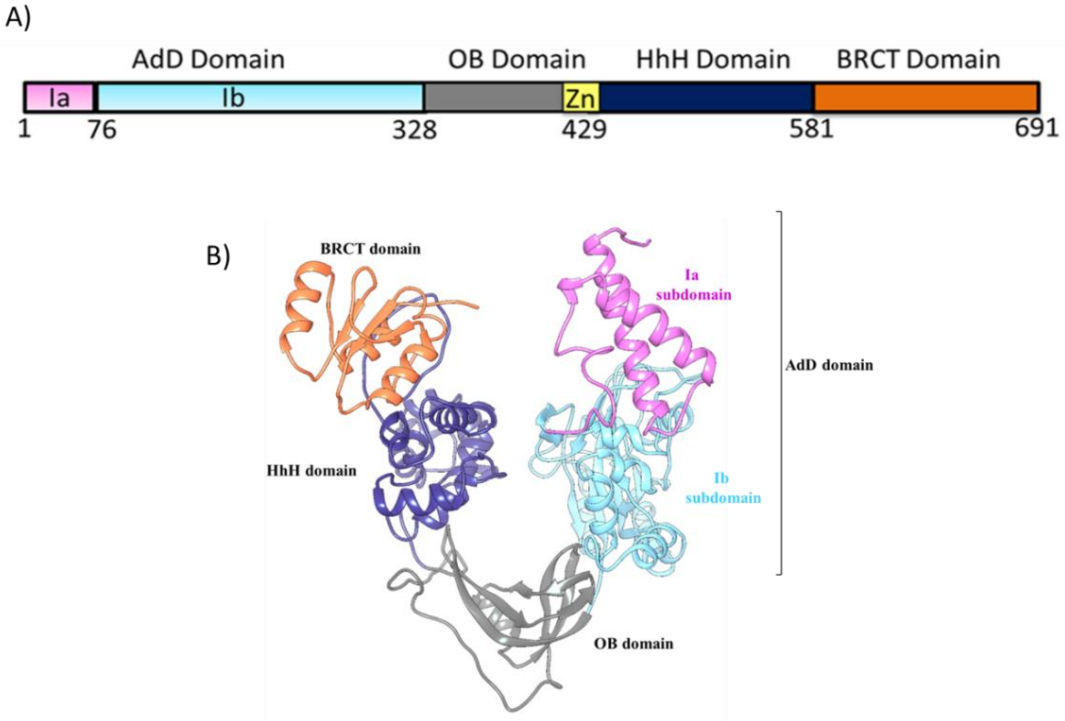


Figure 1.7 Domain organisation of MtbLigA. A) and B) Diagramatic representation of domain architecture of MtbLigA.

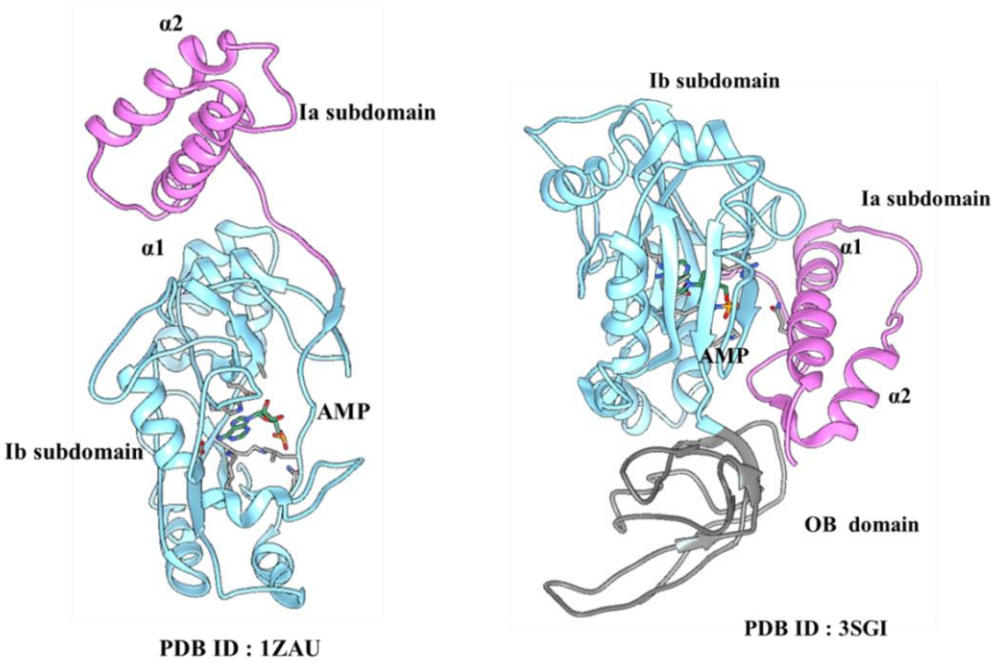


Figure 1.8 Crystal structure of adenylation domain of MtbLigA. Left hand side shows the AdD domain crystal structure with AMP bound in the active site (**PDB ID: 1ZAU**) and right hand side shows the crystal structure of BRCT deleted DNA LigA with Ia subdomain rotated to form 'closed' conformation with AMP in the active site (**PDB ID: 3SGI**).

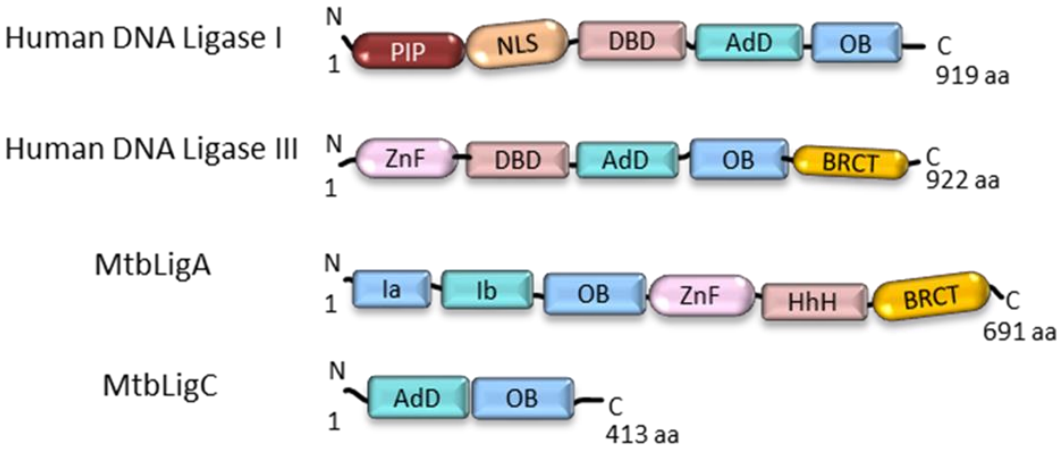


Figure 1.9 Architecture difference in prokaryotic and eukaryotic DNA Ligase. Eukaryotic and mycobacterial DNA ligases have functionally similar domains, but they are jumbled in their linear sequences with other minor differences. AdD is adenylation domain, OB fold is oligomer binding fold, ZnF stands for Zinc finger, HhH represents helix hairpin helix, BRCT denotes BRCA1 C terminal like domain.

1.4 AP endonuclease : a key player of BER and LigA regulator

Class-II AP-endonucleases are the multi-functional, prevalent, rate-limiting enzymes which are extremely vital for the BER pathway. Based on their amino acid sequence homology to *E. coli* endonucleases, these are further divided in two categories: exonuclease III (Exo III) and endonuclease (Endo IV) which are also known as XthA and Nfo families, respectively (Almeida, Soares et al. 2007, Robertson, Klungland et al. 2009). XthA act on duplex DNA as a 3'-5' exonuclease, 3'-phosphodiesterase, 3'-phosphatase and RNaseH (Mol, Hosfield et al. 2000). The 80% of total endonuclease is accounted by the constitituve expression of XthA while Nfo accounts for only 10% of the total endonuclease (Otterlei, Kavli et al. 2000). The crsytal structure of ExoIII and EndoIV had further helped in elucidating the catalytic function of XthA involved in BER. As observed in protein-DNA crystal structure of human and *E. coli* AP-endonuclease (PDB: 1DE8 and 1QUM, respectively) that it electrostatically orients itself onto DNA with preformed DNA binding pocket (**Fig. 2.0 A and B**). The comparative negative potential area is detected in the center of the substrate binding pocket that sets the ion co-ordination site (Harbeck and Peterson 1992, Perkins and Robershotte 1992, Tainer and Friedberg 2000). This orientation of protein adapts non-sequence specific DNA features but is rigid enough for the catalytic centre to become functional only when appropriate substrate come across.

The multi-operational activity of XthA is responsible for sustaining various cellular functions. Its presence is not at all essential for the cell survival, but it is still present ubiquitously in all organisms (Hadi, Coleman et al. 2000, Tell, Fantini et al. 2010). It has been observed in both prokaryotes and eukaryotes that deletion or knock-down of AP endonuclease family proteins results in enhanced sensitivity to DNA damaging agents. It was observed in *E. coli* xthA^{-} knockout strain that XthA deficient strain was found to be exceptionally sensitive to hydrogen peroxide with 20 times higher mortality than the wild type bacteria (Demple and Halbrook 1983). Later it was also found that mutations in mycobacterial AP endonuclease (XthA) compromises the survival of mycobacterium within its host (Puri, Reddy et al. 2014). The nfo mutant of E. coli demonstrated increased sensitivity to alkylating agents like methyl methanesulfonate (MMS), mitomycin C and to oxidants like tert-butyl hydroperoxide and bleomycin but the sensitivity was not comparable to that of xthA mutants. However, the xthA/nfo double mutants were more sensitive to MMS

than either of the single mutants. The xthA/nfo double mutant was found to be unusually sensitive to the lethal effects of gamma ionising radiations. In the present study we also tried to check the essentiality of MtbXthA during the exposure to DNA damaging agents like hyrogen peroxide which exerts the oxidative stress. We checked the lethal dose (LD_{50}) of hydogen peroxide which subsides the growth of *E. coli* XthA knockout strain complemented with MtbXthA to half. To further studies are still going on to zoom in the role of MtbXthA in presence of other DNA damaging agents like MMS, etiopside and camptothecin.

AP-endonucleases hydrolyzes the phosphodiester bond on the 5' end of the abasic site, generating a 3' hydroxyl (3'OH) terminus for extension by a DNA polymerase and a 5' terminus with a 2'deoxyribose 5'phosphate residue (5'dRP) to be removed by an AP lyase. AP-endonucleases functionally interact with several BER components to regulate the pathway as shown in eukaryotes (Thakur, Dhiman et al. 2015). Human APE1 is the counterpart of XthA, and its knockout has been shown to be lethal in mice (Wilson, Takeshita et al. 1997). It stimulates the activities of DNA glycosylase (Kladova, Bazlekowa-Karaban et al. 2018) and FLAP-endonucleases (FEN1) (Dianova, Bohr et al. 2001) for more efficient BER initiation. APE1 also modulates AP lyase and polymerase activities of DNA polymerase β (Liu, Prasad et al. 2007). Further, AP-endonuclease interacts with proliferating cell nuclear antigen (PCNA), suggesting that it has a coordinating role in BER (Kiyonari, Tahara et al. 2009). These activities are drawing attention to its potential as an emerging target for various human diseases. Our lab previous study showed that β-clamp interact with MtbXthA and stimulates the affinity for its substrate (Khanam, Rai et al. 2015). The β-clamp acts as a scaffold protein that recruits other protein participants to the site of damage, we therfore theorize that the β-clamp gets lodged in the damage site with MtbXthA till an AP-Lyase (as yet unidentified in *M. tb*.) enters to remove the 5'-dRP moiety. This would ensue in the freeing of MtbXthA from the damage site and alleviate the activity of DNA polymerase and ligase to take up their conventional roles. Since the *in vitro* effect of MtbXthA-MtbligA interaction was already studied in our lab, so we focused on the strcutural characterization of this bicomplex alone as well as in presence of nicked ds-DNA substrate, which is the subtrate for ligase activity. It was observed that this interaction abolishes ligation activity on nicked DNA substrate and also the product of XthA activity. It was seen in the case of APE1 that it enhanced the DNA Lig I activity against nicked DNA substrate (Ranalli, Tom et al. 2002). The main global difference between APE1 and XthA lies in an ~60-residue N-terminal extension in APE1 compared to XthA (**Fig. 2.1 A and B**). APE1 is known to interact with

other proteins like the XRCC1-DNA Ligase III complex, via residues lying in its N-terminal disordered segment (Vidal, Boiteux et al. 2001). The authors hypothesized that the binding of DNA ligase I to THF-containing substrates is blocked by APE1 while it permits access to nicked DNA substrates, thereby enhancing the efficiency of eukaryotic BER overall and at the same time prevents unwanted 'futile' ligation. So it will be interesting to see how the structural studies justify the *in vitro* results to come out with the conclusion of preventing futile ligation of LigA by XthA during the BER process (Kumar, Bharti et al. 2011). On the other hand, the SAXS derived structure of the MtbLigA-MtbXthA complex ruled out the PIP motif in interactions and instead revealed an interaction interface with significant sequence conservation in MtbXthA.

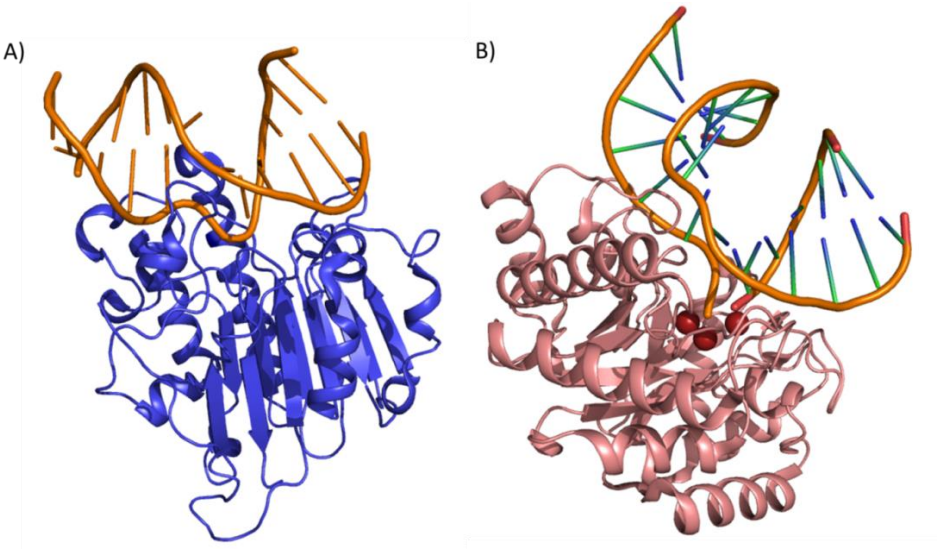


Figure 2.0 AP-endonuclease binding to damaged DNA. **A)** Crystal structure of human apurinic/apyrimidinic endonuclease-1 (ape1) bound to abasic DNA (PDB ID: 1DE8). **B)** crystal structure of *Escherichia coli* endonuclease iv in complex with damaged DNA (PDB ID: 1QUM).

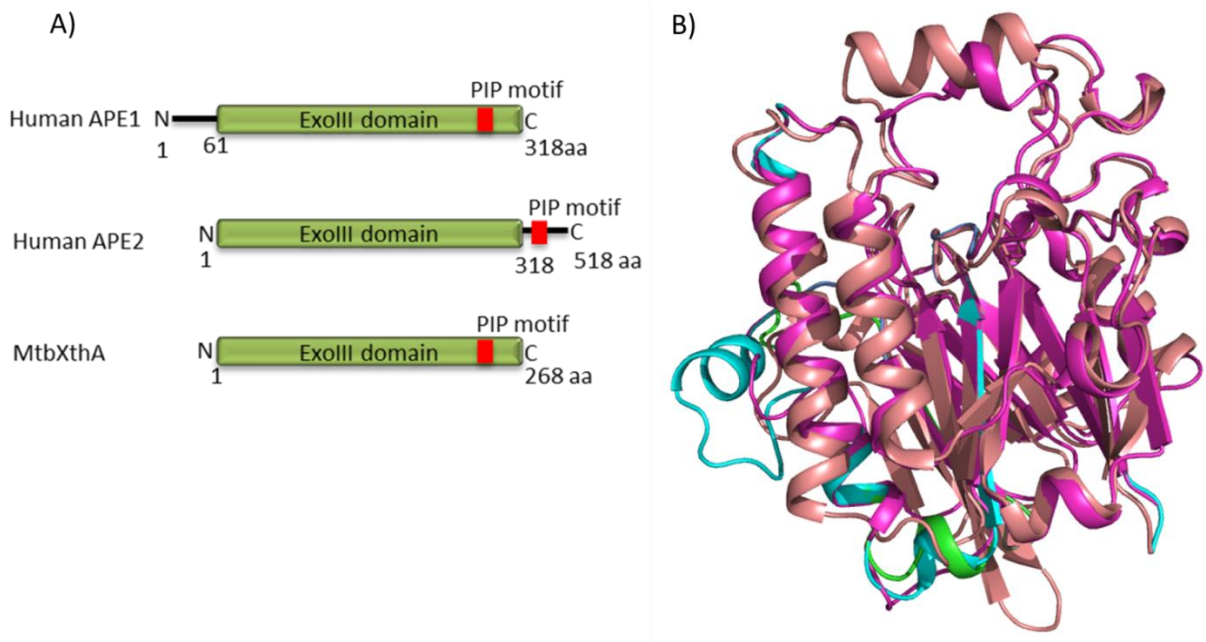


Figure 2.1 Structural comaprison of domain organisation of eukaryotic and prokaryotic AP-endonuclease. A) Architectural difference of prokaryotic and eukaryotic AP endonuclease. Schematic representation to show the structural differences between eukaryotic and bacterial class II AP endonucleases. PIP stands for protein interaction peptide motif that mediates interaction between AP endonucleases and sliding clamps. **B)** Superimposition of various prokaryotic AP endonucleases homologs. *Mycobacterium tuberculosis* XthA (green), *Mycobacterium avium* XthA (cyan), *Mycobacterium bovis* (magenta), *Mycobacterium smegmatis* (yellow) and *E. coli* crystal structure (PDB ID: 1AKO) (salmon) were superimposed.

1.5 BERosomes : role of protein-protein interactions in coordination of BER

Direct physical interactions have been showcased for many proteins participating in BER, and their binding sites have been localized in structural domains (Moor and Lavrik 2018). These direct physical interactions are obeserved in human repair proteins like XRCC1 and DNA glycosylase NEIL1 which interacts with the other short patch (PNKP, Polβ, and LigIIIα) as well as long patch (PNKP, Polδ, FEN1, and LigI) BER proteins. The base excision repair (BER) system evolved in such a way so as to correct the most plethoric DNA damages in cells which is extremely important for maintaing the genomic intergrity. The several enzymes and factors are involved in this multistep pathway which all functions in the coordinated mode so that the repair efficiency is ensured. The coordinated mode is helped by the establishment of protein complexes braced via either direct or indirect DNA-mediated interactions. It had been previously observed that the multiprotein complexes are gathered with the participation of a scaffold protein XRCC1 and poly(ADP-ribose) polymerase 1 regardless of the DNA damage. The PARP1 is also the key regulator of Human BER pathway. Depending on the DNA damage there occurs overall change in the dynamics of the

complex formation at the damaged site. In agreement with the 'passing-the-baton' (or 'substratechanneling') model of the BER process, DNA intermediates are processed in each of the two BER sub-pathways and then the DNA intermediate go through one enzyme to another in a organized and co-ordinated fashion (Wilson and Kunkel 2000). Step-by-step coordination of BER is facilitated by multiple protein–protein interactions (Moor and Lavrik 2018) that involve BER enzymes and accessory proteins. Mostly the biochemical and immunological approaches have been used for the characterization of protein–protein interactions between BER enzymes and other protein factors. These qualitative research have led to the recognition of a several interacting partners and also the particular amino acid sequences that is crucial for protein interaction (Moor, Vasil'eva et al. 2015). Further studies are required to check whether BER proteins are coordinated into constitutive complexes or whether they form successive ephemeral assemblies (Poletto, Legrand et al. 2017). The quantitative investigation including thermodynamic and kinetic studies would help understanding the molecular mechanisms of BER effective interactions occuring in coordinated manner. (**Fig.2.2 A and B**)

The protein-protein interactions are vital for various cellular processes. It is observed quite often that the behavior of individual proteins gets inflected/ regulated when they act as part of complexes. Other important protein interactions governing the BER pathway identified so far include particular complex between DNA glycosylase and FLAP-endonuclease (FEN-1) (Hegde, Theriot et al. 2008), DNA polymerase and DNA ligase (Prasad, Singhal et al. 1996), and between different BER components and XRCC1 (Caldecott, McKeown et al. 1994, Nazarkina Zh, Khodyreva et al. 2007) and PCNA (Gary, Kim et al. 1999, Matsumoto 2001, Xia, Zheng et al. 2005) etc. These findings of complexes manifest that the various steps of BER are regularized and coordinated by the action of multiple protein-protein interactions. The various components of BER machinery receive DNA substrates and relay the DNA products consecutively to the next appropriate active enzyme. This is necessary for preventing the exposure of DNA intermediates to the damaging nucleases action, recombination events, etc. While several crucial participants of mycobacterial BER, like DNA glycosylase including MtbNei (Sidorenko, Rot et al. 2008, Lata, Afsar et al. 2017), AP-endonuclease family protein MtbXthA (Abeldenov, Talhaoui et al. 2015, Khanam, Rai et al. 2015, Khanam, Shukla et al. 2015), DNA polymerase (Arrigo, Singh et al. 2002, Gu, Li et al. 2016), the DNA sliding β-clamp (Gui, Lin et al. 2011, Kukshal, Khanam et al. 2012) and NAD^+ dependent DNA ligase (MtbLigA) (Srivastava,

Dube et al. 2005, Srivastava, Tripathi et al. 2005), have been studied, the protein-interaction network in mycobacterium remains largely unexplored.

Our group is characterizing prokaryotic BER complexes and their roles in the overall DNA repair process. Recently, our lab has identified DNA-mediated interactions between the Mtbβ-clamp and AP endonuclease MtbXthA (Khanam, Rai et al. 2015) (**Fig. 2.4 A and B**). AP endonuclease is the key component of BER pathway, it plays the role of mediator and coordinator in various other steps of the pathway by interacting indirectly or directly with other protein components of the BER pathway and also regulate their enzymatic activites (Borjigin, Arenaz et al. 2012). The DNA sliding clamp also known as beta clamp (βclamp) in bacteria, are toroidal scaffold proteins that act as a processivity factor for DNA polymerases (Wood 1996, Matsumoto 2001, Gu, Li et al. 2016). The eukaryotic homolog of beta clamp is known as PCNA. The clamp has the spectacular ability to interact with multiple partners involved in replication, repair, translesion DNA synthesis etc. and also accelerate their efficiency. In organisms like *P.furiosus*, humans and yeast several studies have showed that their exist physical and functional interaction between AP endonucleases and PCNA clamp and PCNA stimulates the 3'-5' exonuclease as well as 3' phosphatase activities of AP endonuclease (Dianova, Bohr et al. 2001, Kiyonari, Tahara et al. 2009, Roldan-Arjona, Ariza et al. 2019) (**Fig. 2.3 A and B**). Surprisingly PCNA exerts no effect on the endonuclease activity of AP endonuclease. This interactions of PCNA with AP endonuclease resulted in overall increased efficiency of BER pathway. It was also oberved in *in vitro* studies that Ape 1 interacts with DNA ligase and FEN1 (flap endonuclease) and coordinates their activities which results in enhancing the abilities of these two in processing of their substrates for right step by step advancement of the repair pathway (Henricksen, Veeraraghavan et al. 2002). AP endonuclease was also found to enhance the ligation activity of human DNA ligase thus increasing the overall efficiency of BER pathway (Ranalli, Tom et al. 2002). An earlier studies (Moor and Lavrik 2018) about the interactions between human APE1 and DNA Ligase I encouraged us to lookup for similar interactions, if they subsist, in prokaryotes. It fascinated us on account of the large deviations in sequence and domain arrangement between human DNA Ligase I and bacterial LigA. Then again, the domain organization of XthA and APE1 are quite similar, except for ~60-residue N-terminal extension in APE1.

We were curious in finding precise protein interactions that govern the BER pathway in *M. tuberculosis*. As discussed above the interaction between DNA ligase I and PCNA in

humans are important for long patch BER but the analogous interactions between LigA and β-clamp is absent in *M. tuberculosis* (Kukshal, Khanam et al. 2012) suggests that their exist structural differences in mycobacterial repair proteins which may further adds into the deviations in their functional interactions during DNA repair. Since the hypothesis of direct interaction between MtbLigA and Mtb β-clamp failed but the possibility of their transient interaction cannot be eliminated. So we further speculated that there must be some yet unidentified protein factor which may mediate the interaction between the MtbLigA and Mtb β-clamp. A very recent report also presented that mycobacterial Ligase C (LigC) complex interacts with other proteins of the BER setup (Plocinski, Brissett et al. 2017). These findings suggest that a sophisticated regulation of the BER pathway in *Mycobacterium tuberculosis* takes place by formation of multimeric protein complexes at the damaged site. In the present study we demonstrate the direct physical interaction between MtbLigA and MtbXthA by using size exclusion chromatography (SEC), surface plasmon resonance (SPR) and small angle X-ray scattering (SAXS) by using purified recombinant proteins. The MtbLigA and MtbXthA are involved in the terminal and initial stages of bacterial BER respectively and it is quite intersting to find out the how they interact with each other. The complex formation between the two proteins is mediated specifically by a newly identified motif in MtbXthA that interacts with the BRCT domain of MtbLigA. As discussed in previous section that the BRCT domain is implicated to act as a signal transducer element by relaying the signals from damage detectors to the other components of DNA repair machinery via specific protein-protein interactions (Bork, Hofmann et al. 1997, Lim, Park et al. 2001). The functional aspect of the interaction reveals that XthA inhibits the LigA activity. We also have shown earlier that LigA can efficiently act on the product of XthA activity to ligate it. Therefore our results suggest that XthA engages with the C-terminal BRCT domain of LigA to prevent 'futile' ligation of the former's product till an AP-Lyase is recruited to remove the 5'-dRP moiety. These findings offer the mechanistic insights into the role of MtbXthA in preventing MtbLigA from resealing the product created by the action of MtbXthA at abasic sites. Thus we concluded that since the MtbXthA directly interacts with the Mtb β-clamp and MtbLigA, although the latter two donot interact with each other, therefore acting as a bridging molecule between the Mtb β-clamp and MtbLigA. The formation of tri-complex (Mtb β clamp-XthA-LigA) at the damaged site may act as a check on the futile cleavge and ligation cycle in the BER process which if left unattended may cause the derailing of the pathway.

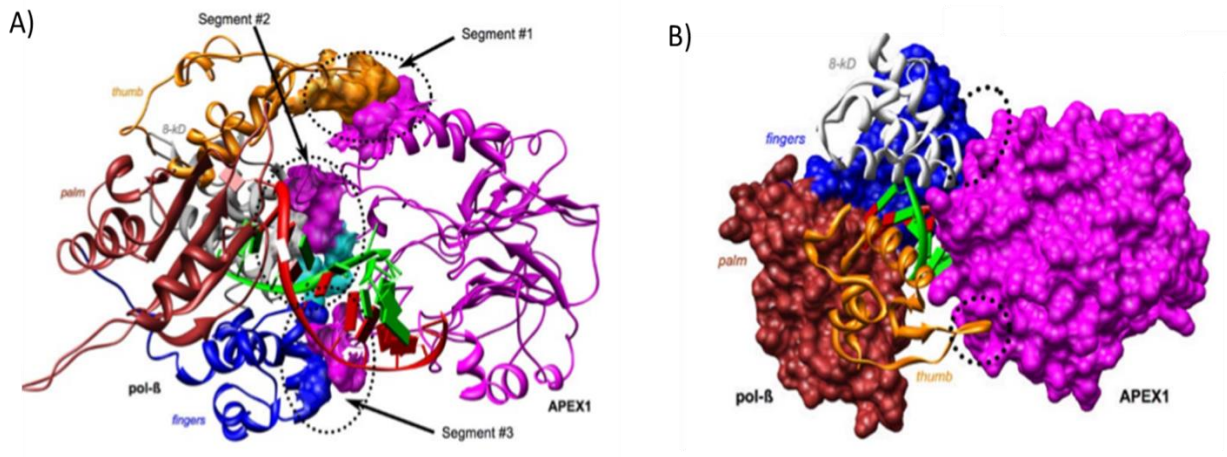


Figure 2.2. The molecular model of AP Endonuclease with partner protein. **A)** Initial 3' complex of APEX1 and pol-β (closed conformation). View of the 3' complex structure. APEX1 is on the right and pol-β is on the left. The area of protein-protein interaction is circled. **B)** Molecular model of human AP Endonuclease APEX1 and DNA Polymerase β complex. The area of protein protein inetraction is circled. Palm, thumb and finger domains of polymerase β are shown in brick, orange and indigo, respectively while APEX1 is shown in magenta. Red and green are the two strands of DNA. Area of interface between the two proteins is circled. Figure adapted from Abyzov et al., 2008.

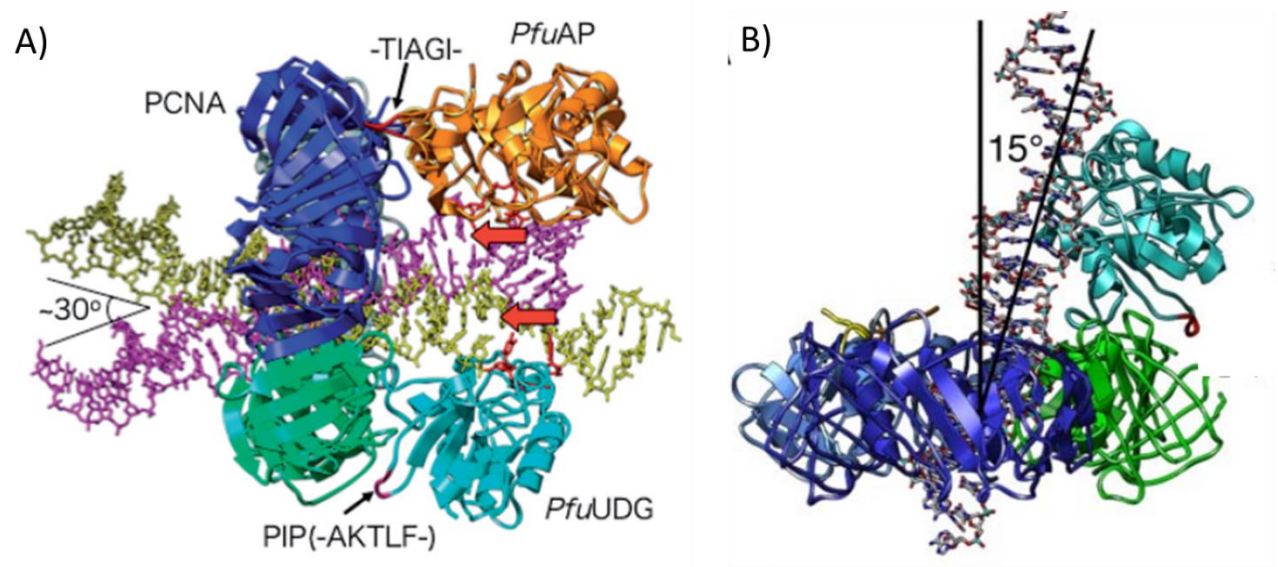


Figure 2.3 The archeal BER complex. A) Molecular model of the *P. furiosus* PfuAPE–PfuPCNA–PfuUDG–DNA complex. The dsDNA bound to PfuAPE is shown in magenta, and that bound to PfuUDG is shown in yellow. The three subunits of PCNA are colored blue, green and gray. The PIP-like motif (152AKTLF156-) of PfuUDG, and the sequence of PfuAP (-23TIAGI27-), which putatively occupies the binding site on another PCNA subunit, are shown in red. Figure adapted from Kiyonari et al., 2009a. **B)** The atomic structural model of the PfuUDG–PfuPCNA–dsDNA complex. When the identified PCNA-binding site in PfuUDG was placed in the PIP-binding site of PfuPCNA, the Thermus thermophilus UDG-bound dsDNA passed through the PCNA ring with a 15◦ inclination from the 3-fold axis of PfuPCNA. Figure adapted from Kiyonari et al., 2009b.

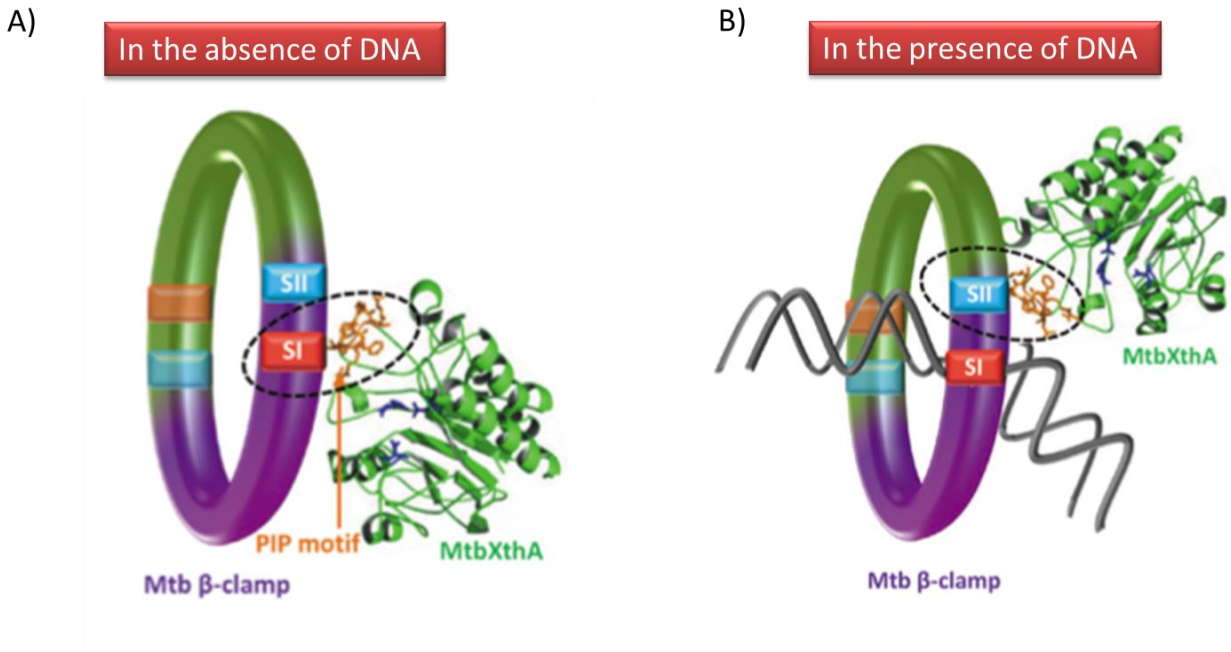


Figure 2.4 Sliding DNA β-clamp exhibits DNA regulated modes of interaction with Class II AP-endonuclease/3'-5' Exonuclease III (XthA). A) In the absence of DNA substrate, the MtbXthA (green ribbon) interacts with the peptide binding groove (PBG) constituted by Subsite I (SI-red box) via PIP motif (orange sticks). B) In the presence of DNA substrate, the mode of interaction changes from subsite I to C-terminal subsite II (SII-blue box). Figure adapted from Khanam et al., 2015.

1.6 Exploiting the mycobacterial BER as drug target

The slow growing aerobic and non-motile mycobactreium tuberculosis (*M. tb*) is the causative agent of deadly disease tuberculosis. The characteristic feature of the bacilli is that beyond the peptidoglycan layer the bacteria have a unique cell wall that has an over constituted content of polysaccharide, lipid, glycolipid and mycolic acids (Kolattukudy, Fernandes et al. 1997). The unusal slow growth rate allows the bacteria to invade with in the host during the latent phase which is also known as dormant phase. During dormant phase the mycobacterium dwells in an inactive state within the infected cells and the metabolic activities shuts down so to maintain the dormant state of bacteria. Dormancy is imposed by the ineffectual execution of host immune response which is able to cut short infection but do not eradicate it entirely. There are chances that dormant bacteria may gets activated during the days of compromised immunity of the host (Ehlers 1999). This persistency of the bacterium causes to become a long lasting nature to the infection caused by *M. tb*. Over the course of time, several mutant strains of the bacterium have evolved that can efficiently resist the pressure caused by drug treatment in the patients suffering from tuberculosis. This has

resulted in the emergence of extremly drug resistant strains that are resistant to the primary-line of anti-TB drugs. These are termed as DR-TB (drug resistant-TB). Multi-drug resistant (MDR) strains show resistance to RIF and INH and xtreme drug resistant (XDR) stains show additional resistance against any of the second line of injectable drugs (SLID: kanamycin, amikacin and capreomycin) and also any fluoroquinolone (Tyedmers, Madariaga et al. 2008). In 2017, 558,000 new cases of MDR-TB and 8.5% of MDR-TB cases had extensively drug resistant TB (XDR-TB) indicates the magnitude of the problem of drug resistance (WHO, 2018). XDR‑TB require the use of third-line of anti‑TB drugs, which are high-priced and frequently have greater side effects than first- or second-line drugs (Gandhi, Moll et al. 2006). The other problematic class of highly drug resistant TB is total drug resistant (TDR-TB) caused by the *M. tuberculosis* strains that are resistant to all available first and second lines of TB drugs (Udwadia and Vendoti 2013). The mycobacteria survives in the harsh oxidative enviornment of the host cell. Its genomic intergrity is maintained and protected by the dedicated DNA repair machinery which is the superior cause of its survival within the host. The whole genome sequencing of various mycobacterial strains has helped us to gain the knowledge of proteins involved in the DNA repair mechanism of this organism (Cole, Brosch et al. 1998). Several reports and experiments has suggested that mycobacterial DNA repair process exist quite differently from other bacteria like *E. coli* and *B. subtilis* (Dos Vultos, Mestre et al. 2009). Although the *M. tuberculosis* genome show high conservation of genes encoding proteins homologous to those involved in *E. coli* base excision and nucleotide excision repair (Mizrahi and Andersen 1998). The protein components constituting the BER and NER pathways in *M. tuberculosis* are enlisted in **Table 1.1** and the schematic representation of the pathways are shown in **Fig. 1.5 and 1.6** respectively. Since the absence of operational mismatch repair (MMR) system in mycobacterium (Springer, Sander et al. 2004), the load of DNA repair get on primarily on BER and NER. In light of high GC content of mycobacterial genome and absence of MMR, the BER and NER pathways come forward as chief candidate responsible to repair oxidative damage of DNA because these two pathways are fundamentally needed for the survival of the microorganism within the unfriendly surroundings of macrophages, they can be useful as effective drug targets. Many studies had demonstrated therapeutic targeting of the mycobacterial NER pathway (UvrABC complex) by a chemical inhibitor 2-(5-amino- 1, 3, 4-thiadiazol2-ylbenzo[f]chromen-3-one (ATBC) at micro molar concentrations (Mazloum, Stegman et al. 2011) and also the BER pathway by targeting glycosylase, endonuclease and DNA ligase. However, further findings

are needed to exploit excision repair system for therapeutic intervention with better inhibition properties at different damaged DNA substrate.

1.6.1 Evaluation of MtbLigA as a drug target

LigA, highly conserved among bacteria, is essential for DNA replication. For the past fifteen years LigA had been treated as a broad-spectrum novel antibacterial target because of its essentiality and marked structural difference with eukaryotic ligases (Tomkinson, Howes et al. 2013). It has especially attracted attention as a target in case of mycobacterium with the ability to battle against multiple drug resistance (Brotz-Oesterhelt, Knezevic et al. 2003, Srivastava, Dube et al. 2005, Srivastava, Dube et al. 2007). The structural studies of several DNA ligases has revealed the mechanism of cofactor and substrate binding and the whole catalysis of ligation reaction (Unciuleac, Goldgur et al. 2017). The structural studies along with the wide range of techniques covering medicinal chemistry, protein biochemistry, pharmacology and molecular biology have prompted intellectual therapeutic or inhibitor design and thus recognition of novel possible antibacterials. The available inhibitors from past fifteen years are against the co-factor binding site of LigA. Ongoing research is to focus on the development of inhibitors which can distuinguish between the ATP-dependent and NAD^+-dependent DNA ligases. Our group had been using the combination of *in silico* screening, bactericidal and *in vitro* assays to identify several classes of inhibitors against LigA. These inhibitors can distinguish between ATP and NAD^+dependent DNA ligases with several folds. Our group had earlier identified novel classes of DNA LigA inhibitors that include Glycosylureids, Glyocsylamines (Srivastava, Tripathi et al. 2005), N-substituted tetra cyclic indole derivatives (Srivastava, Dube et al. 2007), dispirocycloalkanones and (Kukshal, Khanam et al. 2012) hydroxamates. All these inhibitors had competitive mode of action as they compete with NAD+ to bind at the active site. These results were further cross checked by the protein-inhibitor docking in which inhibitor bind at the AMP binding pocket and the aromatic ring of inhibitor mimics the adenosine moiety of NAD^+ while the bulky aliphatic or aromatic moiety which protrudes out from the ring system permeates into the "inhibitor binding" pocket (**Fig. 2.5**) (Yadav, Khanam et al. 2015). These findings suggets the nucleotide metabolism of bacteria as appealing scheme for developing inhibitors.

In *E. fa* the crystal structure (PDB ID: 1TA8) (**Fig. 2.6 Right and Left**) and our present crystal (PDB ID: 6KJM, 6KSC, 6KDU and 6KRH) structures bound with NMN in the Ia subdomain of adenylation domain within the NMN binding pocket opened the new

avenue to target the different site for developing new range of inhibitors. The emerging MDR and XDR strain needs to be tackled with new promising startegy. Therefore we focused our interest on the NMN binding pocket rather than the active site. The Ia subdomain recognises the NAD^+ cofactor by its NMN part and then Ia subdomain rotate around to deliver the cofactor to the catalytic site present in Ib subdomain.

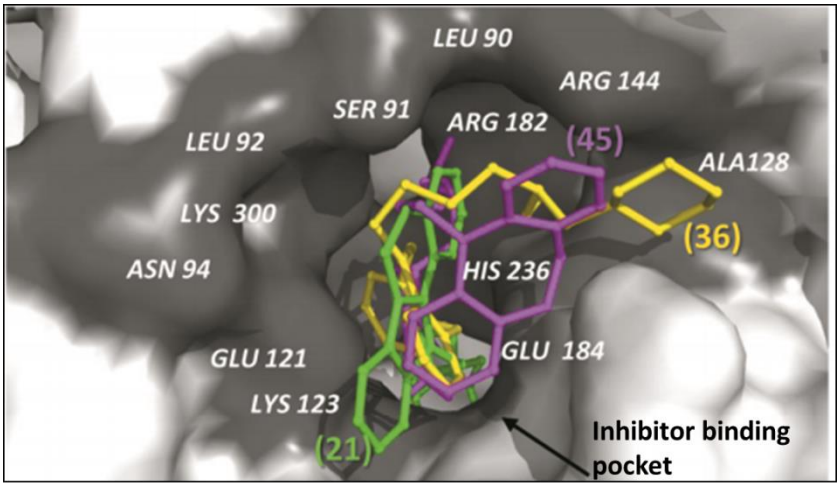


Figure 2.5 Surface representation of DNA Ligase showing inhibitor tricyclic dihydrobenzoxazepine and tetracyclic indole derivatives (pale Green, yellow and magenta) sticks docked into the binding cleft of the protein. The protein has surface representation shown in white and dark grey denotes the substrate binding cleft (Inhibitor binding pocket). Figure adapted from Yadav et al.,2015.

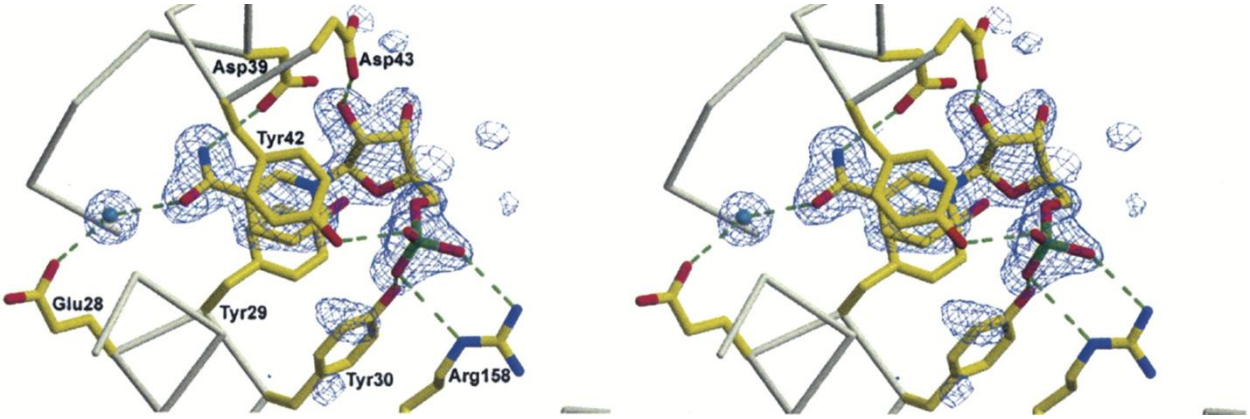


Figure 2.6 The NMN binding by the *E. fa* Ligase shown by stereo view. The interaction between the side chain and NMN are shown by the dash lines. The water molecule is represented by sphere which coordinates between the NMN and protein. The image is adapted from Gajiwala, Pinko et al., 2004.

The NAD^+ then hydrolyses in to AMP and NMN. The AMP gets transferred to enzyme followed by DNA and then NMN get realeased and the ligation is completed. Therefore if the

NMN site is blocked the further three step ligation cycle will automatically get blocked. We did the *in silico* screening of the BioNet fragment library and finally got three fragments which stamp down the overall ligation activity. The crystal structure of adenylation domain with the fragment at NMN binding site (PDB ID: 6KKV) throws the light on the new approach towards developing biotherapeutics.

1.6.2 Evaluating MtbXthA as drug target

AP endonuclease is a fundamental protein within the BER pathway as it works on the huge bulk of AP sites in the cell (Tell, Fantini et al. 2010). Among the human BER components, APE1 has been thoroughly studied for its wide range of cellular functions. In addition of its role as AP endonuclease, APE1 is also known to act as transcription regulator. It performs the redox mediated modification of transcription factors (Demple and Halbrook 1983). Its role in RNA metabolism is yet to be fully elucidated (Carpenter, Corbett et al. 2007). AP endonucleases cleaves the DNA abasic sites resulted by the activity of glycosylases. Due to the requirement of AP endonuclease after the action of almost all DNA glycosylases, and also the ability to repair the replication-blocking abasic sites generated by the alkylating agents, this makes it to be a central player of the pathway. Several preclinical and clinical studies suggested that APE1 can be an appealing drug target for anticancer drug development. CRT0044876 was discovered as a potent and selective APE1 inhibitor. CRT0044876 is particularly inhibitor of the exonuclease III family of enzymes to which APE1 also belongs. It not only inhibits the AP endonuclease but also slows down the 3'-phosphodiesterase and 3'-phosphatase activities of APE1 at low micromolar concentrations (**Fig. 2.7 A and B**) (Madhusudan, Smart et al. 2005). The report had suggested that APE1 imparts the resistance to diverse range of inhibitors and drugs (Thakur, Dhiman et al. 2015).

AP endonucleases are the rate limiting enzymes of BER pathway. They do not merely act as component of BER but also supervises the pathway by interacting with number of component proteins. Our lab had characterized the AP endonuclease homolog of *M. tuberculosis* (Rv0427c) and also determined the factors involved in mediating the substrate recognition. Inspite of the fact that several protein partners of AP endonucleases have been reported from higher eukaryotes but not many studies have been carried out yet on how bacterial AP endonuclease coordinate the BER. In this study we aim to explore the role of *M. tuberculosis* AP endonuclease homolog (MtbXthA) in regulating the intricate protein interaction network constituting the BER pathway. We also aim to target the central player

of the pathway as a novel tool for therapeutic interventions. We checked the essentiality of MtbXthA during the oxidative stress by complementing in *E. coli* XthA knockout strain. Then we also screened the BioNet fragment library and MMV pathogen box against the MtbXthA. After in silico screening, based on the docking scores we checked the *in vivo* effect of these compounds by observing the inhibition of overall growth of bacteria under stress condition. The further studies are going on which includes MtbXthA knockout strain preparation in *M. tb.* H37Ra and then screening of compounds in that strain. Thus this approach may act as a new way to target the mycobacterial BER apparatus.

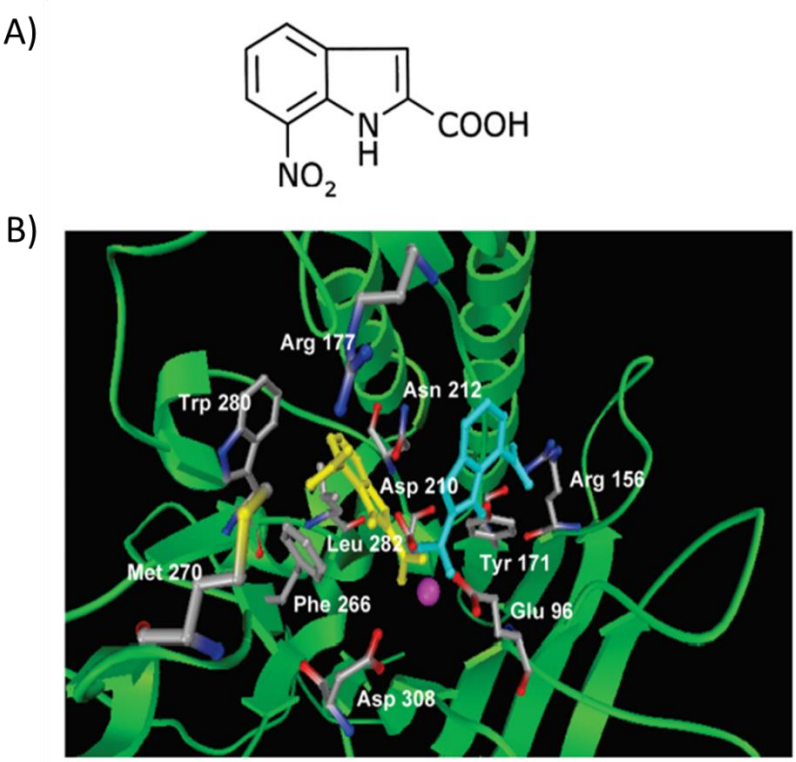


Figure 2.7 Docking of chemical CRT0044876 to the APE1 active site. **A)** CRT0044876 is a chemically known as 7-nitro1H-indole-2-carboxylic acid known to dock at the APE1 active site at two different confirmation. **B)** The two different conformation are indicated in yellow and blue. The interacting residues highlighted and are shown in gray. The image is adapted from Madhusudan, Fiona et al., 2005.

1.6.3 Exploiting protein-protein interaction as a drug target

Multi Drug-Resistant Mycobacterium tuberculosis requires the development of new therapeutics with novel mode of action. Base Excision Repair (BER) pathway mends the DNA damage caused by oxidative stress inside host macrophages. The pathway involves several critical protein-protein interactions. However, these interactions involved in BER

remain unexplored. As discussed above the MtbXthA and Mtb β clamp interact with each other in DNA regulated manner. Another such critical interaction, also characterized by our group involves DNA LigaseA & AP endonuclease II (MtbLigA-XthA). We have determined that in the absence of LigA-XthA interactions, apo-LigA captures the product of XthA for ligation and reverses its action resulting in 'futile' ligation. Also there occurs the formation of tri-complex, where XthA play as brigding molecule between Mtb LigA and beta clamp (Mtb β clamp-XthA-LigA), at the damaged site which act as regulator for futile cleavage and ligation cycle. Therefore by strengthening the LigA-XthA interactions or by disrupting the Mtb β clamp-XthA interaction through specifically designed molecules or peptides represents a novel therapeutic discovery approach where the pathway will be conceptually 'locked' and made inactive (**Fig. 2.8**).

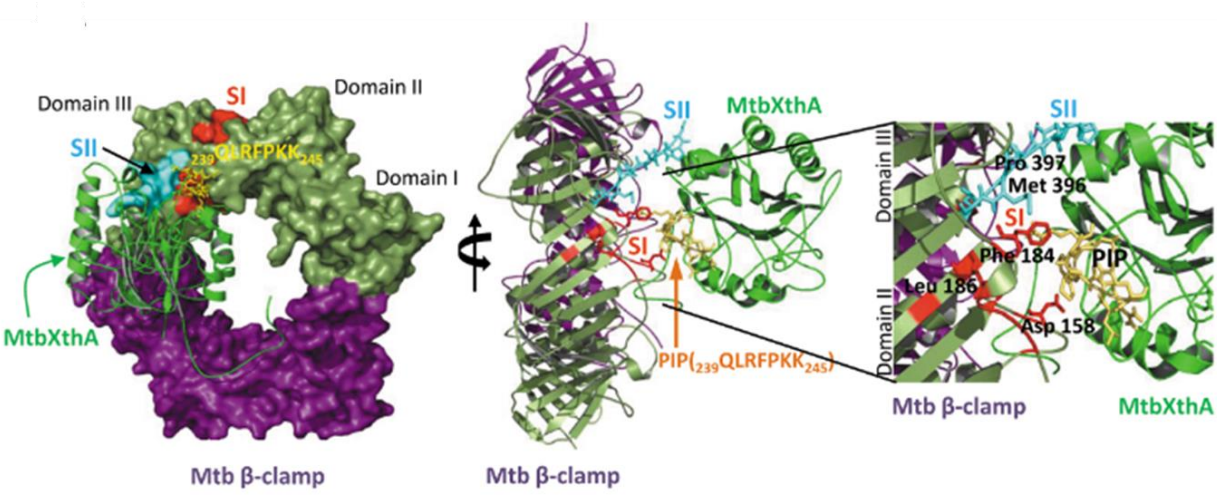


Figure 2.8 The MtbXthA-Mtb βClamp interaction. The β-binding PIP motif in MtbXthA is shown in deep yellow. Two monomers of Mtb β-clamp are shown in pale green and purple. The subsite I and subsite II are shown in red and cyan. The zoomed view shows the interacting resdiues involved.

Chapter 2 | Materials and Methods

Introduction

The materials, experimental methods and approaches during the course of the study are reported in this chapter.

Materials

Company Name	Chemicals procured
M/S Amresco	Isoamyl-alcohol, Potassium chloride, Potassium phosphate monobasic and dibasic, PMSF, Tris base, Chloremphenicol, Kanamycin
M/S Chromous Biotech	Sequencing
M/S Coli genetic stock center, Yale, U.S	*E. coli* Knockout strain of XthA
M/S Fermentas	DNA ladders, Protein molecular weight markers, restriction enzymes and buffers, T4 DNA ligase, dNTPs, T4 polynucleotide kinase, ATP, nuclease free water, IPTG, Taq DNA polymerase
M/S GE Healthcare	glutathione sepharose resin, proteinase K, protein A sepharose, Size exclusion chromatography columns, Mono Q purification coloumns, CM5 chips, SPR buffers including HBSEP+ 10X buffer, 50 mM NaOH, 50 mM Glycine pH 2.0, Bia desorb solution 1 and 2.
M/S Himedia	Agar powder, Luria Bertani broth, tryptone, yeast extract
M/S Integrated DNA technologies (IDT)	PCR primers, FAM labeled oligos , Sequencing
M/S Macherey-Nagel (MN)	DNA gel extraction kit, Plasmid extraction kit
M/S Millipore	Protein concentrator (Centricon), sterile disk filters
M/S New England Biolabs (NEB)	Q5 high fidelity polymerase
M/S Novagen	pET expression plasmids
M/S Qiagen	Ni-NTA Fastflow, Qiagen plasmid mini kit
M/S Qualigens	Acetic acid, acetone, chloroform, hydrochloric acid, isopropanol, methanol, hydrogen peroxide, sodium hydroxide
M/S Santa Cruz	His-probe and GST-probe antibody, goat anti-mice HRP conjugate secondary antibody
M/S Sigma	Agarose, ATP, acrylamide, ammonium sulfate, ammonium persulfate, Bradford reagent, bisacrylamide, BSA, boric acid, bromophenol blue, calcium chloride, coomassie brilliant blue R, DMSO, Diaminobenzidine, DTT, dialysis tubing,

	dialysis tubing closures, disodium hydrogen phosphate, EDTA, ethanolamine, ethidium bromide, formamide, Freund's adjuvant (complete and incomplete), glutaraldehyde, glycerol, glycine, goat antimice HRP conjugate, goat anti-mice HRP conjugate, HEPES-Na, hydrogen peroxide, imidazole, Igepal (Nonidet P-40), lysozyme, magnesium chloride, magnesium acetate, β-mercaptoethanol, nickel(II) sulfate hexahydrate, ponceau stain, potassium chloride, protease inhibitor cocktail, potassium chloride, potassium thiocyanate, potassium iodide, RU7, sodium chloride, sodium dihydrogen phosphate, sodium bicarbonate, sodium chloride, sucrose, sodium acetate, silicon dioxide, sodium dodecyl sulphate, sodium deoxycholate, T2AA, TEMED, triethanolamine, tris-saturated phenol, tween-20, Trizma base, triton X-100, urea, xylene cyanol, skimmed milk powder.
M/S Takara	His60 superflow resin
M/S Tarsons	90 mm Petri dishes, 50 ml and 15 ml falcon tubes, ELISA plates, round bottom polypropylene centrifuge tubes, oakridge tubes.
M/S Thermo Scientific	Oregon maleimide and Alexa fluor-555protein labeling kits, Phusion^{TM} high fidelity polymerase
M/S VWR	Ampicillin

Note: Some specific chemicals are mentioned in the respective chapters.

Methods

2.1 Bioinformatics analysis:

2.1.1 Molecular docking and *in silico* screening of inhibitors against MtbLigA and MtbXthA

The Crystal structure (PDB: 1ZAU) was minimized by using Tripos force field [3] with 500 steps of steepest descent algorithm followed 500 steps of Powell method and 50.00 kcal/mole convergence threshold energy (custom parameters). The flexible docking was performed by Surflex-Dock [4-5] Programme available in SYBYL-X v2.1 [6]. The chosen protein NMN binding site selected to prepare grid using Protomol module. The 2nd Generation BIONET

premium fragment library (1166 compounds) purchased from Key Organics Ltd. (https://www.keyorganics.net/) was screened against defined NMN binding site. The energy minimized MtbXthA molecular model was used for screening of BIONET fragment library and MMV pathogen box against the catalytic site. The Surflex-Dock program available in SYBL-X v2.1 was used for flexible docking.The each active hit was selected by generating 20 maximum poses and 20 maximum conformations per fragment of protein. The final pose of hits with highest docking score were selected which calculated using scoring functions i.e. G-Score, D-Score, PMF-Score and Chem-score.All visualization analysis was done by PyMol v2.2 (https://pymol.org/2/) visualization software.

2.1.2 Homology modeling and protein-protein docking

The homology model of BRCT domain of MtbLigA was generated by using Phyre2, using the structure of the BRCT domain from Human replication factor C large subunit 1 (PDB ID: 2EBU) as a template, which has 99.9% confidence and 99% coverage. For BRCT domain-peptide docking, the minimized model of BRCT domain was used to evaluate interactions with peptides (i) DVR (DVRVGFDGQP) (ii) DGQ (DGQPSWSGKP (iii) YDV (YDVAHVGFD), by using CABS-dock at default parameters and without restricting any binding pocket. CABS-dock allows for large-scale flexibility in peptide and protein to screen the occupying niches with in the protein. 10 docking models were generated and analysed. The best one was selected based on occurrence in docking trajectory. The selected model was analysed by PDBsum to generate the 2D docked conformation of peptide.

Protein-protein docking was carried out using ClusPro (https://cluspro.org) a protein-protein docking webserver. The molecular model of MtbXthA and BRCT domain were used to generate the protein–protein complex. The program generated the presumed complexes by scanning the rotational and translational space between the two proteins, where rigid-body docking was done using the Fast Fourier Transform (FFT) correlation. Among all the complexes generated the best model was used for further analysis. The ClusPro 2 weighed score of the docking model was calculated by the equation:

$$E = 0.40Erep + -0.40Eatt + 600Eelec + 1.00EDARS$$

where Erep and Eatt stands for repulsive and attractive contributions to the van der Waals interaction energy, Eelec is electrostatic energy term and DARS stands for the development of potentials with decoys as the reference state (Kozakov et al., 2006).

2.2 Molecular cloning

All the PCR amplifications, restriction digestions, ligations and transformations were carried out by standard procedures as per Sambrook and Russell, 2001.

2.2.1 PCR

PCR amplifications were carried out using DNAzyme DNA polymerase for routine applications. Q5 high fidelity polymerase was used to amplify gene fragments for cloning into expression vector and for site directed mutagenesis. All the PCR reactions were carried out by standard procedures (Sambrook and Russell, 2001) using genomic DNA or plasmid DNA as template in a PCR System (Eppendorf). The reaction conditions and primer sequences (with restriction enzyme sites highlighted) for all amplified gene segments are given in **Table 2.1** and **Table 2.2.** The products were analyzed on agarose gel in TAE buffer (Appendix) with DNA ladders of relevant range.

2.2.2 Agarose gel based purification of linear DNA

DNA Fragments were resolved on 1 % agarose gel in TAE buffer and the desired DNA bands were excised from the gel and transferred to a sterile micro tube. The DNA was extracted from agarose gel using DNA gel extraction kit following manufacturer's instructions.

2.2.3 Cloning into expression vector

The genes encoding the BRCT domain of MtbLigA were PCR amplified from genomic DNA using primers that contain restriction enzyme site overhangs (**Table 2.1**). The amplified product was electrophoresed on agarose gel and eluted for restriction digestion following which the digested DNA is again eluted from agarose gel. The digested fragments of interest were then ligated into the multiple cloning site (MCS) of expression plasmid of choice (pET23a) using T4 DNA ligase.

2.2.4 Site directed mutagenesis

2.2.4.1 Cloning of AdD domain mutants in AdD domain and in full length MtbLigA

The point mutations of residues lying in the AdD domain of LigA, E22A, E26A, E87A, K123A, K123R and H236Y were generated using pQE60-AdD and pET21d-MtbLigA as template and appropriate primers (**Table 2.1**) and PCR conditions as mentioned (**Table 2.2**). The products were digested by DpnI to remove the template DNA and transformed in *E. coli* DH5α. The mutant clone was then transformed in to the expression host cell.

2.2.4.2 Cloning of BRCT domain mutants in BRCT domain and in full length MtbLigA

Four point mutations, viz. G614I, G621I, G639V, K648Q/K649Q were generated using pET23a-BRCT and pET21d-LigA as template and appropriate primers (**Table 2.1**) and PCR conditions and PCR conditions as mentioned (**Table 2.2**). Expression constructs were sequenced to verify incorporation of mutations mentioned. The confirmed mutant clone was then transformed into the expression host cell.

2.2.4.3 Cloning of AdD domain and BRCT domain mutants in pTrc99A vector

The site directed mutants of AdD and BRCT domain of MtbLigA were cloned in pTrc99A vector for in vivo studies by using pTrc99A-LigA as a template and PCR conditions and primers used as mentioned in the table. The amplified product was then dpnI digested to remove parent strand and then transformed in *E. coli* DH5α.

2.2.4.4 Cloning of MtbXthA^{ID} mutant of MtbXthA

The interaction motif of MtbXthA with MtbLigA $_{104}$DGQPSWSGKP$_{113}$ was mutated to $_{104}$DGQPSWALAP$_{113}$ and called as MtbXthA^{ID}. The MtbXthA^{ID} mutant was generated using pGEX- KG -MtbXthA as template and appropriate primers (**Table 2.1**).

Table 2.1 List of primers used in the study

MtbLigA	Forward	5’-GGA ATT CCA TGG GCT CCC CAG ACG CCG AT-3’
	Reverse	5’-TCA CAA AGC TTT TAC GTT CGT GAC GCG GGT CC-3’
BRCT domain	Forward	5’-GGA ATT CCA TGG GCT CCC CAG ACG CCG AT-3’
	Reverse	5’-ATC GGA TCC GTC ACG CTC GTC GAC CAT TC-3’
E87A	Forward	5’-CAT CTC GCG CGA ATG CTC AGC CTA GAC-3’
	Reverse	5’-CAT TCG CGC GAG ATG GTC GAC GGG-3’
E26A	Forward	5’-GTG CGT GCA CAC CAG TTC CGT TAT-3’
	Reverse	5’-CTG GTG TGC ACG TAC CTC CTC GGC-3’
E22A	Forward	5’- CTG GCC GCA GAG GTA CGT GAG CAC CAG-3’
	Reverse	5’- CAC CTC TGC GGC CAG AGC CTG CCA CTG-3’
H236Y	Forward	5’-ATC TGC GCA GGA CTG GGC CAC GTG GAC-3’
	Reverse	5’-CAG CCC TGC GCA GAT CAT CCG TAG CCG-3’
K123A	Forward	5’-GAG CTA GCA ATC GAC GGC GTC GCG CTG-3’
	Reverse	5’-GTC GAT TGC GAG CTC ACA CAG GTA ATG-3’
K123R	Forward	5’-GAG CTC AGA ATC GAC GGA GTA GCG CTG-3’

	Reverse	5’-GTC GAT TCT GAG CTC ACA CAG GTA ATG-3’
G614I	Forward	5’-ACG CTG GCC ATC CTG ACT ATC GTG GTC-3’
	Reverse	5’- GAT GGT CAG GAT GGC TAG CGT GCG TG-3’
G621I	Forward	5’- GTC ACC ATA TCG CTG ACC GGT TTC TCC-3’
	Reverse	5’- GGT TAG CGABTAT GGT GAC CAC GAT GG-3’
G639V	Forward	5’- ATC GTG GCA CGC GTC GGA AAG GCA GC-3’
	Reverse	5’- GAC ACC GAG CCG GCG GCC TTT CCG ACG CGT GC-3’
KK648/649QQ	Forward	5’- TCG GTG TCG CAG CAG ACC AAC TAT GT-3’
	Reverse	5’- GAC GAC ATA GTT GGT CTG CGA CA-3’
V646S	Forward	5’- GGC TCT GAG TCC AAG AAG ACC AAC TA-3’
	Reverse	5’- CTT CGA CTC CGA GCC TGC AGC CTT GCC-3’
GS660/661LA	Forward	5’- GAC TCC CCG TTA GCC AAG TAC GAC AA-3’
	Reverse	5’- ACC AAC TAT GTC GTC GCC GGA GAC TC-3’
MtbXthA^{ID}	Forward	5’- AGT TGG GCA CTA GCA CCG GAA GTG GCA GCA ACG-3’
	Reverse	5’- CGG TGC TAG TGC CCA GCT GGG CTG GCC-3’
Ligation assay substrate	A1	A1 5’-GTA CGT CGA TCG ATT GGT AGA TCA GTG TCT ATG TAT GTC AGT GAG ATA GTA C-3’
	A2	5’-FAM GTA CTA TCT CAC TGA CAT ACA TAG ACA-3’
	A3	5’-CTG ATC TAC CAA TCG ATC GAC GTA C-3’
SAXS substrate (Nicked DNA substrate)	A1	5’-GGT AGA TCA GTG TCT ATG TAT GTC AGT GAG ATA GTA C-3’
	A2	5’-CTG ATC TAC C-3’
	A3	5’-FAM C TGA CAT ACA TAG ACA-3’
Endonuclease substrate	A1	5’- CCA TTC GTT GTC ATG ACG ACG CTC CGG TAC TCC AGT GTA GCA ATA CGA TTA ATT GAG CTT GCA GGC GCT GTA ATG-3’
	A2	5’(6-FAM)-CAT TAC AGC GCC TGC AAG CTC AAT TAA TCG TAF TGC TAC ACT GGA GTA CCG GAG CGT CGT CAT GAC AAC GAA TGG-3’

Table 2.2 Conditions employed for PCR amplification

	Denaturation temperature (˚C)	Annealing temperature (˚C)*	Extension temperature (˚C)	Polymerase enzyme used
MtbLigA	95	65.8	72	*Taq*
BRCT domain	95	69.8	72	*Taq*
E87A	98	65.7	72	Q5 high fidelity
E26A	98	61.6	72	Q5 high fidelity
E22A	98	62.9	72	Q5 high fidelity
H236Y	98	62.9	72	Q5 high fidelity
K123A	98	65.1	72	Q5 high fidelity
K123R	98	65.1	72	Q5 high fidelity
G614I	98	65.8	72	Q5 high fidelity
G621I	98	69.9	72	Q5 high fidelity
G639V	98	67.4	72	Q5 high fidelity
KK648/649QQ	98	68.2	72	Q5 high fidelity
V646S	98	69.4	72	Q5 high fidelity
GS660/661LA	98	70.7	72	Q5 high fidelity
MtbXthA^{ID} (2-step)	98	72	72	Phusion^{TM} high fidelity

*Optimum annealing temperatures for each of the DNA fragment was identified first by performing gradient PCR

2.5 Expression and purification

2.5.1 Expression and purification of MtbLigA site directed mutants

The His tagged MtbLigA full-length and its deletion mutant MtbLigA1 and AdD domain were expressed and purified to homogeneity as previously described by our laboratory22,28-30. The single site Ia/Ib subdomain mutants E22A, E26A, E87A, K123A/R and H236Y were generated by site directed mutagenesis using pET21d-LigA as template and appropriate primers as indicated in Table S2. The authenticity of expression constructs were verified by DNA sequencing from Chromas biotech. The His-tagged wild type and mutants (deletion/site directed) were overexpressed in the E. coli BL21 (DE3) strain. The mass culture for each protein were grown in of LB medium supplemented with 100µg/ml ampicillin and each culture were grown upto 0.6 optical density of bacterial cell and then induced with 1 mM Isopropyl β-D-1-thiogalactopyranoside for 16h at 22^{0}C at 120 rpm. Cells were harvested by

centrifugation and resuspended in ice-cold Buffer A (50mM Tris-Cl (pH 8.0), 200mM NaCl, 5mM Imidazole). Cells were lysed by sonication followed by centrifugation at 12,000 rpm to remove cell debris. The recombinant mutant proteins were purified by using His60 Ni superflow resin (Takara Bio, Inc.). The column was pre-equilibrated with buffer B (50mM Tris-Cl (pH 8.0), 200mM NaCl). The proteins were eluted with buffer B supplemented with 300mM imidazole. The purified proteins were further subjected to gel filtration through a Superdex-200 10/30 increase column and Superdex-75 coloumn (GE Healthcare) equilibrated with buffer C (50mM HEPES (pH 8.0), 250 mM NaCl and 2mM βME). Eluted fractions were analyzed for purity by SDS-PAGE (Fig. S8) and fractions corresponding to pure protein were pooled and concentrated to 2-15mg/ml, aliquoted, flash frozen and stored at -80^{0}C.

2.5.2 Expression and purification of MtbXthA, MtbXthA^{ID} mutant, MtbLigA and its deletion mutants

Full-length His-tagged MtbLigA and its deletion mutants MtbLigA1, MtbLigA2, MtbLigA3 and BRCT domain alone, were cloned, expressed and purified to homogeneity as previously described by our laboratory (Srivastava, Dube et al. 2005, Khanam, Shukla et al. 2015). Briefly, the clone constructs were expressed in E. coli BL21 (DE3) cells (Novagen) and grown in LB medium containing 0.1 mg ml−1 ampicillin to A600 0.5, followed by isopropyl β-D-thiogalactopyranoside (IPTG) of 1mM induction for 16 hours at 20°C. The cells were harvested by centrifugation at 5000rpm and pellets were resuspended in Buffer A (40 mM Tris-HCl pH 8.0, 150 mM NaCl and 10 mM imidazole, 1 mM PMSF and 0.2 mg ml−1 lysozyme, DNase I) and sonicated. The homogenate was centrifuged at 12,000rpm for 30min and the clarified cell lysates were applied to pre-equilibrated nickel NTA column (Amersham Biosciences). The proteins were eluted using 300 mM imidazole in Buffer A. The eluted fractions containing the protein of interest were pooled and concentrated up to 4-6 mg/ml and subjected to pre-equilibrated (20 mM Tris-HCl pH 8.0, 100mM NaCl) 1 ml MonoQ 5/5 columns on AKTA prime (GE Biosciences). Following 4 column volumes washes with equilibration buffer, a NaCl gradient in 20 mM Tris-HCl pH 8.0 was used to elute the proteins. The resultant fractions were pooled, concentrated and applied over a Superdex 200 size exclusion chromatography (SEC) column (GE Biosciences) in a buffer containing 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5% (v/v) glycerol. Fractions were pooled, concentrated and analysed on 12% SDS-PAGE.

6X-His-tagged and GST-tagged MtbXthA have been cloned, expressed and characterized in our lab as reported earlier (Khanam, Shukla et al. 2015). Overexpression and purification of recombinant MtbXthA/variant proteins was carried out as explained above with minor modifications in cell lysis buffer (40 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.1% N-Lauryl Sarcosine, 10 mM imidazole, 1 mM PMSF and 0.2 mg ml−1 lysozyme, DNase I). The $MtbXthA^{ID}$ mutant wasexpressed in *E. coli* BL21 (DE3) cells (Novagen) and grown in LB medium containing 0.1 mg ml^{-1} ampicillin to A_{600} 0.5, followed by inducing with 0.5mM isopropyl β-D-thiogalactopyranoside (IPTG) for 16 hours at 25°C. The expression constructs were sequenced to verify their integrity. Purification was carried out as discussed above.

2.6 SDS PAGE and Western blotting

For SDS-PAGE protein samples were prepared in 1X Laemmli buffer (Appendix) and boiled at 100°C for 5 min. Samples were centrifuged at 12,000 rpm for 5 min at RT to remove any debris or precipitate. The samples prepared were then resolved on SDS polyacrylamide gels in Tris-Glycine-SDS running buffer (Appendix). After gel was completely resolved it was stained with coomassie brilliant blue R250 (Appendix) for direct visualization of the protein bands. For Western blotting, the resolved SDS-PAGE gel was electro-blotted onto nitrocellulose or PVDF membranes in cold transfer buffer (Appendix) at 100V for 1.5-2 h. Transfer of proteins was confirmed by staining with Ponceau S that was reversed by washing with PBS (Appendix). Blots were blocked with 5% milk in PBS at 4°C overnight or at RT for 1 h. Subsequently, the blot was incubated with the desired dilution of primary antibodies in 2.5% skimmed milk (in PBS) for 2 h at RT/ overnight at 4°C. After five washes with PBS containing 0.05% (v/v) Tween-20 (PBS-T), the blots were incubated with HRPconjugated secondary antibody for another 1.5 h followed by washing with PBS-T. Blots were developed either by chromogenic substrate diamino benzidine (DAB) or by chemiluminescence substrate (ECL kit, GE Healthcare).

2.7 Protein estimation

The protein concentration was determined by using the BCA method (Smith et al.,1985). BCA Protein assay is a detergent-compatible preparation that is based on bicinchoninic acid (BCA) for the colorimetric detection and quantitation of total protein by recording absorbance at 562nm. The color development and absorbance is linear with increasing protein concentrations over the broad working range (20-2000µg/mL). A standard curve was plotted by using BSA as a standard.

2.8 Electrophoretic mobility shift assay (EMSA)

DNA binding by wild type LigA and its site directed mutants was measured by EMSA using 52-bp nicked ds-DNA. 0.66 nM of DNA substrate was incubated with 1 µM and 2.5 µM of wild type and mutant proteins in binding buffer (50mM Tris-Cl (pH 8.0), 10 mM EDTA, 5 mM DTT and 0.1mg/mL BSA) at 25^0C for 30 minutes. After incubation 1X loading dye (15% glycerol and 0.1% Bromophenol blue) was added and the samples were electrophoresed through an 6.5% native PAGE in 0.5X TBE buffer. Substrate without any protein was used as control. Gels were scanned using ImageQuant LAS 4000(GE healthcare).

2.9 Microscopy of DAPI stained bacterial strains.

The exponentially growing culture of *E. coli* temperature sensitive strain GR501 rescued with NAD^+-dependent DNA ligase wild type and mutants, respectively was grown at an A_{600} of 0.8. The culture aliquots of 200 µl were withdrawn and mixed with 10 µl of a DAPI solution (150µg/ml in distilled water). The staining was done for 15 min on ice. The samples were then immobilized on an agar-coated microscope slide and analyzed in a fluorescence microscope (Zeiss) at an excitation wavelength of 360 nm and an emission wavelength of 397 nm. For coating, the slides were dipped into a solution of 2% agar in distilled water.

2.10 Interaction studies

2.10.1 *In Vitro* GST pull down assay

5 µM purified His-tagged MtbLigA, MtbLigA mutants (MtbLigA1, MtbLigA2, MtbLigA3 or BRCT) respectively, were incubated with GST-tagged MtbXthA in 500 µl of assay buffer (50 mM Tris-HCl pH 7.5, 50 mM NaCl, 2 mM DTT, 10% glycerol) at 4˚C for 3 hours. The proteins were then mixed with 50 µl of glutathione sepharose beads (GE Healthcare) with constant rotation at 4˚C for 1 hour. The beads were then washed thrice with assay buffer containing 0.1% Triton X-100. Subsequently beads were boiled in 50 µl of 1X Laemmli buffer and the pulled-down protein complexes were run on 12% SDS-PAGE followed by immune blotting. The blot was probed using anti-His and anti-GST antibodies (Santa Cruz Biotech), respectively.

2.10.2 Fluorescence measurements

To deduce the binding affinities of the cofactor (NAD^+) and nicked DNA to wild type DNA ligase and its mutants , fluorescence titration experiments were conducted using Cary Eclipse spectroflourimeter (Agilent technologies). The binding of cofactor and nicked DNA to protein(s) was monitored by quenching of tryptophan fluorescence at 340 nm after excitation

at 295 nm (slit width 5nm). 2 nM of each protein in the 10 mM Tris-Cl (pH 8.0), 100 mM NaF buffer was titrated with NAD+ (0-80 µM) and nicked DNA (0-12 pM). By plotting the relative fluorescence intensity against the ligand concentration, the apparent binding constant (Kd) was calculated and the data was fitted using single-site binding equation using GraphPad Prism 5.0 (La Jolla California, USA, www.graphpad.com).

$$Y = Bmax * X / (Kd + X)$$

where, Bmax = Maximum binding and Kd = Dissociation constant

A) Protein-protein interactions:

The protein labelling kit (Thermo Scientific) was used to fluorescently label MtbXthA with Oregon green 488 maleimide ($MtbXthA^{OG}$) as per the manufacturer's protocol. The degree of labelling was determined to be 2-4 fluorophores per MtbXthA. Protein-protein interactions were measured by observing the change in fluorescence intensity of fluorescently labelled protein. Titration of purified MtbLigA or BRCT domain was performed **(**0-4.2 µM) in reactions containing 100 nM $MtbXthA^{OG}$ in 500 µl assay buffer (50 mM HEPES-Na pH 8.0, 150 mM NaCl, 2 mM DTT, 8 mM $MgCl_2$ and 0.5 mM EDTA). Reactions were incubated at 4^0C for 30 min and transferred to a 10 mm cuvette using Cary Eclipse Fluorescence spectrophotometer (Agilent Technologies). The fluorescence emission spectra were recorded from 500 to 680 nm with excitation at 494 nm. The relative intensity at 512 nm was plotted against MtbLigA or BRCT domain concentration. For probing non-specific interaction, BSA was used in the control experiment. The apparent K_D for MtbLigA or BRCT domain binding to MtbXthA was determined by fitting the data using single site binding equation (GRAPHPAD PRISM 3, http://www.graphpad.com).

B) Complex disruption assay using FRET:

MtbXthA was labelled with Oregon green 488 maleimide ($MtbXthA^{OG}$) while BRCT domain of MtbLigA was labelled with Alexa flour 555 C2-maleimide ($BRCT^{AF}$) using protein labelling kit (Thermo Scientific). The degree of labelling was found to be 3 fluorophores per MtbXthA and 2 fluorophores per BRCT domain, respectively (determined as per manufacturer's instructions). The interaction between the two labelled proteins was monitored by fluorescence resonance energy transfer (FRET). The donor $MtbXthA^{OG}$ (0.2 µM) and acceptor $BRCT^{AF}$ (2 µM) were incubated in 500 µl FRET buffer (50 mM HEPES-Na pH 8.0, 50 mM NaCl, 2 mM DTT, 8 mM MgCl2 and 0.5 mM EDTA) for 30 min on ice followed by measurement at room temperature using Cary Eclipse Fluorescence

spectrophotometer (Agilent Technologies). Samples were excited at 488 nm and emission spectra were recorded from 495-600 nm. The complex disruption was tested by adding respective peptides and monitoring any change in the FRET efficiency. FRET efficiency (E) was calculated by;

$$E = 1 - Fd'/Fd$$

where Fd' is the emission intensity of the donor in the presence of acceptor and Fd is the emission intensity of the donor alone. The peptides used in the study include (i) DGQ (DGQPSWSGKP) (ii) DVR (DVRVGFDGQP) (iii) YDV (YDVAHVGFD) and (iv) Control peptide (IGQFDLFGV).

2.10.3 Isothermal titration calorimeter (ITC)

AdD domain of MtbLigA was extensively dialyzed overnight at 4^0C in buffer 20 mM HEPES (pH 8.0) and 100 mM NaCl while BRCT domain was extensively dialyzed in 50 mM HEPES-Na pH 8.0, 200 mM NaCl and 2 mM βME. ITC experiments were performed at 25^0C using MicroCal VP-ITC (GE Healthcare) and data analysis was done using Origin 7.0 software with Microcal add-on package. In each ITC measurement cell contained AdD domain (20 µM) and titrated with respective compound contained in syringe (10 mM). For peptide titrations, 0.4 mM of test peptides (i) DGQ (ii) DVR and (iii) YDV were titrated in to 0.05 mM of protein (BRCT domain). Injections of 10 µl of substrate solution were added from a computer-controlled micro syringe at 3 min intervals into the protein solution (cell volume=1.43 ml) with stirring at 394 rpm. All solutions are degassed at 10^0C for 30 minutes under vacuum prior to titration. The enthalpy of reaction (ΔH), binding constant (K) and the stoichiometry value (n) were calculated from measured heat changes ΔH on coalition of inhibitor with protein. The control experiments for each fragment and peptide was performed and subtracted from the corresponding data sets. All experiments were performed with c values $1<c<200$, the quantity $c= K_aM_t$ where M_t is the initial macromolecule concentration. The calibration kit (supplied by manufacturer) was used to calibrate the instrument. Thermodynamic parameters were calculated from the Gibbs free energy equation.

2.10.4 Surface Plasmon Resonance (SPR)

In order to determine the strength of interaction between XthA and DNA ligase, SPR experiments were carried out using Biacore 3000, GE Healthcare. The surface of CM5 sensor

chip was activated with N-hydroxy succinimide (NHS, 0.05M) /N-ethyl-N-(diethylaminopropyl) and carbodiimide (EDC, 0.2 M).The His and GST antibody was immobilised by amine coupling to the activated surface at concentration of 200µg/ml in 10mM sodium acetate buffer at pH=5.0 to a density of 8000 response units. After antibody immobilization the surface was blocked using ethanolamine at pH=7.4 and 50mM NaOH was used as regeneration solution. The GST-XthA / His-M. tb. DNA LigaseA/ His-BRCT domain as a ligand were diluted in running buffer and injected over one flow cell of the anti-GST/ anti-His Ab-coated sensor chip at 5 µl/min until 6000 response units of ligand was immobilized. A second flow cell served as a reference to account for bulk-shift responses and minor, nonspecific interactions with the antiGST/ antiHis Ab, which was subtracted from the binding response. The kinetic experiment was run using application wizard of binding through capturing molecule. The binding for 120 sec after end of each injection was measured at different analyte concentrations which were injected for 3 minutes at flow rate of 20µl/min. The complete regeneration of surface was achieved with two 30 sec injection of 10mM Glycine-HCl pH=2.0. The sensograms were processed by using BIAevaluation software 4.1.1 (GE Healthcare). The steady state equilibrium constant (KD) was determined by fitting the data to a 1:1 Langmuir isotherm.

2.11 Biophysical characterization

2.11.1 Circular Dichroism Measurements

All recombinant SEC purified proteins were dialyzed overnight at 4^0C against buffer (10 mM Tris-Cl (pH 8.0), 100 mM NaF) and centrifuged at 12,000 rpm for 20 minutes at 4^0C to remove any precipitants. All the proteins were diluted to final concentration of 0.18 mg/mL. The spectra were recorded at far UV range (180-260 nm) at 25^0C using 1 mm path length cuvette on (Jasco J-1500) at following experimental settings: sensitivity100 mdeg, scan speed 100 nm/min, bandwidth 1 nm. Each CD spectrum is average of two accumulations and was baseline corrected by subtracting spectrum of corresponding buffer. The experimental data was analysed using K2D3 software.

2.11.2 Size exclusion chromatography of complexes

MtbXthA-MtbLigA complex was prepared by mixing equimolar ratio of MtbXthA and MtbLigA in a buffer containing 50 mM Tris pH 8.0, 200 mM NaCl and 2 mM beta-mercaptoethanol, in absence or presence of nicked DNA substrate (S1). Complexes were incubated for at least 3-4 hours on ice prior to SEC analysis. For MtbLigA-MtbXthA-nicked

DNA complex, equimolar ratio of both the proteins were incubated on ice for 2 hours and then S1 was added followed by further incubation for 30 min on ice such that the final molar ratio becomes 1:1:1. SEC was performed using Superdex 200 increase GL 10/30 column (GE Healthcare) in the same buffer in which complexes were made. The column was calibrated using blue dextran (to determine void volume) and standards Aprotinin (6.51 kDa, 1.3 nm), Ribonuclease (13.7 kDa, 2.4 nm), Ovalbumin (44 kDa, 4.4 nm), Aldolase (158 kDa, 6.62 nm), and Ferritin (440 kDa, 7.99 nm) of known molecular weights and stokes radii.

2.11.3 UV/Vis spectrophotometric analysis of Proteins

All SEC purified samples were subjected to UV/Vis spectrophotometric analysis on Jasco V-750. The 0.5 µM proteins were scanned from the range of 400 to 220 nm. Three consecutive spectra were collected and averaged using Spectral managerTM version II. In the experiments involving complex formation between protein and nicked DNA, spectra were collected for protein-nicked DNA complex fractions eluted in SEC. The 260/280 ratio were calculated at nanodrop (Thermo Scientific NanoDrop spectro).

2.11.4 Small Angle X-Ray Scattering (SAXS)

SAXS in this study was used to gain the insight in to the structural organization and conformational dynamics in the biomolecule in solution form. SAXS experiments were performed at BM29, ESRF (Grenoble, France) and at home source SAXSpace, Anton Paar (CDRI, Lucknow). SAXS measurements were carried out for MtbXthA, MtbLigA, AdD domain, BRCT domain, MtbXthA-MtbLigA and MtbXthA-BRCT domain complex. Studies were also carried out on MtbLigA and MtbXthA-MtbLigA in complex with 27-mer nicked DNA (S1) at equimolar protein/DNA concentrations. The complex was formed by mixing the equimolar concentration of protein-protein/protein-DNA and then centrifuged for 5 min at 10,000 rpm to exclude any aggregated participants. Data were acquired at protein concentrations: 5.3 mg/ml for MtbXthA, 1.3 mg/ml for MtbLigA, 10mg/ml for BRCT domain, 12mg/ml for AdD domain in complex with 4mM cofactor (NAD^+/NMN/AMP), 2.5 mg/ml for MtbLigA-nicked DNA complex, 2.5 mg/ml for MtbXthA-MtbLigA complex and 2.0 mg/ml for MtbXthA-MtbLigA-nicked DNA complex. All proteins were in buffer containing 50 mM Tris pH 8.0, 200 mM NaCl and 2 mM βME. All samples were centrifuged at 12,000 rpm at 4°C for 20 min prior to data collection. Data collection was performed at 10˚C and 10 frames of 1 second for samples were collected at ESRF while 2 frames of 30 minutes at home source. Buffer scattering was measured to generate buffer subtracted

intensity profiles. Data was processed using Primusqt implemented in ATSAS 2.8.1. The forward scattering intensity I(0) and radius of gyration R(g) were estimated using AUTOGNOM which was also used to evaluate the molecular size by plotting pair-distance distribution functions (PDDF). P(r) of scattering data is the representation in real space and reflects the particular particle's shape. P(r) approaches zero at its maximum dimension Dmax. AUTOGNOM results were used to generate 10 independent ab initio models through DAMMIF. The resulting models were further compared using SUPCOMB. The unanimity of the calculated models was represented by the lowest Normalized Spatial Discrepancy (NSD), which was determined by DAMSEL. The analogy of models and superposition was performed with SUPCOMB, CRYSOL and GASBOR. EOM is also used for ensemble modeling of the multi-domain flexible protein, in case of MtbLigA.

2.12 Crystallization

2.12.1 Hanging drop vapour diffusion

It is one of the two, vapour diffusion technique in which protein drops are placed on the siliconized glass coverslip and hanged in inverted position inside the well, containing a higher concentration of the precipitant mixture (crystallization solution). The drop is provided with a closed system by sealing with non-reactive wax, so the chamber becomes airtight and diffusion occurs only in between drop and reservoir solution. Water start diffusing from drop to the reservoir to retain equilibration between the drop and reservoir solute concentration and enters in supersaturating state.

2.12.2 Sample preparation and tray setup

The AdD domain and its mutants were concentrated to around 10-15mg/ml after two-step purification. Before setting up the tray, the protein was centrifuged at 12,000 rpm for 20min to remove any precipitated debris. For co-crystallization, the protein was incubated with cofactor (NAD^+, NMN) or inhibitor (9R0345) for 10minutes at 4˚C. The crystallization condition (0.1 M HEPES-Na pH 7.6, 0.1 M NaCl and 1.5 M ammonium sulfate) was made fresh and filtered through 0.22 µm filter.

2.12.3 X-Ray data collection and processing

2.12.3.1 Data collection

The good looking crystals with proper shape and size were picked in cryoloop (Hampton) and further dipped in 20% glycerol as cryoprotectant. The data collection was done at 100K

temperature. The data was collected at synchrotron Raja Rammana Center for Advance Technology, BL-21, Indore, Madhya Pradesh, India and Elettra, Trieste , Italy.

2.12.3.2 Data collection strategy

The data collection strategy is usually decided after obtaining few diffraction images, diffraction follows a particular pattern (lunes) anything other than this pattern indicates that the crystal is probably not an ideal and having problem like split, more than one crystal or twinning. In the case of multiple crystals, efforts are made to try to find out the spot within the crystal which is single. Spot separation, signal to noise ratio and overloads are some other parameters that have to be considered in data collection. The crystal to detector distance is adjusted to maximize separation of individual spots & the rule of thumb suggests according to the longest cell dimension, detector distance in mm is roughly equal to longest unit cell dimension in angstroms. The choice of exposure time depends on various parameters such as the quality of the crystal, detector distance, etc. Oscillation range affects the overall signal to noise ratio and accuracy of the reflection profile, larger the oscillation range lower the signal upon noise ratio and accuracy of reflection also decreases for increasing oscillation range. So, relatively small oscillation range ($<1.0^0$) expected to improve data quality.

2.12.3.3 Data processing

All the data sets were indexed and integrated using XDS suite (Kabsch, 2010). Following the integration, the process of 'data reduction' was done in order to determine the point-group symmetry of data and the likely space group. The program POINTLESS in CCP4 program suite was used for data reduction. Then after scaling program was used to put all the measurements on the same scale. The program AIMLESS (in CCP4 program suite) was used to average the measurements of symmetry-related reflections and produces the statistics that provide important measures of the data quality. From these analyses, the decision about the quality of the data and whether some measurements should be discarded can be made.

2.12.3.4 Solvent content analysis

Protein crystals are highly fragile in nature due to high water content (27-78% or more) with the most common value being 43% (Matthews, 1968). A significant volume in protein crystal is occupied by the solvent. The solvent content gives an idea about the number of protein subunits in the asymmetric unit of the crystal and calculated by Matthews‟s coefficient either by CCP4 suite, PHENIX (XTRIAGE) or online (MATTPROB).

2.12.3.5 Phasing and structure solution

After data processing the next step is to solve the structure of the protein, for that, we need two components from the data which is collected and processed successfully i.e. structure factor magnitude (Fhkl) and phases (αhkl). The structure factor, the square root of which is the measured Intensity, is a complex number containing information relating to both the amplitude and phase of a wave. Structure factor we can generate by using intensity what we get in data collection but, to do Fourier transformation we need both structure factor and phase.

$$\rho(x,y,z) = 1/V\Sigma hkl\, F\,(h\,k\,l)\, exp\,[-2\,\pi i\,(hx + ky + lz) + i\alpha\,(h\,k\,l)]$$

Where F (h k l) is the structure factor for the corresponding reflection and α (h k l) are the phases for each point. To solve the phase problem which is lost during data collection, we use various methods like Molecular replacement (MR), heavy metal phasing, ab-initio phasing (SAD, MAD) etc.

❖ **Molecular replacement (MR)**

This method is used to obtain an initial set of phases by fitting an identified homologue protein molecule as a search model into the unit cell and further applying rotation and translation along axis a, b, c or x, y, z. In other words, we have a reflection data of unknown protein and the model that we assume approximately similar to the unknown protein. Now the model would be tried optionally to fit inside unknown crystal by applying for all possible orientations and positions until a match between the predicted diffraction & the observed diffraction occurs. We assume the model best fit to the target structure. Now the reflection phases have to be taken from this model to calculate the phase for the unknown structure. The information that we take from known structure does not require similar space group or crystal system. The initial map is calculated with the borrowed phases and experimentally observed amplitudes. The main steps in MR are: Finding the relative orientation of the search model in the unknown crystal cell (Rotation search). Placing this oriented model in the unit cell relative to the crystallographic symmetry elements (Translation search). Assessing the quality of the solution. The three-dimensional structure solutions can be obtained from the simple relation.

$$X = [R]\,M + [T]$$

Where [R] is the appropriate rotation and [T] the required translation to correctly position the model in the unit cell. The Patterson function is the Fourier transform of the squared structure amplitude |F|2 with phases set to zero, does not required phase information and can be calculated directly from the observed data.

2.12.3.6 Model Building and Structure Refinement

After getting phase by MR, the model is further refined and re-built for improved electron density map and better statistics. The refinement cycles are run until the values of R_{merge}, R_{free} reaches to the acceptable range, and there occurs no significant change in the structure. There should be closer agreement between the calculated and observed structure factors. Molecular model coordinates x, y, z and thermal parameters are appropriately altered in order to improve the agreement between the calculated |Fc| and the observed |Fo|. The refinement leads to the best fitted structure with acceptable geometrical (bond length, bond angle and torsion angle) and chemical parameter. The Rfactor (residual factor) is measure of the agreement between the crystallographic model and experimental X-ray diffraction data. It is calculated by the following equation:

$$\mathrm{R} = \mathrm{Fobs}\left| - \right|\mathrm{Fcalc}\left| \right| - \sum \left| Fobs \right|$$

The R_{free} is also computed by the same equation mentioned above and is used to evaluate the possible over refinement of the data also to monitor that the Rfactor is not being artificially reduced by using too many parameters. These are those subsets, which are not included during the structure refinement. R_{free} always remains higher then R factor because model is fitted according to the reflections that contribute to the R factor not that contribute to R_{free} subsets. The PHENIX refinement is used to refine and build model of all the crystal structure discussed in the study.

2.12.3.7 Phenix refinement

Refinement of the structure requires determination of a model parameterization, a refinement target and an optimization method. The parameters for defining the model and target weight optimization is needed to determine the quality of current model which is to be refined. These parameters are further categorized in the classes. The first parameter category is atomic parameters, which include atomic coordinates, Atomic Displacement Parameters (ADPs), model, occupancies and anomalous scattering. Second category is of non-atomic parameters

involving bulk solvent, twinning and anisotropy. Phenix refinement is user-friendly program, executed in main three steps. The first one is, file input (PDB, mtz, sequence file etc.) reading and initialize processing. The second step is execution of process with the given restrain (bulk solvent, XYZ, B factors, TLS) and the final step is to generate output files (PDB, maps and log file containing information about restrain given). The parameters that help to refine structure can be optimized according to our data qualities. These options include refinement target and target weights which quantifies the fit of two basic components i.e. model geometry and experimental X-ray knowledge. PHENIX refinement offer likelihood target that employ weighted least squares approach to all experimental data due to which equal weight is applied to all reflections. Model parameterization includes atomic (atomic coordinates, ADPs, atomic occupancies, anomalous scattering) and non-atomic (bulk solvent, twinning, crystal anisotropy) variables that defines the crystal content and its properties.

Coordinates refinement, refers to the refinement of the atomic position along three orthogonal axes x, y & z and could be carried out in any of the four types i.e. Individual unrestrained, individual restrained, constrained rigid-group (torsion-angle) and pure rigid refinement. The rigid body refinement used at the initial stage of structure refinements. ADP refinement, also known as B-factor, could be the sum of the total motions like local atomic vibration, motion due to rotatable bond, whole residue movement, domain movement, molecule movement and crystal lattice vibrations. Each contributes to the total ADP (U_{total}) could be refined separately. Occupancy refinement refines partial presence or alternative confirmation of atom of residues or ligand at high resolution. We can also choose sets of atoms for the occupancy refinement by using group occupancy refinement option.

Anomalous scattering coefficient refinement; option used, when crystal data contains anomalously scattering atoms. Restraint such as torsion-angle parameterization and specific restrains for coordinates (reference model, secondary structure, Non-crystallographic symmetry (NCS) and Ramachandran plot restrains) utilization may be more essential (Headd et al., 2012) when the data collected in low resolution.

2.13 Refinement strategies and optimization methods

2.13.1 Fourier map and map interpretation

After each round of refinement cycle, manual inspection of map and model is necessary. Real-space refinement of the model was done using 2m|Fo| - D|Fc| and m|Fo| - D|Fc| maps,

are usually contoured at 1.0 and 3.0 respectively, where refers to the RMS (root mean square) deviation in the mean electron density in the maps. These maps are visually examined and model building performed in Coot (Emsley & Cowtan, 2004). The crystallographic Rfree (Brunger et al., 1998) were monitored at each stage to prevent model bias. Waters were added manually by visual inspection or automatically by coot or during the late stage of refinement.

2.13.2 Structure validation

To check the accuracy of the model, validation of that model is very important and it can be done just after each cycle of refinement or webserver. Phenix refinement gives the report of the refined structure with deviation in the standard range of those parameters, which are being refined. Molprobity web server could be used for structure validation. It requires reflection file (having intensities, amplitude and Rfree flags) with PDB file to validate the structure model. Validation has been done under consideration of described aspects i.e. Ramachandran plot and geometric parameters, Rotamer outliers, clash score, the Correlation coefficient between model to map, rotamers and average B- factor for water, protein and ligand molecules.

2.13.3 The Ramachandran plot

This plot was designed by Prof. G.N. Ramachandran, its shows the allowed and disallowed region for protein peptides as a function of torsion angles, ψ & φ two torsion angles in the polypeptide chain, also called Ramachandran angles to describe the rotation of the polypeptide chain around the two bonds on both sides of the Cα atom. Two types of torsion angle present in the backbone of proteins i.e. phi (φ) in (CN-Cα-C) and psi (ψ) in (N-Cα-C-N). Apart from these two major torsion angles, there is another torsion angle considered in Ramachandran plot i.e. omega (ω) torsion angle (O=C-N). The planarity of the peptide bond restricts omega angle to 180^0(Trans) of approximately all main chain peptide bonds. There are defined values for each torsion angle; the molecule could only rotate within that limit. The plot is divided into four quadrants, upper left, upper right, lower left and lower right. Beta-sheet (+ve Ψ and ve Φ), left-handed helix(+ve Ψ and +ve Φ), right-handed α-helix(-ve Ψ and -ve Φ) are placed in the allowed region of upper left, upper right and lower left quadrant, respectively .

2.13.4 Molprobity

MolProbity (Vincent B. Chen et al., 2010) is a structure-validation web service that offers a broad-spectrum solidly based assessment of model quality at both the global and local levels for macromolecules. It uses the Engh and Huber restraints for proteins, nucleic acids, and other molecules in the form of the CCP4 monomer library. It uses REDUCE and PROBE for all-atom contact analysis, RAMALYZE, ROTALYZE, DANGLE, SUITENAME and SILK for other criteria and KiNG for three-dimensional visualization of the structure and its validation markers directly in browser (V. B. Chen et al., 2010). There are fixed default restraints and their associated standard deviation values (Sigma) for bond lengths, bond angles, chiral centers, dihedral (torsion) angles, planar groups (such as aromatic rings), and nonbonded interactions (VDW), which helps during refinement of these parameters. Increasing sigma value indicates a violation of standard set values. It is resolution depended, moderate-to-low resolutions, there should be no outliers, although strained torsion angles are possible. Molprobity provides the complete summaries and graphical representations (two or three dimensional) with charts validation report in result output files.

2.13.5 Rotamer

Rotamer utilizes a huge number of well-fitted protein structures from the databases to evaluate the residues side chi angle and flagged them as an outlier if they reside out of the allowed range of chi angles. It makes a separate window over background distribution of top500 database to show the distribution of chi1-chi2 angles combination for each residues type. For the longer residues, it could show them as an outlier on the basis of values of chi3-chi4 angles even though chi1 and chi2 reside in the favourable region.

2.13.6 Clash score

When two atoms of protein come close enough or overlap each other without making any bond, they cause crashes. The severity of these clashes is measure in the score. In low-resolution protein structure the clash score might be wrong but, if it is a severe clash, the removal of this clash would be very challenging since most of the refinement programs could not refine the structure having severed steric clash (Ramachandran, Kota, Ding, & Dokholyan, 2011). The non-bonded restraints refinement could be done with or without hydrogen by increasing the atomic radius of the non-hydrogenated atom. This estimation applied accurately on a global structural level but generally leaves chemically impossible

geometries accessible. Therefore, phenix calculates clash score after automatically addition of hydrogen to the atoms and run phenix.probe to analyze atom contact. Atoms, which are closer than 0.4Å will be arranged in clash score list and assign the global clash score to the entire structure by dividing the bad overlap by total number of atoms (1000*bad overlap/number of atoms). The clash score should be as low as possible.

2.13.7 Model to map correlation

The real space correlation contains bulk-solvent correction and data scaling and compute likelihood-weighted 2m|Fo|-D|Fc| map. The correlation coefficient value computed from the comparison between 2Fo-Fc map of calculated data and map of the model, which indicates how good the model fit is. In the GUI, lists can be filtered and/or sorted by CC. The residues having score less than 0.8CC are include in the list and to be taken into account manually. Although real-space CC can be a suitable indicator of poorly fit regions of a molecule, it should not be taken as an absolute score, and it is tough to find an ideal limit below which the score should never fall. Few of these will be properly positioned but poorly ordered; not cause any difficulty if their geometry lies in the limit. At the time result evaluation, we should also consider B-factor, occupancy, and absolute density values because these are the indicator of intrinsically disorder part of the molecule. The CC and B-factor of each and every residue are represented by graph form in multi-criterion plot. Just by clicking on the plot we can see the zoomed view of the corresponding residue in coot or PyMOL (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC).

2.13.8 Omit map

Omit map was calculated to remove systematic errors or model biasness , arising out of model due to the initial model used to get structure solution. The omit map calculation comprises subsequent omitting the regions of the structure and refinement. The omit map was generated by using the program „Composite omit map" from Phenix suit (Terwilliger et al., 2008).

2.13.9 Superposition

Superposition of one chain onto the other chain was done by the help of SUPERPOSE (Krissinel & Henrick, 2007) program, which attached with the CCP4 suite. The images were prepared by CCPmg (McNicholas, Potterton, Wilson, & Noble, 2011).

2.14 Screening and Identification of inhibitors against MtbligA

2.14.1 *In vitro* DNA ligation activity and inhibition assay (determination of IC_{50})

5'-FAM labeled 52bp ds-nicked DNA substrate was used to scrutinize the *in vitro* ligation activity and inhibition activity of selected fragments (as discussed in chapter 6) against the MtbLigA (NAD^+-dependent) and bacteriophage T4 DNA ligase (ATP-dependent). For same ligation activity the amount of both enzymes were optimized in the absence of any inhibitor/fragment under the assay conditions. The 0.5 µM of protein (MtbLigA or mutants) / 0.05 U of T4 DNA ligase, 2 nmol of DNA substrate and different concentration of fragments (0-20 µM) in reaction buffer (50 mM Tris-Cl (pH 8.0), 5 mM DTT, 10 mM MgCl2, 10% Me2SO4, 50 µM NAD^+) were incubated for 30 minutes at 25^0C. The reaction was ceased by adding 0.5X stop solution (98% formamide, 10 mM EDTA, 0.15% Xylene Cyanol and 0.15% bromophenol blue). The samples were electrophoresed on 12% 8M UREA-PAGE in 1X TBE as a running buffer. Gels were scanned and quantified using ImageQuant LAS 4000 (GE Healthcare) and ImageQuant TL8.1 software. Since the compounds are soluble in 100% Me_2SO_4 and comprises of about 1/10th volume in reaction mixture, thus 10% Me_2SO_4 was included in all control reactions.

The effectiveness of fragments in inhibiting ligation activity was measured by determining its IC_{50} values. IC_{50} is computed by plotting the relative ligation activity versus inhibitor concentration using GraphPad prism 5.0 software and fitting to the following equation-

$$Vi/V0 = IC50 / (IC50 + [I])$$

Where, V_i and V_0 are rate of ligation in presence and absence of inhibitor respectively, [I] is the inhibitor concentration.

2.14.2 *In vivo* complementation and anti-bacterial assay of MtbLigA (determination of MIC)

A bacterial system consisting of temperature sensitive *E. coli* GR501 strain with lig251 mutation in its LigA gene due to which its growth is normal at permissive temperature 30^0C but hampered at higher non-permissive temperature (43^0C). By complementing MtbLigA, the growth is rescued even at higher temperature. Therefore, we complemented site directed mutants of adenylation domain (E22A, E26A, E87A, K123A, H236Y) and BRCT domain

(G614I, G620L, G639V, KK648/649QQ) and checked the growth at 30^0C and 43^0C by streaking on LB agar plate.

To evaluate the anti-bacterial activity of fragments we also used *E. coli* GR501 strain complemented with ATP-dependent DNA ligase as control. We checked the minimum inhibitory concentration of fragments on both MtbLigA and T4 LigA rescued strains. The MIC values were determined by broth microdilution assays in microtiter plates in a volume of 100 µL. Serial 2-fold dilution of fragment was seeded with inoculums containing approximately 100 colony-forming units/ml in the case of the *E. coli* $ligA^{ts}$ mutant rescued with NAD^+-dependent DNA ligase A, under ambient conditions for 20 h. MICs were deduced as the lowest concentrations of fragment that forbade visible microbial growth. The *E. coli* mutant strain was grown in LB medium containing 25 µg/ml polymyxin B nonapeptide (Sigma-Aldrich) to facilitate penetration of the fragment across the outer membrane.

2.14.3 Time-kill studies

An exponentially growing culture of DNA ligase-null mutant in LB broth were treated at A600 = 0.4 with increasing fragments concentrations (0.5, 1 and 1.5 times of MIC). The viability and effect on growth was inquired by monitoring the absorbance at 600 nm and the number of CFU for 7 hours after the addition of fragment in the culture. For quantification of the CFU, culture aliquots of both the strains were serially diluted in phosphate-buffered saline and plated on LB agar. After incubation for 16 h at 37 °C colonies were visible on the plate.

2.14.4 Mode of Inhibition

By increasing the NAD^+ concentration from 0-50 µM, the saturating substrate concentration for LigA was estimated applying Michaelis-Menten kinetics. The K_m for NAD^+ was determined in 10% Me_2SO_4 using the assay procedure as described above. Varying concentration of NAD^+ from 0-50 µM was used to elucidate the kinetics for different amount of compounds (0, 2, 4, 8, 16, 32 µM). Gels were scanned and quantified as described above. Data were plotted using Michaelis-Menten kinetics in GraphPad Prism 5.0 and K_i values were computed by plotting the apparent Km values against the respective compound concentrations. Using Lineweaver-Burk kinetics, mode of inhibition was determined.

2.14.5 DNA-Inhibitor Interaction

A) Fluorescence assay:

To test the DNA intercalating properties of compounds, detection of fluorescence loss that occur on Ethidium bromide displacement from DNA was measured. The reaction mixture (100 µL) consisting of 8 µg of calf thymus DNA, 0.2 µM of EtBr, 25 mM Tris-Cl (pH 8.0) , 50mM NaCl and 1 mM EDTA. The quenching of EtBr fluorescence on increasing the inhibitor concentration was monitored at excitation and emission wavelength of 485 nm and 612 nm respectively using Cary Eclipse Spectroflourometer (Agilent technologies).

B) Gel shift assay:

100 ng of linear plasmid DNA (PUC 18) was incubated with increasing concentration of compounds in TAE at 25^0C for 30 minutes and DNA shifting was inspected on 1% agarose gel.

2.15 Screening and identification of inhibitors against MtbXthA

2.15.1 *In vivo* complementation of MtbXthA in to *E. coli* XthA knockout strain

The XthA knockout strain of *E. coli* (JW1738) was complemented with MtbXthA and the viability of the complemented strain was checked under oxidative stress. The H_2O_2 was used at lethal dose for stress. The complemented strain was grown till O.D of 0.8 and then 5mM H_2O_2 was supplemented in the culture. The culture was further grown for additional 1hour at 37˚C. After then cells were serially diluted to 10^{-2} and plated on LB agar plates supplemented with appropriate antibiotics. The colonies were counted after incubation of plates for 18 hours at 37˚C. All the experiment was done in dark so to avoid H_2O_2 degradation.

2.15.2 Determination of Lethal dose (LD_{50})

To determine the lethal dose, the complemented and knockout strains were exposed to 0, 1, 2.5, 5, 7.5 and 10 mM of H_2O_2. The statistical analysis is done by two-way ANOVA test, where $*P < 0.05$.

2.15.3 *In vivo* screening of inhibitors

The fragments and compounds of MMV pathogen box were tested at 1µM concentration on the complemented strain under the 5 mM H_2O_2 stress condition. The treated culture is additionally grown for 1 hour at 37˚C and then serially diluted to 10^{-2} concentration and plated on the LB agar plates. The plates were incubated for 18 hours at 37˚C and then colony were counted and scanned.

2.15.4 *In vitro* screening of inhibitors and determination of IC_{50}

IC_{50} is computed by plotting the relative endonuclease activity (normalized response) versus inhibitor concentration using GraphPad prism 5.0 software and fitting to the following equation - $V_i/V_0 = IC_{50} / (IC_{50} + [I])$ where V_i and V_0 are rate of endonuclease cleavage in presence and absence of inhibitor respectively, [I] is the inhibitor concentration.

Chapter 3 | Structural characterization of *Mycobacterium tuberculosis* NAD^+-dependent DNA Ligase A (MtbLigA)

3.1 Introduction

DNA ligases are the fundamental participants in the central dogma of molecular biology including DNA repair and recombination (Soderhall 1976, Sancar 1994).DNA ligases have modular architecture (**Fig. 3.1 A**). NAD^+-dependent DNA ligaseA is a multi-domain flexible protein that adopts distinct conformations throughout the ligation reaction. The adenylation domain (AdD) of MtbLigA has mixed α, β fold that belongs to the nucleotidyl transferase superfamily including mRNA capping enzymes with the conservation of active site (Srivastava, Dube et al. 2005). The structural constitution of AdD domain comprises of Ia subdomain (2 alpha helices) and Ib subdomain (7 beta sheets surrounded by 9 helices). All forms of life share two classes of DNA ligases- one is ATP-dependent which is present in viruses, archaea and eukaryotes while NAD^+-dependent DNA ligases exists mostly in eubacteria. The nick in the DNA is sealed by forming a phosphodiester bond between 3'OH and 5'PO4 group through consecutive nucleotidyl transfer reaction by reconstituting the DNA ligase on nick (Tomkinson, Vijayakumar et al. 2006). DNA ligases are accommodated in toroid shape architecture around the nick by orchestrating an economically efficient platform for DNA ligation (Pascal, O'Brien et al. 2004, Nandakumar, Nair et al. 2007) (**Fig. 3.1 B**). With the help of Small Angle X-ray scattering (SAXS) the structural transitions happening in MtbLigA in the presence of nicked DNA substrate is highlighted in this study. Apo-LigA remains in open extended confirmation with almost negligible interaction between Ib subdomain of AdD and HhH domain and as the encounter occurs with nicked DNA there occurs huge domain swapping in apo-LigA to form toroid architecture around the nick with Ib and HhH domains interacting with each other.

In the first step of ligation reaction, NAD^+ binds to Ia subdomain of adenylation domain via NMN part of NAD+ and correctly position the AMP into the active site (Sriskanda and Shuman 2002, Gajiwala and Pinko 2004). The transfer of AMP from active site lysine to the 5'PO4 nick of the DNA occurs in second step. Finally the nick sealing is accomplished and AMP is released from the active site after the DNA (Shapiro 2014) . There occurs slight variation in sequence alignment and domain organization of ATP and NAD^+-dependent DNA ligases (Kodama, Barnes et al. 1991, Luo and Barany 1996, Sriskanda and Shuman 1998, Mackey, Niedergang et al. 1999, Sriskanda and Shuman 2002). During the ligation reaction steps, both ligases orient themselves differently around the DNA. This co-factor specificity for ATP and NAD^+ is predominately contributed by binding site

constituting structural factors that grip and accommodate the PPi and nicotinamide mononucleotide leaving group, respectively (Hakansson, Doherty et al. 1997, Sriskanda and Shuman 1998, Sriskanda and Shuman 2002, Gajiwala and Pinko 2004, Shuman and Lima 2004). Highly modular structure in NAD^+-dependent DNA ligase assists the distinct structural switches into protein throughout the reaction (Doherty and Suh 2000, Lee, Chang et al. 2000). Nick sensing mechanism is an inherent property in DNA ligase by having strong affinity towards 3'OH and 5'PO4 nick as compare to duplex DNA or gapped DNA, which depends on the covalent adenylation of 5'PO4 of the nick (Ho, Van Etten et al. 1997, Sriskanda and Shuman 1998, Odell and Shuman 1999).

The crystal structures of full length DNA ligase from *T. filiformis*, *E. coli.* and adenylation domains from *M. tuberculosis* and *B. stearothermophilus* shows the different static conformations of Ia subdomain around Ib subdomain of adenylation domain which signifies the biological pertinence of adaptability among these domains (Singleton, Hakansson et al. 1999, Lee, Chang et al. 2000, Srivastava, Tripathi et al. 2005) . With the help of SAXS studies we have mapped the Ia subdomain rotation around Ib subdomain rotation in the adenylated condition while in de-adenylated form the AdD domain remains in open extended conformation with both the subdomains (Ia and Ib) stretched out. The structural and biochemical studies including site directed mutations have flourished the mechanism of adenylation reaction and nick sensing mechanism within the conserved motifs of NAD^+-dependent DNA ligases (Shuman and Ru 1995, Sriskanda, Schwer et al. 1999).

The mechanism of NMN release from active site to prevent the reversal of adenylation reaction is not yet fully determined. Here in this study we have reported the crystal structures of intermediates in adenylation reaction, which contribute to decipher the possible mechanism of NMN release and different conformational forms of AMP and NMN at their respective binding site. There occurs a huge structural transition in DNA ligase in presence of nicked DNA substrate (Nandakumar, Nair et al. 2007). It had been previously reported that K123 is essential for ligation activity and its mutation in to alanine leads to the complete abolish of activity (Shuman 1995). Therefore, by considering the structural understanding we determined the role of E22, E26, E87, H236Y and K123R in the ligation reaction and in the structural and functional integrity of the MtbLigA. We also crystallized the Ia and Ib subdomain mutants (E22A, H236Y) along with active site mutants (K123A/R) and the AdD domain bound with both NMN and AMP to structurally understand the

orientation of the two subdomains in the presence of co-factor during the course of ligation reaction.

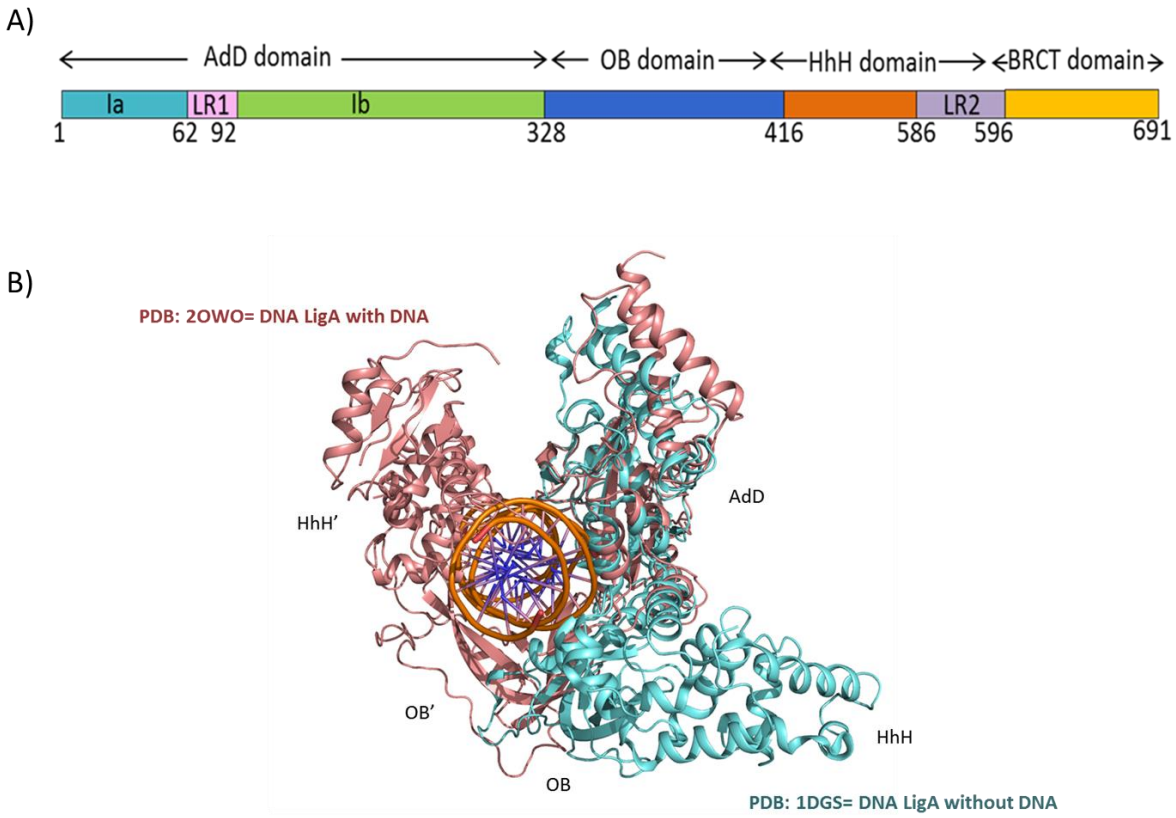


Figure 3.1 Domain organization of MtbLigA and comparison of LigA with DNA of *E. coli* (PDB:2OWO) and without DNA from *T. filiformis* (PDB:1DGS) A) Diagrammatic representation of domain architecture of MtbLigA. Ia (dark cyan) and Ib subdomain (green) of adenylation domain. Oligonucleotide binding (OB) domain (blue), HhH domain (orange) and BRCT domain (dark yellow). **B)** The crystal structure of DNA LigA with DNA (PDB: 2OWO, salmon colour) is superimposed with DNA LigA crystal structure without DNA (PDB: 1DGS, Cyan clour) at Ib subdomain.

3.2 Results

3.2.1 Cloning of AdD domain mutants in AdD domain and Full length MtbLigA, overexpression and purification of recombinant proteins

The His tagged MtbLigA full-length and its deletion mutant MtbLigA1 and AdD domain were expressed and purified to homogeneity as previously described by our laboratory (Srivastava, Dube et al. 2005, Kukshal, Khanam et al. 2012) .The single site Ia/Ib subdomain mutants E22A, E26A, E87A, K123A/R and H236Y were generated by site directed mutagenesis using pET21d-LigA as template and appropriate primers as indicated in **Table 2.2**. The authenticity of expression constructs was verified by DNA sequencing from Chromas biotech (**Fig. 3.2 A**). The His-tagged wild type and mutants (deletion/site directed) were overexpressed in the *E. coli* BL21 (DE3) strain. The mass culture for each protein were grown in of LB medium supplemented with 100µg/ml ampicillin and each culture were grown upto 0.6 optical density of bacterial cell and then induced with 1 mM Isopropyl β-D-1-thiogalactopyranoside for 16h at 22^0C at 120 rpm. Cells were harvested by centrifugation and resuspended in ice-cold Buffer A (50mM Tris-Cl (pH 8.0), 200mM NaCl, 5mM Imidazole). Cells were lysed by sonication followed by centrifugation at 12,000 rpm to remove cell debris. The recombinant mutant proteins were purified by using His60 Ni superflow resin (Takara Bio, Inc.) (**Fig. 3.2 B**).

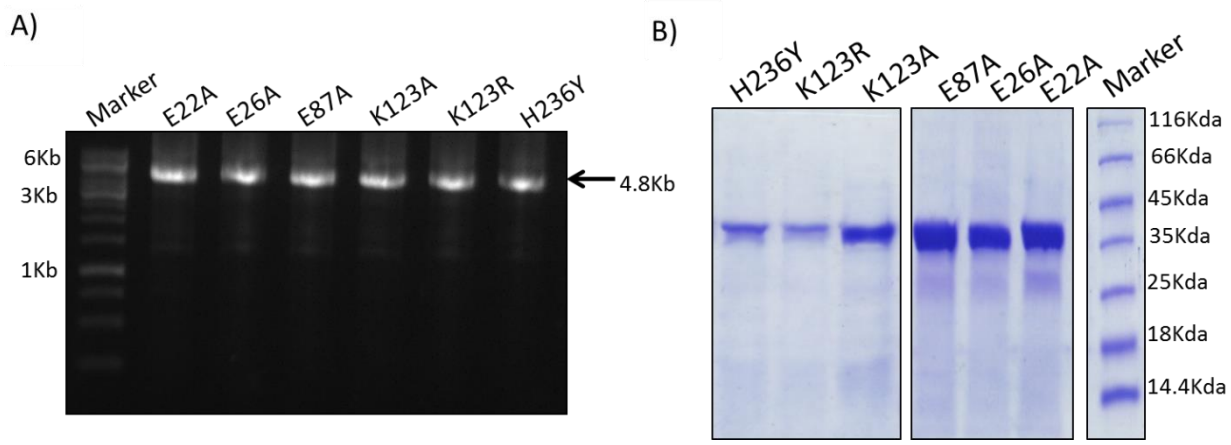


Figure 3.2 Cloning and Purification of AdD domain mutants of MtbLigA. A) PCR mediated site-directed mutagenesis to generate AdD domain mutants of LigA in AdD domain **B)** Ni-NTA purification of AdD domain mutants (35Kda) and the samples were analysed on 12% SDS-PAGE. The arrow shows the size of respective mutant of AdD domain.

3.2.2 Size exclusion chromatography (SEC) purification of recombinant proteins

The column was pre-equilibrated with buffer B (50mM Tris-Cl (pH 8.0), 200mM NaCl). The proteins were eluted with buffer B supplemented with 300mM imidazole. The purified proteins were further subjected to gel filtration through a Superdex-200 10/30 increase column (GE Healthcare) equilibrated with buffer C (50mM Tris-Cl (pH 8.0), 200 mM NaCl and 2mM βME). Eluted fractions were analyzed for purity by SDS-PAGE and fractions corresponding to pure protein were pooled and concentrated to 2-15mg/ml, aliquoted, flash frozen and stored at -80^{0}C. (**Fig. 3.3 A, B and C**).

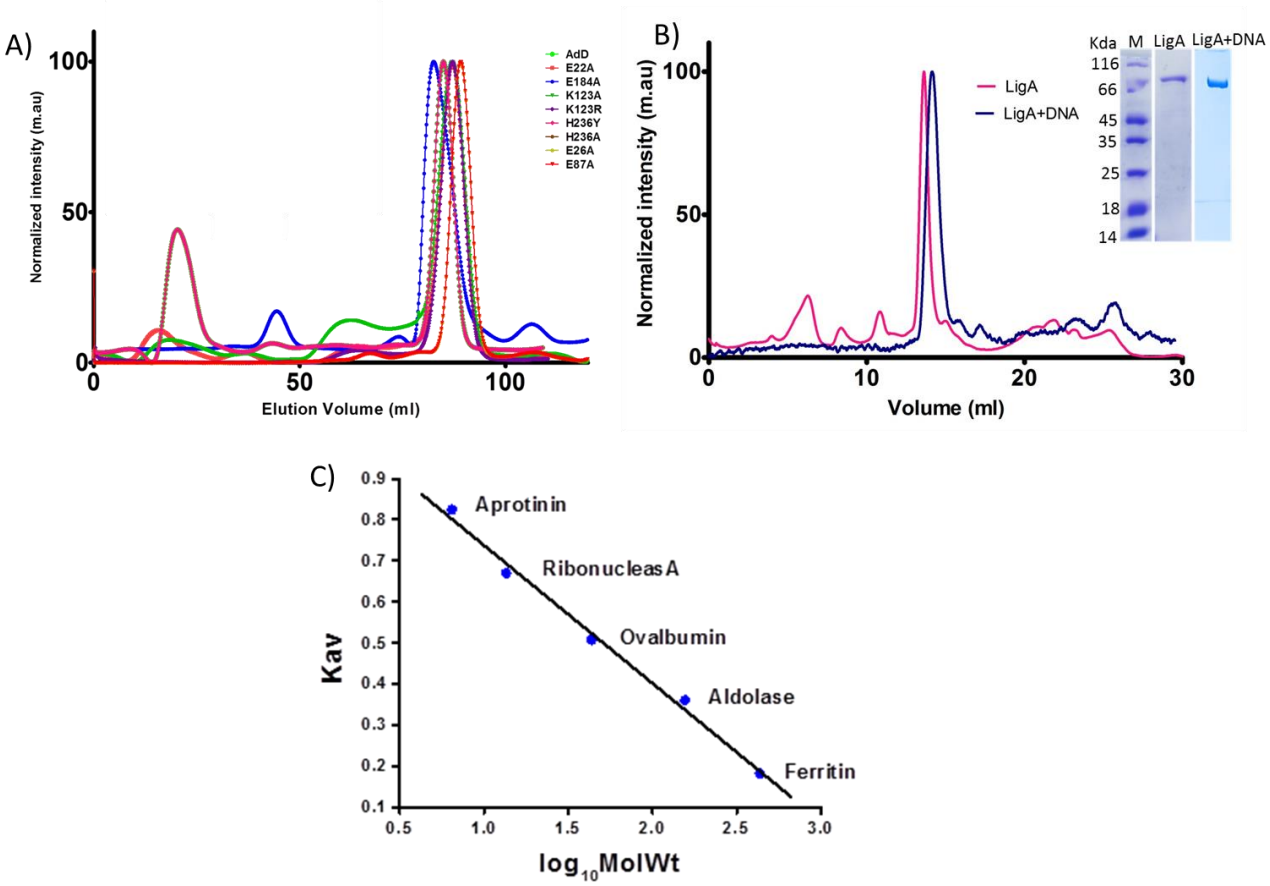


Figure 3.3 Size exclusion chromatography analysis of AdD domain mutants. A) Size exclusion chromatogram of AdD domain mutants using Superdex 200 16/600 coloumn and **B)** apoLigA and LigA in presence of nicked DNA substarte, using Superdex 200 10/300 GL coloumn. The inset SDS-PAGE gel shows the eluted fraction of LigA and LigA+DNA **C)** The graph shows the standard plot using markers aprotinin (6.5 Kda), ribonuclease A (13.7 Kda) ovalbumin (44 Kda), aldolase (158 Kda) and ferritin (440 Kda) run on superdex 200 10/300 GL [GE healthcare].

3.2.3 SAXS analysis of *apo* and DNA bound LigA identifies extended and closed conformations of LigA respectively

We carried out Small-angle X ray scattering (SAXS) analysis involving apo LigA and also in the presence of a 27 bp DNA substrate analog. The SAXS experiments identified a conformational switching between relatively open (SASBDB ID: **SASDDN9**) and closed (SASBDB ID: **SASDDV9**) forms in the absence and presence of DNA substrate homolog respectively. The full-length model of LigA was built for the analysis using crystal structures of homologs and structures solved by our group as template. The pairwise distance distribution function P(r) accurately demonstrated a reduction in maximum dimension (D_{max}) from 16.65 nm (extended conformation) to 13.14 nm (compact conformation) and increased radius of gyration (Rg) from 5.21±0.62 nm to 5.68± 0.47 nm in the presence of 27 bp nicked DNA substrate (**Fig. 3.4 A, Table. 2.1**). The SAXS data were used to generate the ab initio 3D models for extended and closed conformation. The extended structure of MtbLigA fitted the SAXS envelope derived from the average bead model with a χ^2 of 2.93 (**Table. 3.1**). It has been previously reported that on encountering the substrate DNA nick, the OB and HhH domains of DNA ligase rotates to form toroid shaped architecture to facilitate the sequential reactions of ligation. The SAXS data for LigA complexed with nicked DNA supports a toroid shaped model with protruded electron density for BRCT domain (χ^2 of 0.84). The comparative Kratky analysis at lower q values of scattering profiles of both data sets reveals the globularity attained by LigA in the presence of nicked oligo (**Fig. 3.4 B, C**). All samples are monodisperse and free from aggregation as determined by Guinier analysis (**Fig. 3.4 D**). The bead models generated using DAMMIF show the presence of all domains (AdD, OB,HhH and BRCT domain) in the extended conformation (**Fig. 3.5 A I, II**) where the distance between the AdD domain and HhH is prominent which is nearly 2.5 nm and central cavity is about 4.3 nm. Conformational changes in the local spatial disposition of domains around the nicked DNA substrate nick were observed. These movements in domains result in correct orientation of the nick into the active site (**Fig. 3.5 B I, II**). The re-orientation of the domains in the presence of the nicked DNA reduces the distance between AdD and HhH domain to 1.6 nm and associated central cavity with 2.11 nm 'hole'. By superimposing the Ib subdomain to *E. coli* LigA Ib subdomain, the conformational changes in Ia, OB, Zn and HhH domains in presence of nicked DNA was determined and found that Ia subdomain rotates by nearly 147° to possibly help to make kissing contact among the Asp153 (AdD domain) and

Arg464 (HhH domain) just above the nick. The data collection statistics are mentioned in **table 3.2.**

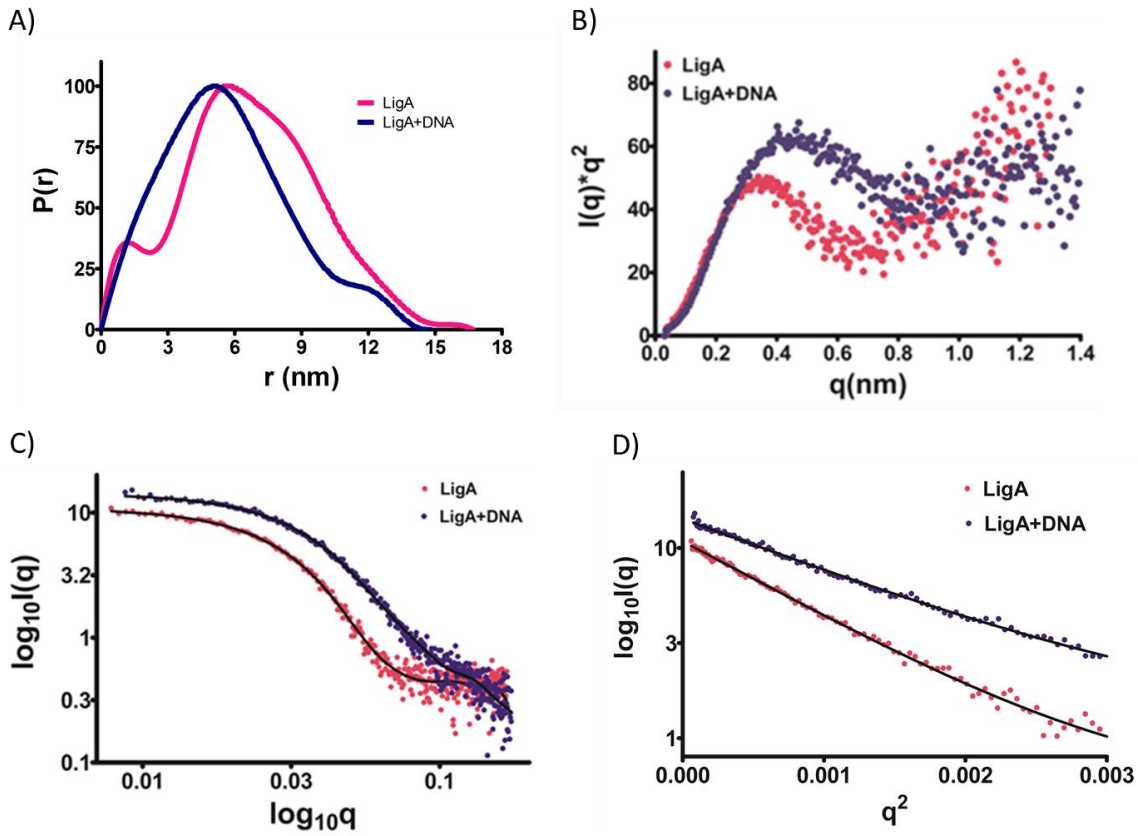


Figure 3.4 Solution structure of MtbLigA. A) The distance distribution function P(r) shows the reduction of maximum dimension (D_{max}) of LigA in the presence of 27mer nicked DNA substrate (royal blue) as compared to the one without DNA (magenta) **B)** The kratky plot also indicates towards the formation of more compact globular structure with DNA (royal blue) as compared to the one without DNA (magenta). **C)** The experimental scattering profile of LigA and LigA+DNA. The black line denotes the fitting of ab initio bead models generated by DAMMIF. **D)** The graph shows the Guinier fitting of the experimental data .

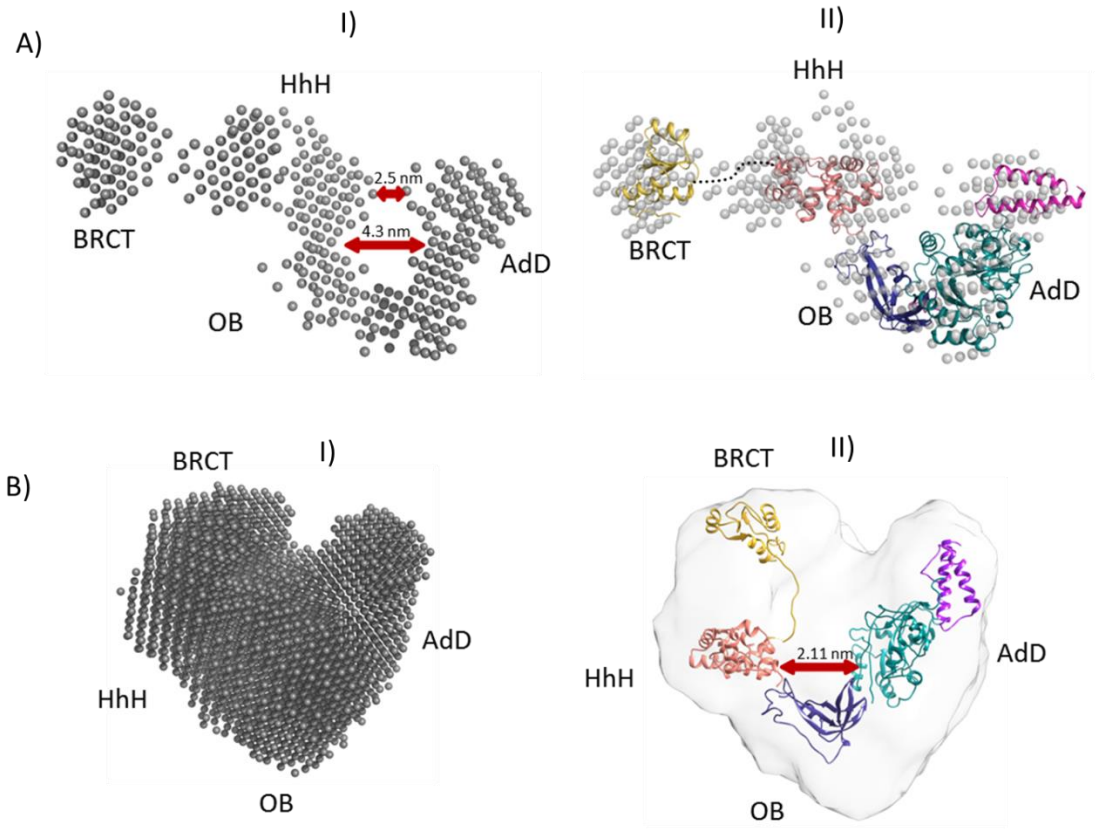


Figure 3.5 Envelope modeling of MtbLigA A) I, Bead models built by DAMMIF also indicates the extended conformation of LigA. The AdD domain is 2.5 nm away from the HhH domain and the central cavity is more wide nearly 4.3 nm. **II,** High resolution extended model of LigA was built using COOT and then fitted by SUPCOMB into the SAXS envelope. The missing residues are shown in dots. **B). I,** Closed conformation of LigA in the presence of nicked DNA as observed in the bead model generated by DAMMIF. The distance between AdD domain and HhH domain reduced to 1.6 nm and the diameter of central cavity reduces to 2.1 nm. **II,** The high resolution closed model of LigA+DNA was fitted into SAXS envelop by SUPCOMB. The ensemble model was generated by EOM.

Table 3.1 The SAXS statistics

Sample	Dammif		Supcomb	Crysol	EOM
	χ^2	NSD	NSD	χ^2	χ^2
LigA	0.79	1.082±0.043	2.93	-	2.76
LigA+ DNA	0.64	0.850±0.046	3.10	1.14	2.05

NSD- Normal spatial discrepancy

Table 3.2: SAXS data collection and statistics of apo-LigA and LigA+DNA

	***apo*LigA - SASDDN9**	**LigA+DNA complex – SASDDV9**
Data Collection Parameters		
X-ray source	BM29, ESRF	BM29, ESRF
Detector	Pilatus 1D	Pilatus 1D
Wavelength (nm)	0.081	0.081
Exposure time per frame	1 sec	1 sec
No. of frames collected	10	10
Concentration (mg/ml)	1.3	2.5
Measurement temperature (^{0}C)	10	10
Structural parameters		
I(0) [from Guinier]	10.86±0.11	14.5±0.13
Rg (nm) [from Guinier]	5.22±0.08	4.49±0.07
I(0) [from P(r)]	10.82	14.5
Rg (nm) [from P(r)]	5.20	4.49
D_{max} (nm)	16.7	14.79
V^{porod} (nm^3)	870	262.05
Modelling parameters		
Symmetry	P1	P1
Chi^2 (Gasbor)	0.79	0.65
Model fitting	SUPCOMB	SUPCOMB

3.2.4 SAXS studies of cofactor induced domain rotation in AdD domain

The SAXS analysis of de-adenylated form of AdD domain (SASBDB ID: **SASDEU5**) and when exposed to excess of NAD^+ (SASBDB ID: **SASDG65**) where NMN and AMP are bound in their respective sites) demonstrates the expanded form of AdD domain with entirely relaxed Ia subdomain aside from active site in Ib subdomain. The maximum dimension (D_{max}) of de-adenylated AdD domain is 8.84 nm and 8.89 nm respectively. In agreement with the present crystal structure of AdD domain, on the exposure of NAD^+ to de-adenylated AdD domain causes the rotation of Ia subdomain near to Ib subdomain and reduces the Dmax to 6.17 nm in the overall architecture of AdD domain (SASBDB ID: **SASDEC6**) (**Fig. 3.6 A**). The de-adenylated form shows less globular behavior as compare to adenylated form as observed in Kratky analysis and this behavior manifest the change in shape of AdD domain which is tempted by NAD^+ (**Fig. 3.6 B**). The Gasbor and Dammif analysis of buffer subtracted scattering profiles shows extended density for Ia subdomain (**Fig. 3.7 A I, II, III and C I, II, III**) and more globular in the presence of NAD^+ (**Fig. 3.7 B I, II, III**) which are best fitted with 0.7 and 0.84 chi^2 values, respectively. The Rg value for extended conformation is 2.43 ± 0.05 nm and 2.49 ± 0.05 nm while for closed conformation is 2.37 ±0.13 nm (**Table. 3.3**). The scattering and Guinier analysis shows that protein samples are mono-disperse and devoid of aggregation (**Fig.3.6 C, D**). The solution studies of AdD domain verifies NAD^+ stimulated movement of Ia subdomain for the adenylation of active site K123 and organizes structure for serial stairs of DNA ligation.

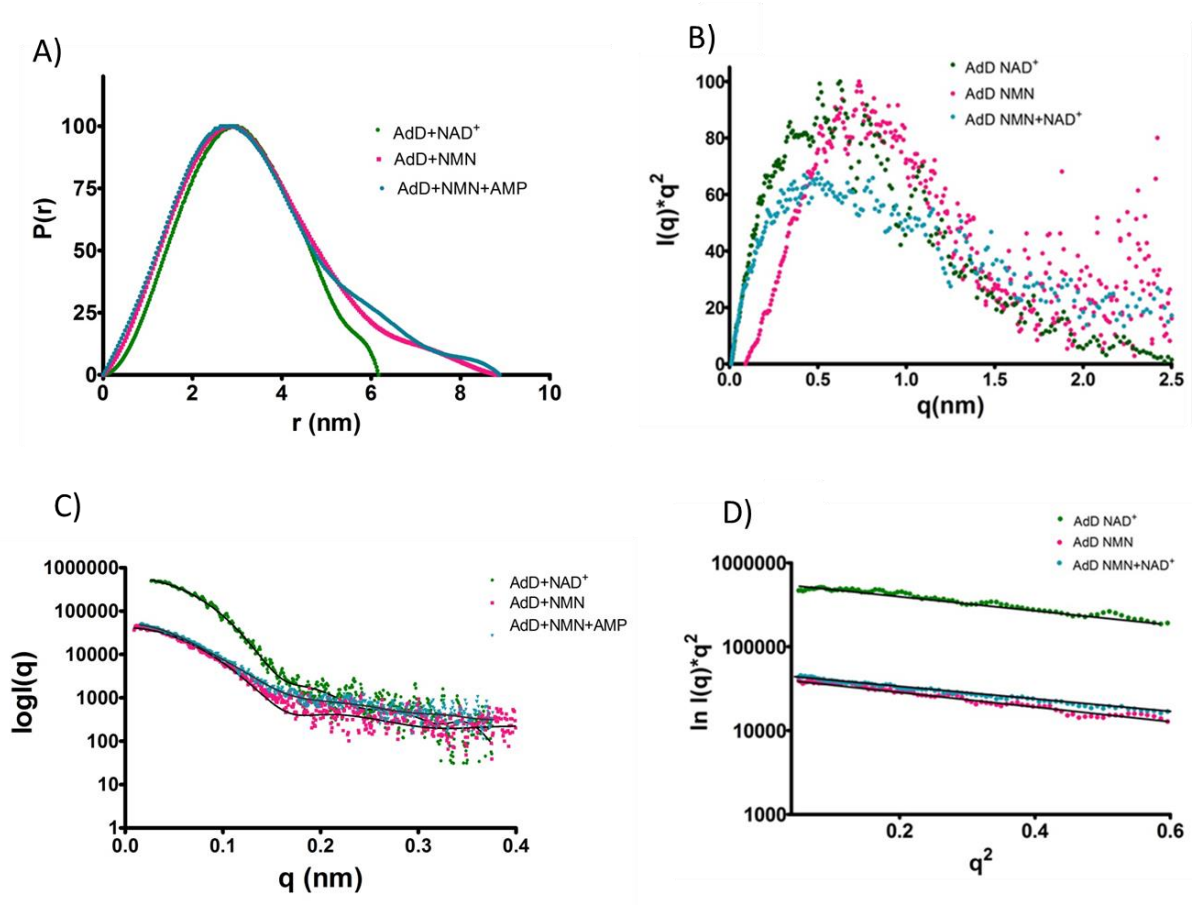


Figure 3.6 Solution structure of AdD domain. A) The distance distribution function P(r) shows the reduction of maximum dimension (D_{max}) of AdD domain in presence of NAD^+ (green) while extended in the presence of NMN (pink) and NMN, AMP (blue). **B)** The kratky plot also indicates towards the formation of more compact globular structure with NAD^+ (green) as compared to the one with NMN (pink) and NMN, AMP (blue). **C)** The experimental scattering profile. The black line denotes the fitting of ab initio bead models generated by SUPCOMB. **D)** The graph shows the Guinier fitting of the experimental data.

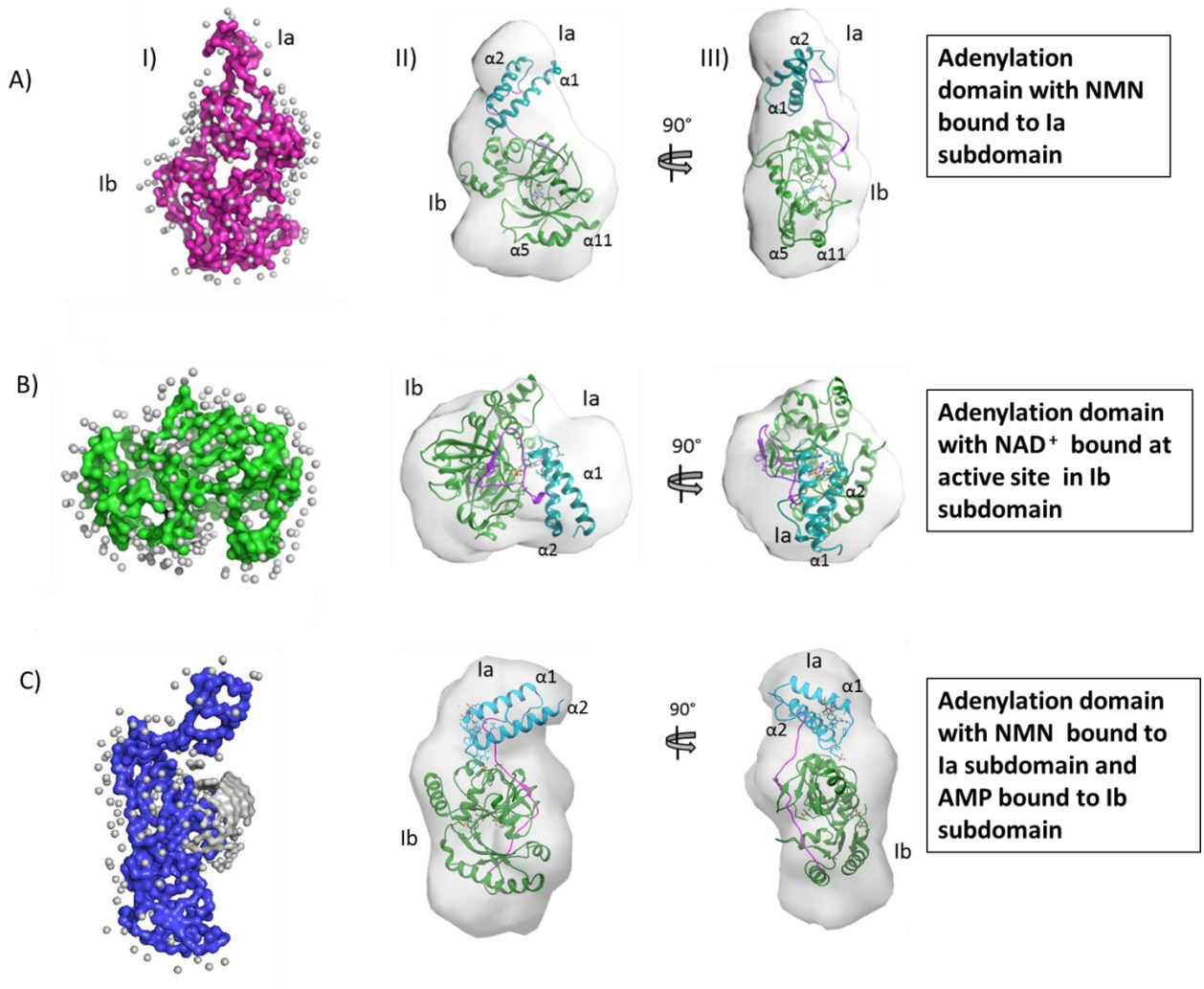


Figure 3.7 The comparative analysis of AdD domain rotation during the adenylation reaction bound to cofactors. A) I, The Gasbor analysis shows the extended conformation of AdD domain (magenta). **II and III,** Envelope model of de-adenylated AdD domain generated using DAMMIF and the higher resolution structure were fitted using SUPCOMB. **B) I,** Ia subdomain rotation is observed in AdD domain in the presence of NAD+. The Gasbor model (green) shows the globular compact structure of adenylated form of AdD domain. **II and III,** In agreement with the Gasbor model, the DAMMIF models were also showing the compact behavior and the high resolution structures were fitted using SUPCOMB**. C) I**, The gasbor analysis shows the extended confirmation again observed when NMN and AMP are bound at their respective binding site (blue). **II and III**, DAMMIF envelop and high resolution structures were fitted using SUPCOMB.

Table 3.3: SAXS data collection and statistics of AdD domain

	AdD+NMN - SASDEU5	**AdD+NAD$^+$**- SASDEC6	**AdD+NMN+AMP**- SASDG65
Data Collection Parameters			
X-ray source	SAXSspace, CDRI	SAXSspace, CDRI	SAXSpace, CDRI
Detector	HPC Mythen2 R1K	HPC Mythen2 R1K	HPC Mythen2 R1K
Wavelength (nm)	0.154	0.154	0.154
Exposure time per frame	1800 sec	1800 sec	1800sec
No. of frames collected	2	2	2
Concentration (mg/ml)	12.5	12.5	12.5
Measurement temperature (^{0}C)	10	10	10
Structural parameters			
I(0) [from Guinier]	42835.10±527.73	599573±22300	51321.1±611.42
Rg (nm) [from Guinier]	2.43±0.05	2.37±0.13	2.49±0.05
I(0) [from P(r)]	42780	595300	51080
Rg (nm) [from P(r)]	2.48	2.35	2.50
D_{max} (nm)	8.84	6.17	8.89
V^{porod} (nm^3)	69	94.26	68.10
Modelling parameters			
Symmetry	P1	P1	P1
Chi^2 (Gasbor)	0.41	0.23	0.21
Model fitting	SUPCOMB	SUPCOMB	SUPCOMB
NSD	2.6	2.3	2.6

3.2.5 Crystal structures of AdD domain of MtbLigA-

a) AdD domain crystal structure with NMN and AMP bound (PDB: 6KJM)

The first step of adenylation reaction starts with the binding of NAD^+ to open conformation of Ia subdomain. We determined the crystal structure of adenylation domain with NMN as well as AMP bound (**Fig. 3.8 A, B**, PDB: **6KJM**) where NMN is present within the two helices of Ia subdomain as previously shown in *E. fa.* and in *B. su* (Gajiwala and Pinko 2004). The residues in two helices (α1 and α2) play crucial role in stabilization of NMN ring by making direct contact with it. The NMN ring is sandwiched between Y31 from helix 1 and F44 from helix 2. The hydroxyl group of Y32 is in hydrogen bonding range with the phosphate group of NMN. Y31 and Y32 are within the Van der Waal's interaction range and Y32 orients Y31 for its optimal interaction with nicotinamide ring of NMN. D45 (helix 2) is within the H-bonding range (3.45 Å) from 2'OH group of ribose ring (**Fig. 3.8 B, I**). I39 and D41 which are present in the loop connecting the two helices makes H-bond with amide N of the NMN with distance of 3.0 Å and 3.06 Å respectively. In place of S36 in *E. fa.,* a non-polar isoleucine (I39) is evolved in MtbLigA for NMN stabilization. A water molecule present in the site is involved in the H-bonding with the D41 (distance 2.75 Å). G72 forms two H-bond acting as acceptor for O1 and O2 oxygen's of phosphate group of NMN with distance of 3.03 Å for O1P and 2.97 Å for O2P. In *E. fa.* R158 makes direct bonding with NMN but in MtbLigA, A161 is present which do not interact directly with NMN. Thus NMN is stabilized in binding pocket mostly by interactions involved with nicotinamide and ribose ring. The canonical niche of NAD^+ binding is evolved in such a way that whole NAD^+ molecule can reconcile themselves within binding pocket with utmost stabilization interaction with NMN in Ia subdomain as compared to AMP. The AMP is transferred to the active site lysine where it adapts '*anti*' or '*syn*' conformations. In this structure (PDB: **6KJM**) AMP adapts '*anti*' conformation which is stabilized by base stacking interactions of H236 and V298 with adenine ring of AMP (**Fig. 3.8 B, II**). K123 and K324 are involved in H-bonding with O3P of AMP (2.86 Å and 3.10 Å, respectively) and N94 forms non-bonded contacts with O1P of AMP. E184 which is placed above H236 forms hydrogen bond with the 3'O of ribose sugar. AMP is further stabilized by several H-bonding interactions with residues involving E121, I124 and K300 in the distances of 2.84, 3.24 and 3.10 Å respectively with AMP. The LIGPLOT analysis of the interaction involving NMN and AMP are depicted in **Figure 3.9**.

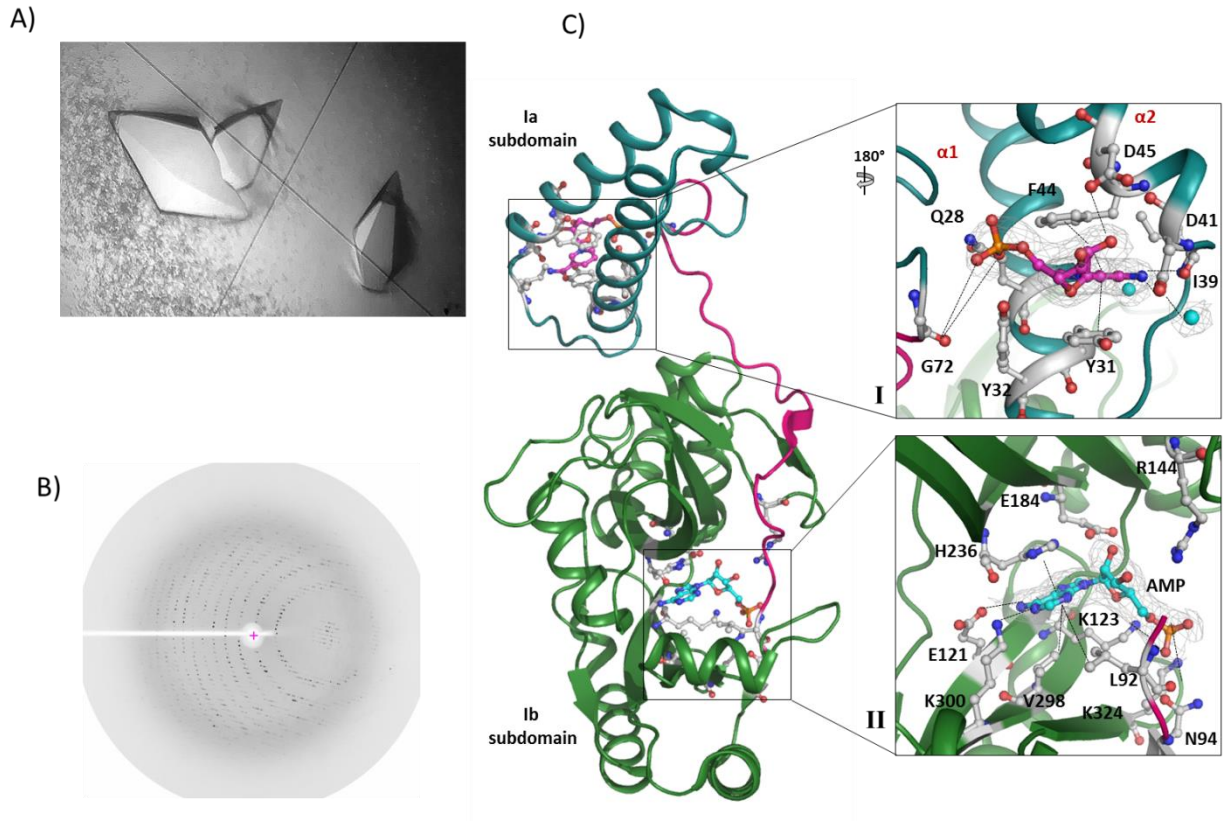


Figure 3.8 Crystal structure of Adenylation domain with NMN and AMP bound at their respective binding sites. A) The tetragonal shaped crystal of AdD domain obtained after one week at room temperature in condition 0.1M HEPES Na pH7.6, 0.1M NaCl and 1.5M Ammonium sulfate. **B)** The diffraction image was generated by Imosflm. **C)** Crystal structure of AdD domain (8-328 aa) (PDB: **6KJM**) showing the open conformation of Ia, Ib subdomains. The α1 and α2 helices are shown in dark cyan and Ib subdomain in green. I, zoomed view of the NMN binding pocket with residues responsible for stabilization of NMN II, zoomed view of AMP binding active site with AMP bound in '*anti*' confirmation. The figure is generated using PyMOL 0.99rc6.

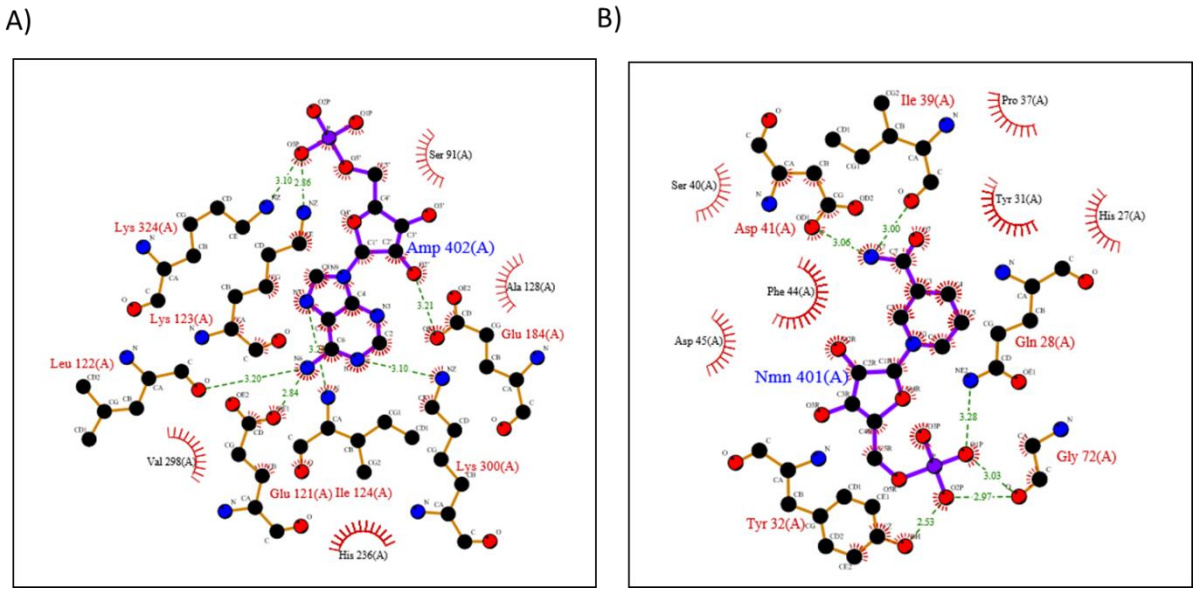


Figure 3.9 LIGPLOT analysis of AdD domain crystal structure done using PDBsum Pictorial database of 3D structures in the Protein Data Bank. A) LIGPLOT of interaction involving cofactor AMP bound in active site. K123 and K324 forms H-bond with phosphate group of AMP. E184 forms H-bond with ribose sugar of AMP. Other residues involved in H-bonding are L122, E121, K300. **B)** LIGPLOT analysis of interactions involving cofactor NMN bound in Ia subdomain NMN binding site. The nicotinamide ring is base stacked between F44 and Y31. D41 and I39 forms H-bond with the amide N of the nicotinamide ring.

b) Crystal structure of active site mutants K123A and K123R

Further, we crystallized the active site mutants K123A and K123R in the presence of NMN and NAD^+, respectively (**Fig. 3.10 A, B**). The K123A crystal structure (PDB: **6KSC**) shows the well-defined NMN density between the two helices of Ia subdomain. In this structure, we found dual occupancy of NMN within the open conformation of Ia subdomain (**Fig. 3.11 A**). Due to the dual occupancy of phosphate group of NMN, it is projected away by 3.6 Å and 4.8 Å as compared to the NMN bound AdD domain crystal structure (PDB: **6KJM**). The nicotinamide ring forms base stacking interaction with the F44 and Y31. I39 and D41 each form two H-bonds with amide N of NMN. The Y32 is not in appropriate orientation to maximize the Van der Waal's interaction with Y31 and is tilted by 14.1° towards outside which may facilitate the release of NMN from the binding pocket (**Fig. 3.11 A, I**). As compared to the AdD domain crystal structure (PDB: **6KJM**), a water molecule is present near G72 instead of D41 and another water molecule is present near phosphate group of NMN which interacts with O16 of phosphate group of NMN in first occupancy and O17 in second occupancy (3.15 Å and 3.35 Å respectively). The absence of water molecule near G72

in 6KJM structure may be due the presence of phosphate group of NMN which is involved in the H-bonding with G72 therefore barely leaving space for any water to adjust. The three different occupancies (1- for NMN of AdD domain crystal structure, 2A and 2B- for dual occupancy in K123A NMN bound crystal structure) for NMN as shown in the surface model clearly demonstrate the displacement of water molecule from the site in the AdD crystal structure due clash with phosphate group of NMN (**Fig. 3.11 A, II**). The exclusive conservation of K123 throughout the DNA ligases attracts the attention to mutate it into positively charged amino acid like arginine without much changing the length of side chain. The presence of '*syn*' AMP at active site indicates the successful adenylyl transfer of AMP to the K123R mutant (PDB: **6KSD**). E184, a conserved residue important for NAD^+ recognition is engaged in direct H-bonding with 2'O of ribose sugar. R144 act as H-bond donor (2.9 Å) to O19 of phosphate group of AMP. Other major residues involved in H-bonding are- E121, L122, I124, and K300. K300 plays the major role in determination of specificity of adenine moiety. V298 is involved in non-bonded interaction with AMP **(Fig. 3.11 B**). K123R mutant is not able to ligate the nicked substrate may be due to the steric clashes at the active site during the accommodation of DNA (**Fig. 3.11 C**). The LIGPLOT analysis of both mutant crystal structures is shown in the **Figure 3.12**. The surface analysis of K123R mutant shows shallower active (5.4 Å wide) site as compare to wild type active site (7.15 Å wide) (**Fig. 3.13 A,B**).

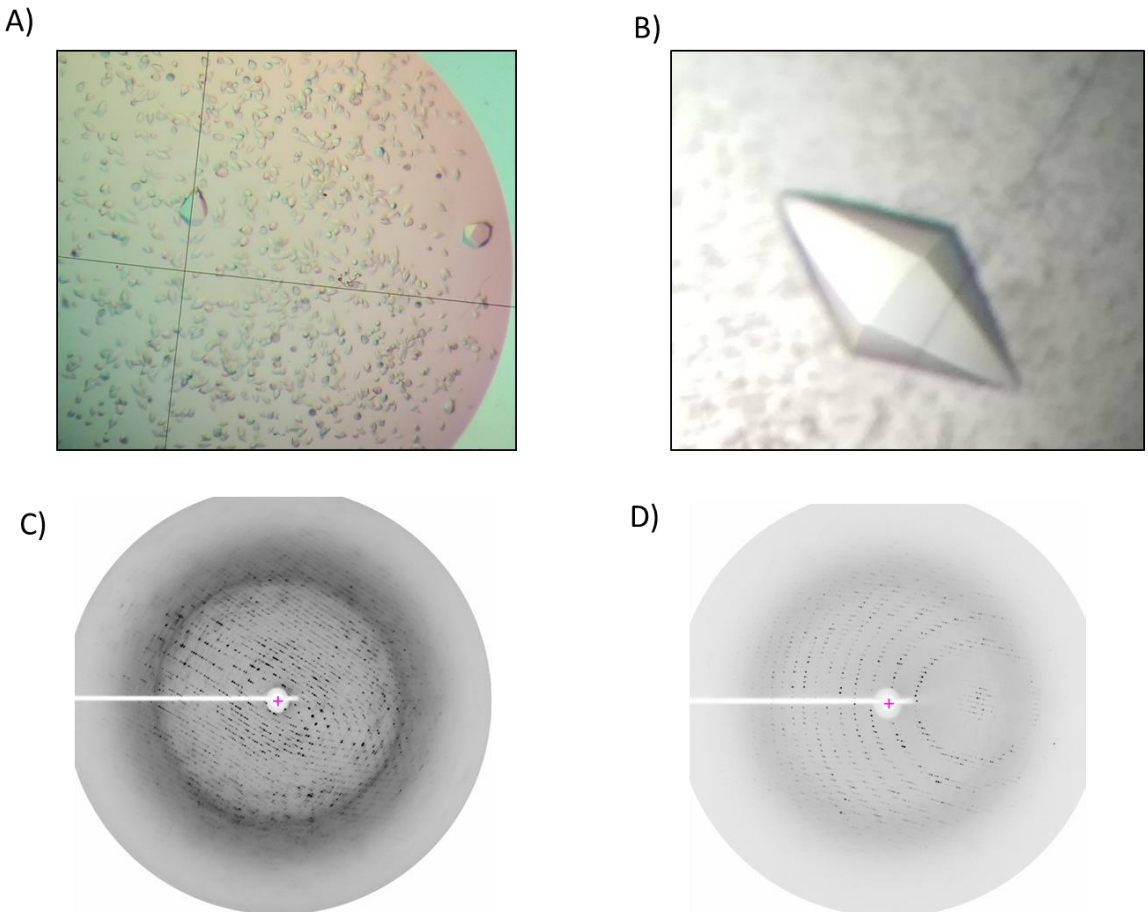
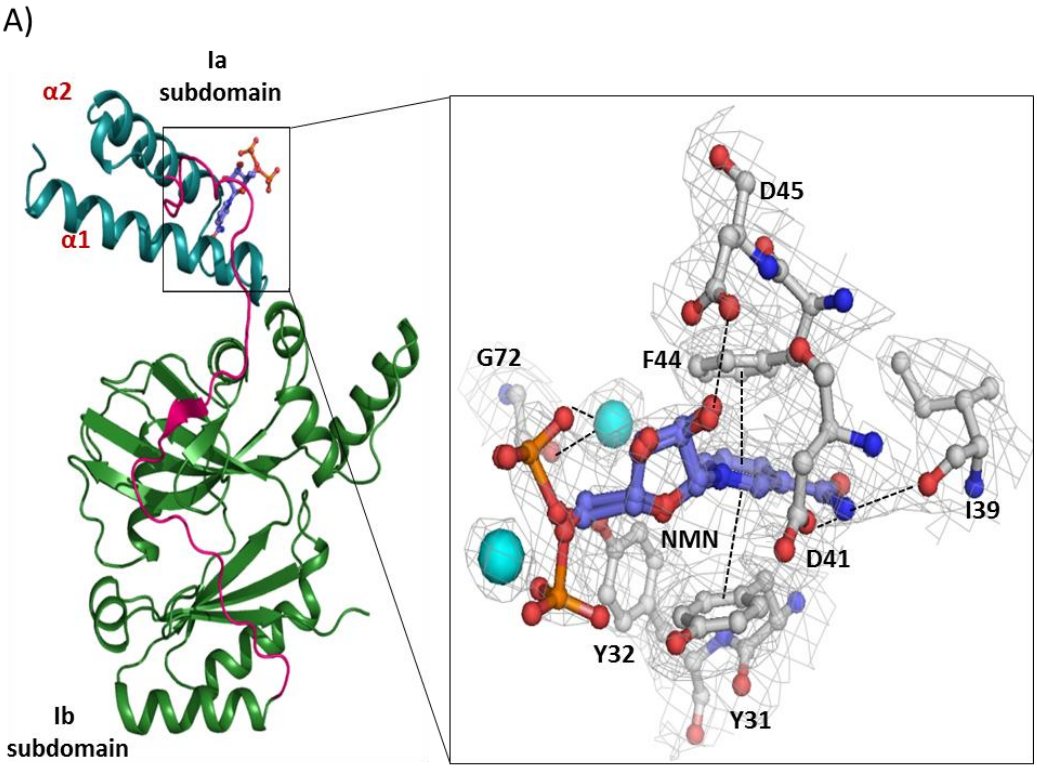
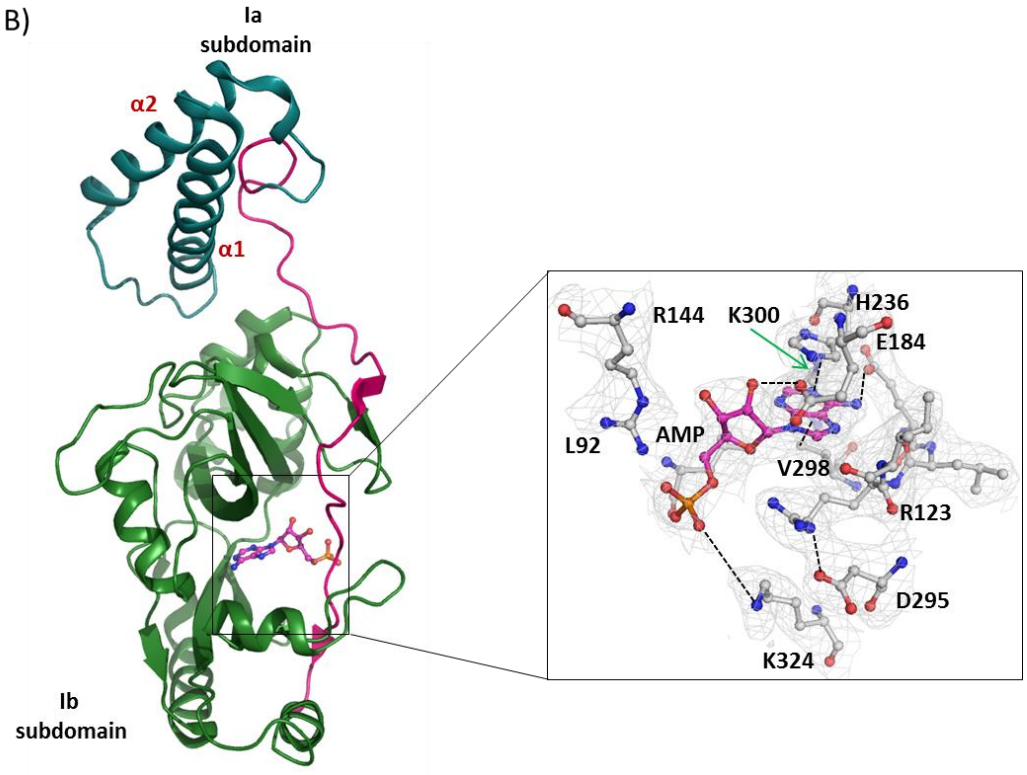


Figure 3.10 The crystals and diffraction pattern of active site mutants. A) The crystals of active mutant K123A **B)** The crystal of active site mutant K123R. Both the crystals are crystallized in the same condition as AdD domain crystal. **C and D)** The diffraction pattern of K123A and K123R generated from Imosflm.


A)
Ia
subdomain
α2
α1
Ib
subdomain
D45
G72
F44
I39
NMN
D41
Y32
Y31


B)
Ia
subdomain
α2
α1
Ib
subdomain
R144
K300
H236
E184
L92
AMP
V298
R123
D295
K324

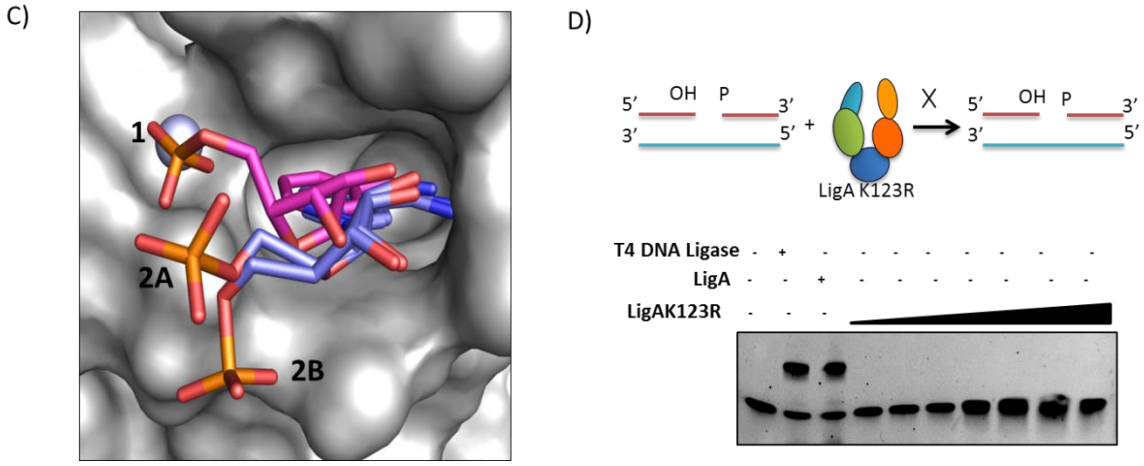


Figure 3.11 Crystal structure of active site mutant K123A and K123R. A) The crystal structure of K123A mutant shows the different conformation of NMN into the binding pocket. The residues Y31, Y32, I39, D41, F44 and D45 plays important role in NMN stabilization. **II,** The Phosphate group of NMN with dual occupancy is protruded outside the pocket (PDB: **6KSD**) in comparison with AdD NMN bound AdD domain (PDB: 6KJM). **B)** The AMP binding pocket of K123R (PDB: **6KSC**). The adenine ring is stacked between H236 and V298 by base stacking interaction. L92, K324 stabilizes the phosphate group while E184 serve as adenine specificity determinant. The Figures are generated from PyMOL 0.99rc6. **C)** The in vitro activity of K123R mutant which shows its inefficiency to ligate the nicked oligo. Lane 1- control, lane 2-T4 DNA Ligase Lane 3- Wild type LigA, Lane 4 to 10-LigA K123R mutant.

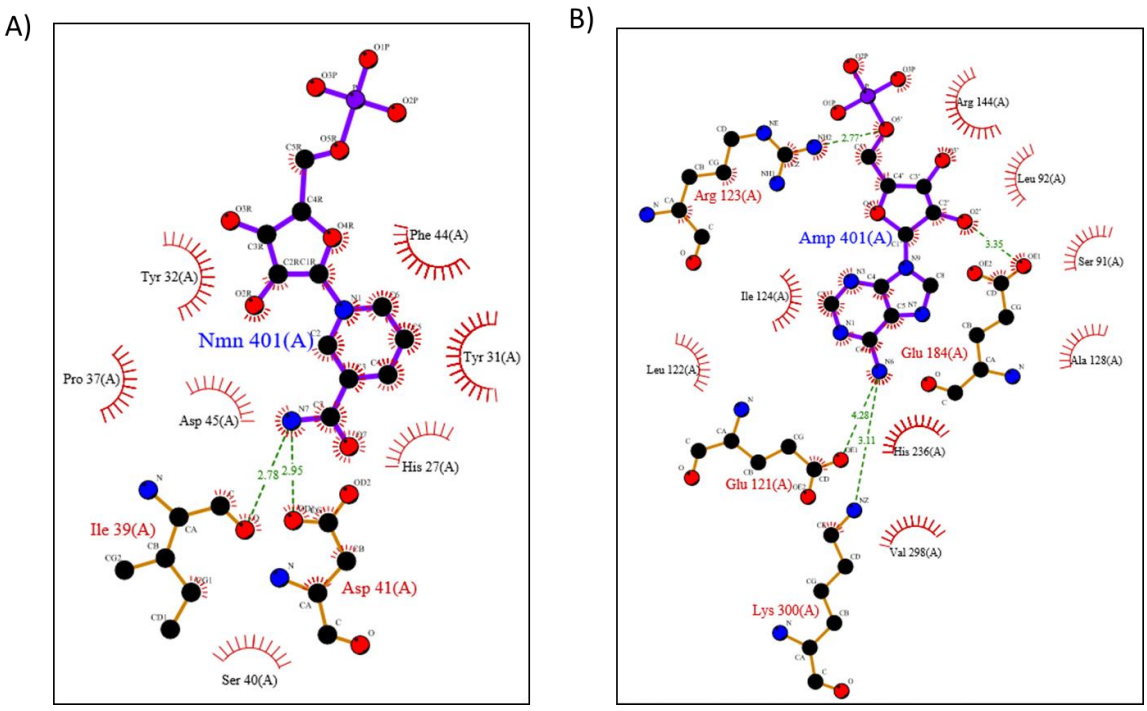


Figure 3.12 LIGPLOT analysis of active site mutants of AdD domain crystal structure done using PDBsum Pictorial database of 3D structures in the Protein Data Bank. A) LIGPLOT analysis of interactions involving cofactor NMN bound in Ia subdomain NMN binding site in K123A mutant (PDB: **6KSD**). The nicotinamide ring is base stacked between F44 and Y31. D41 and I39 forms H-bond with the amide N nicotinamide ring. **B)** LIGPLOT of interaction involving cofactor AMP bound in active site in K123R mutant (PDB: **6KSC**). K123 forms H-bond with phosphate group of AMP. AMP exist in '*syn*' conformation. E184 forms H-bond with ribose sugar of AMP. Other residues involved in H-bonding are E121 and K300.

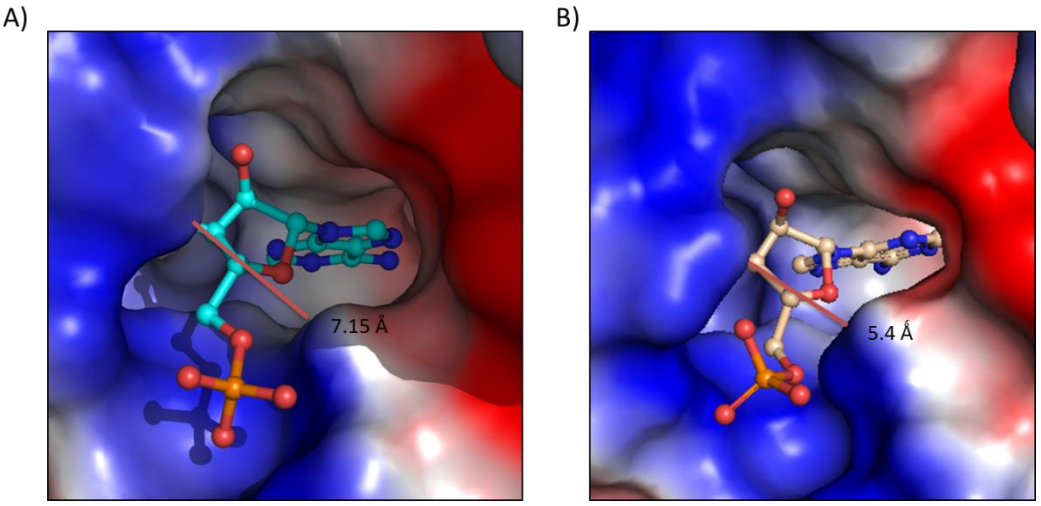


Figure 3.13 Comparison of active site between wild type AdD domain and active site mutated K123R AdD domain. A) The wide pocket of wild type active site **B)** Narrowing of the active site due to the mutation of K123R and may be the reason of reduced activity.

c) Crystal structure of Ia and Ib subdomain mutants E22A and H236Y

The consequences of structure guided mutation of Ia subdomain residues E22, E26 and E87 to alanine are further tried to determine by crystallization. We got poor diffracting crystals of E26A and E87A although E22A crystals (**Fig. 3.14 A**) diffracted well at resolution 2.2 Å (PDB: **6KDU**) and we found one different conformation for NMN and covalently bound AMP to the K123. NMN is stacked between the F44 and Y31. The amide O of nicotinamide ring is further stabilized by water mediated interaction with I39. The O1 of D45 orients in such a way that it can form hydrogen bond with the 3'O of ribose sugar (**Fig. 3.15 A**). AMP is stabilized in '*anti*' conformation in the active site. K123 is involved in covalent bonding with phosphate group of AMP (1.30 Å). Other residues stabilizing AMP are E121 and L122 by forming H-bond with amide N of adenine moiety and N94 by acting as H-bond acceptor to O20 of phosphate group of AMP (**Fig. 3.15 B**).

The conservation of tyrosine in active site to facilitate the base stacking interaction with the AMP is present in most of the prokaryotes, while it is replaced by histidine only in *M. tuberculosis*. So we mutated the H236 to Y236 and crystallized in the presence NAD^+ (PDB: **6KRH**) (**Fig. 3.14 B**). The crystal structure shows the density for NMN and AMP in their respective sites. The NMN binding site residues are in the same orientation as seen in the E22A crystals (PDB: **6KDU**) and most of the active site residues are in the same conformation as in the MtbLigA1 (PDB: **3SGI**) structure except R144 which is involved in interaction with O2 of ribose sugar of AMP while in H236Y mutant structure R144 goes apart around 63°. AMP is in '*anti*' conformation and K123 forms covalent bond with phosphate group of AMP (**Fig. 3.15 C**). The LIGPLOT analysis of E22A and H236Y mutant with interactions involving cofactors bound in their respective sites are depicted in **Figure 3.16 and 3.17**. By analyzing all the NMN bound structures, we found that NMN adapts a series of conformation in the binding pocket. These local occupancies of NMN show that the phosphate group leaves the binding niche first (**Fig. 3.18**). The data collection and refinement statistics of all the crystal structures are mentioned in **Table 3.4**.

The different conformation of AMP in their respective crystal structures of AdD domain and the residues involved in the interaction are mentioned in **Table 3.5.**

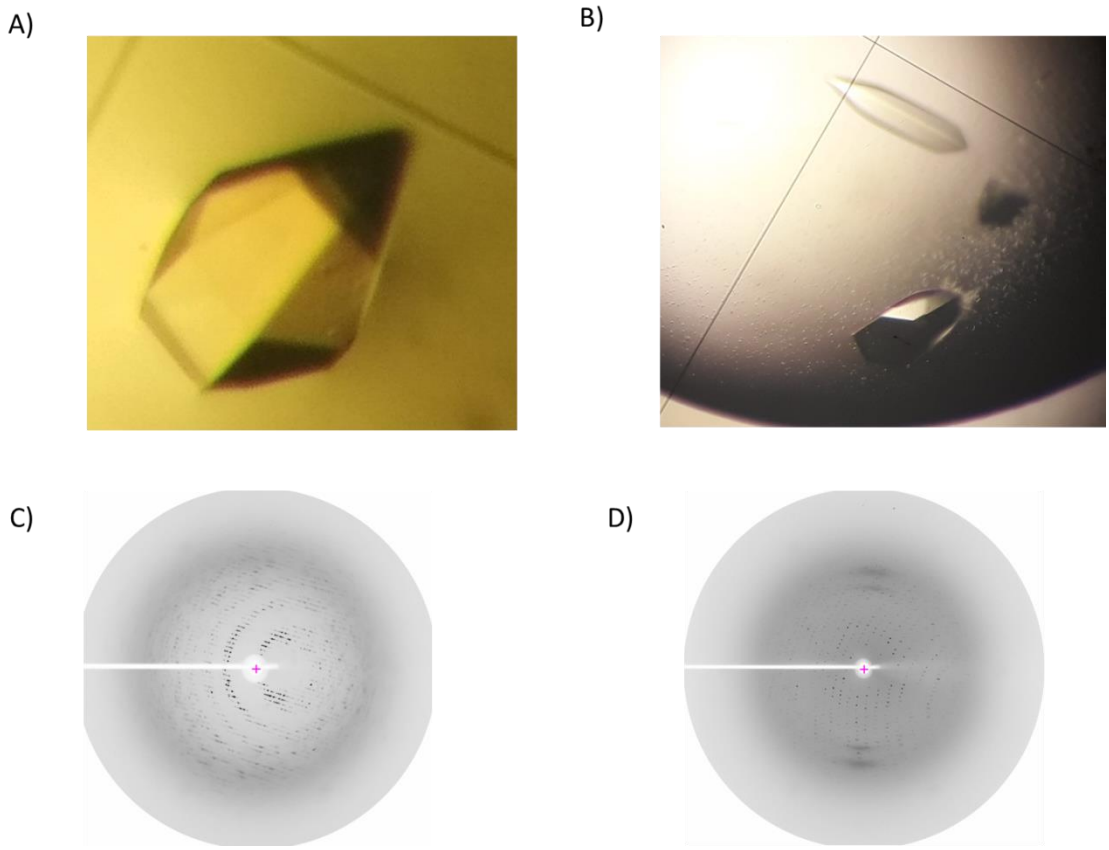


Figure 3.14 Crystals and diffraction pattern of Ia and Ib subdomain mutants of AdD domain. A) The crystals of Ia subdomain mutant E22A which is critical for interaction with Ib subdomain during the adenylation reaction. **B)** The crystal of H236Y mutant that helps in base stacking of AMP. Both the crystals are crystallized in the same condition as AdD domain crystal. **C and D)** The diffraction image of E22A and H236Y generated from Imosflm.

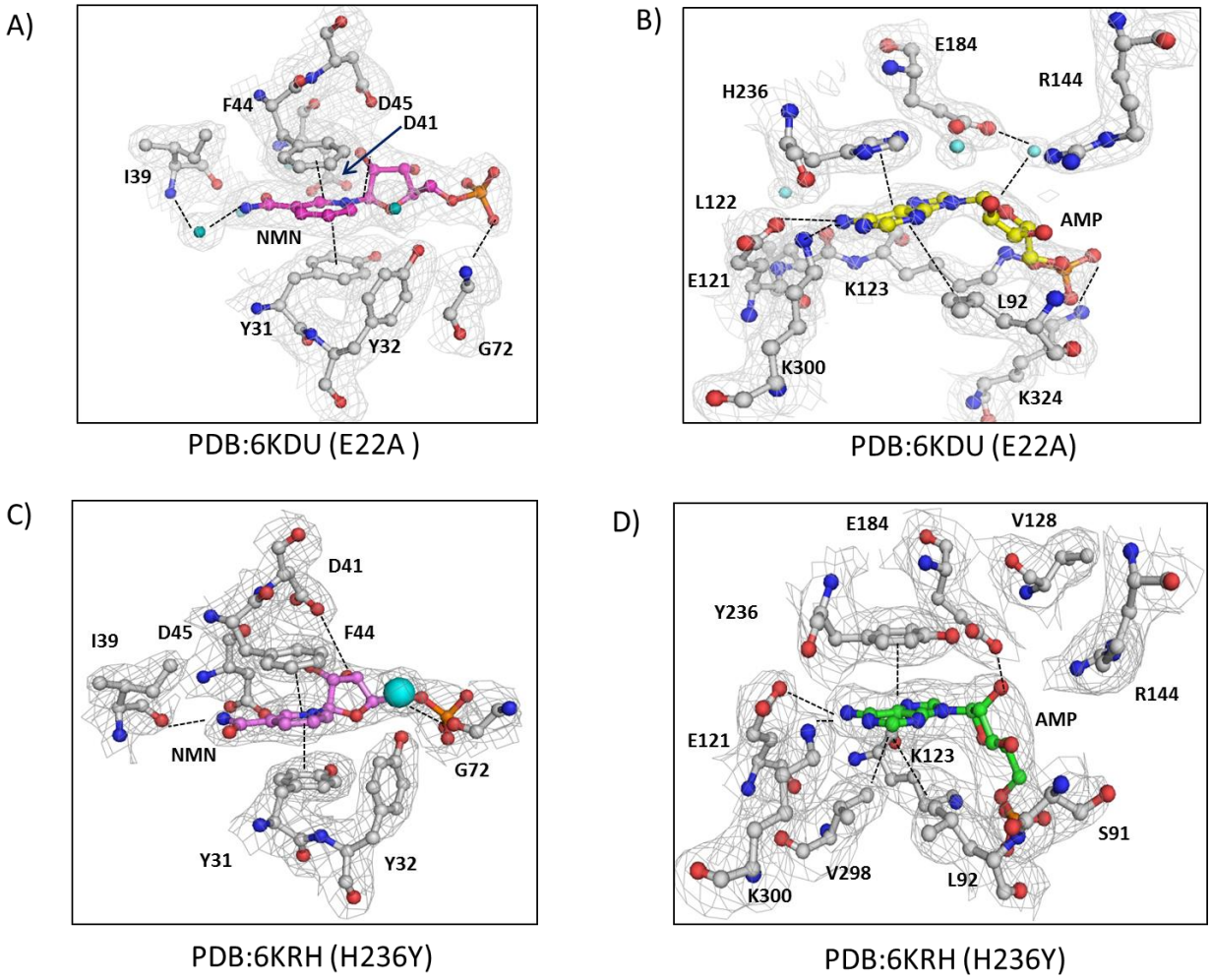


Figure 3.15 The zoomed view of active site in E22A (PDB: 6KDU) and H236Y (PDB: 6KRH) mutant crystal structure. A) The crystal structure of NMN bound E22A mutant shows the one more different conformation of NMN into the binding pocket. The I39 forms a water mediated hydrogen bond with the N of nicotinamide ring of NMN. The D45 forms hydrogen bond with the 3' OH of ribose sugar. Y32 occupies a position in such a way that it can form hydrogen bond with the oxygen NMN phosphate group (PDB: **6KDU**). **B)** The covalently bound AMP is stabilized into the binding pocket. The E184 forms an water mediated interaction with the O of ribose sugar. K300 and E121 are within the range of hydrogen bonding with the amide N. **C)** The crystal structure of NMN bound H236Y mutant (PDB: **6KRH**). A water-mediated interaction is seen with G72. In addition, water present near I39 in case E22A mutant is absent in this crystal structure. **D)** The AMP binding pocket of H236Y crystal structure demonstrates the efficient binding of AMP at the active site (PDB: **6KRH**).

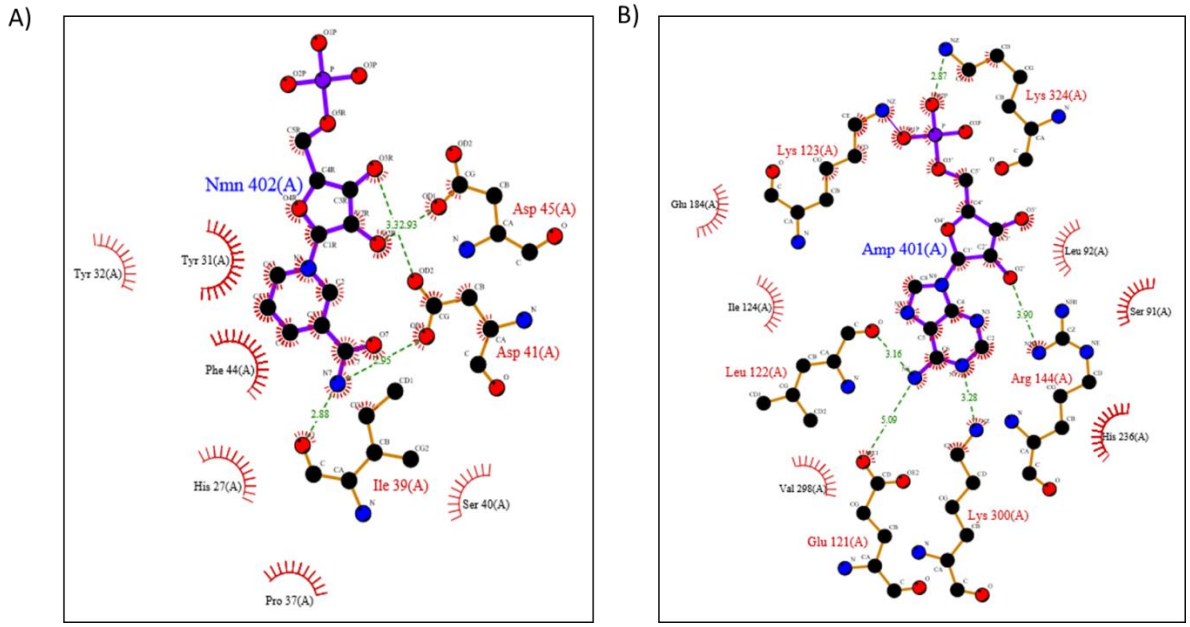


Figure 3.16 LIGPLOT analysis of Ia subdomain mutant E22A of AdD domain crystal structure (PDB: 6KDU) done using PDBsum Pictorial database of 3D structures in the Protein Data Bank. A) LIGPLOT analysis of interactions involving cofactor NMN bound in Ia subdomain NMN binding site. The nicotinamide ring is base stacked between F44 and Y31. D41 and I39 forms H-bond with the amide N nicotinamide ring. **B)** LIGPLOT of interaction involving cofactor AMP bound in active site. K123 forms covalent with phosphate group of AMP and AMP is in '*anti'* form. E184 forms H-bond with ribose sugar of AMP. Other residues involved in H-bonding are E121 and K300.

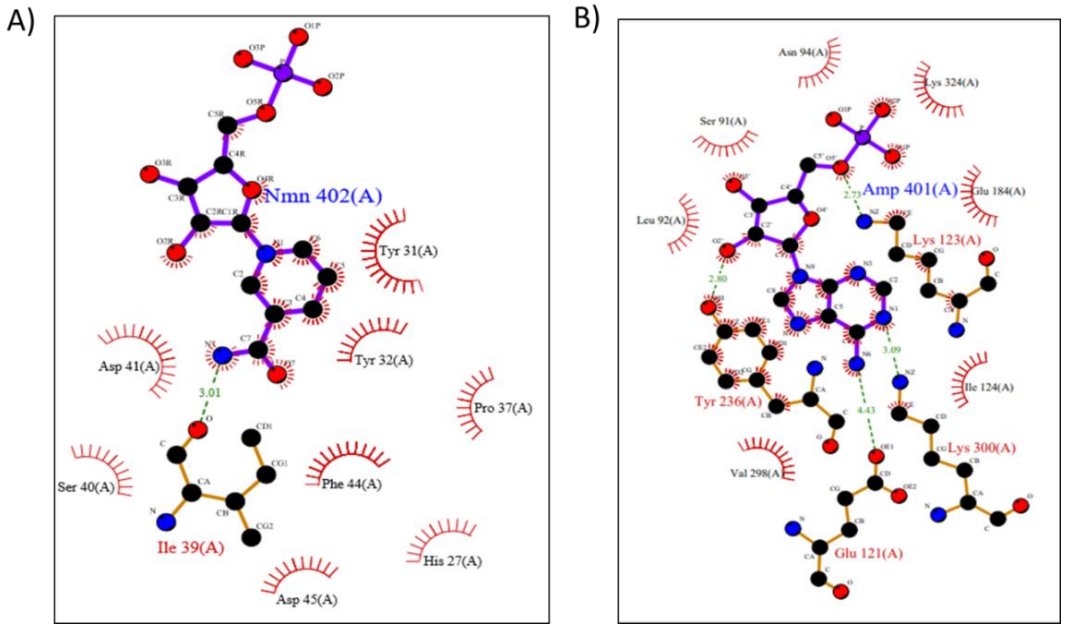


Figure 3.17 LIGPLOT analysis of Ib subdomain mutant H236Y of AdD domain crystal structure (PDB: 6KRH) done using PDBsum Pictorial database of 3D structures in the Protein Data Bank. A) LIGPLOT analysis of interactions involving cofactor NMN bound in Ia subdomain NMN binding site Ia subdomain. The nicotinamide ring is base stacked between F44 and Y31. D41 and I39 forms H-bond with the amide N nicotinamide ring. **B)** LIGPLOT of interaction involving cofactor AMP bound in active site. K123 forms covalent bond with phosphate group of AMP. AMP exist in '*anti'* conformation. E184 forms H-bond with ribose sugar of AMP. Other residues involved in H-bonding are E121 and K300.

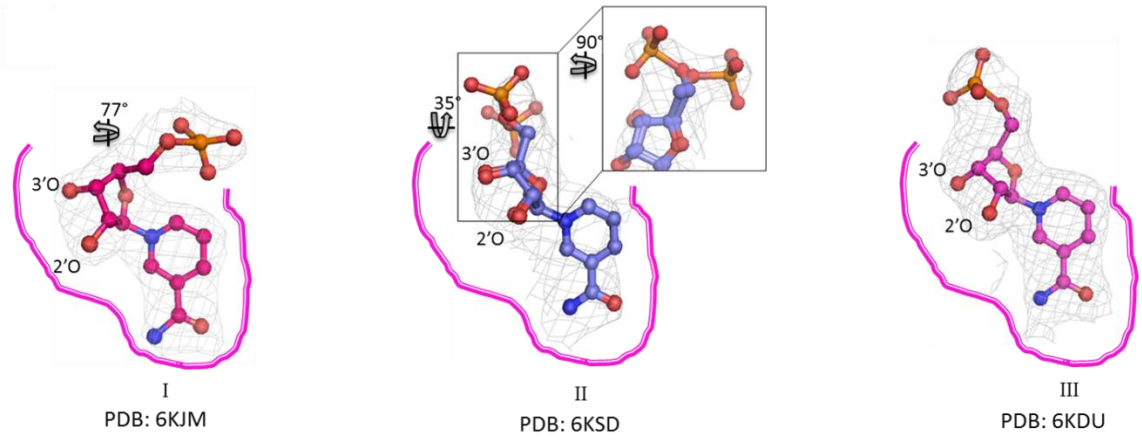


Figure 3.18 Comparison of NMN conformation in wild type AdD domain (PDB: 6KJM), active site mutant K123A (PDB: 6KSD) and Ia subdomain mutant E22A (PDB: 6KDU) . The three different conformations of NMN (**6KJM**, **6KSD** and **6KDU**) in the binding pocket indicates that phosphate group leaves the binding niche first.

3.5.6 Serial remodeling of active site on the basis of crystal and solution structures

The intermediates of ligation reaction from previously reported crystal structure of AdD of *E. fa* (Gajiwala and Pinko 2004) and *M. tb* (Srivastava, Tripathi et al. 2005) exhibit only first adenylation step but do not evidences the mechanism of NMN departure from active site and also the successive transitions happening into the basic architecture of AdD domain during the course of reaction. By comparison, of crystal and solution structures here we can plausibly demonstrate the sequential adenylation of K123. The crystal structure (PDB: **6KSC)** supported by SAXS study of de-adenylated form of AdD domain demonstrate the expanded form with entirely relaxed Ia subdomain aside from active site in Ib subdomain (SASBDB ID: **SASDEC6**) (**Fig. 3.19 A**). The gasbor analysis of buffer subtracted scattering profile shows the extended density with best fit of chi^2 0.41. In *E. fa.* when NMN bound crystals were exposed to AMP, the crystal packing changed from C2 to $P2_1$, and obtained a new electron density map indicating the difference in the position of Ia subdomain (Gajiwala and Pinko 2004). This adenylation reaction intermediate (PDB: **1TAE**) along with the solution structure (SASDBD ID: **SASDEC6**) showed the rotation of Ia subdomain around 180° to deliver AMP to active site (**Fig. 3.19 B**)(Gajiwala and Pinko 2004). The 2Fo-Fc map establishes the evidences the clear density of NAD^+ in the active site. The insertion of NAD^+ into the active site causes a conformational switching in Ib subdomain to correctly make a platform for AMP transfer to active site lysine by phosphoamide bond formation at low pH conditions. The maximum dimension (D_{max}) reduces from 8.84 nm to 6.17 nm (**Fig. 3.19 E**). The scattering profile indicates the globularity with best fit of chi^2 0.23. The crystal structures are superimposed on SAXS density by SUPCOMB with NSD values of 2.6 and 2.3 for de-adenylated and adenylated AdD domain respectively. After transfer of AMP from NAD^+ to active site lysine, Ia subdomain rotates almost 54° and flips both helices (α1 and α2) by 42.9° (PDB: **3SGI**) as compared to NMN bounded helices where the crossing angle is 35.8° to nearby environment (**Fig. 3.19 C**). This novel intermediate in concomitant steps of reaction suggests the mechanism to keep off NMN from active site to prevent the reversal of reaction. The E87 present at loop region between Ia and Ib subdomain forms an ionic interaction with H240. The conservation of this ionic interaction in all reported conformation of NAD^+-dependent ligases may serve as a loop locker pair required for the rotation of Ia subdomain around the Ib subdomain during the course of adenylation reaction. As compared to fully

closed conformation of Ia subdomain, the partially closed conformation is stabilized by extensive ionic interactions with Ib subdomain. E22, E26, and E87 from proximal region of Ia subdomain poses themselves within the ionic bond range with R315, R144 and K357 of Ib subdomain, respectively. R34 is present 3.66 Å apart from the oxygen of D146 and may undergo ionic stabilization for maintaining the partially closed conformation (**Fig. 3.20 A**). The NAD^+ bound AdD crystal structure of *E. fa.* (PDB: **1TAE**) shows only 4 ionic interactions between Ia and Ib subdomain as compared to MtbLigA1 crystal structure (PDB: **3SGI**) having 7 ionic interactions between two domains (**Fig. 3.20 A, B**). The crystal structure is superimposed on SAXS density with NSD value of 2.4. The Ia subdomain again rotates 171° to attain the subsequent conformation to deliver AMP to 5'PO_4 of the nick with bound NMN at Ia subdomain (PDB: 6KJM). The solution structure (SASDBD ID: **SASDG65**) also validates the fact that before the entry of nicked DNA, the Ia subdomain leaves the vicinity of Ib subdomain. The scattering profile also indicates the stretched out structure with fit of 0.21 chi^2 and D_{max} of 8.8 nm. The NSD value of superimposition is 2.6 (**Fig. 3.19 D**). The kratky analysis of all the solution structures shows the structural changes occurring during the reaction (**Fig. 3.19 B**) and guinier analysis indicates that all the protein samples are homogenous and monodisperse (**Fig. 3.19 C**). The linker region (62-92 amino acid) between the Ia subdomain and Ib subdomain is showing the intact electron density during adenylation reaction of K123 (**Fig. 3.21 A, B, C**). The stepwise fluctuation within the interacting capacity between Ia and Ib subdomain demonstrates the different biologically applicable conformational states of AdD domain throughout the successive steps of DNA ligation reaction. The SAXS data collection statistics are mentioned in **table 3.3.**

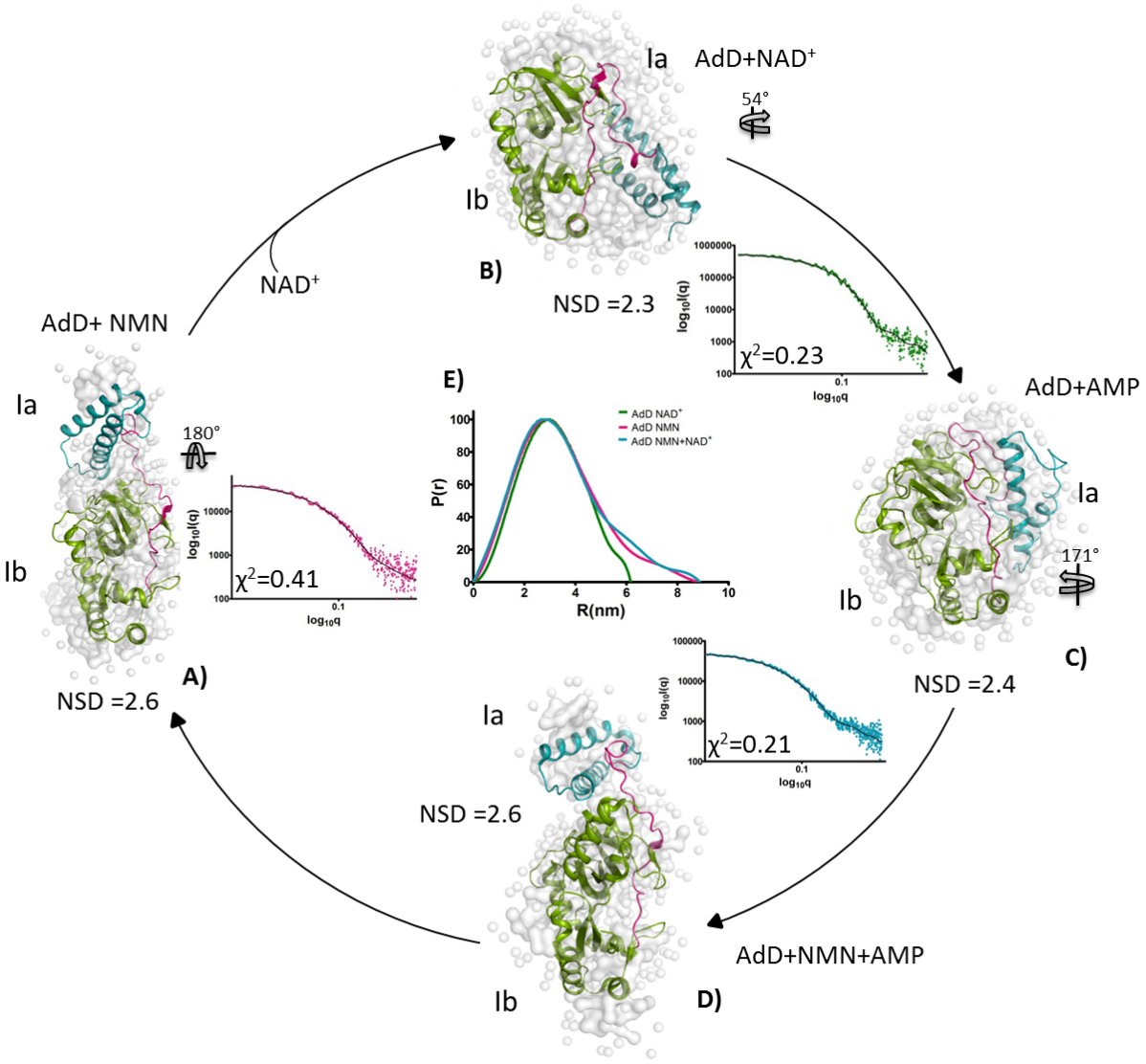


Figure 3.19 Serial remodeling of active site on the basis of crystal and solution structure. A) Crystal structure (PDB: **6KJM**) of the de-adenylated form of AdD domain superimposed on SAXS envelope using SUPCOMB with NSD of 2.6. The scattering profile indicates the open extended form with gasbor fit of chi2 0.41. **B)** In presence of NAD^+, Ia subdomain rotates around Ib subdomain (180˚) to deliver AMP to active site. The crystal structure of *E. fa* AdD (PDB: **1TAE**) is superimposed on SAXS density with NSD 2.3. The scattering curve of adenylated AdD domain indicates the globularity with gasbor fit chi2 of 0.23 .**C)** The Ia subdomain again rotates around 54˚ to keep off the NMN from active site in order to prevent reaction reversal. The crystal structure of adenylated AdD domain (PDB: **3SGI**) is superimposed on SAXS density with NSD of 2.4. D, after adenylating the K123 of Ib subdomain, Ia again goes back in extended confirmation by rotating around 171˚ and crystal structure with bound NMN and AMP (PDB: **6KJM**) is superimposed on SAXS density with NSD of 2.6 and scattering with gasbor fit of chi2 0.21. E, The P(r) obtained by SAXS analysis indicates the different positions of Ia subdomain with Dmax 6.17 nm (green) for adenylated NAD+ bound AdD domain and 8.84 nm and 8.9 nm (pink and dark cyan) for de-adenylated NMN and AMP+NMN bound AdD domain respectively.

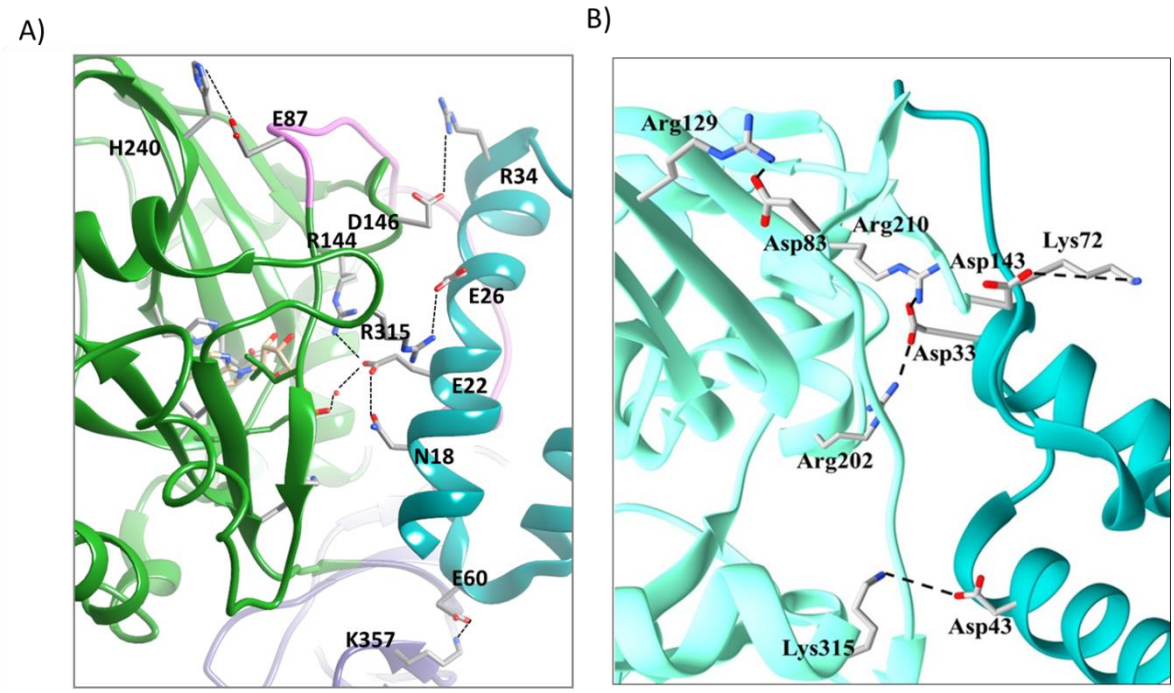


Figure 3.20 The stabilization of particular conformation of Ia subdomain near the Ib subdomain. **A)** PDB: 3SGI. The E22 and E26 and R34 forms hydrogen bonds with R144, R315 and D146. **B)** PDB: 1TAE shows the four ionic interactions between R129 and D83, R202 and D33, K315 and D43, K72 and D143.

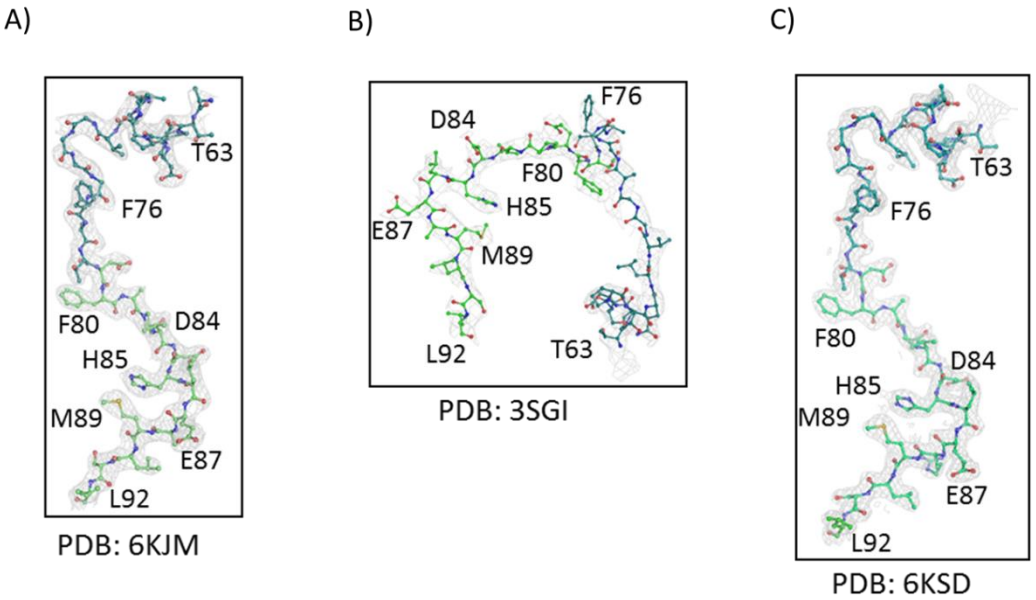


Figure 3.21 Well defined electron density contoured at 1.0 sigma was observed for crystallographic snapshots of BRCT deleted MtbLigA and AdD domain bound with NMN and AMP (PDB: 3SGI and 6KJM respectively) and active site mutant K123A (PDB: 6KSD). A, B and C) The loop between Ia and Ib subdomain (62-92 amino acid) remains intact during the adenylation reaction as shown by the continuous electron density.

Table 3.4: X-ray data collection and statistics

Data collection	6KJM-**Wildtype AdD domain**	6KSD-**Active site mutant K123A**	6KSC-**Active site mutant K123R**	6KDU-**Ia subdomain mutant E22A**	6KRH-**Ib subdomain mutant H236Y**
Space group	P61 2 2	P61 2 2	P61 2 2	P61 2 2	P21 2 2
Cell dimension	a=95.73,b=95.73, c= 201.35 α= 90, β= 90, γ= 120	a=95.88,b=95.88, c= 200.04 α= 90, β= 90, γ= 120	a= 95.23, b= 95.23,c=201.611 α= 90, β= 90, γ= 120	a=96.168,b=96.168, c= 200.835 α= 90, β= 90, γ= 120	a=95.55,b=95.55, c= 200.59 α= 90, β= 90, γ= 120
Wavelength (nm)	0.9795	0.9795	0.9795	0.9795	0.9795
Resolution range (Å)	43.23-2.2	46.557-2.5	36.27-2.40	43.23-2.2	47.77-2.6
Completeness (%)	100	99.9	99.1	100	100
Multiplicity	8.0	10.2	11.4	19.2	10.6
R_{merge}	0.111	0.110	0.098	0.072	0.134
I/σI	15.9	13.2	14.5	32.4	17.7
Refinement					
Reflections	28493	17255	21664	28693	17390
Total no. of atoms-					
Protein	4821	2378	4731	4821	4919
Ligand	AMP- 34 NMN-35	AMP- 35	AMP- 34	AMP- 33 NMN-35	AMP-36 NMN-38
sulfate	19	91	9	14	23
water	104	106	22	102	66
R_{work}	0.2184	0.2434	0.272	0.2105	0.2410
R_{free}	0.2460	0.3034	0.308	0.245	0.2712
R.m.s.d Bond length (Å)	0.004	0.008	0.011	0.018	0.012
Bond Angle (°)	1.057	0.930	0.75	1.886	1.189
Ramachandran Favoured Allowed Outlier	96.55 3.45 0.00	93.17 5.97 0.31	92.16 7.84 0.00	96.24 3.76 0.00	93.73 3.45 2.82
Average B, all atoms ($Å^2$)	38.97	67.0	73.23	48.72	50.71

Table 3.5 Selected properties of cofactor AMP in its respective complexes with AdD domain of MtbLigA

Structure	PDB ID	Ligand	Conformation	Active site lysine adenylated	Contact with Glu184	Residues	AdD domain conformation
AdD domain	6KJM	AMP	*anti*	No	Yes (3.2Å)	E121,L122, K123, I124, E184, K300, K324	open
AdD domain active site mutant K123R	6KSC	AMP	*syn*	No	Yes (3.35Å)	E121, R123, E184, K300	open
AdD domain Ia subdomain mutant E22A	6KDU	AMP	*anti*	Yes	No	E121, L122, K300, K324	open
AdD domain Ib subdomain mutant H236Y	6KRH	AMP	*anti*	Yes	No	Y236, K300	open

3.3 Discussion

Structures reported here open the new avenues to structurally understand the step-wise adenylation reaction accompanied by domain flipping. The hydrolysis of phosphodiester bond between AMP and NMN part of NAD^+ takes place in presence of water molecule at the active site. This hydrolysis is succeeded by the adaptation of an *anti*-AMP or *syn*-AMP

conformation by forming covalent bond or H-bond with K123. The imperative essentiality of reaction is to reconcile structural adjustment in such a way that NMN should get released from the active site. The altered volume of NMN binding pocket interacting residues is interpreted by the rotation and broadening of α1 and α2 helices of Ia subdomain. Based on previous understanding about NAD^+-dependent DNA ligases, we translated present crystal structures and solution structures that decode the reaction intermediate of adenylation reaction during the step I of ligation. The presence of nicked DNA substrate at nearly physiological condition induces the topological adaptability in LigA to create a platform for ligation. The toroid forming property of DNA ligases are also conserved from ATP to NAD^+-dependent DNA ligases (Pascal, O'Brien et al. 2004, Nandakumar, Nair et al. 2007).

Despite the different domain organization, the ATP and NAD^+-dependent DNA ligase have nearly the same mechanism of reaction at step2 and step3 of the ligation reaction. They both have different mode of action for the adenylation of active site lysine. NAD^+-dependent DNA ligases have Ia subdomain which is probably evolved in such a way that it can rotate around the Ib subdomain to facilitate the step1 in a very distinct manner. This huge domain rotation is possible due to the presence of nearly 20 amino acid loop between Ia and Ib subdomain which is shortened in ATP-dependent DNA ligases(Pascal, O'Brien et al. 2004). We have also highlighted the role of important residues (E22, E26, E87, K123 and H236) during the first step of the reaction. The mutation of these amino acids showed reduced ligation activity that might be due to altered transfer of AMP to the active site lysine or the 5' PO_4 of the nicked DNA.

With the help of several crystal structures reported earlier (Odell and Shuman 1999, Singleton, Hakansson et al. 1999, Lee, Chang et al. 2000, Nandakumar, Nair et al. 2007, Srivastava, Dube et al. 2007) and crystal as well as solution structures in this study reveals the mechanistic conformational alterations associated with cofactor binding during adenylation. We captured the adenylated and de-adenylated AdD domain via crystallization and Small angle X-ray scattering. By blending the results from both techniques, the Ia subdomain rotation around Ib subdomain and cofactor induced active site remodeling can be visualized. The huge change in maximum dimension during the adenylation gives clear insight about the mechanism of domain rotation and flipping in order to prevent the reaction reversal by keeping the local NMN away from active site.

Chapter 4 | Mutational analysis of full length MtbLigA

4.1 Introduction

The MtbLigA is a highly modular structure with 691 amino acid consisting of N-terminal adenylation domain which is further subdivided in to Ia and Ib subdomain followed by oligomer binding domain (OB), Zn finger, helix-hairpin-helix (HhH) and the C-terminal BRCT domain (**Fig. 4.1**). As explained in the previous chapter the adenylation domain of LigA adapts different conformation during the course of reaction. The crystal structure of Ia subdomain mutant E22A (PDB: 6KDU) depicts the essentiality of this residue during the adenylation step to adapt closed confirmation during step 2. It is also previously studied that deleting the Ia subdomain leads to abolish the adenylation step without impacting the phosphodiester bond formation at the pre-adenylated nick (Timson and Wigley 1999, Kaczmarek, Zaniewski et al. 2001, Wang, Zhu et al. 2009). The residues E26 and E87 also play the major role during initial adenylation steps by making ionic interactions with Ib subdomain. Ib subdomain active site residue K123 when mutated to alanine completely loses its ability to ligate the nicked DNA substrate (Zhu and Shuman 2005). We also mutated it to arginine to check it's *in vitro* as well as *in vivo* consequences. Further the residue H236, involved in base stacking interaction with adenine moiety of AMP, was mutated to tyrosine and checked its effect. It is also reported that on deleting the C-terminal domain there occurs marked reduction in the ligation activity (Timson and Wigley 1999, Zhu and Shuman 2005, Poidevin and MacNeill 2006, Wang, Nair et al. 2008). Since it cannot perform the second and third step of the reaction. More effort is required to absolutely decode the role of this domain in detail. The deletion of this domain resulted in the threefold reduction in *E. coli* LigA activity (Wilkinson, Smith et al. 2005) although the *T. filiformis* LigA possess activity even after its deletion (Jeon, Shin et al. 2004). In contrary to the above case no activity, either *in vitro* or *in vivo*, is noticed in the case of the *M. tuberculosis* enzyme. The same situation exist in the enzymes from the viral source e.g the active enzyme from *A. moorie* lacks this domain (Sriskanda, Moyer et al. 2001) and *M. Sanguinipes* (Lu, Tong et al. 2004), while the LigA from *mimivirus* possesses a BRCT domain, essential to its activity (Benarroch and Shuman 2006). We therefore here report the site directed mutagenesis performed on the highly conserved residues of BRCT domain (G614, G639, G621 and KK648/649) (**Fig. 4.2**) which further emphasizes on its functional role in different steps of nick sealing. The molecular docking of short oligo with the BRCT domain also demonstrates the residues involved in the interaction with the oligo and also facilitates the efficient ligation (**Fig. 4.3 A, B**). To further highlight the role of residues of AdD domain and BRCT domain we also complemented them

in *E. coli* GR501 strain lacking the functional LigA. We checked the effect of these mutations on the survivability at higher non-permissive temperature.

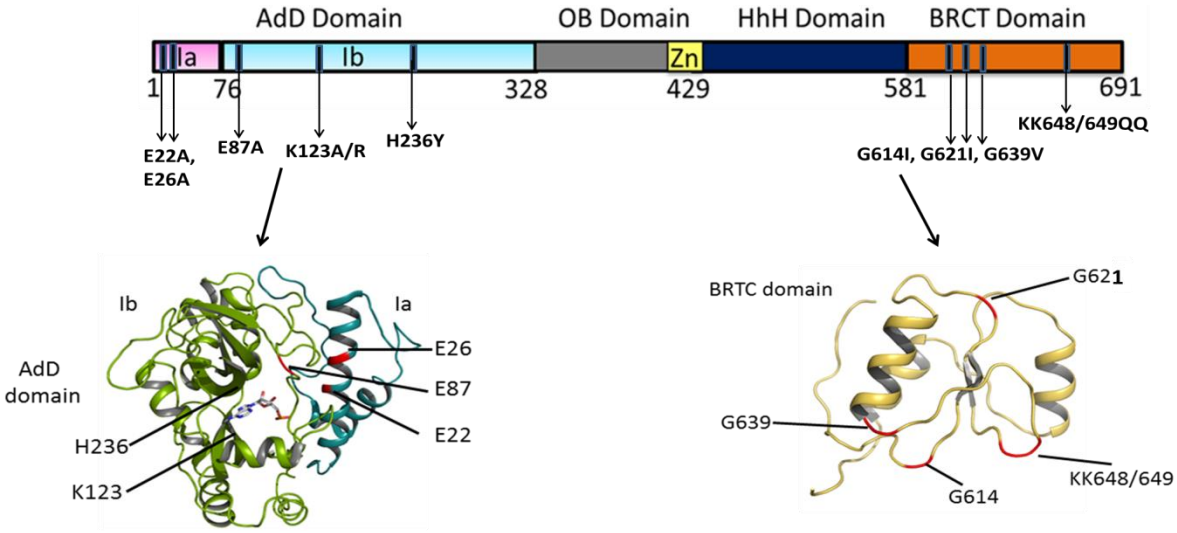


Figure 4.1 The diagrammatic representation of MtbLigA marked with mutated residues. The chosen residues are highlighted in red colour in cartoon model of AdD and BRCT domain.

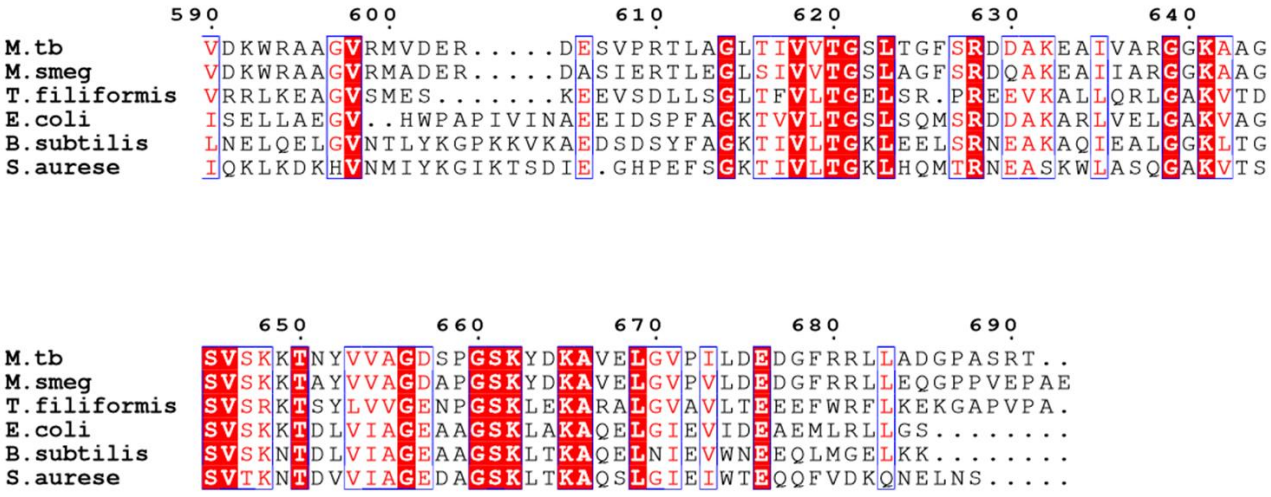


Figure 4.2 Sequence alignment of BRCT domain of respective NAD$^+$-dependent DNA ligases. The protein sequence of NAD$^+$-dependent DNA ligases from several eubacterial species were aligned using Clustal Omega. The alignment studies predicted the highly conserved residues (highlighted in red).

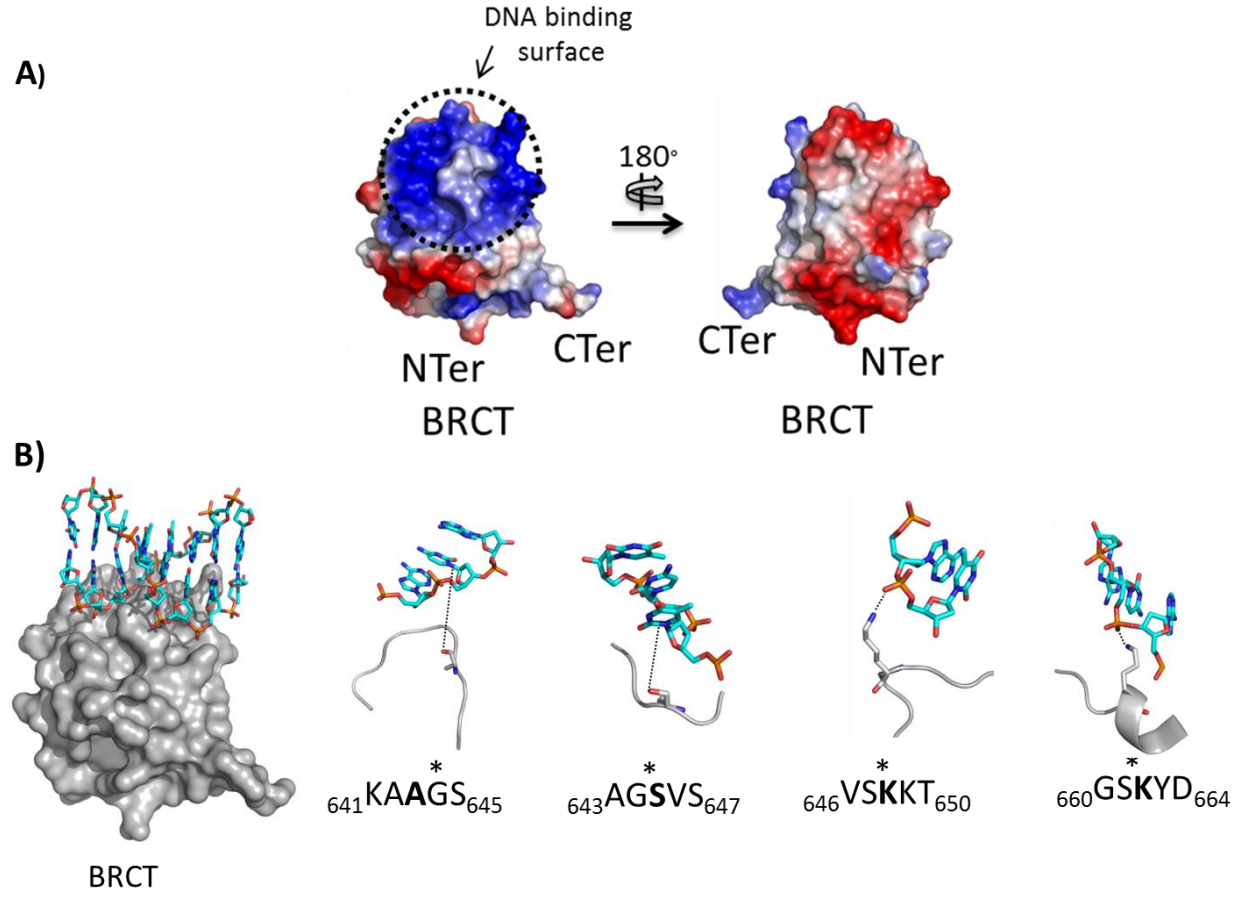


Figure 4.3 Identification of DNA binding region in BRCT. A) The positively charged surface (Blue) on the BRCT domain shows the residues which binds with DNA. **B)** The molecular docking of short oligo with BRCT domain depicts the residues playing crucial role by binding with DNA during ligation reaction.

4.2 Results

4.2.1 Cloning and purification of AdD domain (E22A, E26A ,E87A,K123A/R and H236Y) and BRCT domain mutants (G614I, G621I, G639V and KK648/649QQ) of MtbLigA

In order to characterize the crucial residues from AdD/BRCT domain of MtbLigA in ligation reaction, the mutants were generated by PCR-mediated site directed mutagenesis using pET21d-LigA as template (**Fig. 4.4 A, B**). The clones were verified by sequencing (Chromus Biotech and IDT) and double restriction digestion (**Fig. 4.4 C, D**). For *in vivo* complementation experiments, the mutants were generated by using the pTrc99a-LigA as the template. The primers used in site-directed mutagenesis are mentioned in the table. The expression of mutants is induced in *E. coli* BL21(DE3) cells and after induction with 1mM

IPTG the culture was grown at 25˚C at 120 rpm for 18 hours. The affinity purification using Ni-NTA beads was carried out using the lysate of cells overexpressing these mutants. The protein fractions were collected at different gradient of imidazole and finally eluted at 300mM. All the fractions were evaluated by running on 12% SDS PAGE gel (**Fig. 4.5 A, B**). The purified fractions were pooled and dialyzed before subjecting to size exclusion chromatography.

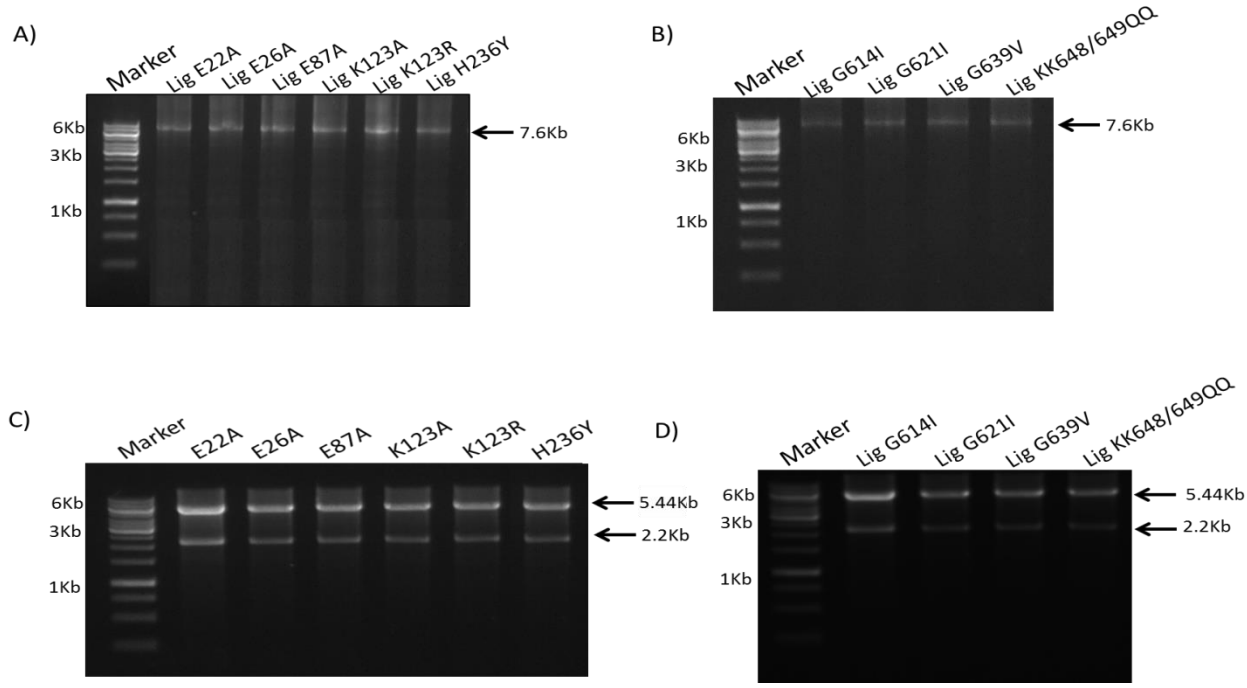


Figure 4.4 Cloning of AdD and BRCT domain mutants of MtbLigA. The PCR mediated site directed mutagenesis to generate **A)** AdD domain mutants of LigA. The arrow indicated the amplified product (7.6Kb). **B)** and BRCT domain mutants of LigA. **C and D)** The restriction digestion of the AdD domain mutant and BRCT domain mutant clone. The arrow indicates the size of vector pET21d (5.44 Kb) and Insert mutant (2.2 Kb).

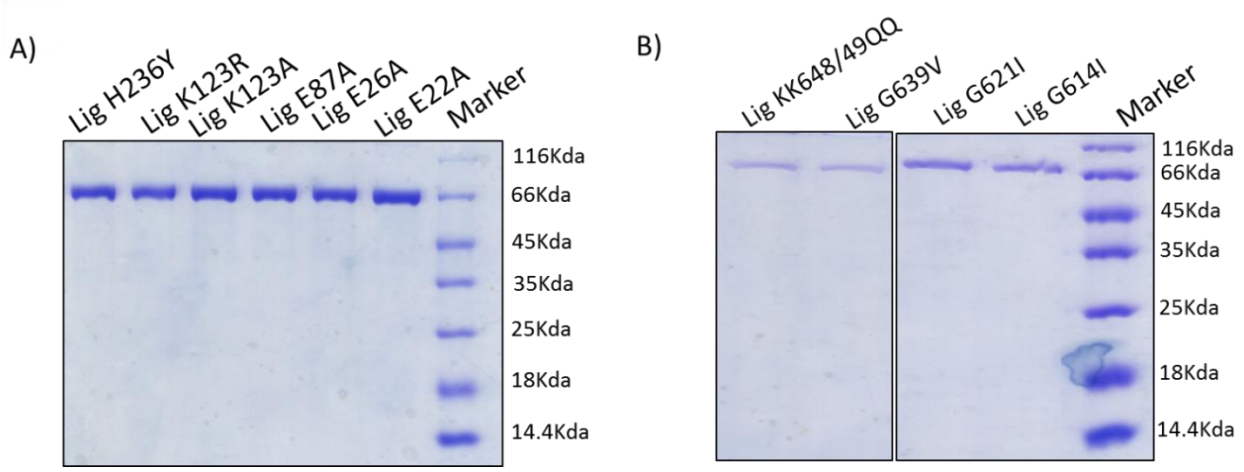


Figure 4.5 The Ni-NTA affinity Purification of A) AdD domain and **B)** BRCT domain mutants. The samples were run on 12% SDS-PAGE.

4.2.2 Size exclusion chromatography and Circular Dichroism of proteins

The purified proteins were further subjected to gel filtration through a Superdex-200 10/30 increase column (GE Healthcare) equilibrated with buffer C (50mM HEPES (pH 8.0), 250 mM NaCl and 2mM βME) (**Fig. 4.6 A**). Eluted fractions were analyzed for purity by SDS-PAGE and fractions corresponding to pure protein were pooled and concentrated to 2-15mg/ml, aliquoted, flash frozen and stored at -80^{0}C. For performing circular dichroism, all recombinant SEC purified proteins were dialyzed overnight at 4^{0}C against buffer (10 mM Tris-Cl (pH 8.0), 100 mM NaF) and centrifuged at 12,000 rpm for 20 minutes at 4^{0}C to remove any precipitants. All the proteins were diluted to final concentration of 0.18 mg/ml. The spectra were recorded at far UV range (180-260 nm) at 25^{0}C using 1 mm path length cuvette on (Jasco J-1500) at following experimental settings: sensitivity100 mdeg, scan speed 100 nm/min, band width 1 nm. Each CD spectrum is average of two accumulations and was baseline corrected by subtracting spectrum of corresponding buffer (**Fig. 4.6 B**). The experimental data was analyzed using K2D3 software.

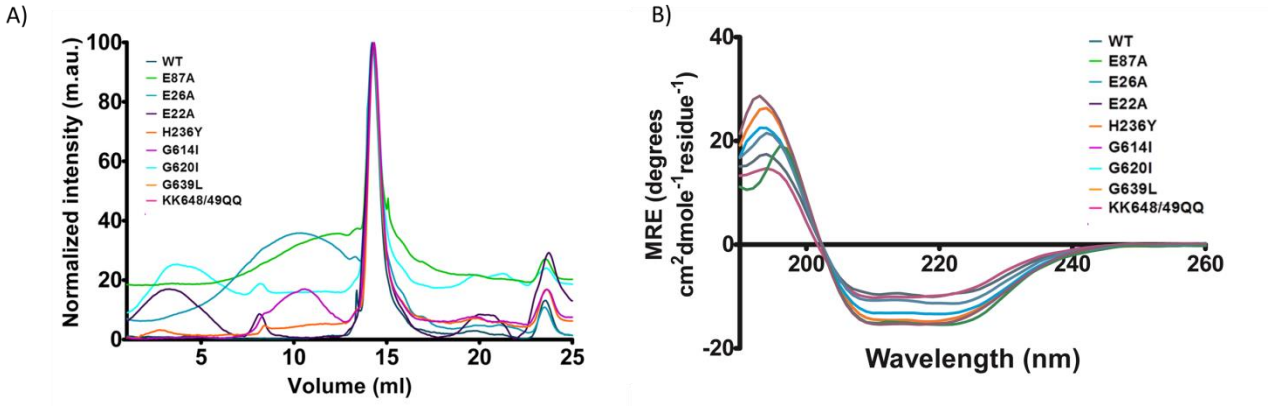


Figure 4.6 Size exclusion chromatographic and CD spectroscopic analysis of MtbLigA mutants of AdD and BRCT domain. A) The size exclusion chromatogram of AdD domain mutants (E22A, E26A, E87A and H236Y) and BRCT domain mutants (G614I, G621I, G639V and KK648/649 QQ). **B)** CD spectra of wilt type MtbLigA (WT) and mutants shows the no change in secondary structure occurs due to mutation.

4.2.3 Effect of mutation on binding affinity with cofactor and DNA

The Ia subdomains residues (E22, E26, E87 and H236Y) play major role during the course of adenylation reaction by making ionic interactions with Ib subdomain and facilitates the stabilization of required conformational adaptability, so we targeted these residues for alanine scanning to check their role in overall ligation activity. Similarly the highly conserved residues in BRCT domain mainly glycine residues G614, G621 were mutated to isoleucine and G639 to valine. The binding affinity of WT and mutants for NAD^+ was determined by intrinsic florescence and found that WT has 2.7 µM Kd value while E22A, E26A, E87A, H236Y, G614I, G621I and G639V showed 387 µM, 83.8 µM, 28.27 µM, 152 µM, 51 µM, 152 µM, ND, respectively (**Fig. 4.7 A**). The Kd value of WT LigA for nicked DNA appeared in 2.52 nM while G639 showed the lowest binding ability with 70.52 nM Kd value (**Fig. 4.7 B**). The electrophoretic mobility shift assay deciphers the decreased binding capacity for these mutants with reduced intensity of shifted bands (**Fig. 4.7 C**).

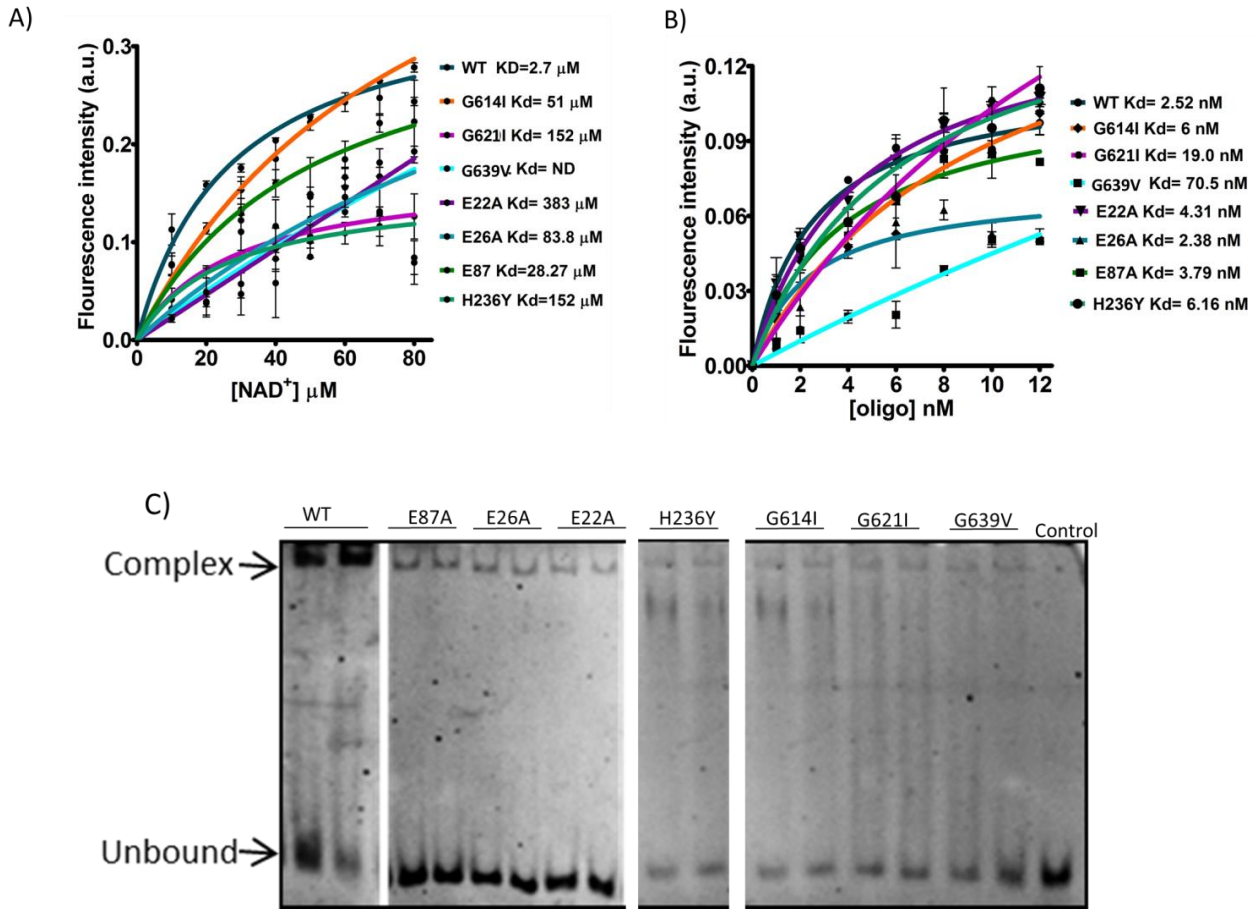


Figure 4.7 Determination of dissociation constant (Kd) of WT MtbLigA and mutants with A) cofactor NAD^+ **and B)** nicked DNA oligo. WT MtbLigA is denoted in dark cyan, E22A in dark purple, E26A in slate blue , E87A in green , H236Y in sea green while BRCT domain mutants are color coded as G614I in orange, G621I in light purple and G639V in cyan. **C)** Gel shift assay of WT and all mutants in the presence 5'FAM labeled 52mer nicked ds DNA substrate. The arrows indicate the LigA-DNA and mutant-DNA complex and the free unbound probe, respectively.

4.2.4 Effect of mutation on *in vitro* ligation activity

During adenylation reaction Ia subdomain rotates around Ib subdomain in the sequential way and mutating Ia subdomain residues with alanine might disrupts the interactions formed by these residues with the Ib subdomain therefore restricting the close proximity of Ia with Ib which is mandatory for proper adenylation. The extent of ligation by wild type LigA and each mutant was evaluated as a function of LigA and specific activities were normalized to the wild type values (defined as 100%). There occurs marked reduction in ligation activity of BRCT domain mutants. G639V and G621I showed the maximum reduction in activity of around 98% and 87% respectively while G614I shows 80% reduced activity with respect to wild type. Whereas 89% and 64% reduction of ligation activity observed in E22A, H236Y

and E26A mutants, respectively. E87A mutant is 61% as active as wild type (**Fig. 4.8 A, B and C**).

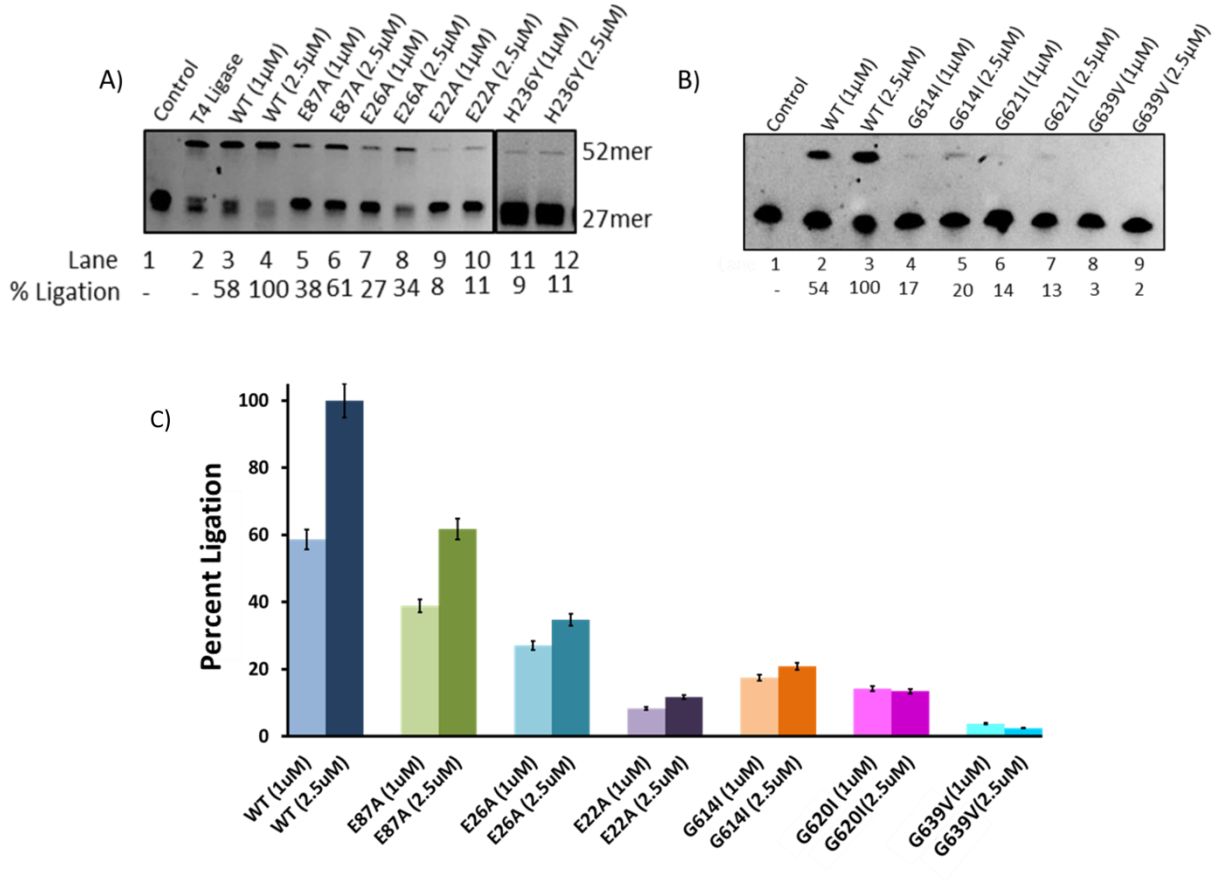


Figure 4.8 The extent of ligation measured for A) AdD domain mutants and **B)** BRCT domain mutants in two concentrations (1 µM and 2.5 µM) and compared with WT (considered 100%). Lane 1 is control without any enzyme and lane 2 is positive control with T4 DNA ligase. **C)** The Bar graph depicts the percent ligation of all mutants in comparison with WT. The G639V mutant shows the least percent ligation among all mutants.

4.2.5 Effect of mutations on *in vivo* survival of bacteria in LigA deficient *E. coli* GR501 strain

The consequences of these mutants on the temperature sensitive strain of ligA deficient *E. coli.* GR501 strain is evaluated thereafter. The wild type MtbLigA is able to rescue the ligA deficient *E. coli*. GR501 cells at non-permissive temperature stress of 43°C. The complementation with the BRCT deleted mutant (MtbLigA1) is not able to rescue the cells (Srivastava, Dube et al. 2007). We further came up with the mutation of more specific amino acids crucial for ligation reaction and their complementation assay. Alanine scanning of Ia

subdomain residues (E22, E26, and E87) shows delayed growth at 43°C as compare to wild type. There is no growth in the active site mutant K123A and the substitution of H236 with Y is also not able to rescue the cells in the temperature stress condition despite of the not much change in the active site as shown in the crystal structure. The mutation of actively involved amino acids like G614I, G621I, G439L and KK648/649QQ in binding with DNA, are not able to efficiently ligate the nicked oligo as shown in *in vitro* assays. The complementation with these mutants is also not able to rescue the temperature sensitive *E. coli* strain as compared to wild type. The G614I mutation is most lethal and KK648/649QQ mutant after that. The G621I and G639V mutant are also not able to rescue as compare to wild type LigA (**Fig. 4.9**).

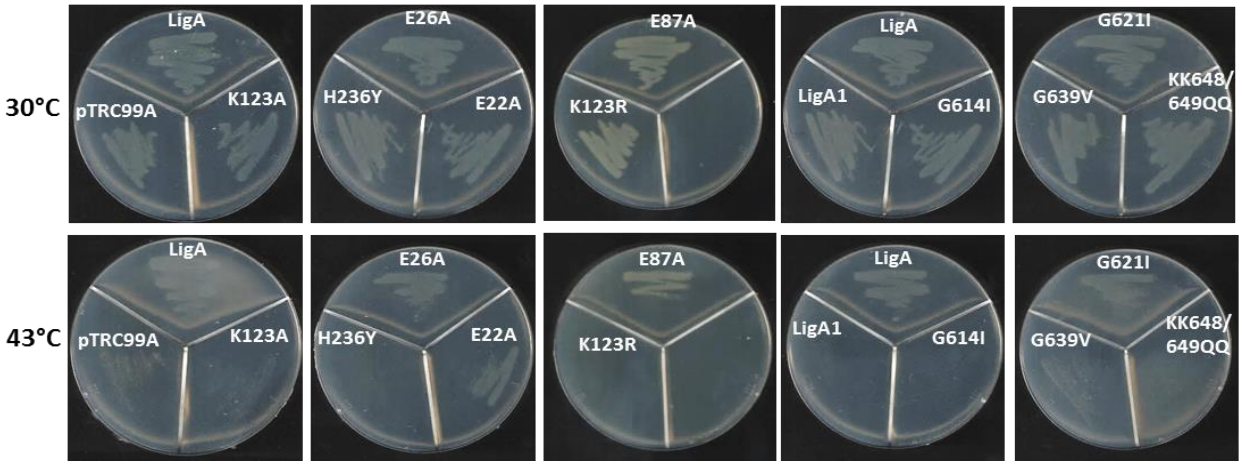
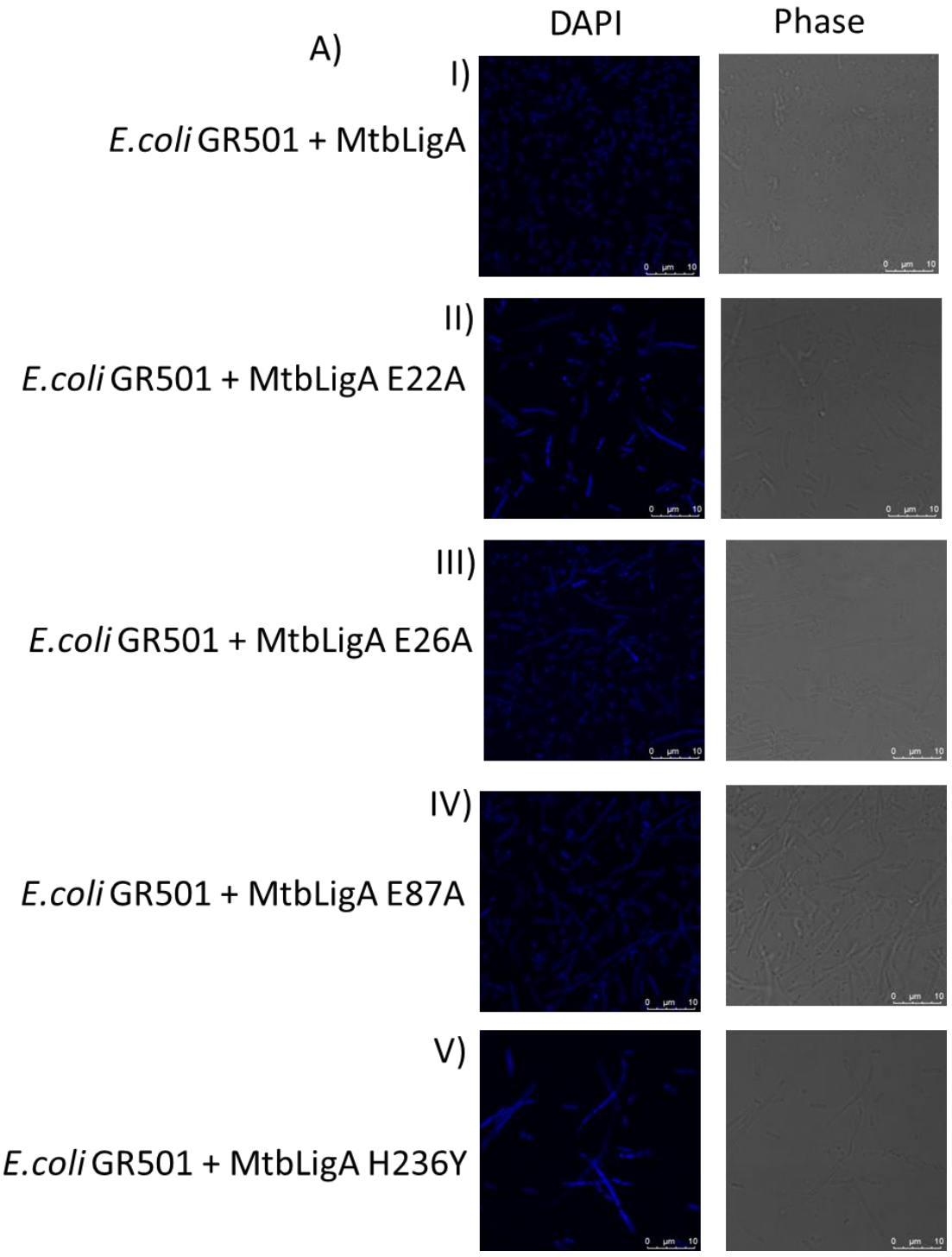


Figure 4.9 The *in vivo* complementation in temperature sensitive strain of *E. coli.* GR501 which lacks functional DNA LigA. First row shows the viability of strain at permissive temperature 30°C, wild type and mutants which are cloned into pTRc99 are viable at the same temperature. In second row, at higher non-permissive 43°C, the complementation with the wild type is rescuing the cells with the temperature stress, while the E22A, E26A, E87A, G621I and G639V shows delayed growth at 43^0C while active site mutants K123A/R, H236Y, LigA1, G614I and double mutant KK648/49QQ are lethal.

4.2.6 Microscopic examination of *E. coli* GR501 strain rescued with LigA mutants

We also visualized the effect of E22A, E26A, E87A, H236Y from Ia subdomain of AdD domain and G614I, G621I and G639V from BRCT domain in full length LigA on the morphology of *E. coli* GR501 strain rescued with NAD^+-dependent DNA ligase mutants (**Fig. 4.10 A, B**). The microscopic examination of the mutant cells showed a remarkable change in the morphology of the microbe when compared to the wild type (**Fig. 4.10 A, I**). The staining of DNA with 4',6-diamidino-2-phenylindole (DAPI), revealed a notable phenotype of defective cell division viz. filament formation in the mutant cells. This result is

in line with the inhibition of cell division due to chromosome degradation and impaired Okazaki fragment repair, while allowing cell growth(Brotz-Oesterhelt, Beyer et al. 2005). The phenotype observed in the mutant cells are similar to that observed in the presence of Novobiocin, a known DNA gyrase inhibitor which results in rapid inhibition of DNA synthesis and eventually impair cell division causing formation of filamentous cells (Johnson, Yao et al. 2005).



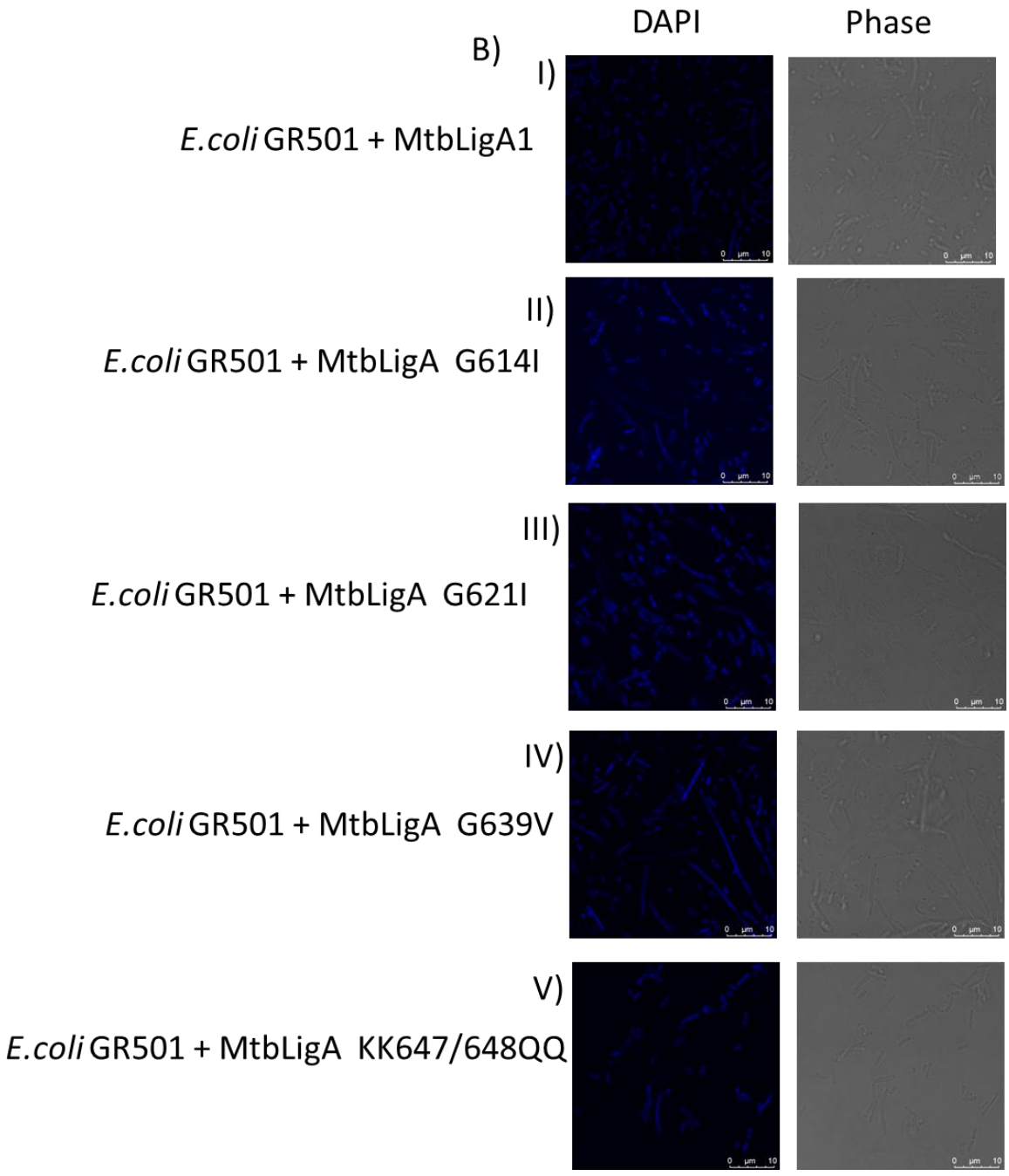


Figure 4.10 The mutation at critical residues of AdD domain and BRCT domains induce the filament formation in bacterial cell expressing NAD^+-dependent ligase. Microscopic examination of *E. coli* GR501 strain rescued with **A)** WT LigA and AdD domain mutants **B)** with BRCT deleted LigA1 mutant and BRCT domain mutants. The right panel shows the light field and left panel shows fluorescence microscopy after staining with DNA specific flourophore DAPI.

4.3 Discussion

Despite the different domain organization, the ATP and NAD^+-dependent DNA ligase have nearly the same mechanism of reaction at step II and step III of the ligation reaction. They both have different mode of action for the adenylation of active site lysine. NAD^+-dependent DNA ligases have Ia subdomain which is probably evolved in such a way that it can rotate around the Ib subdomain to facilitate the step I in a very different manner. This huge domain rotation is possible due to the presence of nearly 20 amino acid loop between Ia and Ib subdomain which is shortened in ATP-dependent DNA ligases (Pascal, O'Brien et al. 2004). We have highlighted the role of E22, E26, E87 and H236 during the first step of the reaction. The alanine scanning of E22, E26 and E87 shows reduced ligation activity which might be due to altered transfer of AMP to the active site lysine while mutation of H236 to tyrosine also shows the marked reduced ligation activity which may be due to the improper stacking and orientation of AMP in the active site. The BRCT domain is not reported in the crystal structure of *E. co*. DNA LigaseA with nicked oligo (Nandakumar, Nair et al. 2007). Surprisingly the NMR shifts of the BRCT domain with DNA (oligo) shows that BRCT domain interacts with oligo by mainly with lysine residues present only on one side of the surface. So by observing these outcomes of our analysis we concluded that BRCT may have a transient interaction with oligo during the last step of the ligation. We have further zoomed in the role of BRCT domain in the ligation reaction by mutating the highly conserved amino acids and found that G614I, G621I, G639V and KK648/49QQ mutant have severe effect on ligation. These amino acids are present nearly the K641, K648, K632 and S645. So by introducing the long side chain amino acids near these interacting residues may interfere with the binding of BRCT domain to oligo due to steric clashes.

Chapter 5 | Structural Characterization of MtbXthA and MtbLigA Complex

5.1 Introduction

DNA is exposed to spontaneous hydrolytic degradation and attacked by various metabolic byproducts, such as reactive oxygen/nitrogen radicals and alkylating agents, as well as numerous environmental agents (Lindahl 1993, Chatterjee and Walker 2017). This genome instability is counteracted by the intensive DNA repair system. Since it is well known that *Mycobacterium tuberculosis* can thrive in the host for years and this capability is imparted by its commendable DNA repair system that mends the DNA damages caused by the uncongenial oxidative environment inside the host cell (Ehrt and Schnappinger 2009, Stallings and Glickman 2010). There are several repair pathway in *M. tb* but due to the absence of mismatch repair pathway, the base excision repair (BER) pathway plays crucial role during the oxidative DNA damaging responses and ensures the correction of most abundant damages (Mirazhi and Cole et al). The BER pathway is conserved throughout the kingdoms of life, with some enzymes being found in all major taxa, while others restricted to one or several branches of the life tree (Aravind, Walker et al. 1999, Eisen and Hanawalt 1999). The pathway commences with the identification of damaged base by DNA glycosylases (Dizdaroglu, Coskun et al. 2017, Mullins, Rodriguez et al. 2019) which results in formation of intact or cleaved AP site. Accumulation of AP site may prove to be deleterious to the bacterial survival since unprocessed AP sites are highly mutagenic and blocks the replication and transcription machinery (Boiteux and Guillet 2004). Then according to the 'passing the baton' model, the AP sites act as substrate for the AP endonuclease which process AP site by incising the phosphodiester backbone of DNA, the key step of the pathway (Whitaker and Freudenthal 2018). The pathway is then completed by the action of downstream enzymes DNA polymerase that fills the bases and finally DNA ligase seals the nick (Krokan and Bjoras 2013). On the contrary, of being tacked together into one stable multisubunit complex, these enzymes pass the repair intermediates between them in a highly coordinated manner. Thus the BER pathway orchestrate in a coordinated fashion that is mediated by specific protein-protein interactions (Wallace 2014, Endutkin, Yudkina et al. 2019) .

AP-endonuclease, is the central player of the pathway, incises the immediate 5' end of the AP site resulting in generation of 3'-OH and a 5'-dRP motif. The 5'-dRP is then to be removed by the AP lyase enzyme. *M. tb* AP endonuclease II (MtbXthA) is a multifunctional enzyme also comprising of 3'-5' exonuclease, 3'-phosphodieesterase, 3'-phosphatase and

RNase H activity (Mol, Hosfield et al. 2000). The importance of enzyme was also revealed by the report that showed the mutation in mycobacterial AP endonuclease, XthA, resulted in compromised survival of mycobacterium with in the host (Puri, Reddy et al. 2014). The studies in eukaryotes have focused on the functional interaction of AP endonuclease with other BER components to regulate the pathway (Thakur, S., Dhiman, M., et al. 2015). Human APE I is known to interact with bi-functional glycosylase FEN-1, DNA polymerase β, DNA Lig I to modulate their activities for more effective repair. Further, AP-endonuclease interacts with proliferating cell nuclear antigen (PCNA), suggesting that it has a coordinating role in BER (Sukhanova, Khodyreva et al. 2005). Finally BER ceases with the action of DNA ligases, which catalyze the formation of phosphodiester bonds between 5'-phosphoryl and 3'-OH termini within a DNA strand (Shuman 2009). In *Mycobacterium tuberculosis*, NAD^+-dependent DNA ligase (LigA) is essential enzyme and has a modular architecture. Previous studies had shed the light on the role of BRCT domain which makes direct interactions with the DNA substrate thus helps the enzyme to form toroid shape around the nick for ligation (Lee, Chang et al. 2000). In eukaryotes, the BRCT domain, via specific protein-protein interactions, acts as a DNA damage signal transducer element that conveys signals from the damage sensors to other components of DNA repair machinery (Sheng, Zhao et al. 2011).

The various studies in eukaryotes and prokaryotes suggest direct interactions between the enzymes catalyzing different, generally consecutive, steps of the BER (Moor and Lavrik 2018). The channeling of substrate takes place between the enzymes, which could be essential for prevention of DNA substrate to the detrimental action of nucleases. The mycobacterial BER components are characterized well but the protein-protein interaction network is not yet fully explored. In the present study, we focused on those unexplored interactions that regulate the pathway. Earlier our group identified that there occurs no interaction between Mtbβ clamp and and MtbLigA contrary to what was identified in humans and *E. coli* (Vidal, Boiteux et al. 2001). Although the possibility of transient interaction between the two cannot be denied. Later our group further identified a new complex between the MtbXthA and Mtb βclamp in a DNA regulated manner (Khanam, Rai et al. 2015). A recent report also demonstrated that mycobacterial Ligase C (LigC) complex interacts with proteins of the BER apparatus (Plocinski, Brissett et al. 2017). Thus we hypothesized that there must be a unidentified factor (possibily MtbXthA) which could act as bridge between the two (Mtbβ clamp and MtbLigA). These findings indicate that a sophisticated regulation of the BER pathway in *M. tuberculosis* takes place by formation of multimeric protein complexes *viz.* 'Berosomes'.

In this study we have identified the direct physical interaction between the MtbXthA and MtbLigA by using size exclusion chromatography (SEC) and surface plasmon resonance (SPR) using the purified recombinant proteins. Through SAXS analysis, we also structurally characterized MtbXthA-LigA and MtbXthA-BRCT domain complexes in solution. The complex formation between the two proteins is mediated specifically by a newly identified motif in MtbXthA that interacts with the BRCT domain in MtbLigA. We also identified the residues in BRCT domain playing crucial role in interaction with XthA. The interaction of XthA with LigA inhibits the latter's activity. Our group has also shown that LigA can efficiently act on the product of XthA activity to ligate it. Our present structural and previous biochemical results suggest that XthA engages with the C-terminal BRCT domain of LigA to prevent 'futile' ligation of the former's product until an AP-Lyase is recruited to remove the 5'-dRP moiety. These findings offer the mechanistic insights into the role of MtbXthA in preventing MtbLigA from resealing the DNA termini with 3'OH and 5'dRP moieties created by the action of MtbXthA at abasic sites.

5.2 Results

5.2.1 Size exclusion chromatography (SEC) and UV-Vis Spectral analysis

To probe the direct physical interaction between the MtbXthA and MtbLigA we performed size exclusion chromatography (SEC) of the individual proteins and their complexes. The standard protein markers were used to calculate the molecular mass and Stoke's radii (**Fig. 5.1 A and B, Table 5.1**). Fractions corresponding to the peaks were collected and analysed by SDS-PAGE (**Fig. 5.1 D**). The SDS-PAGE shows that proteins were homogenously eluted without any contamination as observed on the gel with single band while in case of complexes the bands of corresponding protein partners in complex formation was observed. MtbXthA, elutes with an apparent molecular weight 39.8 KDa (**Fig. 5.1 D, green peak**), while MtbLigA elutes at approximately 121.04 KDa (**Fig. 5.1 D, orange peak**). The complex between the two was formed by mixing them together in stoichiometric ratio of 1:1, MtbLigA and MtbXthA were co-eluted in the molecular mass range ~106.5 KDa (**Fig. 5.1 D, purple peak**). The UV-Vis spectra of the eluted peaks showed the absence of any contaminating DNA (**Fig. 5.2 A, B and C**). To further ascertain the role of BRCT in complex formation we incubated BRCT deleted MtbLigA1 mutant and MtbXthA for complex formation but both failed to interact and both proteins were eluted separately (**Fig. 5.1 D, blue peaks**) with no evident shifting in their individual elution volumes. Interestingly, purified alone BRCT

domain (**Fig. 5.1 C**) co-eluted with MtbXthA (**Fig. 5.1 D, red peak**). These results indicate that, a DNA independent binary complex form between MtbXthA and MtbLigA, and that this interaction is mediated by the BRCT domain of the latter.

In previous chapter, the detail discussion about the structural conformation of apoMtbLigA which remains in extended conformation and changes to closed compact toroid shaped conformation around the nick in presence of DNA to complete the ligation reaction. Therefore, it will be interesting to see how interaction with MtbXthA will effect this conformational change in MtbLigA. When the MtbXthA, MtbLigA and 27-mer nicked DNA substrate (S1) were mixed in equimolar ratio of 1:1:1, we observed that the eluted peak shows the presence of DNA in the complex in UV-Vis spectral analysis (**Fig. 5.2 B, red dotted**). The result indicates that MtbLigA probably to adopt a compact conformation in the presence of DNA with stokes radii of 3.19 nm as compared to the apoLigA with stokes radii of 4.02 nm. However, in the presence of MtbXthA, the Stoke's radius of the tri-complex increases to 3.79 nm indicating that addition of MtbXthA resulted in the formation of a bigger oligomeric complex. (**Fig. 5.1 D, red dotted peak**). The SDS-PAGE analysis of the eluted peak shows the presence of both MtbXthA and MtbLigA indicating the complex formation between the two proteins (**Fig. 5.1 D**). The UV-Vis spectra of the eluted peak also confirmed the presence of DNA in the fraction volume (**Fig. 5.2 A and B, pink and magenta**). The SEC results overall show that MtbXthA-MtbLigA-DNA forms a ternary complex.

The incision by MtbXthA at an abasic site creates a nicked DNA intermediate on which downstream proteins including MtbLigA act to repair the DNA damage. We therefore wanted to study the interaction between the two proteins in the presence of nicked DNA. We first performed SEC analysis of MtbLigA incubated with a nicked DNA substrate analog (S1). The apparent molecular weight of the DNA-LigA complex was ~112.27 KDa with a reduced Stoke's radii of 3.19 nm compared to MtbLigA alone (4.02 nm) (**Fig. 5.1 D, yellow dotted peak and Table 5.1**). The UV-Vis spectral analysis of the eluted peak shows the presence of DNA in the complex (**Fig. 5.2 A-C, slate blue**). The result suggests that MtbLigA likely adopts a compact conformation in the presence of DNA. The UV-Vis spectra of the eluted peak also confirmed the presence of DNA in the fraction volume (**Fig. 5.2 C, magenta**). The SEC results overall show that MtbXthA-MtbLigA-DNA forms a ternary complex.

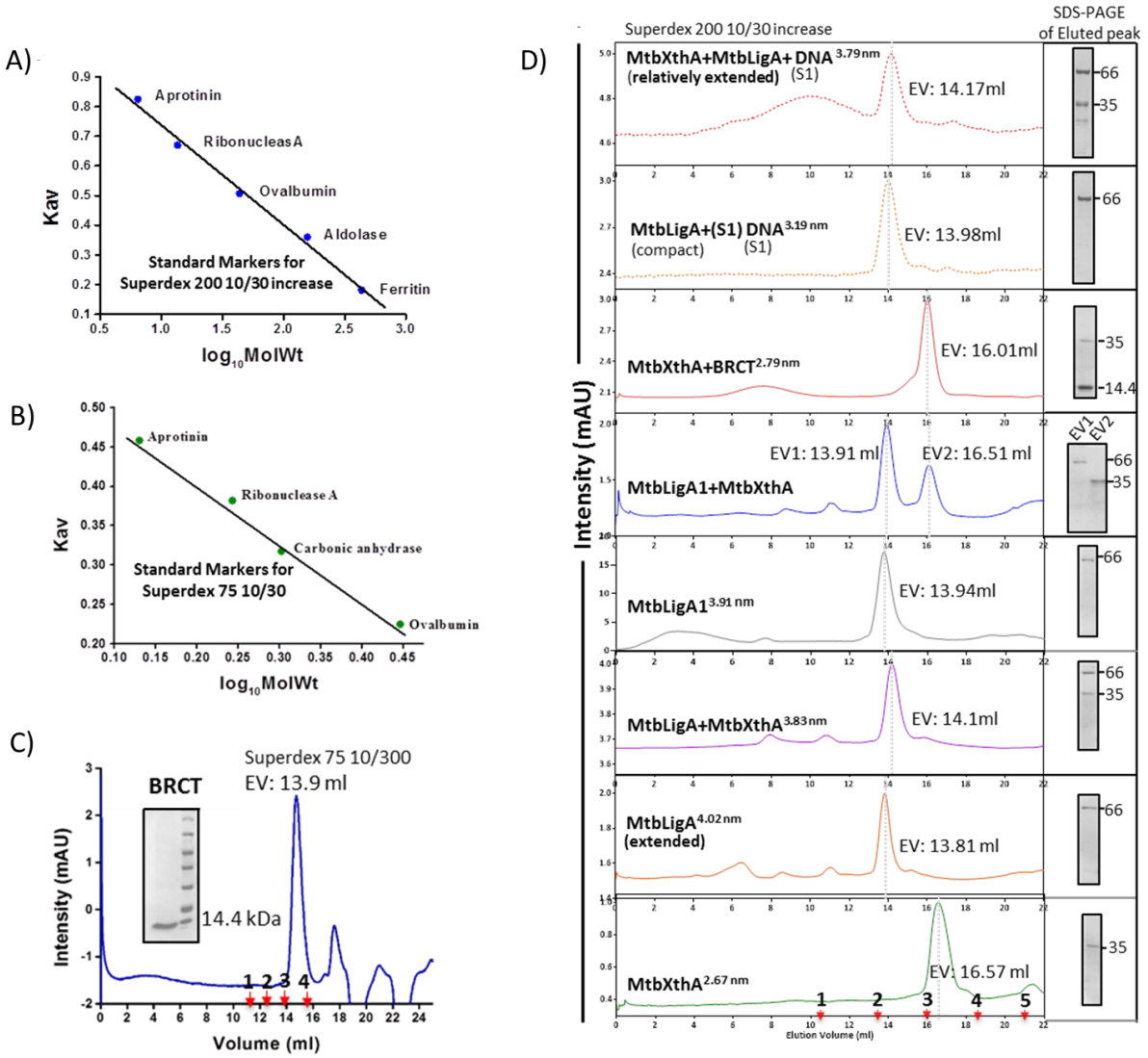


Figure 5.1 Size exclusion chromatographic analysis. A) Standard curve for calibration of S200 10/30 increase size exclusion column. **B)** Standard curve for calibration of Superdex 75 10/30 size exclusion column. The columns were calibrated using blue dextran (to determine void volume) and standards Aprotinin (6.51 kDa, 1.3 nm), Ribonuclease (13.7 kDa, 2.4 nm), Carbonic anhydrase (29 kDa, 2.01 nm), Ovalbumin (44 kDa, 4.4 nm), Aldolase (158 kDa, 6.62 nm), and Ferritin (440 kDa, 7.99 nm) of known molecular weights and stokes radii. **C)** SEC profile of BRCT domain Superdex 75 10/30 column. The fraction corresponding to the eluted peak was analysed by 10% SDS-PAGE (shown as inset). **D)** SEC profile of proteins/complexes as indicated. The elution volume (EV) for proteins/complexes is shown on the side of respective eluted peaks. The fraction corresponding to the eluted peaks of proteins/complexes was analysed by 12% SDS-PAGE (shown as inset on right). The molecular weight marker (kDa) is indicated on each gel image. The stoke's radii calculated is depicted as superscript to indicated labelling in the graphs. The SEC profile for protein in complex with nicked DNA complex (S1) is show in dotted lines. The elution of standard proteins is marked as red arrows at the bottom of the chromatogram as 1: Ferritin (440 kDa), 2: Aldolase (158 kDa), 3: Ovalbumin (44 kDa), 4: Ribonuclease A (13.7 kDa) and 5: Aprotinin (6.5 kDa) for superdex 200 10/30 increase column.

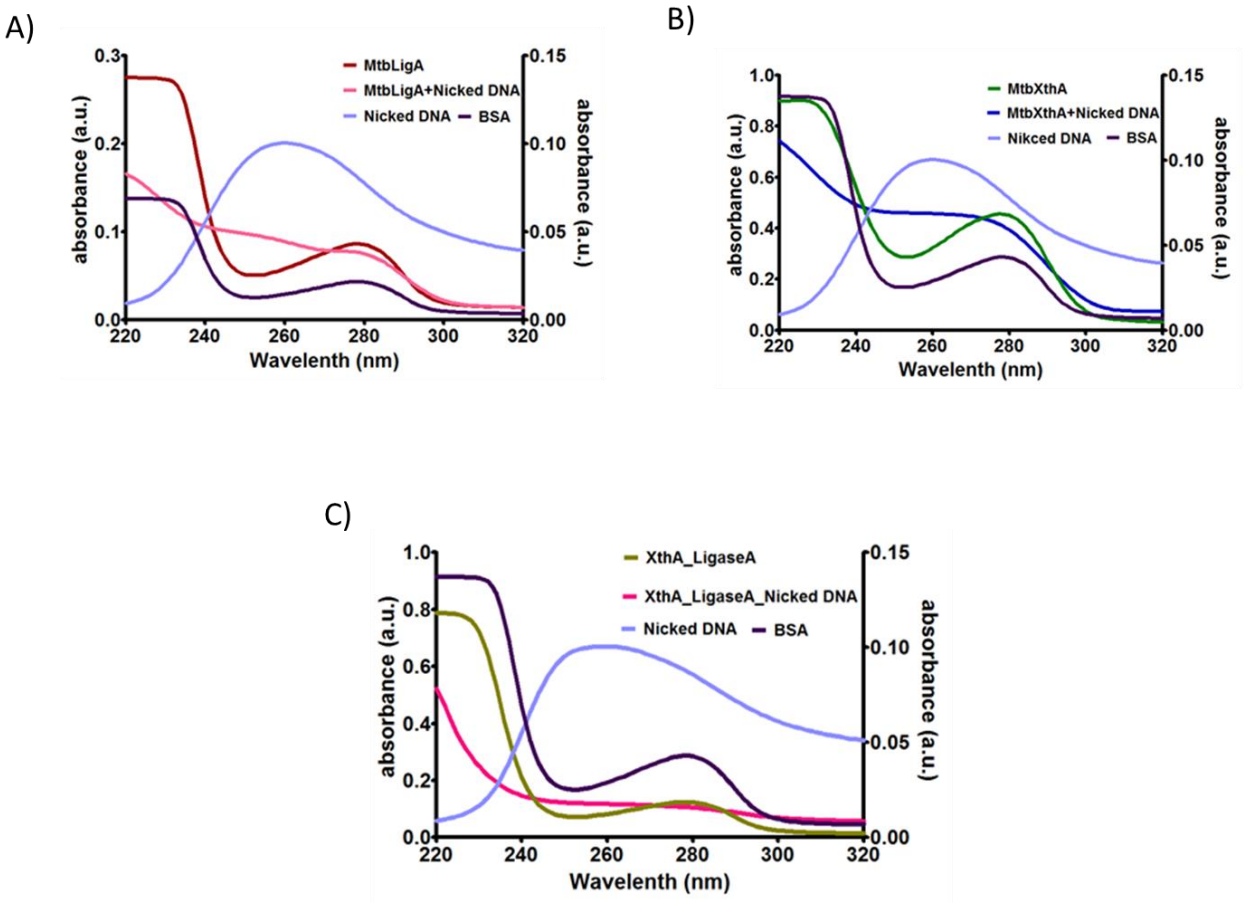


Figure 5.2 UV-Visible spectral analysis of eluted peak fractions of proteins/complexes in SEC. A) UV/Vis spectra of apoMtbLigA (brown) and MtbLigA in complex with nicked DNA (pink) indicates the presence of DNA in the eluted fraction. **B)** The UV/Vis spectra of MtbXthA in complex with nicked DNA (navy blue) indicates that XthA donot bind with nicked DNA while alone XthA (green) shows peak maxima at 280 nm. **C)** The complex between MtbXthA-LigA (olive) shows the peak maxima at 280 nm while MtbXthA-LigA complex in presence of nicked DNA also indicate the peak at 260 nm thus indicating the formation of ternary complex (magenta). The BSA (dark purple) shows peak maxima at 280 nm and nicked ds-DNA showing peak maxima at 260nm (slate blue) are used as controls.

Table 5.1 Size exclusion elution values for protein and complexes

Protein/Protein complex	Theoretical Molecular weight (KDa)	Average Molecular weight (KDa) by SEC	Stoke's Radii (nm)	260/280 ratio
MtbXthA	35	39.81	2.67	0.51
MtbLigA	77	121.04	4.02	0.57
BRCT domain	11	22.30	2.14	0.53
MtbLigA1	66	115.80	3.91	0.56
MtbXthA+BRCT domain	46	45.68	2.79	0.54
MtbXthA+MtbLigA	112	106.45	3.83	0.59
MtbLigA+Nicked DNA	101	112.27	3.19	0.82
MtbXthA+MtbLigA+Nicked DNA	136	103.20	3.79	1.03

*The 260/280 ratios were measured by nanodrop for eluted fractions.

5.2.2 *In vitro* pull down assay

The essentiality of BRCT domain for effective ligation draws our attention to structurally characterize this physical interaction. The extensive point mutational analysis of conserved residues in the BRCT domain which may facilitate the physical interaction with XthA signifies the importance of conserved G614, G620, G639, KK648/649, V646, GS660/661 situated at the surface of BRCT domain (**Fig. 5.3 A and B**) initially confirmed by *in vitro* GST-pull down assay.

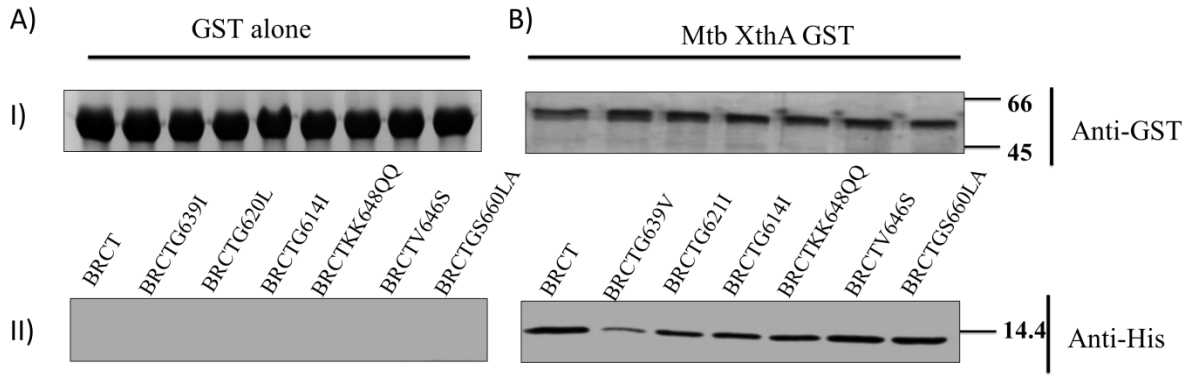


Figure 5.3 Pull down assay to probe MtbLigA-MtbXthA interactions using BRCT domain point mutants. **A, I)** GST alone or **B, I)** GST-tagged WT MtbXthA was incubated with His-tagged BRCT domain or BRCT domain site directed mutants BRCTG614I, BRCTKK648/649QQ, BRCTG639V, BRCTV646S, BRCTGS660/661LS, BRCTG621I and mixed with glutathione sepharose beads. The immunoblot analysis of the pulled down complexes was carried out using appropriate antibodies as detailed in Methods.

5.2.3 Surface plasmon resonance

The MtbLigA and BRCT domain binding affinities with MtbXthA are estimated using surface plasmon resonance and the KD values of 67.7 nM and 66.1 nM respectively are determined (**Fig. 5.4 A and B**). The substitution of Gly614 with Ile614 in MtLigA and BRCT domain reduces the affinity towards MtbXthA by nearly 10-12 folds with KD value of 781 nM (**Fig. 5.5 A**). Gly614 orchestrates their side chain by forming a hydrogen bond with the Arg100 in the XthA. The hydrogen bond interaction between the conserved 648KK649 residues in BRCT domain with the conserved Asp106 and Gln108 residues in XthA is also diminished by substitution of Ile in place of Gly614. The affinity is also get reduced on the mutation of G621L and G639I and KD values observed are 152 nM and 2.77 µM respectively (**Fig. 5.5 B and C**). There occurs a remarkable decrease in the flexibility of MtbLigA in presence of XthA in contrast to its flexible conformation around the nicked DNA. This strong-arm relationship also causes a domain swapping in Class-II endonuclease that leads to the reduced endonuclease activity. The results are summarized in **table 5.2** and docked model of BRCT domain and XthA are shown in **Fig. 5.5 D** with mutated residues on BRCT domain highlighted in pink.

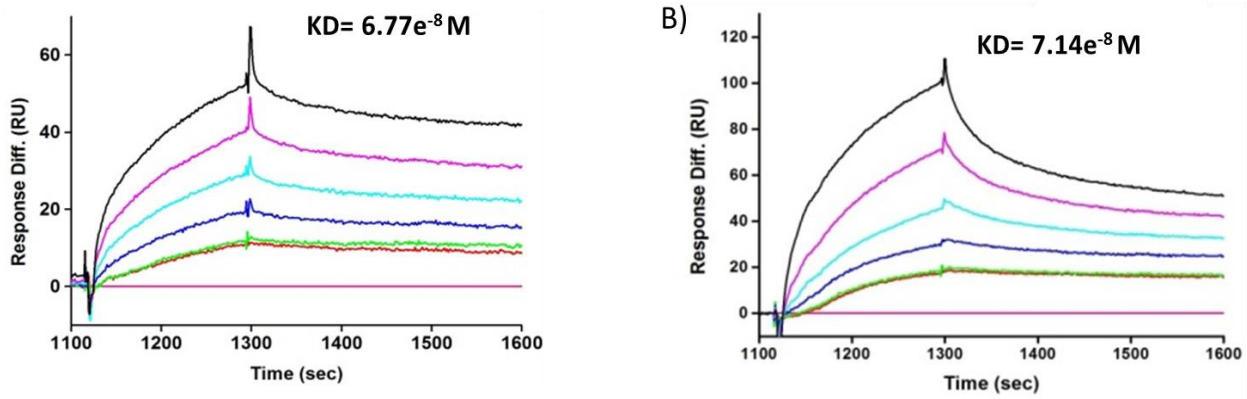


Figure 5.4 Surface plasmon resonance to determine dissociation constant (KD) between **A)** MtbXthA-GST and LigA **B)** MtbXthA-GST and BRCT domain. The KD value of MtbXthA-LigA/BRCT domain is 67.7 nM and 66.1 nM, respectively.

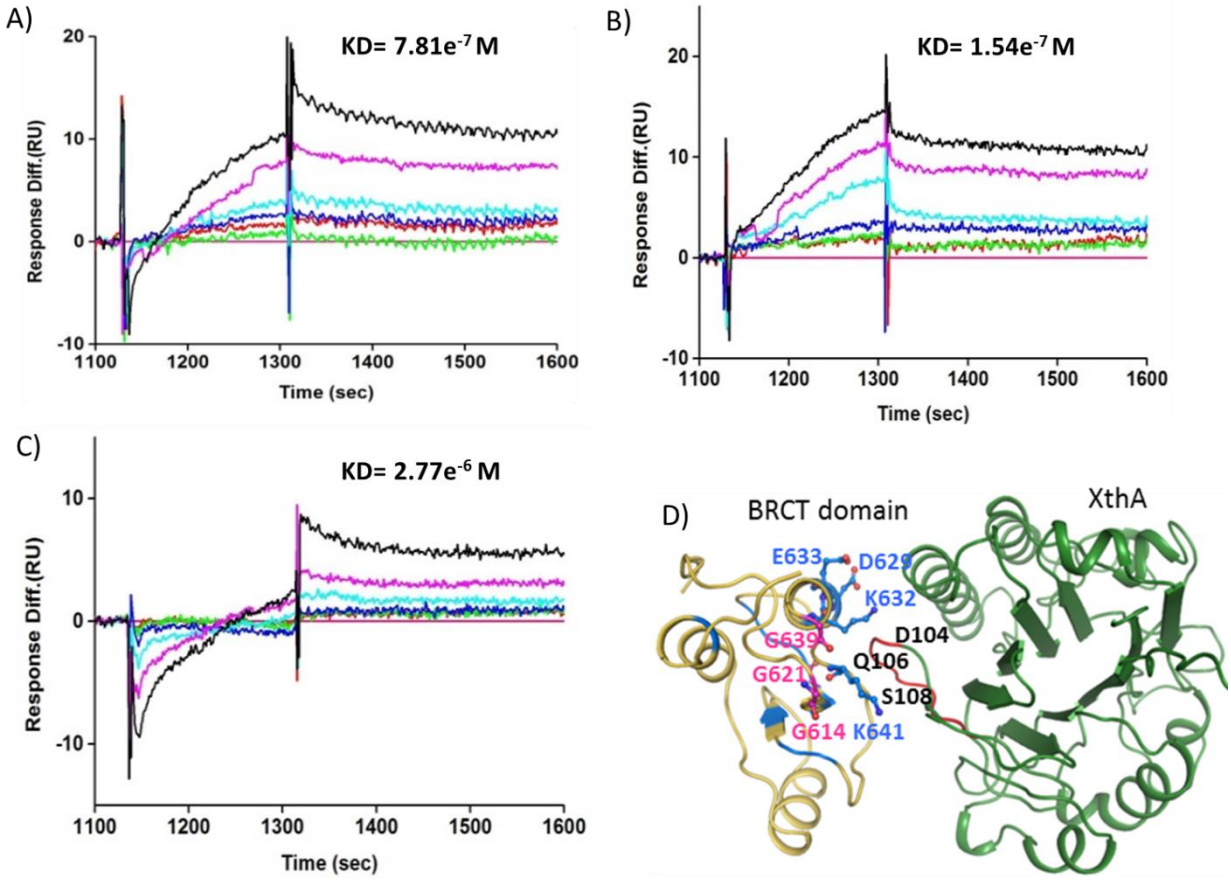


Figure. 5.5 The interaction analysis between MtbXthA and site directed mutants of BRCT domain in LigA full length. The K_D of interaction between **A)** MtbXthA and MtbLigA^{G614I} is 781 nM. **B)** MtbXthA and MtbLigA^{G621I} is 154 nM and **C)** MtbXthA and MtbLigA^{G639V} is 2.77 µM. **D)** The docked model of BRCT domain and XthA showing the binding of $_{104}$DGQ$_{112}$ motif of XthA near the G639, K632, K641 and D632 residues of BRCT domain.

Table 5.2 The summary of surface plasmon resonance analysis of interaction between MtbXthA and LigA/BRCT domain and MtbXthA and site directed mutants of BRCT domain in LigA.

S.No	Ligand	Analyte	KD (M)	Rmax
1.	XthA-GST	MtbLigA	$6.77*10^{-8}$	49.1
2.	XthA-GST	BRCT domain	$7.14*10^{-8}$	74
3.	XthA-GST	$MtbLigA^{G614I}$	$7.81*10^{-7}$	24.9
4.	XthA-GST	$MtbLigA^{G620I}$	$1.54*10^{-7}$	16
5.	XthA-GST	$MtbLigA^{G639V}$	$2.66*10^{-6}$	27.6

5.2.5 Structural characterization of MtbXthA-MtbLigA and MtbXthA-BRCT domain complex

To get an insight into the structure of the MtbXthA-MtbLigA complex, small-angle X-ray scattering (SAXS) was used to first characterize the solution behavior of individual proteins MtbXthA and MtbLigA. While MtbXthA behaved as a globular protein with a radius of gyration (Rg) 2.37±0.00 nm and Dmax 7.3 nm (**Fig. 5.6 A, 5.7 A and Table 5.3**), MtbLigA adopted an extended conformation with large Rg 5.2±0.08 nm and Dmax 16.7 nm as already discussed in the previous chapter.

The SEC analysis estimated a stoichiometric ratio of 1:1 interaction between MtbXthA-MtbLigA, therefore we studied the behavior of MtbXthA-MtbLigA complex in solution at equimolar ratios. MtbXthA-MtbLigA complex appeared multi-domain and flexible in nature with 6.66 ±0.02 nm Rg values and large Dmax 23.86 nm (**Fig. 5.6 B, 5.7 B and Table 5.3 and 5.4**). The Kratky plot for the MtbXthA-MtbLigA complex was parabolic (**Fig.5.8 B, iii purple**) and the linearity of Guinier plot at lower q values suggests that the protein complex was devoid of any aggregation (**Fig. 5.8 B, ii**). The $\chi 2$ values obtained for the experimental data and the SAXS profile calculated from the modeled structure supports the quality of the complex model (**Table 5.3, 5.4**). Importantly, the interaction interface in the

SAXS model shows that the extended MtbLigA engages with MtbXthA through its BRCT domain (**Fig. 5.7 E**).

We also carried out the SAXS analysis of BRCT domain of MtbLigA in complex with MtbXthA to examine the mode of interaction between the two in solution (**Fig.5.7 D**). We found that the BRCT domain alone exhibits an Rg=1.00±0.004 nm. The MtbXthA-BRCT domain complex exhibits increased Rg=2.2±0.2 nm indicating the formation of complex in solution (**Fig. 5.7 E**). The maximum dimension, Dmax, of the complex also increases. The excellent fit with a chi^2 value of 0.203 supports the quality of the model of the complex (**Table 5.3**). The Guinier analysis shows the absence of inter-particle interactions in the protein samples (**Fig. 5.8 D, E ii**). The SAXS experiments demonstrate that the BRCT domain of MtbLigA engages with MtbXthA at a different region compared to the interacting protein interaction peptide (PIP) motif reported by our group earlier for interactions of MtbXthA with the Mtbβ-clamp (Khanam, Rai et al. 2015).

5.2.6 MtbXthA, MtbLigA forms a ternary complex with nicked DNA

We next analyzed the MtbXthA-MtbLigA complex in the presence of nicked DNA substrate (S1) using SAXS (**Fig. 5.7 C**). The addition of DNA into the MtbXthA-MtbLigA complex introduced disorder at near qmin values as reflected by Kratky plot of the complex (Rg 4.52±0.45 nm) (**Fig. 5.8,C iii, teal**). Ternary complex exhibits a reduced Dmax (17.06 nm) compared to the binary MtbXthA-MtbLigA complex (23.86 nm) indicating adoption of more compact structure (**Table 5.3 and 5.4**). While the Kratky plot for MtbXthA-MtbLigA complex was parabolic (**Fig. 5.8 B, iii**), inclusion of DNA in the complex resulted in an upward trend in the plot, towards higher angles reflecting increased disorder (**Fig. 5.8 C,iii**). The SAXS model of ternary complex showed an interaction between MtbXthA and BRCT domain of MtbLigA. In contrast to the compact conformation adopted by MtbLigA in its complex with nicked-DNA, it adopts a relatively extended conformation while interacting with XthA, both in the presence and in the absence of nicked DNA. The linearity of Guinier plot at lower q values suggests that the proteins and complexes were devoid of any aggregation (**Fig. 5.8 C, ii**). The $\chi 2$ values obtained for the experimental data and the SAXS profile calculated from the modeled structure supports the quality of the modeled complex (**Table 5.3, 5.4**).

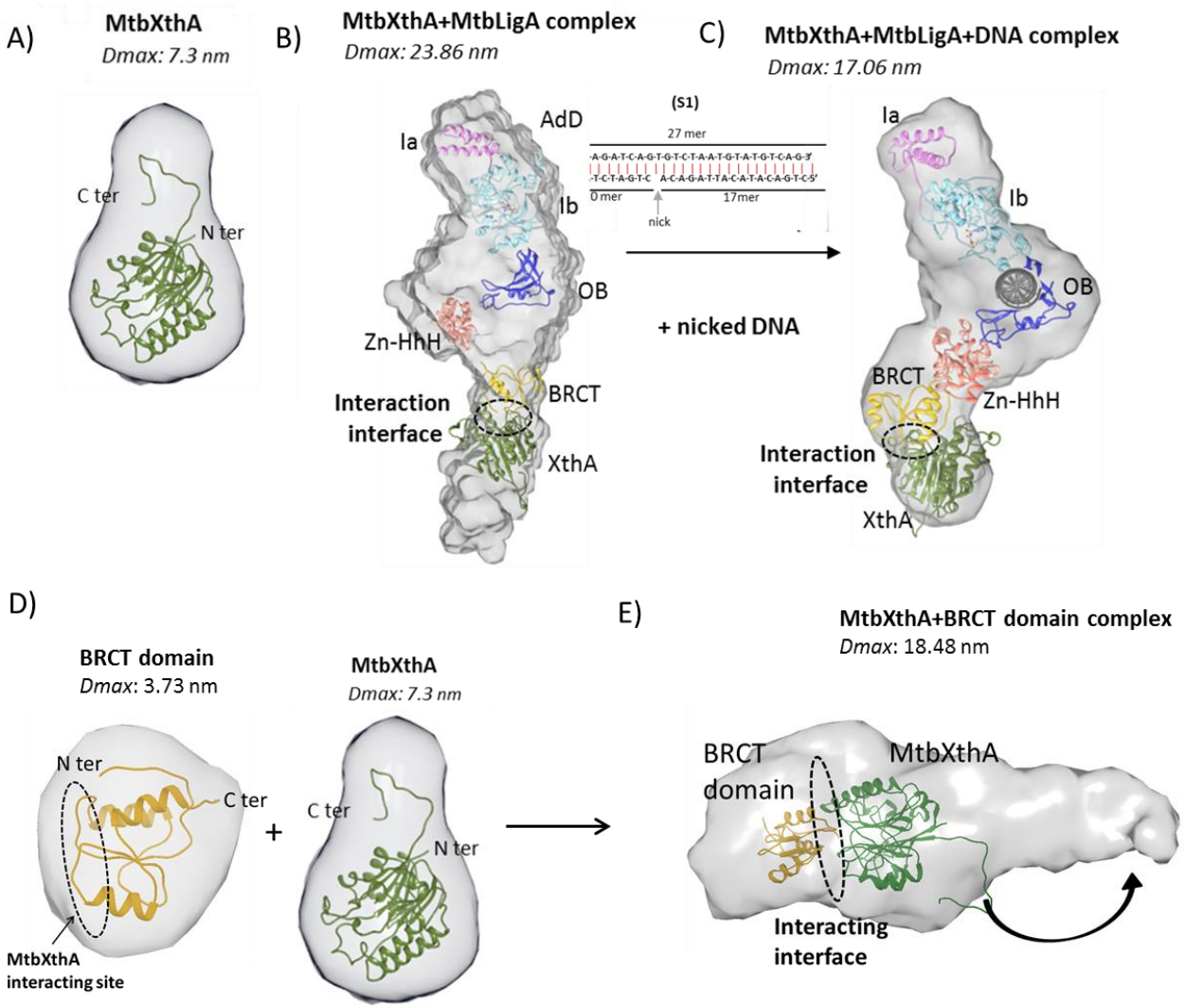


Figure 5.6 Small angle X-Ray scattering (SAXS) envelope of protein and complexes fitted with homology models. **A)** MtbXthA alone **B)** MtbXthA+MtbLigA binary complex. The complex has 'extended' conformation where MtbLigA BRCT domain interacts with MtbXthA. The interaction interface is marked with dotted circle. **C)** MtbXthA+MtbLigA+DNA ternary complex. SAXS shows MtbLigA only 'partially closed' around DNA as compared to MtbLigA-DNA complex. The BRCT domain is seen interacting with MtbXthA with interaction interface marked by dotted circle. **D)** BRCT domain and **E)** BRCT domain in complex with MtbXthA. The interacting interface is shown in dotted circle.

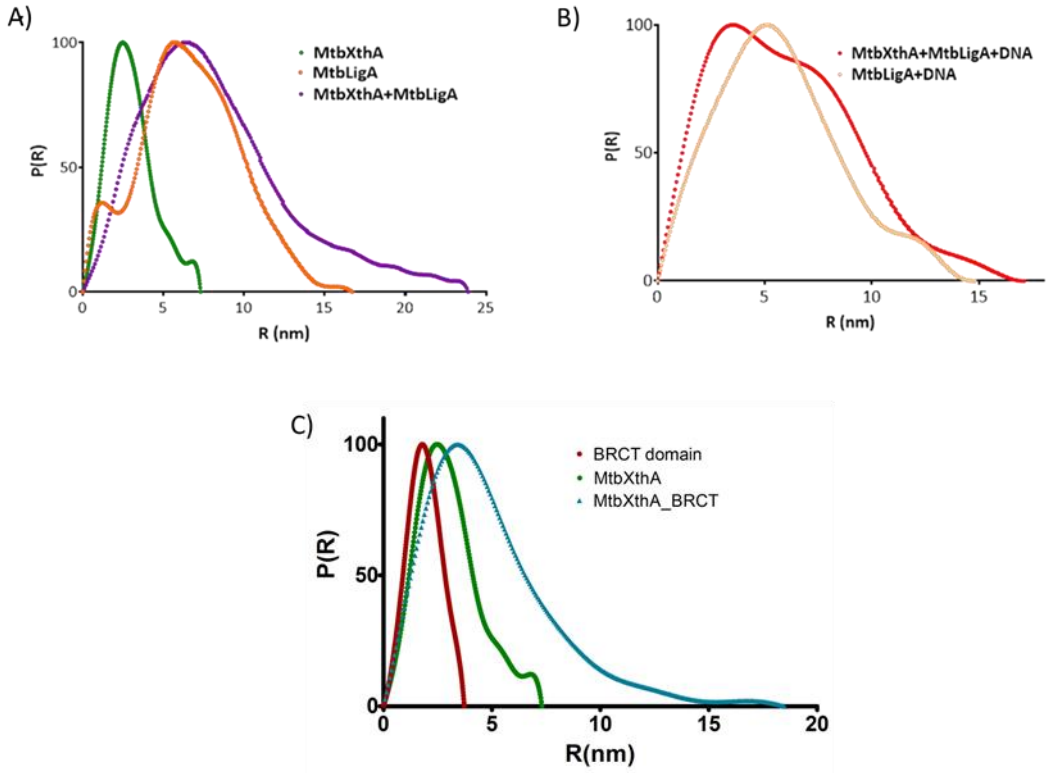


Figure 5.7 Normalized Pairwise interatomic distance distribution P(r) graphs. A) Normalized Pairwise interatomic distance distribution P[r] function of MtbXthA, MtbLigA and MtbXthA+MtbLigA complex demonstrates an increase in Dmax value for the complex. **B)** Normalized Pairwise interatomic distance distribution P[r] function for MtbLigA+DNA and MtbXthA+MtbLigA+DNA. **C)** Normalized Pairwise interatomic distance distribution P[r] function for BRCT domain, XthA and XthA-BRCT domain complex.

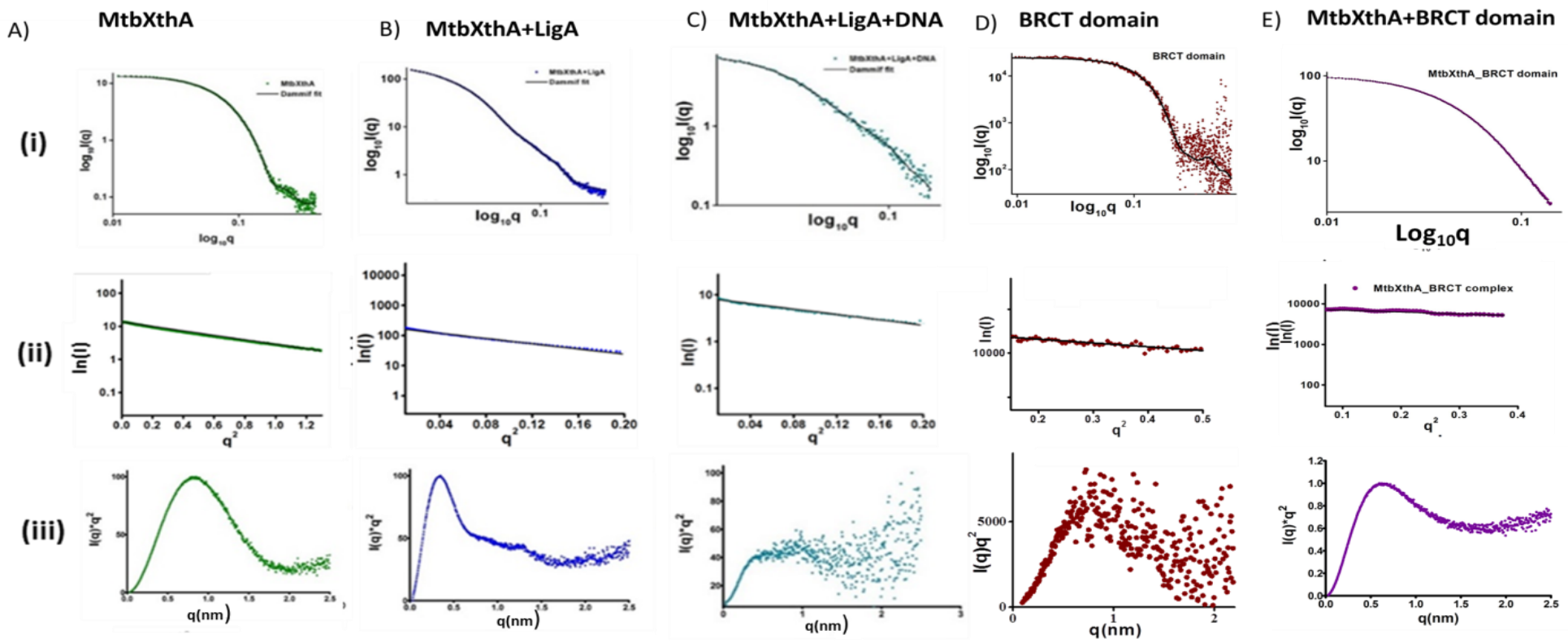


Figure 5.8 Small angle X-Ray scattering. (i and ii) Scattering and Guinier plots (iii) Normalised Kratky plots indicating degree of disorder of respective protein and complexes. A) MtbXthA B) MtbXthA+MtbLigA binary complex C) MtbXthA-MtbLigA-DNA ternary complex D) BRCT domain E) MtbXthA-BRCT domain complex.

Table 5.3 SAXS data collection and statistics

	MtbXthA	MtbXthA+LigA	MtbXthA+LigA+Nicked ds-DNA	BRCT domain	MtbXthA + BRCT domain
SASBDB ID	**SASDDD8**	**SASDEW3**	**SASDEX3**	**SASDD88**	**SASDDQ8**
Data collection parameters					
X-ray source	BM29,ESRF	BM29,ESRF	BM29,ESRF	Saxspace, CDRI	BM29,ESRF
Detector	Pilatus 1D	Pilatus 1D	Pilatus 1D	Dectris Mythen 2R1K	Pilatus 1D
Wavelength (nm)	0.081	0.081	0.081	0.154	0.081
Exposure time per frame	1sec	1sec	1sec	30min	1sec
No. of frames collected	10	10	10	2	10
Concentration (mg/ml)	5.3	2.5	2.0	10	4.6
Measurement temperature (°C)	10	10	10	10	10
Structural parameters					
I(0) [from Guinier]	13.63±0.012	197.66±0.46	8.15±0.073	25148±58.86	98.53±0.10
Rg (nm) [from Guinier]	2.37±0.00	6.24±0.02	4.52±0.45	1.58±0.02	3.69±0.01
I(0) [from P(r)]	13.54±0.012	201.6±0.479	8.317±0.099	24290±11.39	100.6±0.20
Rg (nm) [from P(r)]	2.35±0.003	6.66±0.0213	4.788±0.089	1.44±0.047	4.05±0.025
Dmax (nm)	7.3	23.86	17.06	3.7	18.48
Modeling Parameters					
Symmetry	P1	P1	P1	P1	P1
Anisotropy	Unknown	Unknown	Unknown	Unknown	Unknown
Modelling iterations	10	10	10	10	10
Chi^2	0.5765	1.02	0.8048	0.789	0.968

Table 5.4 SAXS data fitting and validation

Sample	Dammif		Supcomb	Crysol
	χ^2	NSD	NSD	χ^2
MtbXthA	0.5765	0.514±0.06	1.94	-
MtbXthA+MtbLigA	1.02	0.645±0.041	2.94	-
MtbXthA+MtbLigA+DNA	0.8048	0.695±0.02	2.8	1.951
BRCT domain	0.789	0.484	-	0.917
MtbXtha+BRCT domain	0.968	0.577	-	0.262

NSD- normalized spatial discrepancy

5.2.7 Identification of a conserved MtbXthA motif that mediates MtbXthA-MtbLigA interactions

5.2.7.1 Homology modeling and Protein-protein docking

The interaction between the identified peptides in MtbXthA and the BRCT domain of MtbLigA was further validated by *in silico* molecular docking. For BRCT domain- peptide docking, the minimized model of BRCT domain was used to evaluate interactions with peptides (i) DGQ (DGQPSWSGKP and (ii) YDV (YDVAHVGFD), by using CABS-dock at default parameters and without restricting any binding pocket. CABS-dock allows for large-scale flexibility in peptide and protein to screen the occupying niches with in the protein. The 10 different docking models were generated and analyzed. The best one was selected based on occurrence in docking trajectory. The selected model was analyzed by PDBsum to generate the 2D docked conformation of peptide. The BRCT protein-MtbXthA peptide docking predictions suggested that the DGQ peptide stably binds BRCT by forming numerous favorable hydrogen bonds and hydrophobic interactions (**Fig. 5.9 A**), whereas the YDV peptide was observed to have very poor affinity towards BRCT domain marked by only weak interactions (**Fig. 5.9 B**).

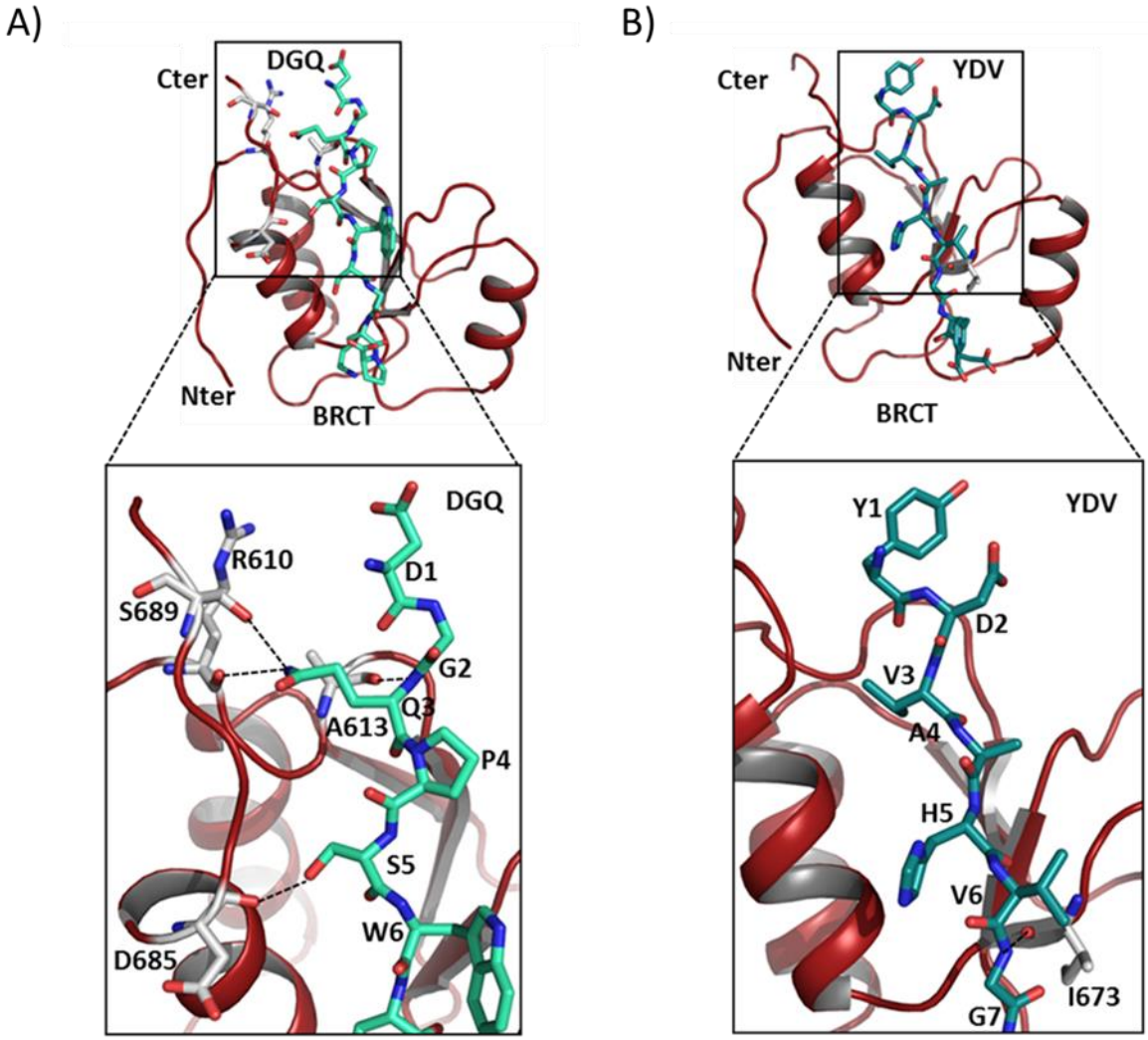


Figure 5.9 Modelling of protein-peptide interaction by molecular docking. The MtbXthA derived peptides **A)** DGQ **B)** YDV shown as green sticks, were docked against molecular model of MtbLigA BRCT domain (deep red) to get structural insight into the interaction between MtbXthA and BRCT domain of MtbLigA. The interaction sites are shown in zoom-in view in panels below. Hydrogen bonds are depicted as black dotted lines.

5.2.7.2 Isothermal titration calorimetric assay to determine XthA derived peptides binding with BRCT domain of MtbLigA

It is clear from the above results that the C-terminal BRCT domain of MtbLigA interacts with MtbXthA. Correspondingly, we examined the SAXS derived MtbXthA-MtbLigA model to identify the potential region on MtbXthA involved in mediating the interactions of the complex. An examination of the interaction interface showed that the BRCT domain of MtbLigA engages with an amino acid stretch $_{104}$**DGQPSWSGKP**$_{113}$ in MtbXthA (**Fig. 5.10**). We then designed a peptide corresponding to the interaction interface region in MtbXthA and performed Isothermal Titration Calorimetery (ITC) to determine its binding with the BRCT domain of MtbLigA. The test peptides used were: DGQ ($_{104}$**DGQPSWSGKP**$_{113}$) located on the interaction interface and a control YDV ($_{75}$**YDVAHVGFD**$_{83}$), present on an adjoining loop in proximity but away from the interaction interface. The DGQ peptide bound to BRCT domain with apparent Kd of 2.47 µM, while the YDV peptide showed minimal interactions (Kd of 88.49 µM) (**Fig. 5.11 A and C**). Peptides titrated against only the buffer did not show any non-specific binding of peptides (**Fig. 5.11 B and D**). The thermodynamic analysis of DGQ peptide binding suggests the occurrence of favorable hydrogen and hydrophobic interactions (**Table 5.5**).

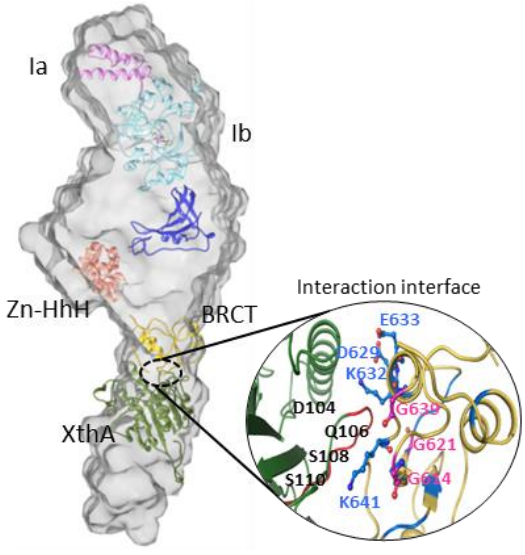


Figure 5.10 Validation of interaction interface between MtbXthA and MtbLigA. SAXS envelope of MtbXthA-MtbLigA complex with fitted homology model showing interaction interface between MtbXthA and BRCT domain of MtbLigA. The BRCT domain and MtbXthA are represented in yellow and green, respectively. The residues (marked) forming contacts at the interface are shown in sticks. The hydrogen bonds are shown as black dotted lines.

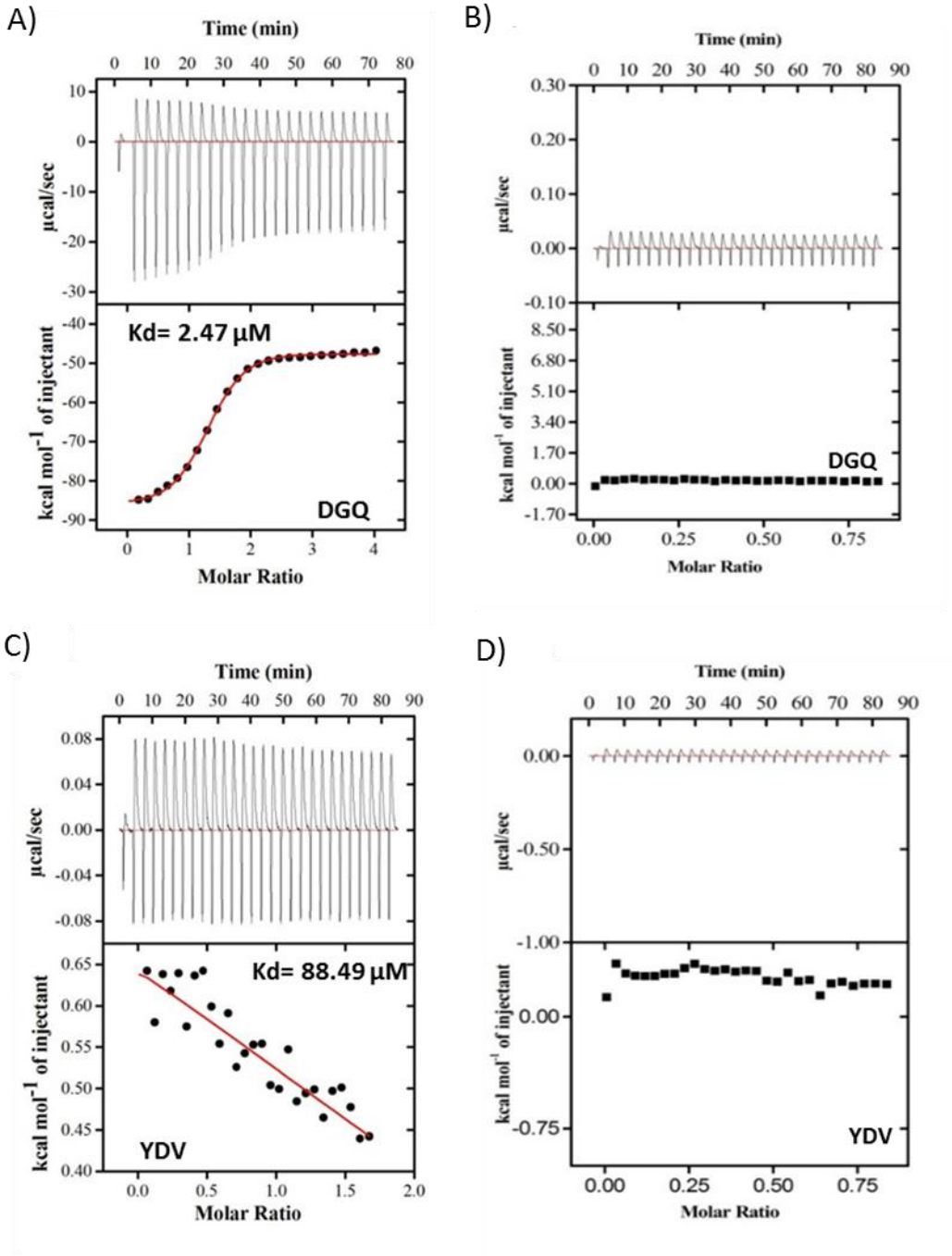


Figure 5.11 Isothermal calorimetry (ITC): The binding isotherms for **A)** DGQ peptide: located on interaction interface and **C)** control YDV peptide, with the BRCT domain of MtbLigA are shown. Both raw ITC isotherms (top panels) and data after integration and normalization (bottom panels) are depicted. The solid lines (red) in the bottom panels show the fits with a one-site binding model. The control experiments involving interaction between **B)** DGQ and **D)** YDV peptide with reaction buffer.

Table 5.5 Thermodynamic parameters for ITC experiments.

Ligand / Concentration (mM)	BRCT (µM)	N	KD (µM)	ΔG (kcal mol^{-1})	ΔH (kcal mol^{-1})	-TΔS (kcal mol^{-1})
DGQ (0.4)	5	2.44	2.4752	-4.987295	-9.958	-14.94529
YDV (0.4)	5	4.19	88.491	-5.63434	0.9137	-6.54804

5.2.8 Disruption assay using FRET and determination of interaction kinetic analysis between MtbXthA mutant ($MtbXthA^{ID}$) and WT MtbLigA using SPR

We next performed peptide mediated disruption assays against the MtbXthA-BRCT complex. Fluorescently labelled $MtbXthA^{OG}$ and $BRCT^{AF}$ (Alexa fluor labelled) were used to monitor complex formation using FRET (**Fig. 5.12 A**). FRET between $MtbXthA^{OG}$ (donor) and $BRCT^{AF}$ (acceptor) was measured in the absence (**Fig. 5.12 B and C; purple**) and presence of peptides (DGQ-orange and YDV-green). Inclusion of DGQ peptide in the reaction mixture resulted in a marked decrease in FRET indicative of inhibition of MtbXthA-BRCT complex formation (orange), while the peptide YDV resulted in a poor decrease in FRET efficiency (green). No reduction in FRET was observed on inclusion of a negative control peptide (IGELLSPSLFGV) (grey). The result supports that the sequence motif, $_{104}$**DGQPSWSGKP**$_{113}$ in MtbXthA is important for forming a complex with MtbLigA. After SAXS derived examination we concluded that BRCT of MtbLigA engages with the $_{104}$**DGQ**PSW**SGK**P$_{113}$ amino acid stretch in MtbXthA. To further validate this point we tried to mutate the first and last three amino acid of the stretch to $_{104}$DGQPSW**ALA**P$_{113}$ and there occurs 10 times enhancement in the KD value of $MtbXthA^{ID}$ mutant and wild-type LigA (**Fig. 5.12 D**).

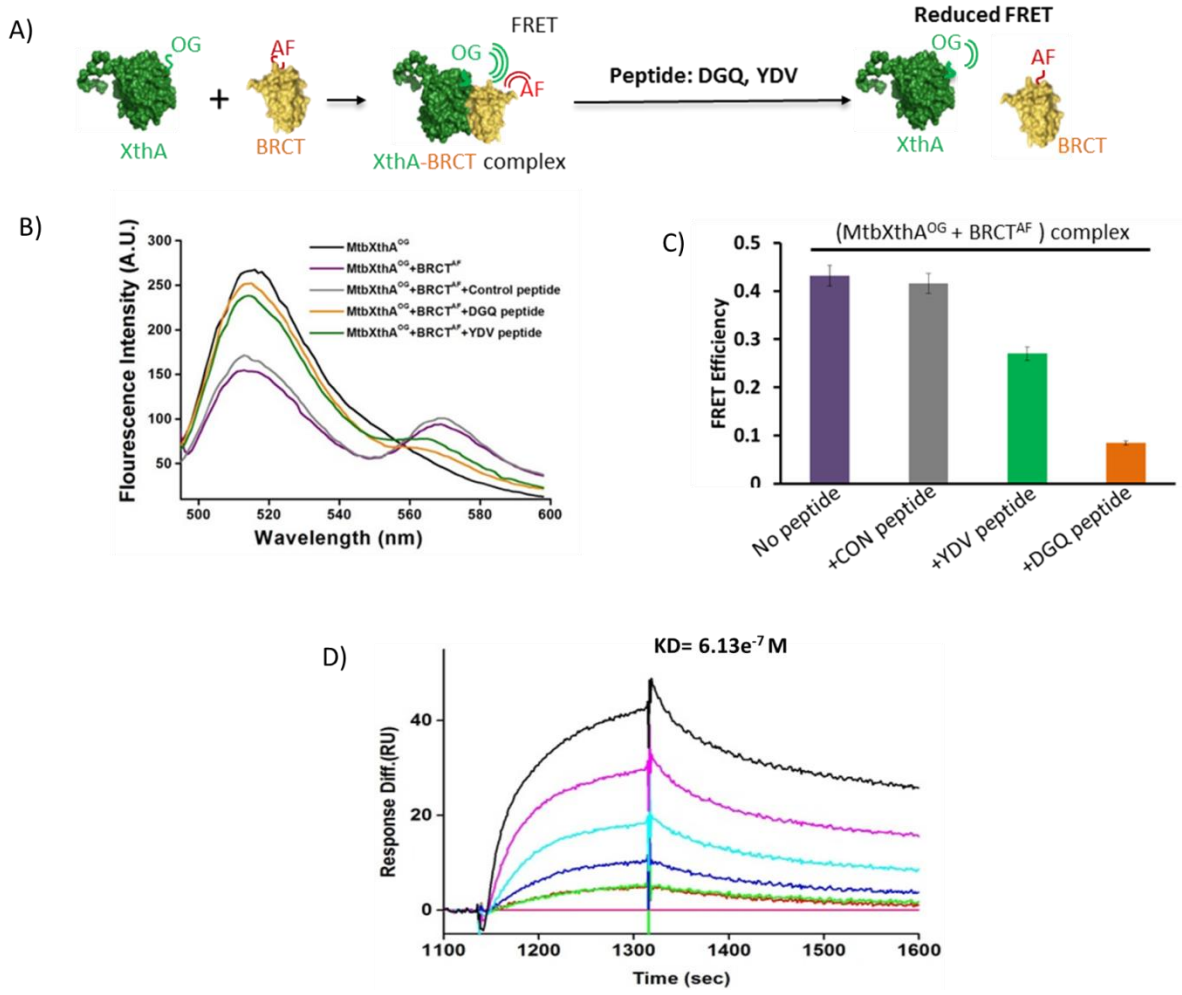


Figure 5.12 Disruption assay A) Schematic representation of the Fluorescence resonance energy transfer (FRET) experiment to probe disruption of the MtbXthA-BRCT domain complex by the respective peptides. **B)** FRET analysis showed that the MtbXthA derived peptides disrupt the interaction between Oregon green 488 maleimide-labelled MtbXthA ($MtbXthA^{OG}$) and Alexa fluor 488-labelled BRCT domain ($BRCT^{AF}$). A decrease in FRET occurred in the presence of DGQ peptide (orange). The YDV (green) moderately disrupted FRET efficiency while control peptide (grey) could not elicit disruptive effects. **C)** Bar graphs showing FRET efficiency of respective peptides depicted as: YDV (green), DVR (orange), DGQ (deep green) and control peptide (grey). The purple bar represents the control without peptides. Error bars correspond to standard deviation of two independent experiments. D) Surface plasmon resonance to determine dissociation constant (KD) between $MtbXthA^{ID}$-GST and MtbLigA. The KD value of MtbXDS and MtbLigA is 613 nM.

5.3 Discussion

Not much is known about protein interactions regulating BER in bacteria. Recently, LigC, an ATP-dependent mycobacterial ligase, was shown to function as a scaffolding protein in excision repair pathways during the stationary phase and also that it interacts with core BER components (Plocinski, Brissett et al. 2017). The authors have reported interactions of LigC with XthA as well (Plocinski, Brissett et al. 2017). Despite large structural differences that exist between mycobacterial and human proteins and between LigA and LigC, we found that MtbLigA and MtbXthA form a complex *in vivo*. The *in vitro* assays using recombinant proteins demonstrate the existence of direct physical interactions between MtbLigA and MtbXthA. We have also demonstrated that this interaction is mediated by BRCT domain of MtbLigA. Although there are various protein interactions mediated by BRCT domains which are important for BER in eukaryotes (Masson, Niedergang et al. 1998, Taylor, Wickstead et al. 1998, Dulic, Bates et al. 2001, Vidal, Boiteux et al. 2001), to our knowledge this is the first report where the BRCT domain of NAD^+-dependent DNA ligase has been shown to mediate a protein-protein interaction in bacteria.

The XRCC1-DNA Ligase III complex interacts with residues at the N-terminal disordered tail of an otherwise well-structured protein, APE1 (Gorman, Morera et al. 1997, Mol, Izumi et al. 2000, Beernink, Segelke et al. 2001). However, sequence alignment and *in silico* homology modeling showed that such an N terminal extension is absent in MtbXthA (Khanam, Rai et al. 2015). While we have reported the presence of a clamp interacting protein interaction peptide (PIP) motif in MtbXthA (Khanam, Rai et al. 2015), it was of interest to find the region in MtbXthA through which it interacts with MtbLigA. Examination of the SAXS model of the XthA-BRCT domain complex reveals an interaction interface. Sequence analysis results of the interaction interface shows that MtbXthA harbors a conserved motif '$_{104}$DGQPSW**SGK**P$_{113}$' at the interaction surface that may mediate interactions with LigA. The motif is positioned on a loop in XthA protein, extending away from the protein surface and can conceivably interact with partner proteins.

AP-endonucleases are known to interact and stimulate various BER components. APE1 increases DNA glycosylase turnover by displacing it from its product (Parikh, Mol et al. 1999, Waters, Gallinari et al. 1999). FEN1 and DNA polymerase β are also stimulated by APE1 (Bennett, Wilson et al. 1997, Ranalli, Tom et al. 2002). Additionally, APE1 enhances

the product formation by DNA LigaseI on nicked DNA substrate with 3'-OH and 5'-phosphoryl termini (Ranalli, Tom et al. 2002). We explored the influence of MtbXthA on the activity of LigA, and to our surprise found that MtbXthA inhibits the ligation activity of MtbLigA by several folds. This was unusual as AP-endonucleases have otherwise been found to increase the overall efficiency of BER. We showed that the decrease in ligase activity was neither due to the intrinsic 3'-5'exonuclease activity of MtbXthA on the DNA substrate nor due to mutual competition to bind to the substrate. Our previous report demonstrated that the BRCT domain is required for efficient ligation by MtbLigA and the BRCT deleted mutant of MtbLigA is incapable of forming a complex with nicked DNA (Srivastava, Dube et al. 2007). Additionally, structural studies of *T. fi* DNA ligase showed that LigA adopts a toroidal conformation in the presence of DNA where the BRCT domain forms extensive interactions with adenylation domain and acts as a gate which regulates DNA binding and release (Lee, Chang et al. 2000). Our work suggests that the interactions of MtbXthA with BRCT domain of MtbLigA precludes the latter from substrate encircling, binding and making important contacts with the adenylation domain. This results in an inefficient DNA binding and the observed reduction in nick ligation by MtbLigA. Interaction of MtbXthA with BRCT domain renders MtbLigA trapped in a relatively 'open' inefficient conformation. The observation is further validated by size exclusion analysis as there occurs reduction in stokes radii of MtbLigA in presence of nicked DNA substrate as compared to the one in absence of DNA substrate indicating towards the formation of compact closed conformation. SAXS studies involving MtbXthA and BRCT domain alone as well as both in complex shows that interaction is mediated through BRCT domain.

MtbXthA acts at the 5'-side of the AP-site and generates a nicked DNA with 3'-OH and 5'-dRP moieties at the incised termini. The 5'-dRP moiety from pre-incised AP-DNA is removed by AP-Lyase to generate suitable DNA ends for further processing (Bogenhagen and Pinz 1998). While, ATP ligases like mitochondrial DNA ligase, T4 DNA ligase and DNA Ligase I, are active on DNA substrates with 5' incised AP-site bearing AP-Lyase activity (Bogenhagen and Pinz 1998), no such activity has been yet been reported for bacterial NAD^+-Ligase. Earlier our group hypothesized that MtbLigA might be active at the incised abasic site and also found that MtbLigA could efficiently ligate nicks adjacent to 3'-OH and 5'-THF or 5'-dRP DNA termini. This

clearly suggests that MtbLigA can directly reverse the action of MtbXthA at abasic site by resealing the nicked DNA with 5'-dRP residue. Therefore, the detected inhibition of the MtbXthA AP-endonuclease activity by MtbLigA can be attributed to re-ligation of the incised abasic DNA generated by MtbXthA. This religation step would be a futile step during AP-site DNA repairs process and could be detrimental for cells by locking the 5'-dRP residue back into the DNA backbone.

The present work supports a new level of coordination in mycobacterial BER pathway orchestrated by MtbXthA. We propose that MtbXthA keeps a check on MtbLigA ligase activity by engaging with its BRCT domain and keeping it in an inactive 'open' conformation. This restricts DNA binding of LigA and formation of important contacts with the adenylation domain necessary for effective ligation. MtbXthA restrains MtbLigA to protect incised abasic DNA against futile resealing activity of MtbLigA. Our previous study showed β-clamp increases the affinity of MtbXthA for its substrate (Khanam, Rai et al. 2015). Since β-clamp acts as a scaffold protein that recruits protein participants to the site of damage, we speculate that the β-clamp stays at the damage site with MtbXthA till an AP-Lyase (as yet unidentified) is recruited to remove the 5'-dRP moiety. This would result in the release of MtbXthA from the damage site and facilitate the action of DNA polymerase and ligase to resume their normal functions**.**

Chapter 6 | Identification of Small Molecule Fragments as Inhibitor of MtbLigA

6.1 Introduction

M. tuberculosis is a worldwide health menace to human population. The emergence of multi drug resistance (MDR) and extensively drug resistance (XDR) strains of this deadly pathogen opens up the demand to formulate novel drugs or biosimilars with high degree of accuracy and precision. Bacterial NAD^+-dependent DNA ligase has emerged as a new candidate for therapeutic intervention because of its essential nature and different architecture and cofactor requirement from the host counterpart. Although potent LigA inhibitors like arylamine compounds and chloroquine derivatives have been identified, they have the disadvantage of binding non-specifically to DNA and not being able to compete with NAD^+-binding (Ciarrocchi, MacPhee et al. 1999). Earlier our group had reported diverse compound families, which inhibit MtbLigA with several fold specificity compared to ATP-dependent ligases including the human DNA ligase I (Srivastava, Dube et al. 2007, Yadav, Khanam et al. 2015). Up till now the active site of the enzyme was in focus for inhibitor screening but the crystal structure with NMN bound in Ia subdomain of AdD domain opens up the new approach of screening the future inhibitors against NAD^+-dependent DNA ligases. The residues Y31, Y32, F44, I39, D41 and D45 constitute the NMN binding site. By molecular screening of fragment library (1166 fragments), we found 14 fragments showing higher affinity towards the NMN binding pocket of MtbLigA than ATP-dependent DNA ligases like Human Ligase1 on the basis of their docking score. Docking and modeling studies attributed the specificity of these inhibitors to their ability to mimic NMN-enzyme interaction. In summary, the *in vitro* and *in vivo* assays demonstrated that these compounds could distinguish between NAD^+-and ATP-dependent ligases and that the *in vivo* mode of their action is primarily through inhibition of the essential LigA.

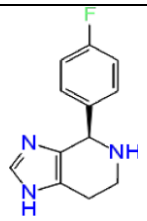
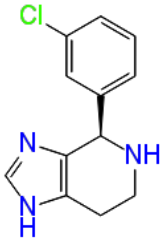
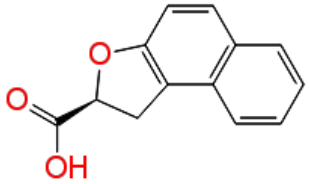
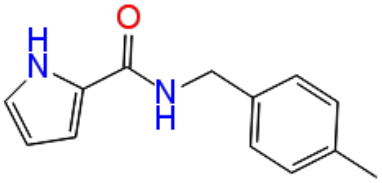
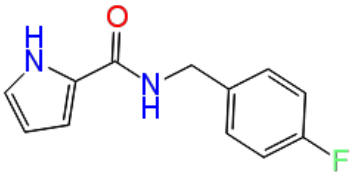
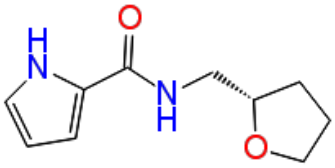
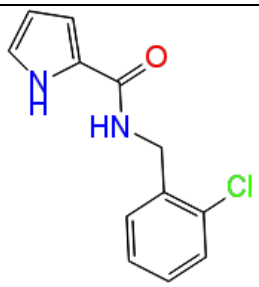
6.2 Results

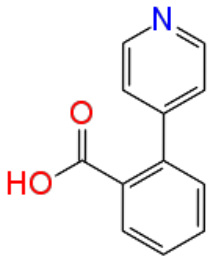
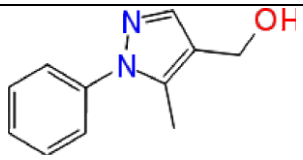
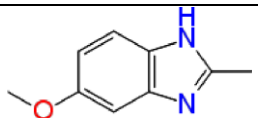
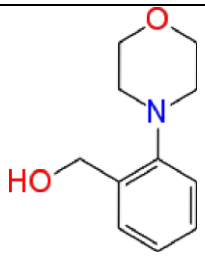
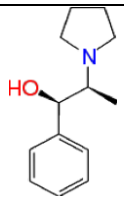
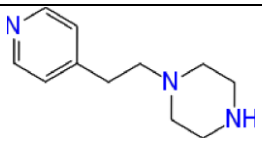
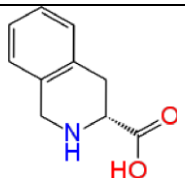
6.2.1 *In silico* screening of the Fragments against NMN binding pocket of MtbLigA

The Crystal structure (PDB: **1ZAU**) was minimized by using Tripos force field with 500 steps of steepest descent algorithm followed 500 steps of Powell method and 50.00 kcal/mole convergence threshold energy (custom parameters). The flexible docking was performed by Surflex-Dock Program available in SYBYL-X v2.1. (https://www.certara.com/pressreleases/ certara-enhances- sybyl-x-drug-design-and discovery-software-suite/). The chosen protein NMN binding site selected to prepare grid using Protomol module. The 2nd Generation BIONET premium fragment library (1166 compounds) purchased from Key Organics Ltd. (https://www.keyorganics.net/) was screened against defined NMN binding site. The each active hit was selected by generating 20 maximum poses and 20 maximum conformations per

fragment of protein. The final pose of hits with highest docking score were selected which calculated using scoring functions i.e. G-Score, D-Score, PMF-Score and Chem-score. All visualization analysis was done by PyMol v2.2 (https://pymol.org/2/) and Chimera visualization software. By molecular screening of fragment library (1166 fragments), we found 14 fragments showing higher affinity towards the NMN binding pocket of MtbLigA than ATP-dependent DNA ligases like Human Ligase1 on the basis of their docking score (**Table 6.1**). The docking results indicated that all the inhibitor could fit well in NMN binding site pocket showing maximum interaction. This docking result highlighted some key residues that were consistently participating in the binding of all the inhibitors within the actives site. GLN28, TRY31, TRY32, ILE39, ASP41, and GLY72 were predicted from PyMol to form hydrogen-bonding interactions with the inhibitor in several docking conformations. Few of the docked compounds showed significantly strong binding while others demonstrated low affinity with fewer interactions. This may be attributed due to the differences in functional groups in the compounds.

Table 6.1 *In silico* screening of fragments from BioNet library

S. No.	Scaffold ID	Scaffold structure	Docking score
1.	1R-0678		6.5
2.	1R-0690		6.1
3.	1R-0197		4.9
4.	9R-0345		6.4
5.	9R-0342		6.1
6.	9R-0337		7.1
7.	9R-0344		6.2

8.	PS-5707		6.2
9.	PS-4282		5.4
10.	PS-3417		6.3
11.	FS-3025		6.3
12.	FS-2278		6.3
13.	FS-1570		6.2
14.	FS-3738		6.1

*S. No., serial number; Binding energies for respective compounds were calculated by SYBYL-X v2.1.

6.2.2 *In vitro* enzymatic assay and DNA- inhibitor interaction assay

To access the potential of the fragments as specific inhibitor of MtbLigA, we performed assays where all the fragments were examined at concentrations from 2-20 µM against purified MtbLigA and T4 DNA ligase in ligation assays. This aided us to screen out the potential fragments that could discriminate between NAD^+ and ATP dependent DNA ligases for further analysis. Of the fourteen fragments, 1R-0678, 9R-0345 and FS-3025 were able to distinguish better. The rest fragments had affinity towards ATP ligases as they do not affect the ligation activity. After *in vitro* screening of all 14 fragments we got three fragments which

showed higher affinity towards the NMN binding site with IC_{50} of 5.4 µM for 1R-0678, 7.4 µM for 9R-0345 and 8.6 µM for FS-3025 (**Fig. 6.1 A, B, C and D**). By using standard enzyme kinetics we evaluated the mode of inhibition of the three fragments 1R-0678, 9R-0345 and FS-3025 with respect to NAD^+. The in vitro enzyme kinetic analysis of these fragments shows that they inhibit the enzyme activity in competitive mode with inhibitory constant (K_i) 16.17 µM for 1R-0678, 10.108 µM for 9R-0345 and 15.53 µM for FS-3025 (**Fig. 6.2 A, B and C**) as also visualized in double reciprocal plot (Fig.6.2 D, E and F). We next performed ethidium bromide displacement assays for the selected fragments in the current study to access their interactions with DNA, if any. It was earlier reported that aryl amino compounds, a class of DNA ligase inhibitors, generally intercalate with DNA, which might affect their inhibitory potencies (Ciarrocchi, MacPhee et al. 1999). The maximum concentration of the fragment (0, 50 and 250 µM) was added in the reaction mixture and the fluorescence spectrum was observed (**Fig. 6.3 A, B and C**). The EtBr displacement assay and gel shift assay shows that the fragments do not interact with the linearized pUC18 DNA molecule (**Fig. 6.3 D**).

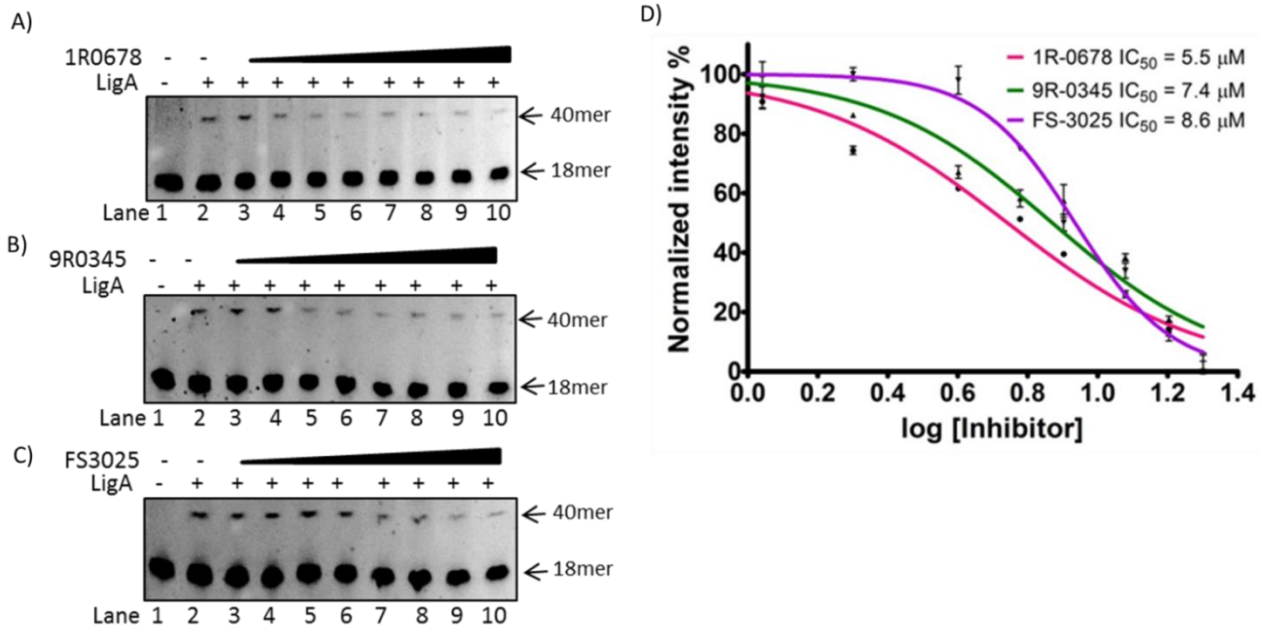


Figure 6.1 *In vitro* inhibition of ligation activity of LigA in presence of fragments. A) 1R-0678, **B)** 9R-0345 and **C)** FS-3025. Lane 1 is control devoid of enzyme and lane 2 is positive control with WT LigA without any inhibitor fragment. All the three fragments were taken in increasing concentration range of 2-20 µM. **D)** The gels were quantified to calculate IC50 of each fragment and 1R-0678 (pink) shows the least inhibitory concentration of 5.4 µM and 9R-0345 (slate blue), FS-3025 (purple) shows the IC50 of 7.4 µM and 8.6 µM respectively. Error bars corresponds to standard deviation from two independent experiments.

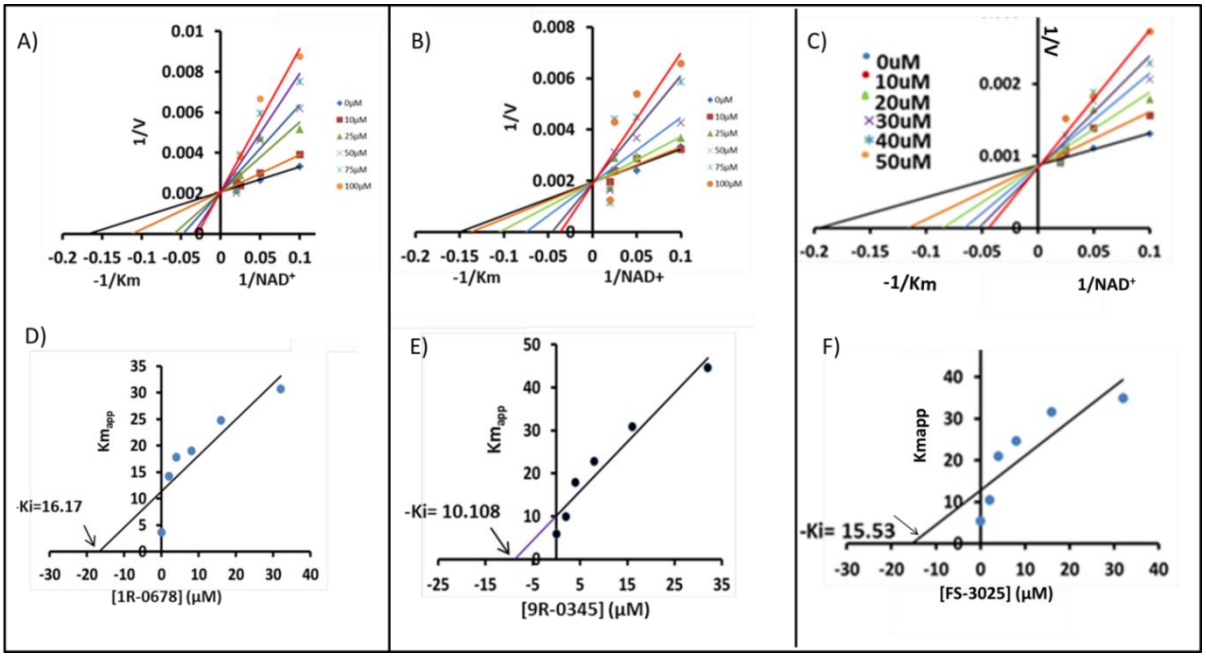


Figure 6.2 Graphs representing competitive mode of inhibition and determination of inhibition constant. A, B and C) The double reciprocal plot indicates competitive binding to NAD^+ for 1R-0678 , 9R-0345 and FS-3025. **D, E and F)** Determination of mode of inhibition in the presence of varying concentration of NAD^+ (0-50 µM) with increasing concentration of 1R-0678, 9R-0345 and FS-3025 (0, 2, 4, 8,16, 32 µM) and inhibition constant calculation (Ki) of 16.17 for 1R-0678 , 10.108 for 9R-0345 and 15.53 for FS-3025.

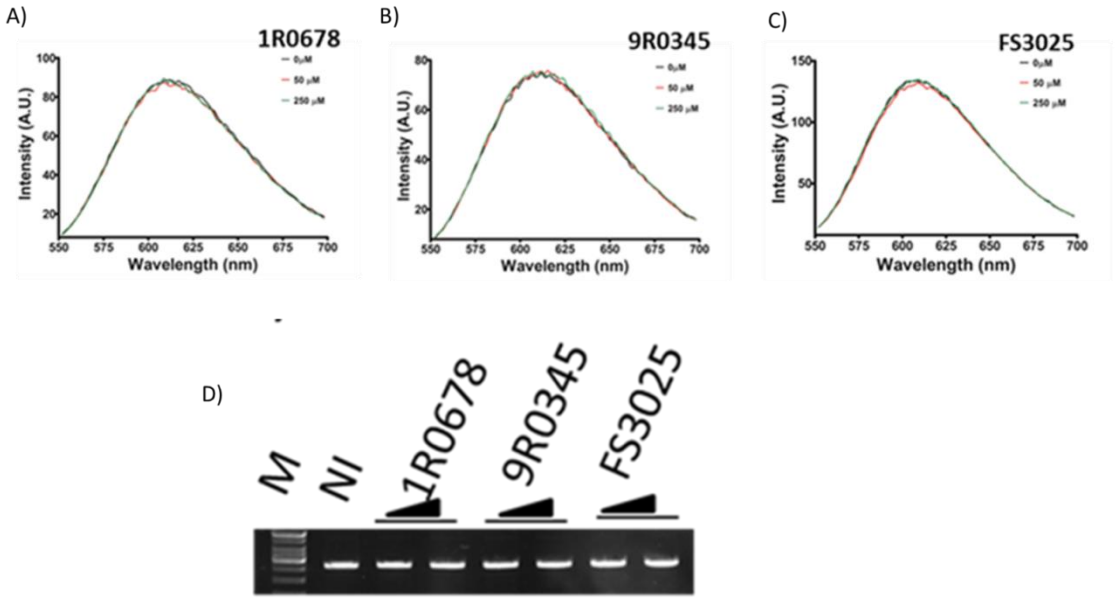


Figure 6.3 DNA-inhibitor interaction assay. A, B and C) Fluorescence spectra to fragment 1R-0678, 9R-0375 and FS-3025. **D)** Gel shift assay to probe DNA-inhibitor interaction.

6.2.3 *In vivo* antibacterial assay

Two bacterial systems were utilized in order to assess the *in vivo* inhibition of NAD^+-dependent DNA ligases as previously reported (Srivastava, Dube et al. 2005, Srivastava, Tripathi et al. 2005). First system used in the assay was a temperature sensitive strain *E. coli* GR501 that carries a lig251 mutation in its LigA gene due to which it grows normally at 30˚C but growth is strongly hampered at higher temperatures. This deficiency is surmounted by complementing it with NAD^+ or ATP dependent ligases (Brotz-Oesterhelt, Knezevic et al. 2003, Lavesa-Curto, Sayer et al. 2004). Consequently this strain has been used to study the *in-vivo* specificity of inhibitors for LigA. We subsequently used pTrc99A-based systems involving MtbLigA and T4Lig in this strain (Ren et al. 1997). The *E. coli* GR501 strain possessing only the pTrc99A plasmid demonstrated much augmented sensitivity to the compounds compared with the corresponding ligase-complemented strains due to the remnant ligase activity in the mutant strain relative to the growth-rescued strains, which harbor a high copy number of the overexpressed ligase used to retrieve them. The bactericidal activity of these fragments demarcated 1R-0678 and 9R-0345 fragment to be potentially active with minimum inhibitory concentration (MIC) of 20 µg/ml and 60 µg/ml (**Table 6.2**). The time kill study also shows that 1R-0678 and 9R-0345 exhibited augmented sensitivity to GR501 strain harboring MtbLigA (**Fig. 6.4 A and B**). These results certify that 1R-0678 and 9R-0345 has higher particularity to NAD^+-dependent ligases and suggests that observed antibacterial activities are due to *in vivo* inhibition of LigA (**Fig. 6.4 C**).

Table 6.2 Antibacterial activity of fragments

Fragments	MIC (µg/ml)		
	E. coli GR501+ *pTrc99A*	*E. coli* GR501+ Mtb NAD^+ LigA	*E. coli* GR501+ T4 DNA LigA
1R-0678	0.4	20	113
9R-0345	0.6	60	132
FS-3025	0.6	>250	180

*The MIC values are given in µg/ml. The strains used in the study are explained in the text.

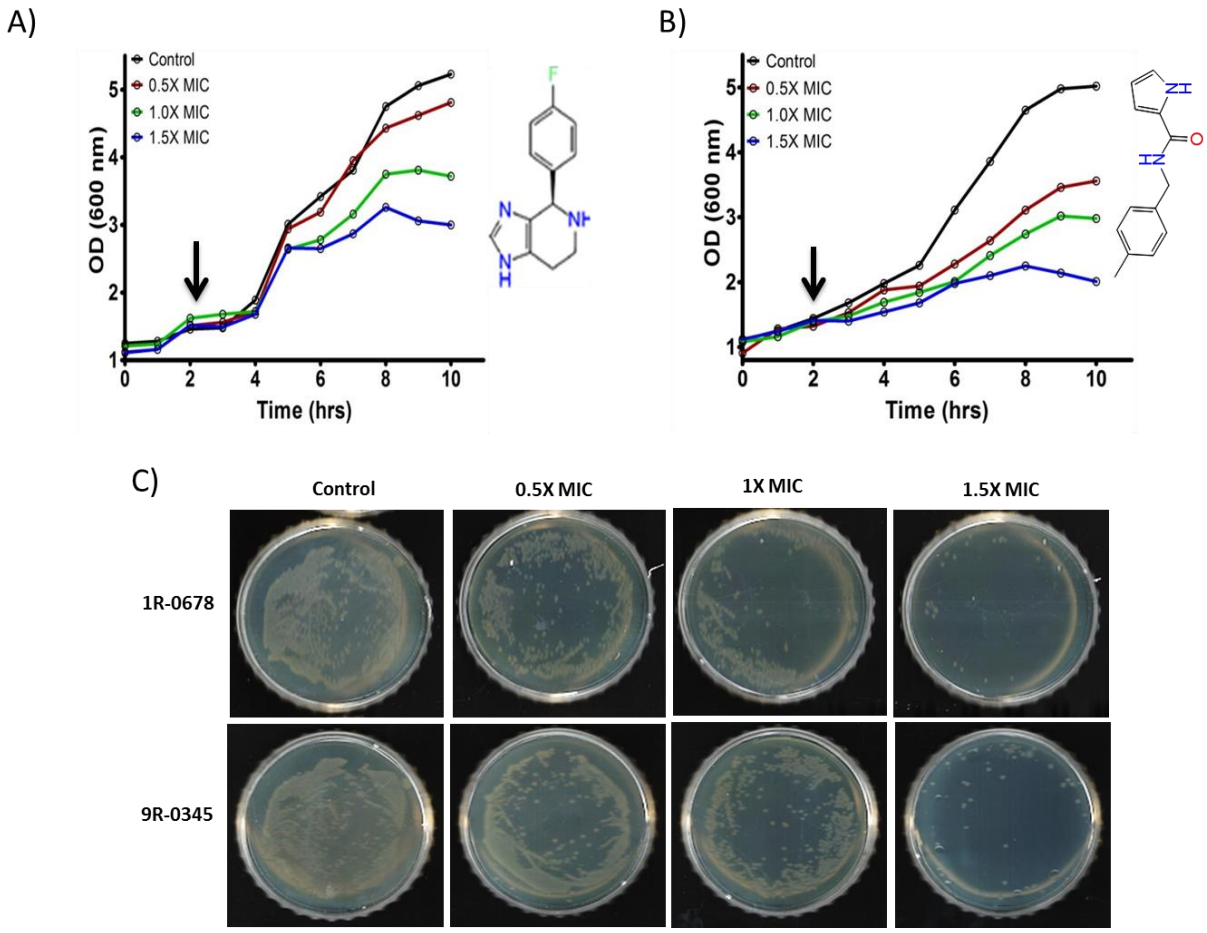


Figure 6.4 Antibacterial activity of fragments. A and B) Bactericidal activity of fragment 1R-0678 and 9R-0345 respectively. Effect on growth as reflected in changes in $O.D_{600}$ nm *E. coli* GR501 on exposure to fragments in µg/ml representing 0.5 to 1.5 times the MIC values. The arrow indicates the point at which fragment was added. The structure of respective fragment is shown beside the graphical representation of the experiment. **C)** Comparative viability as shown by surviving CFU. The cells were plated at dilution ratio of 10^{-3} on LB agar plate. The control and amount of added fragment as multiples of MIC are also indicated.

6.2.4 Thermodynamic analysis of fragment binding with AdD domain of MtbLigA

We further continued to ascertain the thermodynamic properties of the selected fragments. The 1R-0678 and 9R-0345 scaffold binds well in the NMN binding pocket with the maximum favorable hydrogen bonding stabilization interaction with 26.4 µM and 20.4 µM KD value, respectively (**Fig. 6.5 A I and II, B I and II**) while FS-3025 shows less affinity 1037.74 µM KD value (**Fig. 6.5 C I and II**). The thermodynamics parameters of fragments binding with AdD domain is tabulated in table 6.3.The enthalpy of reaction (ΔH), binding constant (K) and the stoichiometry value (n) were calculated from measured heat changes ΔH

on coalition of inhibitor with protein (**Fig. 6.5 D**). The control experiments for each fragment was performed and subtracted from the corresponding data sets. ITC data was evaluated using Origin software 7.0 associated with VP-ITC software.

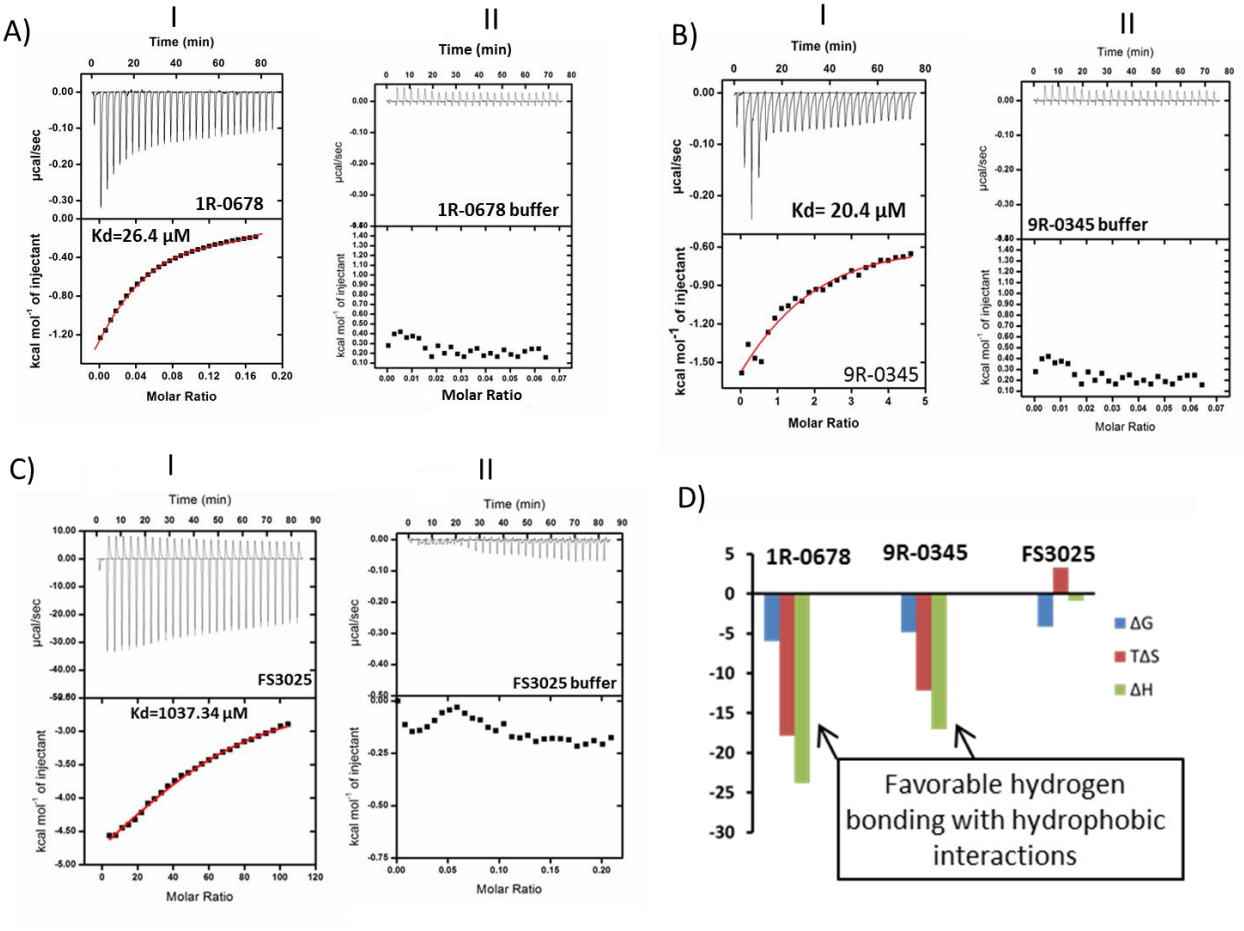


Figure 6.5 Thermodynamic study of interaction between fragments and AdD domain of MtbLigA. A) I, Thermodynamic study of interaction between AdD domain with fragments 1R-0678 by Isothermal titration calorimetery gives the Kd value of 26.4 µM. **II,** Control experiment of 1R-0678 titration with buffer. **B) I,** ITC of fragment 9R-0345 and AdD domain gives the Kd value of 20.4 µM. **II,** Control experiment of 9R-0345 titration with buffer. **C)** Binding affinity of FS3025 with AdD is 1037.34 µM **II,** and its control with buffer. **D)** Thermodynamic parameter comparison of the three fragments.

Table 6.3 ITC data and thermodynamics of fragment binding with adenylation domain

Ligand/concentration (mM)	Adenyaltion domain (µM)	N	KD (µM)	ΔG (kcal mol^{-1})	ΔH (kcal mol^{-1})	-TΔS(kcal mol^{-1})
1R-0678 (1 mM)	10	0.968	202.4	-4.83	-17	-12.16
9R-0345 (1 mM)	10	0.845	26.4	-5.96	-23.8	-17.83
FS-3025 (1 mM)	10	0.830	1037.34	-4.13	-0.860	3.2

N= Stoichiometry, KD= Dissociation constant, ΔG= Gibbs free energy, ΔH= Binding Enthalpy, ΔS= Entropy, T= temperature (K)

6.2.5 Co-crystal structure of 9R-0345 and fragment with AdD domain of MtbLigA

After the confirmation of inhibitory activity of these fragments, we tried co-crystallizing them with adenylation domain of MtbLigA and 9R-0345 was successfully co-crystallized. We got poor diffracting co-crystals in case of 1R-0678. The 9R-0345 co-crystal yielded clear electron density for the fragment in the NMN binding site. Y31 and F44 are involved in base stacking interaction with benzene ring of the fragment (PDB: **6KKV**). D45 is not directly involved in the interaction with NMN in contrast to that observed in case of other NMN bound crystal structure (PDB: **6KSC**). Instead backbone of D41 and I39 forms hydrogen bonding with C12 and C16 of benzene ring of fragment. Other residues involved in non-bonded interactions are S40, P37 and H27 (**Fig. 6.6 A and Fig.6.7**). The Ramachandran plot depicts the 94.67 % residues in the preferred region while 5.33 % in the allowed region while no residue in the outliers (**Fig. 6.6 B**)

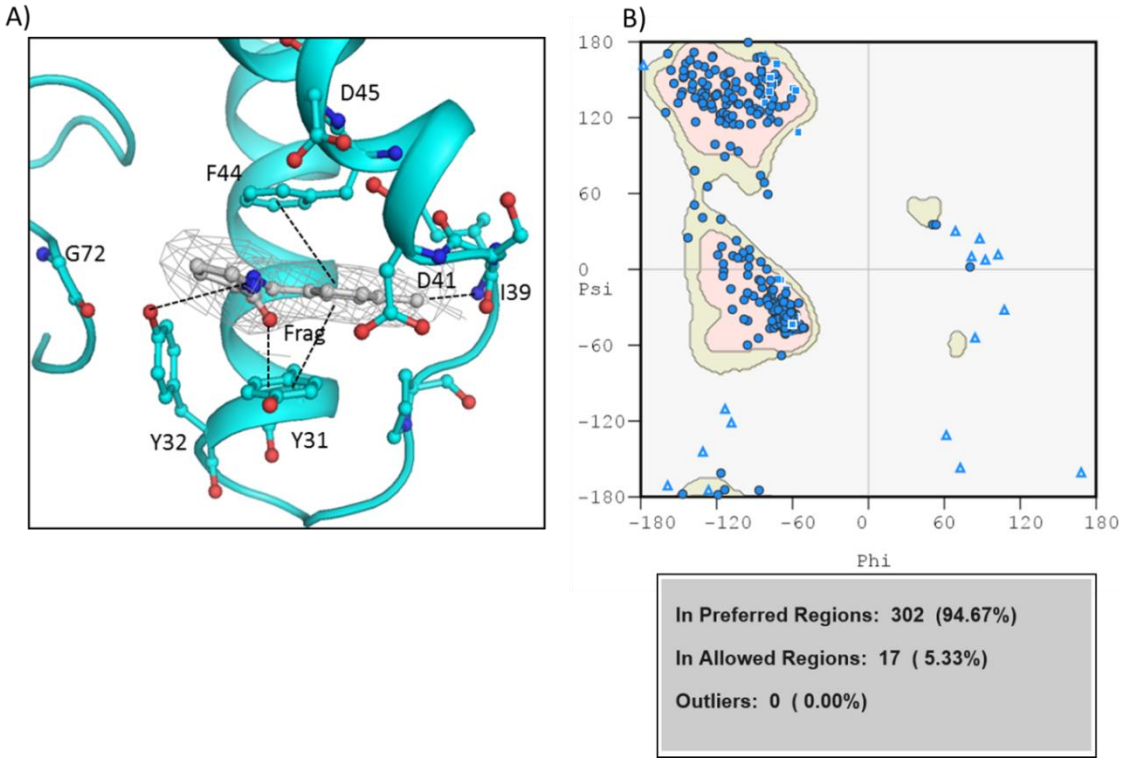


Figure 6.6 Electron density and Ramachandran plot of fragment. A) The 9R-0345 bound co-crystal structure of AdD domain shows the binding of fragment in between the two helices of Ia subdomain. **B)** The Ramachandran plot generated with SAVES.

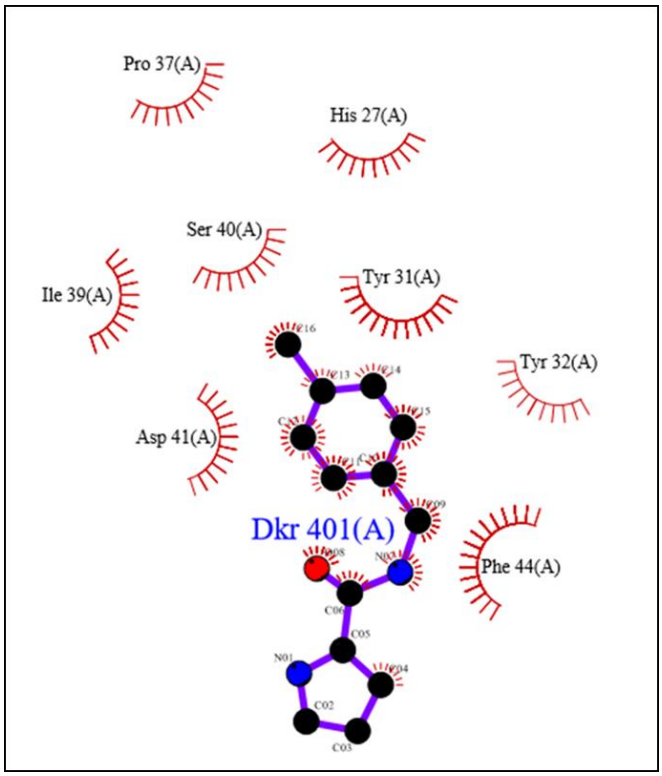


Figure 6.7 The LIGPLOT analysis - the ligand bound in the NMN binding pocket between the two helices of Ia subdomain of AdD domain of LigA.

Table 6.4 Data collection statistics

Data collection	**6KKV- AdD domain bound with 9R0345 fragment**
Space group	P61 2 2
Cell dimension	a=95.40,b=95.40, c= 201.28 α= 90, β= 90, γ= 119.99
Wavelength (nm)	0.9795
Resolution range (Å)	43.10-2.56
Completeness (%)	100
Multiplicity	11.5
R_{merge}	0.134
I/σI	17.7
Refinement	
Reflections	18196
Total no. of atoms-	
Protein	4870
Ligand	9R-0345- 28
sulfate	9
water	31
R_{work}	0.2340
R_{free}	0.2780
R.m.s.d **Bond length (Å)** **Bond Angle (◦)**	 0.003 0.666
Ramachandran **Favoured** **Allowed** **Outlier**	 94.67 5.33 0.00
Mean B value	51.38

6.3 Discussion

The current need of anti-tuberculosis drugs impulses us to identify novel scaffolds. The present crystal structures of AdD domain of MtbLigA are subjected to screen some natural fragment scaffolds. The 9R-0345 occupies the whole NMN binding active site as it mimics the shape of NMN while 1R-0678 shows similarity with the nicotinamide ring of NMN. The FS-3025 scaffold is too short but it still shows *in vitro* inhibition of ligation activity by binding deep in the pocket, although fails to show any bactericidal activity. We have also determined the thermodynamic behavior of binding of these scaffolds with the adenylation domain. 9R and 1R are stabilized by favorable hydrogen and hydrophobic interactions while the FS scaffold binding is mainly dominated by hydrogen bonding. The crystal structure of adenylation domain with 9R fragment opens up the new ways of targeting ligase by focusing on NMN binding site. In conclusion, we have presented the crystal structures including MtbLigA1 (AdD and OB domain), adenylation domain and other Ia/Ib domain mutants bound with their cofactor which defines the missing understanding of step1 of the ligation reaction. Mutational and fragment screening analysis against MtbLigA may be taken forward for the drug improvement with high degree of fidelity.

Chapter 7 | Identification of Small Molecule Fragments as Inhibitor of MtbXthA

7.1 Introduction

Several reports suggest the stamping out of bacterial cells when exposed to reactive oxygen species. The reactive oxygen intermediates (ROIs) such as superoxide (O_2), hydrogen peroxide (H_2O_2) and hydroxyl radicals (OH) directly damage the DNA, proteins and lipids thereby subsequently killing the organism (Farr and Kogoma 1991). There exist two line of defense against the damage caused by ROIs in bacterial cells. The first line of defense includes the enzymes catalase and superoxide dismutase, which directly detoxicate these radicals. The second line of defense includes the repair mechanism that repairs the damage caused by these ROIs and also cleanses the already damaged components by cleaving or degrading them further so as to avoid the cell death. The DNA repair enzymes actively participate in this defense mechanism (Davies and Lin 1988, Hornback and Roop 2006). In *Mycobacterium tuberculosis,* base excision repair (BER) pathway plays the major role in the repair and removal of these oxidative lesions in DNA (Haring, Rudiger et al. 1994). DNA glycosylases recognizes the lesion and removes the nitrogenous base leaving the apurinic / apyrimidic (AP) site. Then comes the endonuclease (XthA) that cuts the phosphate backbone at 5'-PO4 and leaves the free 3'-OH on the adjacent nucleotide for action of the following enzymes viz. DNA polymerase I and DNA ligase. Thus this how the lesion get repaired in the nutshell. The XthA accounts for 90% of endonuclease activity in the bacteria (Khanam, Shukla et al. 2015). The rest of the 10% of the endonuclease activity is performed by endonuclease IV, which is encoded by nfo gene (Moeller, Setlow et al. 2011). Earlier reports in *E. coli* had already hinted towards the importance of this pathway as the knockout mutant of XthA and double knockout of XthA and nfo gene leads to the hypersensitivity of the strain *in vitro* under the oxidative stress (Imlay and Linn 1987, Hagensee and Moses 1989). It has also been manifested that BER plays a predominant role in protecting *Salmonella enterica* and *Serovar typhimurium* from oxidative killing by murine macrophages (Richardson, Soliven et al. 2009). To determine if the xthA gene product plays a role in oxidative defense in *M. tuberculosis*, an XthA knockout in *E. coli* and the similar knockout complemented with MtbXthA were exposed to H_2O_2 at final concentrations of 1, 2.5, 5, 7.5 and 10 mM. The knockout (JW1738) was much more susceptible to killing by H_2O_2 under these conditions than the MtbXthA rescued strain (JW1738+MtbXthA). The introduction of MtbXthA restored the resistance of this strain to H_2O_2. Further to get, the depth of this resistance by MtbXthA complemented strain; the effect of oxidative stress was checked on the MtbXthA active site mutants, which also displayed the sensitivity just like the knockout one. Moreover,

we also screened *in silico*, *in vitro* and *in vivo* effect of few compounds that targets the endonuclease activity of MtbXthA. We got few with IC_{50} in micro-molar range, and further optimization might give more convincing results to be treated as more potent antibacterial.

7.2 Results

7.2.1 Complementation assay- MtbXthA confers resistance from hydrogen peroxide induced effects

The aerobic metabolism generates various reactive oxygen species as their byproduct. These byproducts attack the DNA resulting in various types of lesions including AP sites. Since AP sites are highly genotoxic for cells therefore must be removed through the dedicated BER pathway. The pathway requires the AP endonuclease to clean up these AP sites. *M. tb* possesses conserved AP endonuclease viz. MtbXthA which perform 90% of the endonuclease activity (Khanam, Shukla et al. 2015). Therefore, we determined whether MtbXthA is responsible for conferring protection against oxidative stress to *E. coli* cells deficient in XthA. We found that the exclusive absence of MtbXthA increased the susceptibility of *E. coli* XthA knockout to hydrogen peroxide (**Fig. 7.1**). We complemented MtbXthA in to *E. coli* XthA knockout (JW1738) and exposed the cells to the various concentration of H_2O_2. To determine the lethal dose, the complemented and knockout strains were exposed to 0, 1, 2.5, 5, 7.5 and 10 mM of H_2O_2. In addition, we found the LD_{50} to be 5 mM of H_2O_2 (**Fig. 7.2 A**). Then we also checked the survival percentage that clearly showed the effect of oxidative stress on both strains. The values are the mean ± SD (standard deviation) of three independent experiments in duplicate. Together, these results support the hypothesis that in the absence of endonuclease (ΔXthA) which plays an important role in counteracting the genotoxic effects of H_2O_2 in growing cells (**Fig. 7.2 B**).

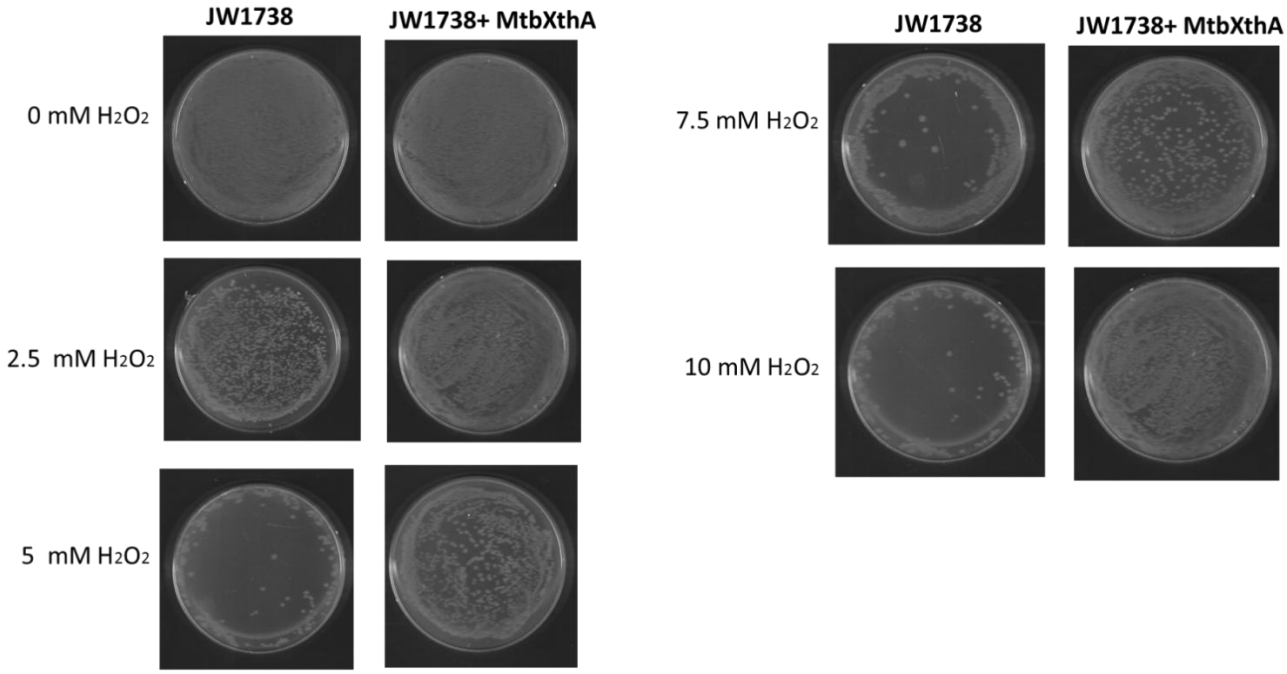


Figure 7.1 Complementation assay in the *E. coli* XthA knockout strain (JW1738). The JW1738+ MtbXthA is the knockout strain rescued with MtbXthA complementation. To the sensitivity towards oxidative stress both strains were subjected to increasing concentration of H_2O_2 (0, 2.5, 5, 7.5 and 10 mM). The cells after treatment ($O.D_{600}$ of 0.8) were serially diluted to 10^{-2} and plated on LB agar plate supplemented with kanamycin and ampicillin antibiotic.

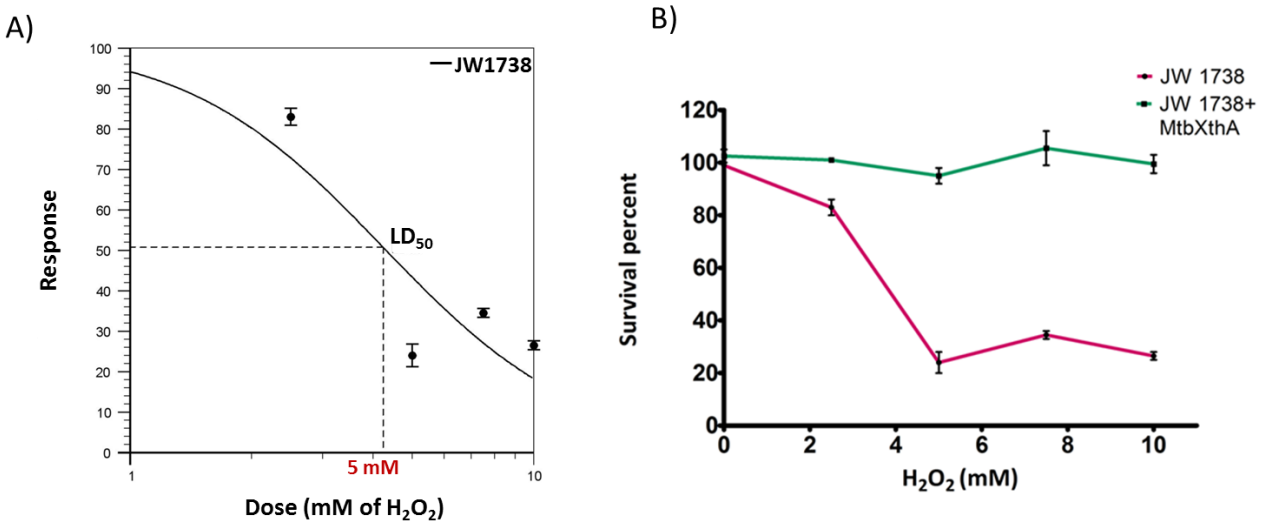


Figure 7.2 Lethal dose and survival percent determination from the complementation assay under stress. A) Survival curve of the knockout strain (JW1738) when subjected to increasing concentrations of H_2O_2 (0–10 mM) for 30 min. The lethal dose (LD_{50}) is 5mM of H_2O_2. The values are the mean ± SD of three independent experiments in duplicate. The statistical analysis is done by two-way ANOVA test, where *P < 0.05. **B)** Complementation assay in the *E. coli* ΔxthA strain (JW1738). A survival increase in plates with H_2O_2 was observed in the presence of MtbXthA in the *E. coli* knockout strain. Undiluted cultures and serial dilutions (10−1, 10−2, 10−3 and 10−4) from cultures (OD_{600} of 0.8) were plated on LB agar containing different concentrations of H_2O_2.

7.2.2 Complementation of MtbXthA catalytic mutant (E67A/D251A and Y137S)-

To further check the role of catalytic residues of MtbXthA in the repair of oxidative lesions, we complemented active site mutants (E67A/D251A and Y137S) of XthA in to the knockout strain (ΔXthA) and treated them at zero concentration of H_2O_2 and at lethal dose of H_2O_2 i.e 5 mM. All three mutants showed sensitivity equivalent to that of the knockout strain. Thus this complementation assay further confirms that endonuclease activity is essential for the repair of oxidative lesions (**Fig. 7.3 A and B**).

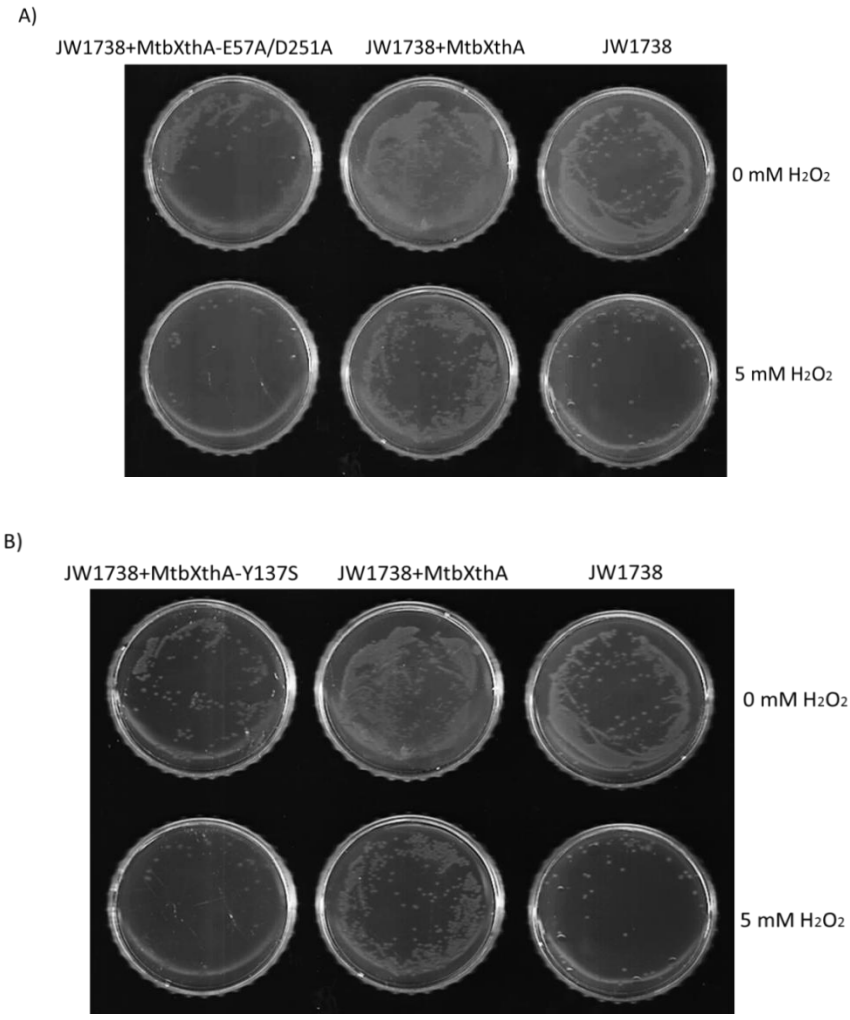
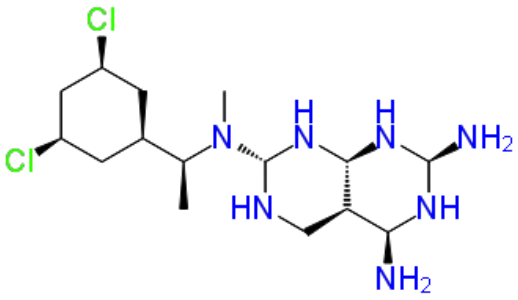
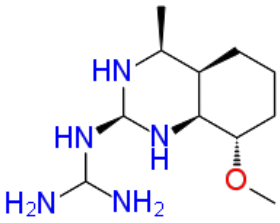
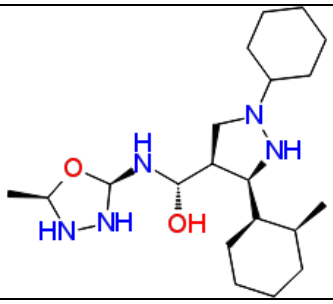
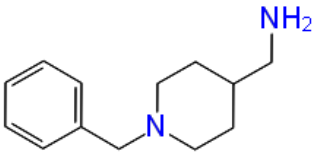
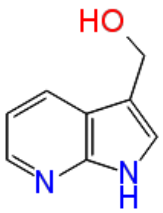
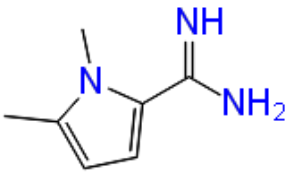


Figure 7.3 Complementation assay in the *E. coli* XthA knockout strain (JW1738) with active site mutants of MtbXthA. A) JW1738 was complemented with XthA double mutant E57A/D251A and exposed to the lethal dose of 5mM H_2O_2. **B)** JW1738 complemented with Y137S mutant.

7.2.3 *In silico* screening of MtbXthA against the active site residues-

After checking the role of active site residues of MtbXthA, we further did *in silico* screening of the MMV and BioNet library against the active site of MtbXthA (**Table 7.1**). The XthA model was prepared by Phyre2 and structure was minimized by using Tripos force field with 500 steps of steepest descent algorithm followed 500 steps of Powell method and 50.00 kcal/mole convergence threshold energy (custom parameters). The flexible docking was performed by Surflex-Dock Programme available in SYBYL-X v2.1. The chosen protein active sites selected to prepare grid using Protomol module by defining around active site region. The second Generation BIONET premium fragment library (1166 compounds) purchased from Key Organics Ltd. (https://www.keyorganics.net/) and MMV pathogen box library was screened against defined active site. The each active hit was selected by generating 20 maximum poses and 20 maximum conformations per fragment of protein. The final pose of hits with highest docking score were selected which calculated using scoring functions i.e. G-Score, D-Score, PMF-Score and Chem-score. All visualization analysis was done by PyMol v2.2 visualization software (**Fig. 7.1 A-H**)

Table 7.1 *In silico* screening of fragments from BioNet library and from MMV pathogen box

S. No	Scaffold ID	Scaffold Structure	Docking Score	IC_{50}(μM)
1.	MMV689437_000		6.862	26.53
2.	MMV000907_000		6.072	38.92
3.	MMV020320_000		6.724	19.07
4.	FS-2025_000		4.32	50.20
5.	GE-0726_000		4.56	29.52
6.	SS-4078_000		4.38	44.19

*S. No., serial number; Binding energies for respective compounds were calculated by SYBYL-X v2.1.

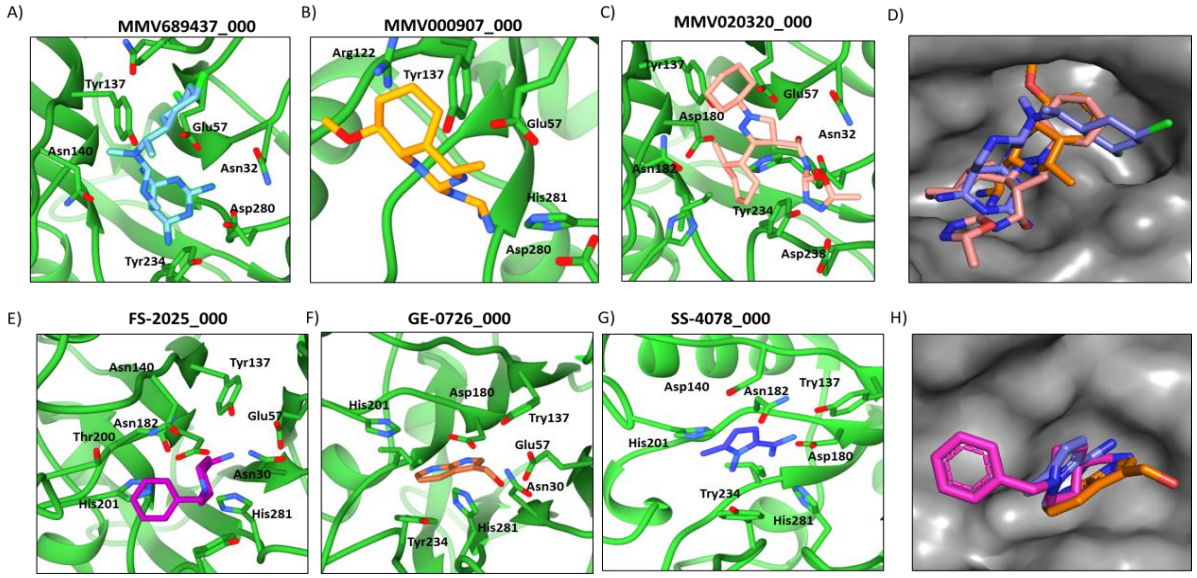


Figure 7.1 Docked confirmations of the inhibitors. The inhibitors are shown in sticks **A)** MMV689437_000 (cyan), **B)** MMV000907 (orange), **C)** MMV023020 (salmon), **E)** FS-2025 (magenta), **F)** GE-0726 (fire brick) and **G)** SS-4078 (blue) at the binding cleft of XthA. **D and H)** Surface representation of XthA showing MMV and fragment inhibitors docked into the binding cleft of protein.

7.2.4 *In vitro* screening of the scaffolds against MtbXthA endonuclease activity

After *in silico* screening, to quickly identify the potentially useful scaffolds, we carried out screening of the selected scaffolds for its *in vitro* inhibitory potency against the endonuclease activity of MtbXthA (**Fig. 7.2 A I, II, III and B I, II, III**). The scaffold MMV020320_000 showed the IC_{50} of 19.06 µM among the other scaffold thus indicating towards the good inhibitory potency. The rest scaffolds showed IC_{50} as 26.53 µM, 38.92 µM, 50.20 µM, 29.52 µM and 44.19 µM for MMV689437_000, MMV000907_000, FS-2025_000, GE-0726_000 and SS-4078_000 respectively (**Fig. 7.3 A and B**) The inhibitor scaffolds concentration used in the experiment was 0, 0.1, 0.2, 0.3, 0.6, 0.8, 1.0, 1.2 and 1.4 µM and XthA concentration was 1.5 µM. The graph pad prism 5.0 was used to calculate the IC_{50} of the respective scaffold.

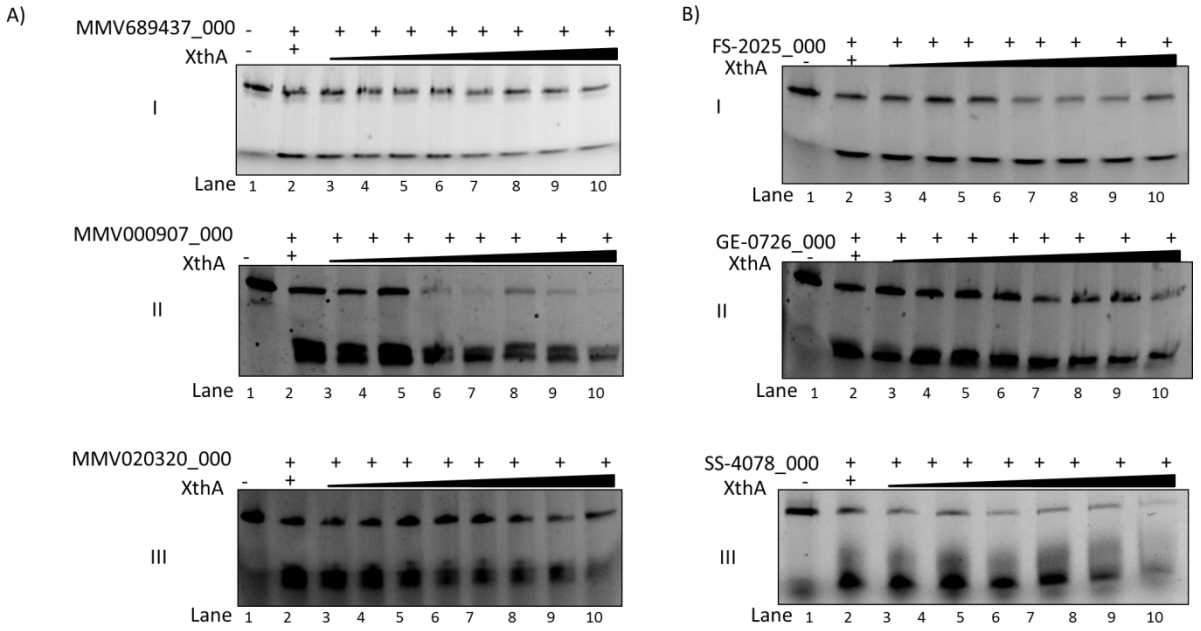


Figure 7.2 *In vitro* inhibition of endonuclease activity of XthA in presence of screened scaffolds from A) MMV pathogen box MMV689437_000, MMV000907_000 and MMV020320_000 1R-0678, 9R-0345 and FS-3025. **B)** From Bionet library of fragments, FS_2025_000, GE-0726_000 and SS-4078_000. Lane 1 is control devoid of enzyme and lane 2 is positive control with WT MtbXthA without any inhibitor. All the inhibitors were taken in increasing concentration range of 0.1-1.4 µM.

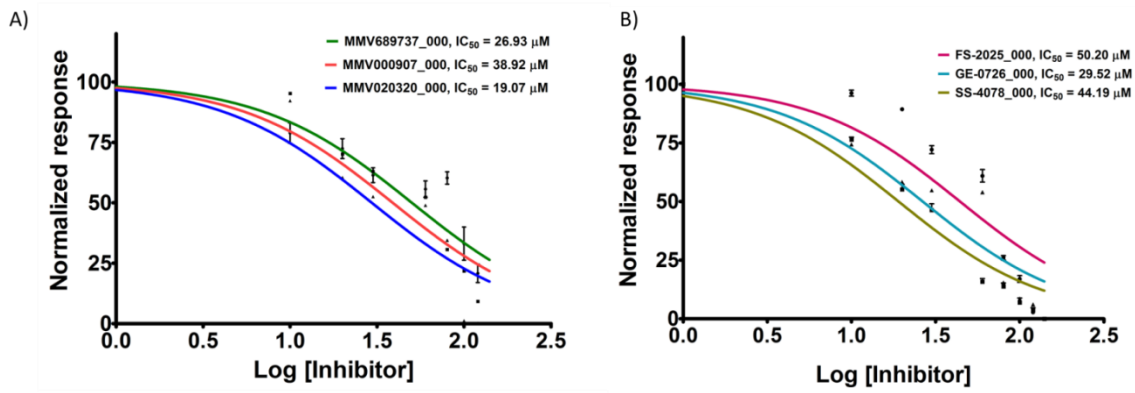


Figure 7.3 The gels were quantified to calculate IC_{50} of each inhibitor. **A)** MMV020320_000 (blue) shows the least inhibitory concentration of 19.07 µM and MMV689737_000 (green) and MMV000907_000 (red) shows the IC_{50} of 26.93 µM and 38.92 µM, respectively. **B)** GE-0726_000 (cyan) shows the least IC_{50} of 29.52 µM among the fragments from Bionet while FS-2025_000 and SS-4078_000 shows the IC_{50} of 50.20 µM and 44.19 µM, respectively. Error bars corresponds to standard deviation from two independent experiments.

7.2.5 *In vivo* screening of the inhibitor scaffold

Further these scaffolds at concentration of 1 µM were tested *in vivo* on MtbXthA complemented Knockout strain at 5 mM concentration of H_2O_2. The colonies were observed after 16 hrs of the treatment followed by plating on LB agar plates with respective antibiotic (**Fig. 7.4 I-VIII**). The *in vivo* results were quite surprising as compared to that of *in vitro* screening results. *In vivo* results suggest that may be apart from endonuclease some other factors like *in vivo* protein interactions are also taking count during the stress. The further analysis of these scaffolds on non-pathogenic mycobacterial strain is under study and their results might open a new avenue for developing the therapeutics. The initial screening showed that MMV020320_000 is the most lethal followed by MMV68947_000 in the MMV library while GE-0726_000 showed the maximum lethality among the BioNet library compounds.

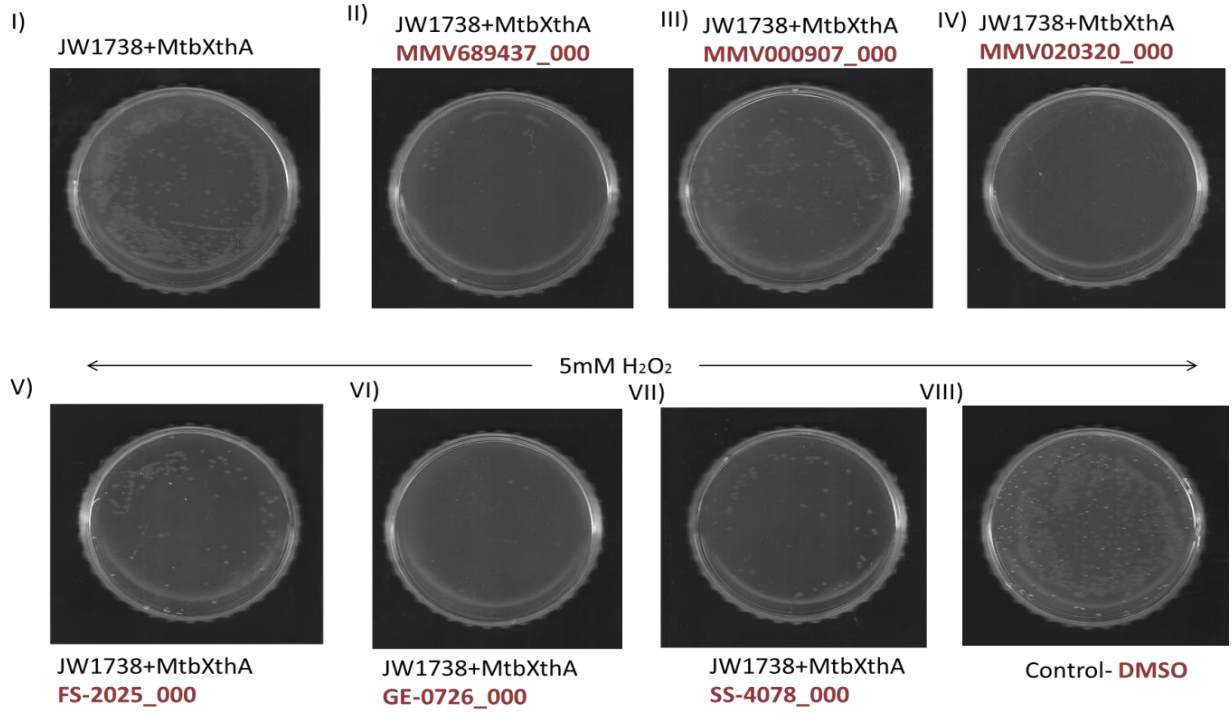


Figure 7.4 *In vivo* screening of the inhibitor scaffolds. I) JW1738+MtbXthA. The complemented strain in presence of inhibitor scaffold **II)** MMV689737_000 **III)** MMV000907_000 **IV)** MMV020320_000 **V)** FS-2025_000 **VI)** GE-0726_000 and **VII)** SS-4078_000. All the fragments were tested at 1 µM concentration in the presence of 5mM of H_2O_2.

7.3 Discussion

The *in vitro* and *in vivo* results obtained in this study suggests that the product of the xthA gene is an genuine AP endonuclease that takes part in BER in *Mycobacterium tuberculosis*. Moreover, like its *E. coli* counterpart, XthA-1 appears to play a crucial function in protecting this bacterium from oxidative impairment to its DNA. The capability of the *Mycobacterium tuberculosis* to resist the oxidative shooting down pathways of host macrophages has been suggested to be of critical importance in their power to create and cultivate chronic infections in their mammalian hosts. Although it was observed in laboratory, assays that the corresponding gene product is not required for wild-type virulence. The results here plausibly explain the essentiality of XthA gene for maintain the genomic integrity. We also screened few scaffolds which show the *in vitro* inhibition of MtbXthA in micro molar range as well as good visible inhibition in in vivo. The further optimization and using non-pathogenic mycobacterium strain (H37Ra) will help to resolve the need of more advanced therapeutics.

References

Abeldenov, S., I. Talhaoui, D. O. Zharkov, A. A. Ishchenko, E. Ramanculov, M. Saparbaev and B. Khassenov (2015). "Characterization of DNA substrate specificities of apurinic/apyrimidinic endonucleases from Mycobacterium tuberculosis." DNA Repair (Amst) 33: 1-16.

Abeldenov, S., I. Talhaoui, D. O. Zharkov, A. A. Ishchenko, E. Ramanculov, M. Saparbaev and B. Khassenov (2015). "Characterization of DNA substrate specificities of apurinic/apyrimidinic endonucleases from Mycobacterium tuberculosis." DNA Repair (Amst) 33: 1-16.

Almeida, M. C., S. F. Soares, P. R. Abreu, L. M. Jesus, L. C. Brito and M. Bernardo-Filho (2007). "Protective effect of an aqueous extract of Harpagophytum procumbens upon Escherichia coli strains submitted to the lethal action of stannous chloride." Cell Mol Biol (Noisy-le-grand) 53 Suppl: OL923-927.

Almeida, M. C., S. F. Soares, P. R. Abreu, L. M. Jesus, L. C. Brito and M. Bernardo-Filho (2007). "Protective effect of an aqueous extract of Harpagophytum procumbens upon Escherichia coli strains submitted to the lethal action of stannous chloride." Cell Mol Biol (Noisy-le-grand) 53 Suppl: OL923-927.

Aravind, L., D. R. Walker and E. V. Koonin (1999). "Conserved domains in DNA repair proteins and evolution of repair systems." Nucleic Acids Res 27(5): 1223-1242.

Arrigo, C. J., K. Singh and M. J. Modak (2002). "DNA polymerase I of Mycobacterium tuberculosis: functional role of a conserved aspartate in the hinge joining the M and N helices." J Biol Chem 277(3): 1653-1661.

Arrigo, C. J., K. Singh and M. J. Modak (2002). "DNA polymerase I of Mycobacterium tuberculosis: functional role of a conserved aspartate in the hinge joining the M and N helices." J Biol Chem 277(3): 1653-1661.

Barbin, A., R. Wang, P. J. O'Connor and R. H. Elder (2003). "Increased formation and persistence of 1,N(6)-ethenoadenine in DNA is not associated with higher susceptibility to carcinogenesis in alkylpurine-DNA-N-glycosylase knockout mice treated with vinyl carbamate." Cancer Res 63(22): 7699-7703.

Barbin, A., R. Wang, P. J. O'Connor and R. H. Elder (2003). "Increased formation and persistence of 1,N(6)-ethenoadenine in DNA is not associated with higher susceptibility to carcinogenesis in alkylpurine-DNA-N-glycosylase knockout mice treated with vinyl carbamate." Cancer Res 63(22): 7699-7703.

Barnes, D. E. and T. Lindahl (2004). "Repair and genetic consequences of endogenous DNA base damage in mammalian cells." Annu Rev Genet 38: 445-476.

Barnes, D. E. and T. Lindahl (2004). "Repair and genetic consequences of endogenous DNA base damage in mammalian cells." Annu Rev Genet 38: 445-476.

Beernink, P. T., B. W. Segelke, M. Z. Hadi, J. P. Erzberger, D. M. Wilson, 3rd and B. Rupp (2001). "Two divalent metal ions in the active site of a new crystal form of human

apurinic/apyrimidinic endonuclease, Ape1: implications for the catalytic mechanism." J Mol Biol 307(4): 1023-1034.

Benarroch, D. and S. Shuman (2006). "Characterization of mimivirus NAD+-dependent DNA ligase." Virology 353(1): 133-143.

Benarroch, D. and S. Shuman (2006). "Characterization of mimivirus NAD+-dependent DNA ligase." Virology 353(1): 133-143.

Bennett, R. A., D. M. Wilson, 3rd, D. Wong and B. Demple (1997). "Interaction of human apurinic endonuclease and DNA polymerase beta in the base excision repair pathway." Proc Natl Acad Sci U S A 94(14): 7166-7169.

Bogenhagen, D. F. and K. G. Pinz (1998). "The action of DNA ligase at abasic sites in DNA." J Biol Chem 273(14): 7888-7893.

Boiteux, S. and J. P. Radicella (1999). "Base excision repair of 8-hydroxyguanine protects DNA from endogenous oxidative stress." Biochimie 81(1-2): 59-67.

Boiteux, S. and J. P. Radicella (1999). "Base excision repair of 8-hydroxyguanine protects DNA from endogenous oxidative stress." Biochimie 81(1-2): 59-67.

Boiteux, S. and M. Guillet (2004). "Abasic sites in DNA: repair and biological consequences in Saccharomyces cerevisiae." DNA Repair (Amst) 3(1): 1-12.

Borjigin, M., P. Arenaz and B. Stec (2012). "Chinese hamster AP endonuclease operates by a two-metal ion assisted catalytic mechanism." FEBS Lett 586(3): 242-247.

Borjigin, M., P. Arenaz and B. Stec (2012). "Chinese hamster AP endonuclease operates by a two-metal ion assisted catalytic mechanism." FEBS Lett 586(3): 242-247.

Bork, P., K. Hofmann, P. Bucher, A. F. Neuwald, S. F. Altschul and E. V. Koonin (1997). "A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins." FASEB J 11(1): 68-76.

Bork, P., K. Hofmann, P. Bucher, A. F. Neuwald, S. F. Altschul and E. V. Koonin (1997). "A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins." FASEB J 11(1): 68-76.

Bowater, R. and A. J. Doherty (2006). "Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining." PLoS Genet 2(2): e8.

Bowater, R. and A. J. Doherty (2006). "Making ends meet: repairing breaks in bacterial DNA by non-homologous end-joining." PLoS Genet 2(2): e8.

Brotz-Oesterhelt, H., D. Beyer, H. P. Kroll, R. Endermann, C. Ladel, W. Schroeder, B. Hinzen, S. Raddatz, H. Paulsen, K. Henninger, J. E. Bandow, H. G. Sahl and H. Labischinski (2005). "Dysregulation of bacterial proteolytic machinery by a new class of antibiotics." Nat Med 11(10): 1082-1087.

Brotz-Oesterhelt, H., D. Beyer, H. P. Kroll, R. Endermann, C. Ladel, W. Schroeder, B. Hinzen, S. Raddatz, H. Paulsen, K. Henninger, J. E. Bandow, H. G. Sahl and H. Labischinski (2005). "Dysregulation of bacterial proteolytic machinery by a new class of antibiotics." Nat Med 11(10): 1082-1087.

Brotz-Oesterhelt, H., I. Knezevic, S. Bartel, T. Lampe, U. Warnecke-Eberz, K. Ziegelbauer, D. Habich and H. Labischinski (2003). "Specific and potent inhibition of NAD+-dependent DNA ligase by pyridochromanones." J Biol Chem 278(41): 39435-39442.

Brotz-Oesterhelt, H., I. Knezevic, S. Bartel, T. Lampe, U. Warnecke-Eberz, K. Ziegelbauer, D. Habich and H. Labischinski (2003). "Specific and potent inhibition of NAD+-dependent DNA ligase by pyridochromanones." J Biol Chem 278(41): 39435-39442.

Brotz-Oesterhelt, H., I. Knezevic, S. Bartel, T. Lampe, U. Warnecke-Eberz, K. Ziegelbauer, D. Habich and H. Labischinski (2003). "Specific and potent inhibition of NAD+-dependent DNA ligase by pyridochromanones." J Biol Chem 278(41): 39435-39442.

Caglayan, M. and S. H. Wilson (2015). "Oxidant and environmental toxicant-induced effects compromise DNA ligation during base excision DNA repair." DNA Repair (Amst) 35: 85-89.

Caglayan, M. and S. H. Wilson (2015). "Oxidant and environmental toxicant-induced effects compromise DNA ligation during base excision DNA repair." DNA Repair (Amst) 35: 85-89.

Caldecott, K. W., C. K. McKeown, J. D. Tucker, S. Ljungquist and L. H. Thompson (1994). "An interaction between the mammalian DNA repair protein XRCC1 and DNA ligase III." Mol Cell Biol 14(1): 68-76.

Caldecott, K. W., C. K. McKeown, J. D. Tucker, S. Ljungquist and L. H. Thompson (1994). "An interaction between the mammalian DNA repair protein XRCC1 and DNA ligase III." Mol Cell Biol 14(1): 68-76.

Callebaut, I. and J. P. Mornon (1997). "From BRCA1 to RAP1: a widespread BRCT module closely associated with DNA repair." FEBS Lett 400(1): 25-30.

Callebaut, I. and J. P. Mornon (1997). "From BRCA1 to RAP1: a widespread BRCT module closely associated with DNA repair." FEBS Lett 400(1): 25-30.

Carpenter, E. P., A. Corbett, H. Thomson, J. Adacha, K. Jensen, J. Bergeron, I. Kasampalidis, R. Exley, M. Winterbotham, C. Tang, G. S. Baldwin and P. Freemont (2007). "AP endonuclease paralogues with distinct activities in DNA repair and bacterial pathogenesis." EMBO J 26(5): 1363-1372.

Carpenter, E. P., A. Corbett, H. Thomson, J. Adacha, K. Jensen, J. Bergeron, I. Kasampalidis, R. Exley, M. Winterbotham, C. Tang, G. S. Baldwin and P. Freemont (2007). "AP endonuclease paralogues with distinct activities in DNA repair and bacterial pathogenesis." EMBO J 26(5): 1363-1372.

Chatterjee, N. and G. C. Walker (2017). "Mechanisms of DNA damage, repair, and mutagenesis." Environ Mol Mutagen 58(5): 235-263.

Chatterjee, N. and G. C. Walker (2017). "Mechanisms of DNA damage, repair, and mutagenesis." Environ Mol Mutagen 58(5): 235-263.

Chayot, R., B. Montagne, D. Mazel and M. Ricchetti (2010). "An end-joining repair mechanism in Escherichia coli." Proc Natl Acad Sci U S A 107(5): 2141-2146.

Chayot, R., B. Montagne, D. Mazel and M. Ricchetti (2010). "An end-joining repair mechanism in Escherichia coli." Proc Natl Acad Sci U S A 107(5): 2141-2146.

Cho, E. A., M. J. Prindle and G. R. Dressler (2003). "BRCT domain-containing protein PTIP is essential for progression through mitosis." Mol Cell Biol 23(5): 1666-1673.

Cho, E. A., M. J. Prindle and G. R. Dressler (2003). "BRCT domain-containing protein PTIP is essential for progression through mitosis." Mol Cell Biol 23(5): 1666-1673.

Ciarrocchi, G., D. G. MacPhee, L. W. Deady and L. Tilley (1999). "Specific inhibition of the eubacterial DNA ligase by arylamino compounds." Antimicrob Agents Chemother 43(11): 2766-2772.

Ciarrocchi, G., D. G. MacPhee, L. W. Deady and L. Tilley (1999). "Specific inhibition of the eubacterial DNA ligase by arylamino compounds." Antimicrob Agents Chemother 43(11): 2766-2772.

Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M. A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead and B. G. Barrell (1998). "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence." Nature 393(6685): 537-544.

Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, A. Krogh, J. McLean, S. Moule, L. Murphy, K. Oliver, J. Osborne, M. A. Quail, M. A. Rajandream, J. Rogers, S. Rutter, K. Seeger, J. Skelton, R. Squares, S. Squares, J. E. Sulston, K. Taylor, S. Whitehead and B. G. Barrell (1998). "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence." Nature 393(6685): 537-544.

Cromie, G. A., J. C. Connelly and D. R. Leach (2001). "Recombination at double-strand breaks and DNA ends: conserved mechanisms from phage to humans." Mol Cell 8(6): 1163-1174.

Cromie, G. A., J. C. Connelly and D. R. Leach (2001). "Recombination at double-strand breaks and DNA ends: conserved mechanisms from phage to humans." Mol Cell 8(6): 1163-1174.

Crowley, D. J., I. Boubriak, B. R. Berquist, M. Clark, E. Richard, L. Sullivan, S. DasSarma and S. McCready (2006). "The uvrA, uvrB and uvrC genes are required for repair of ultraviolet light induced DNA photoproducts in Halobacterium sp. NRC-1." Saline Systems 2: 11.

Crowley, D. J., I. Boubriak, B. R. Berquist, M. Clark, E. Richard, L. Sullivan, S. DasSarma and S. McCready (2006). "The uvrA, uvrB and uvrC genes are required for repair of ultraviolet light induced DNA photoproducts in Halobacterium sp. NRC-1." Saline Systems 2: 11.

Davies, K. J. and S. W. Lin (1988). "Degradation of oxidatively denatured proteins in Escherichia coli." Free Radic Biol Med 5(4): 215-223.

Davies, K. J. and S. W. Lin (1988). "Degradation of oxidatively denatured proteins in Escherichia coli." Free Radic Biol Med 5(4): 215-223.

de Ory, A., C. Carabana and M. de Vega (2019). "Bacterial Ligase D preternary-precatalytic complex performs efficient abasic sites processing at double strand breaks during nonhomologous end joining." Nucleic Acids Res 47(10): 5276-5292.

de Ory, A., C. Carabana and M. de Vega (2019). "Bacterial Ligase D preternary-precatalytic complex performs efficient abasic sites processing at double strand breaks during nonhomologous end joining." Nucleic Acids Res 47(10): 5276-5292.

de Vega, M. (2013). "The minimal Bacillus subtilis nonhomologous end joining repair machinery." PLoS One 8(5): e64232.

de Vega, M. (2013). "The minimal Bacillus subtilis nonhomologous end joining repair machinery." PLoS One 8(5): e64232.

Demple, B. and J. Halbrook (1983). "Inducible repair of oxidative DNA damage in Escherichia coli." Nature 304(5925): 466-468.

Demple, B. and J. Halbrook (1983). "Inducible repair of oxidative DNA damage in Escherichia coli." Nature 304(5925): 466-468.

Dianova, II, V. A. Bohr and G. L. Dianov (2001). "Interaction of human AP endonuclease 1 with flap endonuclease 1 and proliferating cell nuclear antigen involved in long-patch base excision repair." Biochemistry 40(42): 12639-12644.

Dianova, II, V. A. Bohr and G. L. Dianov (2001). "Interaction of human AP endonuclease 1 with flap endonuclease 1 and proliferating cell nuclear antigen involved in long-patch base excision repair." Biochemistry 40(42): 12639-12644.

Dizdaroglu, M., E. Coskun and P. Jaruga (2017). "Repair of oxidatively induced DNA damage by DNA glycosylases: Mechanisms of action, substrate specificities and excision kinetics." Mutat Res 771: 99-127.

Doherty, A. J. and S. W. Suh (2000). "Structural and mechanistic conservation in DNA ligases." Nucleic Acids Res 28(21): 4051-4058.

Dos Vultos, T., O. Mestre, T. Tonjum and B. Gicquel (2009). "DNA repair in Mycobacterium tuberculosis revisited." FEMS Microbiol Rev 33(3): 471-487.

Dos Vultos, T., O. Mestre, T. Tonjum and B. Gicquel (2009). "DNA repair in Mycobacterium tuberculosis revisited." FEMS Microbiol Rev 33(3): 471-487.

Dulic, A., P. A. Bates, X. Zhang, S. R. Martin, P. S. Freemont, T. Lindahl and D. E. Barnes (2001). "BRCT domain interactions in the heterodimeric DNA repair protein XRCC1-DNA ligase III." Biochemistry 40(20): 5906-5913.

Ehlers, S. (1999). "Immunity to tuberculosis: a delicate balance between protection and pathology." FEMS Immunol Med Microbiol 23(2): 149-158.

Ehlers, S. (1999). "Immunity to tuberculosis: a delicate balance between protection and pathology." FEMS Immunol Med Microbiol 23(2): 149-158.

Ehrt, S. and D. Schnappinger (2009). "Mycobacterial survival strategies in the phagosome: defence against host stresses." Cell Microbiol 11(8): 1170-1178.

Eisen, J. A. and P. C. Hanawalt (1999). "A phylogenomic study of DNA repair genes, proteins, and processes." Mutat Res 435(3): 171-213.

Endutkin, A. V., A. V. Yudkina, V. S. Sidorenko and D. O. Zharkov (2019). "Transient protein-protein complexes in base excision repair." J Biomol Struct Dyn 37(17): 4407-4418.

Farr, S. B. and T. Kogoma (1991). "Oxidative stress responses in Escherichia coli and Salmonella typhimurium." Microbiol Rev 55(4): 561-585.

Farr, S. B. and T. Kogoma (1991). "Oxidative stress responses in Escherichia coli and Salmonella typhimurium." Microbiol Rev 55(4): 561-585.

Frosina, G., P. Fortini, O. Rossi, F. Carrozzino, G. Raspaglio, L. S. Cox, D. P. Lane, A. Abbondandolo and E. Dogliotti (1996). "Two pathways for base excision repair in mammalian cells." J Biol Chem 271(16): 9573-9578.

Frosina, G., P. Fortini, O. Rossi, F. Carrozzino, G. Raspaglio, L. S. Cox, D. P. Lane, A. Abbondandolo and E. Dogliotti (1996). "Two pathways for base excision repair in mammalian cells." J Biol Chem 271(16): 9573-9578.

Gajiwala, K. S. and C. Pinko (2004). "Structural rearrangement accompanying NAD+ synthesis within a bacterial DNA ligase crystal." Structure 12(8): 1449-1459.

Gajiwala, K. S. and C. Pinko (2004). "Structural rearrangement accompanying NAD+ synthesis within a bacterial DNA ligase crystal." Structure 12(8): 1449-1459.

Gandhi, N. R., A. Moll, A. W. Sturm, R. Pawinski, T. Govender, U. Lalloo, K. Zeller, J. Andrews and G. Friedland (2006). "Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa." Lancet 368(9547): 1575-1580.

Gandhi, N. R., A. Moll, A. W. Sturm, R. Pawinski, T. Govender, U. Lalloo, K. Zeller, J. Andrews and G. Friedland (2006). "Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa." Lancet 368(9547): 1575-1580.

Gary, R., K. Kim, H. L. Cornelius, M. S. Park and Y. Matsumoto (1999). "Proliferating cell nuclear antigen facilitates excision in long-patch base excision repair." J Biol Chem 274(7): 4354-4363.

Gary, R., K. Kim, H. L. Cornelius, M. S. Park and Y. Matsumoto (1999). "Proliferating cell nuclear antigen facilitates excision in long-patch base excision repair." J Biol Chem 274(7): 4354-4363.

Gilson, T., A. E. Greer, A. Vindigni, G. Ketner and L. A. Hanakahi (2012). "The alpha2 helix in the DNA ligase IV BRCT-1 domain is required for targeted degradation of ligase IV during adenovirus infection." Virology 428(2): 128-135.

Gilson, T., A. E. Greer, A. Vindigni, G. Ketner and L. A. Hanakahi (2012). "The alpha2 helix in the DNA ligase IV BRCT-1 domain is required for targeted degradation of ligase IV during adenovirus infection." Virology 428(2): 128-135.

Gorman, M. A., S. Morera, D. G. Rothwell, E. de La Fortelle, C. D. Mol, J. A. Tainer, I. D. Hickson and P. S. Freemont (1997). "The crystal structure of the human DNA repair endonuclease HAP1 suggests the recognition of extra-helical deoxyribose at DNA abasic sites." EMBO J 16(21): 6548-6558.

Gu, S., W. Li, H. Zhang, J. Fleming, W. Yang, S. Wang, W. Wei, J. Zhou, G. Zhu, J. Deng, J. Hou, Y. Zhou, S. Lin, X. E. Zhang and L. Bi (2016). "The beta2 clamp in the Mycobacterium tuberculosis DNA polymerase III alphabeta2epsilon replicase promotes polymerization and reduces exonuclease activity." Sci Rep 6: 18418.

Gu, S., W. Li, H. Zhang, J. Fleming, W. Yang, S. Wang, W. Wei, J. Zhou, G. Zhu, J. Deng, J. Hou, Y. Zhou, S. Lin, X. E. Zhang and L. Bi (2016). "The beta2 clamp in the Mycobacterium tuberculosis DNA polymerase III alphabeta2epsilon replicase promotes polymerization and reduces exonuclease activity." Sci Rep 6: 18418.

Guarne, A., S. Ramon-Maiques, E. M. Wolff, R. Ghirlando, X. Hu, J. H. Miller and W. Yang (2004). "Structure of the MutL C-terminal domain: a model of intact MutL and its roles in mismatch repair." EMBO J 23(21): 4134-4145.

Guarne, A., S. Ramon-Maiques, E. M. Wolff, R. Ghirlando, X. Hu, J. H. Miller and W. Yang (2004). "Structure of the MutL C-terminal domain: a model of intact MutL and its roles in mismatch repair." EMBO J 23(21): 4134-4145.

Gui, W. J., S. Q. Lin, Y. Y. Chen, X. E. Zhang, L. J. Bi and T. Jiang (2011). "Crystal structure of DNA polymerase III beta sliding clamp from Mycobacterium tuberculosis." Biochem Biophys Res Commun 405(2): 272-277.

Gui, W. J., S. Q. Lin, Y. Y. Chen, X. E. Zhang, L. J. Bi and T. Jiang (2011). "Crystal structure of DNA polymerase III beta sliding clamp from Mycobacterium tuberculosis." Biochem Biophys Res Commun 405(2): 272-277.

Hadi, M. Z., M. A. Coleman, K. Fidelis, H. W. Mohrenweiser and D. M. Wilson, 3rd (2000). "Functional characterization of Ape1 variants identified in the human population." Nucleic Acids Res 28(20): 3871-3879.

Hadi, M. Z., M. A. Coleman, K. Fidelis, H. W. Mohrenweiser and D. M. Wilson, 3rd (2000). "Functional characterization of Ape1 variants identified in the human population." Nucleic Acids Res 28(20): 3871-3879.

Hagensee, M. E. and R. E. Moses (1989). "Multiple pathways for repair of hydrogen peroxide-induced DNA damage in Escherichia coli." J Bacteriol 171(2): 991-995.

Hagensee, M. E. and R. E. Moses (1989). "Multiple pathways for repair of hydrogen peroxide-induced DNA damage in Escherichia coli." J Bacteriol 171(2): 991-995.

Hakansson, K., A. J. Doherty, S. Shuman and D. B. Wigley (1997). "X-ray crystallography reveals a large conformational change during guanyl transfer by mRNA capping enzymes." Cell 89(4): 545-553.

Hakansson, K., A. J. Doherty, S. Shuman and D. B. Wigley (1997). "X-ray crystallography reveals a large conformational change during guanyl transfer by mRNA capping enzymes." Cell 89(4): 545-553.

Harbeck, C. and L. Peterson (1992). "Elephants dancing in my head: a developmental approach to children's concepts of specific pains." Child Dev 63(1): 138-149.

Harbeck, C. and L. Peterson (1992). "Elephants dancing in my head: a developmental approach to children's concepts of specific pains." Child Dev 63(1): 138-149.

Haring, M., H. Rudiger, B. Demple, S. Boiteux and B. Epe (1994). "Recognition of oxidized abasic sites by repair endonucleases." Nucleic Acids Res 22(11): 2010-2015.

Haring, M., H. Rudiger, B. Demple, S. Boiteux and B. Epe (1994). "Recognition of oxidized abasic sites by repair endonucleases." Nucleic Acids Res 22(11): 2010-2015.

Hegde, M. L., C. A. Theriot, A. Das, P. M. Hegde, Z. Guo, R. K. Gary, T. K. Hazra, B. Shen and S. Mitra (2008). "Physical and functional interaction between human oxidized

base-specific DNA glycosylase NEIL1 and flap endonuclease 1." J Biol Chem 283(40): 27028-27037.

Hegde, M. L., C. A. Theriot, A. Das, P. M. Hegde, Z. Guo, R. K. Gary, T. K. Hazra, B. Shen and S. Mitra (2008). "Physical and functional interaction between human oxidized base-specific DNA glycosylase NEIL1 and flap endonuclease 1." J Biol Chem 283(40): 27028-27037.

Henricksen, L. A., J. Veeraraghavan, D. R. Chafin and R. A. Bambara (2002). "DNA ligase I competes with FEN1 to expand repetitive DNA sequences in vitro." J Biol Chem 277(25): 22361-22369.

Henricksen, L. A., J. Veeraraghavan, D. R. Chafin and R. A. Bambara (2002). "DNA ligase I competes with FEN1 to expand repetitive DNA sequences in vitro." J Biol Chem 277(25): 22361-22369.

Ho, C. K., J. L. Van Etten and S. Shuman (1997). "Characterization of an ATP-dependent DNA ligase encoded by Chlorella virus PBCV-1." J Virol 71(3): 1931-1937.

Hornback, M. L. and R. M. Roop, 2nd (2006). "The Brucella abortus xthA-1 gene product participates in base excision repair and resistance to oxidative killing but is not required for wild-type virulence in the mouse model." J Bacteriol 188(4): 1295-1300.

Hornback, M. L. and R. M. Roop, 2nd (2006). "The Brucella abortus xthA-1 gene product participates in base excision repair and resistance to oxidative killing but is not required for wild-type virulence in the mouse model." J Bacteriol 188(4): 1295-1300.

Imlay, J. A. and S. Linn (1987). "Mutagenesis and stress responses induced in Escherichia coli by hydrogen peroxide." J Bacteriol 169(7): 2967-2976.

Imlay, J. A. and S. Linn (1987). "Mutagenesis and stress responses induced in Escherichia coli by hydrogen peroxide." J Bacteriol 169(7): 2967-2976.

Jeon, H. J., H. J. Shin, J. J. Choi, H. S. Hoe, H. K. Kim, S. W. Suh and S. T. Kwon (2004). "Mutational analyses of the thermostable NAD+-dependent DNA ligase from Thermus filiformis." FEMS Microbiol Lett 237(1): 111-118.

Jeon, H. J., H. J. Shin, J. J. Choi, H. S. Hoe, H. K. Kim, S. W. Suh and S. T. Kwon (2004). "Mutational analyses of the thermostable NAD+-dependent DNA ligase from Thermus filiformis." FEMS Microbiol Lett 237(1): 111-118.

Jeon, H. J., H. J. Shin, J. J. Choi, H. S. Hoe, H. K. Kim, S. W. Suh and S. T. Kwon (2004). "Mutational analyses of the thermostable NAD+-dependent DNA ligase from Thermus filiformis." FEMS Microbiol Lett 237(1): 111-118.

Johnson, E. P., S. Yao and D. R. Helinski (2005). "Gyrase inhibitors and thymine starvation disrupt the normal pattern of plasmid RK2 localization in Escherichia coli." J Bacteriol 187(10): 3538-3547.

Johnson, E. P., S. Yao and D. R. Helinski (2005). "Gyrase inhibitors and thymine starvation disrupt the normal pattern of plasmid RK2 localization in Escherichia coli." J Bacteriol 187(10): 3538-3547.

Junop, M. S., G. Obmolova, K. Rausch, P. Hsieh and W. Yang (2001). "Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair." Mol Cell 7(1): 1-12.

Junop, M. S., G. Obmolova, K. Rausch, P. Hsieh and W. Yang (2001). "Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair." Mol Cell 7(1): 1-12.

Kaczmarek, F. S., R. P. Zaniewski, T. D. Gootz, D. E. Danley, M. N. Mansour, M. Griffor, A. V. Kamath, M. Cronan, J. Mueller, D. Sun, P. K. Martin, B. Benton, L. McDowell, D. Biek and M. B. Schmid (2001). "Cloning and functional characterization of an NAD(+)-dependent DNA ligase from Staphylococcus aureus." J Bacteriol 183(10): 3016-3024.

Kaczmarek, F. S., R. P. Zaniewski, T. D. Gootz, D. E. Danley, M. N. Mansour, M. Griffor, A. V. Kamath, M. Cronan, J. Mueller, D. Sun, P. K. Martin, B. Benton, L. McDowell, D. Biek and M. B. Schmid (2001). "Cloning and functional characterization of an NAD(+)-dependent DNA ligase from Staphylococcus aureus." J Bacteriol 183(10): 3016-3024.

Khanam, T., A. Shukla, N. Rai and R. Ramachandran (2015). "Critical determinants for substrate recognition and catalysis in the M. tuberculosis class II AP-endonuclease/3'-5' exonuclease III." Biochim Biophys Acta 1854(5): 505-516.

Khanam, T., A. Shukla, N. Rai and R. Ramachandran (2015). "Critical determinants for substrate recognition and catalysis in the M. tuberculosis class II AP-endonuclease/3'-5' exonuclease III." Biochim Biophys Acta 1854(5): 505-516.

Khanam, T., A. Shukla, N. Rai and R. Ramachandran (2015). "Critical determinants for substrate recognition and catalysis in the M. tuberculosis class II AP-endonuclease/3'-5' exonuclease III." Biochim Biophys Acta 1854(5): 505-516.

Khanam, T., N. Rai and R. Ramachandran (2015). "Mycobacterium tuberculosis class II apurinic/apyrimidinic-endonuclease/3'-5' exonuclease III exhibits DNA regulated modes of interaction with the sliding DNA beta-clamp." Mol Microbiol 98(1): 46-68.

Khanam, T., N. Rai and R. Ramachandran (2015). "Mycobacterium tuberculosis class II apurinic/apyrimidinic-endonuclease/3'-5' exonuclease III exhibits DNA regulated modes of interaction with the sliding DNA beta-clamp." Mol Microbiol 98(1): 46-68.

Kim, K., S. Biade and Y. Matsumoto (1998). "Involvement of flap endonuclease 1 in base excision DNA repair." J Biol Chem 273(15): 8842-8848.

Kim, K., S. Biade and Y. Matsumoto (1998). "Involvement of flap endonuclease 1 in base excision DNA repair." J Biol Chem 273(15): 8842-8848.

Kitsera, N., M. Rodriguez-Alvarez, S. Emmert, T. Carell and A. Khobta (2019). "Nucleotide excision repair of abasic DNA lesions." Nucleic Acids Res.

Kitsera, N., M. Rodriguez-Alvarez, S. Emmert, T. Carell and A. Khobta (2019). "Nucleotide excision repair of abasic DNA lesions." Nucleic Acids Res.

Kiyonari, S., S. Tahara, M. Uchimura, T. Shirai, S. Ishino and Y. Ishino (2009). "Studies on the base excision repair (BER) complex in Pyrococcus furiosus." Biochem Soc Trans 37(Pt 1): 79-82.

Kiyonari, S., S. Tahara, M. Uchimura, T. Shirai, S. Ishino and Y. Ishino (2009). "Studies on the base excision repair (BER) complex in Pyrococcus furiosus." Biochem Soc Trans 37(Pt 1): 79-82.

Kladova, O. A., M. Bazlekowa-Karaban, S. Baconnais, O. Pietrement, A. A. Ishchenko, B. T. Matkarimov, D. A. Iakovlev, A. Vasenko, O. S. Fedorova, E. Le Cam, B. Tudek, N. A. Kuznetsov and M. Saparbaev (2018). "The role of the N-terminal domain of human apurinic/apyrimidinic endonuclease 1, APE1, in DNA glycosylase stimulation." DNA Repair (Amst) 64: 10-25.

Kladova, O. A., M. Bazlekowa-Karaban, S. Baconnais, O. Pietrement, A. A. Ishchenko, B. T. Matkarimov, D. A. Iakovlev, A. Vasenko, O. S. Fedorova, E. Le Cam, B. Tudek, N. A. Kuznetsov and M. Saparbaev (2018). "The role of the N-terminal domain of human apurinic/apyrimidinic endonuclease 1, APE1, in DNA glycosylase stimulation." DNA Repair (Amst) 64: 10-25.

Klungland, A. and T. Lindahl (1997). "Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1)." EMBO J 16(11): 3341-3348.

Klungland, A. and T. Lindahl (1997). "Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1)." EMBO J 16(11): 3341-3348.

Kodama, K., D. E. Barnes and T. Lindahl (1991). "In vitro mutagenesis and functional expression in Escherichia coli of a cDNA encoding the catalytic domain of human DNA ligase I." Nucleic Acids Res 19(22): 6093-6099.

Kolattukudy, P. E., N. D. Fernandes, A. K. Azad, A. M. Fitzmaurice and T. D. Sirakova (1997). "Biochemistry and molecular genetics of cell-wall lipid biosynthesis in mycobacteria." Mol Microbiol 24(2): 263-270.

Kolattukudy, P. E., N. D. Fernandes, A. K. Azad, A. M. Fitzmaurice and T. D. Sirakova (1997). "Biochemistry and molecular genetics of cell-wall lipid biosynthesis in mycobacteria." Mol Microbiol 24(2): 263-270.

Krokan, H. E. and M. Bjoras (2013). "Base excision repair." Cold Spring Harb Perspect Biol 5(4): a012583.

Kukshal, V., T. Khanam, D. Chopra, N. Singh, S. Sanyal and R. Ramachandran (2012). "M. tuberculosis sliding beta-clamp does not interact directly with the NAD+-dependent DNA ligase." PLoS One 7(4): e35702.

Kukshal, V., T. Khanam, D. Chopra, N. Singh, S. Sanyal and R. Ramachandran (2012). "M. tuberculosis sliding beta-clamp does not interact directly with the NAD+-dependent DNA ligase." PLoS One 7(4): e35702.

Kumar, P., S. K. Bharti and U. Varshney (2011). "Uracil excision repair in Mycobacterium tuberculosis cell-free extracts." Tuberculosis (Edinb) 91(3): 212-218.

Kumar, P., S. K. Bharti and U. Varshney (2011). "Uracil excision repair in Mycobacterium tuberculosis cell-free extracts." Tuberculosis (Edinb) 91(3): 212-218.

Lata, K., M. Afsar and R. Ramachandran (2017). "Biochemical characterization and novel inhibitor identification of Mycobacterium tuberculosis Endonuclease VIII 2 (Rv3297)." Biochem Biophys Rep 12: 20-28.

Lata, K., M. Afsar and R. Ramachandran (2017). "Biochemical characterization and novel inhibitor identification of Mycobacterium tuberculosis Endonuclease VIII 2 (Rv3297)." Biochem Biophys Rep 12: 20-28.

Lavesa-Curto, M., H. Sayer, D. Bullard, A. MacDonald, A. Wilkinson, A. Smith, L. Bowater, A. Hemmings and R. P. Bowater (2004). "Characterization of a temperature-sensitive DNA ligase from Escherichia coli." Microbiology 150(Pt 12): 4171-4180.

Lavesa-Curto, M., H. Sayer, D. Bullard, A. MacDonald, A. Wilkinson, A. Smith, L. Bowater, A. Hemmings and R. P. Bowater (2004). "Characterization of a temperature-sensitive DNA ligase from Escherichia coli." Microbiology 150(Pt 12): 4171-4180.

Lee, J. Y., C. Chang, H. K. Song, J. Moon, J. K. Yang, H. K. Kim, S. T. Kwon and S. W. Suh (2000). "Crystal structure of NAD(+)-dependent DNA ligase: modular architecture and functional implications." EMBO J 19(5): 1119-1129.

Lee, J. Y., C. Chang, H. K. Song, J. Moon, J. K. Yang, H. K. Kim, S. T. Kwon and S. W. Suh (2000). "Crystal structure of NAD(+)-dependent DNA ligase: modular architecture and functional implications." EMBO J 19(5): 1119-1129.

Lee, J. Y., J. Chang, N. Joseph, R. Ghirlando, D. N. Rao and W. Yang (2005). "MutH complexed with hemi- and unmethylated DNAs: coupling base recognition and DNA cleavage." Mol Cell 20(1): 155-166.

Lee, J. Y., J. Chang, N. Joseph, R. Ghirlando, D. N. Rao and W. Yang (2005). "MutH complexed with hemi- and unmethylated DNAs: coupling base recognition and DNA cleavage." Mol Cell 20(1): 155-166.

Lim, D., H. U. Park, L. De Castro, S. G. Kang, H. S. Lee, S. Jensen, K. J. Lee and N. C. Strynadka (2001). "Crystal structure and kinetic analysis of beta-lactamase inhibitor protein-II in complex with TEM-1 beta-lactamase." Nat Struct Biol 8(10): 848-852.

Lim, D., H. U. Park, L. De Castro, S. G. Kang, H. S. Lee, S. Jensen, K. J. Lee and N. C. Strynadka (2001). "Crystal structure and kinetic analysis of beta-lactamase inhibitor protein-II in complex with TEM-1 beta-lactamase." Nat Struct Biol 8(10): 848-852.

Lindahl, T. (1993). "Instability and decay of the primary structure of DNA." Nature 362(6422): 709-715.

Lindahl, T. (1993). "Instability and decay of the primary structure of DNA." Nature 362(6422): 709-715.

Liu, Y., R. Prasad, W. A. Beard, P. S. Kedar, E. W. Hou, D. D. Shock and S. H. Wilson (2007). "Coordination of steps in single-nucleotide base excision repair mediated by apurinic/apyrimidinic endonuclease 1 and DNA polymerase beta." J Biol Chem 282(18): 13532-13541.

Liu, Y., R. Prasad, W. A. Beard, P. S. Kedar, E. W. Hou, D. D. Shock and S. H. Wilson (2007). "Coordination of steps in single-nucleotide base excision repair mediated by apurinic/apyrimidinic endonuclease 1 and DNA polymerase beta." J Biol Chem 282(18): 13532-13541.

Lu, J., J. Tong, H. Feng, J. Huang, C. L. Afonso, D. L. Rock, F. Barany and W. Cao (2004). "Unique ligation properties of eukaryotic NAD+-dependent DNA ligase from Melanoplus sanguinipes entomopoxvirus." Biochim Biophys Acta 1701(1-2): 37-48.

Lu, J., J. Tong, H. Feng, J. Huang, C. L. Afonso, D. L. Rock, F. Barany and W. Cao (2004). "Unique ligation properties of eukaryotic NAD+-dependent DNA ligase from Melanoplus sanguinipes entomopoxvirus." Biochim Biophys Acta 1701(1-2): 37-48.

Luo, J. and F. Barany (1996). "Identification of essential residues in Thermus thermophilus DNA ligase." Nucleic Acids Res 24(15): 3079-3085.

Mackey, Z. B., C. Niedergang, J. M. Murcia, J. Leppard, K. Au, J. Chen, G. de Murcia and A. E. Tomkinson (1999). "DNA ligase III is recruited to DNA strand breaks by a zinc finger motif homologous to that of poly(ADP-ribose) polymerase. Identification of two functionally distinct DNA binding regions within DNA ligase III." J Biol Chem 274(31): 21679-21687.

Mackey, Z. B., C. Niedergang, J. M. Murcia, J. Leppard, K. Au, J. Chen, G. de Murcia and A. E. Tomkinson (1999). "DNA ligase III is recruited to DNA strand breaks by a zinc finger motif homologous to that of poly(ADP-ribose) polymerase. Identification of two functionally distinct DNA binding regions within DNA ligase III." J Biol Chem 274(31): 21679-21687.

Madhusudan, S., F. Smart, P. Shrimpton, J. L. Parsons, L. Gardiner, S. Houlbrook, D. C. Talbot, T. Hammonds, P. A. Freemont, M. J. Sternberg, G. L. Dianov and I. D. Hickson (2005). "Isolation of a small molecule inhibitor of DNA base excision repair." Nucleic Acids Res 33(15): 4711-4724.

Madhusudan, S., F. Smart, P. Shrimpton, J. L. Parsons, L. Gardiner, S. Houlbrook, D. C. Talbot, T. Hammonds, P. A. Freemont, M. J. Sternberg, G. L. Dianov and I. D. Hickson (2005). "Isolation of a small molecule inhibitor of DNA base excision repair." Nucleic Acids Res 33(15): 4711-4724.

Masson, M., C. Niedergang, V. Schreiber, S. Muller, J. Menissier-de Murcia and G. de Murcia (1998). "XRCC1 is specifically associated with poly(ADP-ribose) polymerase and negatively regulates its activity following DNA damage." Mol Cell Biol 18(6): 3563-3571.

Matsumoto, Y. (2001). "Molecular mechanism of PCNA-dependent base excision repair." Prog Nucleic Acid Res Mol Biol 68: 129-138.

Matsumoto, Y. (2001). "Molecular mechanism of PCNA-dependent base excision repair." Prog Nucleic Acid Res Mol Biol 68: 129-138.

Mazloum, N., M. A. Stegman, D. L. Croteau, B. Van Houten, N. S. Kwon, Y. Ling, C. Dickinson, A. Venugopal, M. A. Towheed and C. Nathan (2011). "Identification of a chemical that inhibits the mycobacterial UvrABC complex in nucleotide excision repair." Biochemistry 50(8): 1329-1335.

Mazloum, N., M. A. Stegman, D. L. Croteau, B. Van Houten, N. S. Kwon, Y. Ling, C. Dickinson, A. Venugopal, M. A. Towheed and C. Nathan (2011). "Identification of a chemical that inhibits the mycobacterial UvrABC complex in nucleotide excision repair." Biochemistry 50(8): 1329-1335.

Mishina, Y., E. M. Duguid and C. He (2006). "Direct reversal of DNA alkylation damage." Chem Rev 106(2): 215-232.

Mishina, Y., E. M. Duguid and C. He (2006). "Direct reversal of DNA alkylation damage." Chem Rev 106(2): 215-232.

Mizrahi, V. and S. J. Andersen (1998). "DNA repair in Mycobacterium tuberculosis. What have we learnt from the genome sequence?" Mol Microbiol 29(6): 1331-1339.

Mizrahi, V. and S. J. Andersen (1998). "DNA repair in Mycobacterium tuberculosis. What have we learnt from the genome sequence?" Mol Microbiol 29(6): 1331-1339.

Moeller, R., P. Setlow, M. Pedraza-Reyes, R. Okayasu, G. Reitz and W. L. Nicholson (2011). "Role of the Nfo and ExoA apurinic/apyrimidinic endonucleases in radiation resistance and radiation-induced mutagenesis of Bacillus subtilis spores." J Bacteriol 193(11): 2875-2879.

Moeller, R., P. Setlow, M. Pedraza-Reyes, R. Okayasu, G. Reitz and W. L. Nicholson (2011). "Role of the Nfo and ExoA apurinic/apyrimidinic endonucleases in radiation resistance and radiation-induced mutagenesis of Bacillus subtilis spores." J Bacteriol 193(11): 2875-2879.

Mol, C. D., D. J. Hosfield and J. A. Tainer (2000). "Abasic site recognition by two apurinic/apyrimidinic endonuclease families in DNA base excision repair: the 3' ends justify the means." Mutat Res 460(3-4): 211-229.

Mol, C. D., D. J. Hosfield and J. A. Tainer (2000). "Abasic site recognition by two apurinic/apyrimidinic endonuclease families in DNA base excision repair: the 3' ends justify the means." Mutat Res 460(3-4): 211-229.

Mol, C. D., T. Izumi, S. Mitra and J. A. Tainer (2000). "DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination [corrected]." Nature 403(6768): 451-456.

Moor, N. A. and O. I. Lavrik (2018). "Protein-Protein Interactions in DNA Base Excision Repair." Biochemistry (Mosc) 83(4): 411-422.

Moor, N. A. and O. I. Lavrik (2018). "Protein-Protein Interactions in DNA Base Excision Repair." Biochemistry (Mosc) 83(4): 411-422.

Moor, N. A., I. A. Vasil'eva, R. O. Anarbaev, A. A. Antson and O. I. Lavrik (2015). "Quantitative characterization of protein-protein complexes involved in base excision DNA repair." Nucleic Acids Res 43(12): 6009-6022.

Moor, N. A., I. A. Vasil'eva, R. O. Anarbaev, A. A. Antson and O. I. Lavrik (2015). "Quantitative characterization of protein-protein complexes involved in base excision DNA repair." Nucleic Acids Res 43(12): 6009-6022.

Moore, J. K. and J. E. Haber (1996). "Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in Saccharomyces cerevisiae." Mol Cell Biol 16(5): 2164-2173.

Moore, J. K. and J. E. Haber (1996). "Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in Saccharomyces cerevisiae." Mol Cell Biol 16(5): 2164-2173.

Mullins, E. A., A. A. Rodriguez, N. P. Bradley and B. F. Eichman (2019). "Emerging Roles of DNA Glycosylases and the Base Excision Repair Pathway." Trends Biochem Sci 44(9): 765-781.

Nandakumar, J., P. A. Nair and S. Shuman (2007). "Last stop on the road to repair: structure of E. coli DNA ligase bound to nicked DNA-adenylate." Mol Cell 26(2): 257-271.

Nandakumar, J., P. A. Nair and S. Shuman (2007). "Last stop on the road to repair: structure of E. coli DNA ligase bound to nicked DNA-adenylate." Mol Cell 26(2): 257-271.

Nandakumar, J., P. A. Nair and S. Shuman (2007). "Last stop on the road to repair: structure of E. coli DNA ligase bound to nicked DNA-adenylate." Mol Cell 26(2): 257-271.

Nash, R. A., K. W. Caldecott, D. E. Barnes and T. Lindahl (1997). "XRCC1 protein interacts with one of two distinct forms of DNA ligase III." Biochemistry 36(17): 5207-5211.

Nash, R. A., K. W. Caldecott, D. E. Barnes and T. Lindahl (1997). "XRCC1 protein interacts with one of two distinct forms of DNA ligase III." Biochemistry 36(17): 5207-5211.

Nazarkina Zh, K., S. N. Khodyreva, S. Marsin, J. P. Radicella and O. I. Lavrik (2007). "Study of interaction of XRCC1 with DNA and proteins of base excision repair by photoaffinity labeling technique." Biochemistry (Mosc) 72(8): 878-886.

Nazarkina Zh, K., S. N. Khodyreva, S. Marsin, J. P. Radicella and O. I. Lavrik (2007). "Study of interaction of XRCC1 with DNA and proteins of base excision repair by photoaffinity labeling technique." Biochemistry (Mosc) 72(8): 878-886.

Odell, M. and S. Shuman (1999). "Footprinting of Chlorella virus DNA ligase bound at a nick in duplex DNA." J Biol Chem 274(20): 14032-14039.

Odell, M. and S. Shuman (1999). "Footprinting of Chlorella virus DNA ligase bound at a nick in duplex DNA." J Biol Chem 274(20): 14032-14039.

Otterlei, M., B. Kavli, R. Standal, C. Skjelbred, S. Bharati and H. E. Krokan (2000). "Repair of chromosomal abasic sites in vivo involves at least three different repair pathways." EMBO J 19(20): 5542-5551.

Otterlei, M., B. Kavli, R. Standal, C. Skjelbred, S. Bharati and H. E. Krokan (2000). "Repair of chromosomal abasic sites in vivo involves at least three different repair pathways." EMBO J 19(20): 5542-5551.

Parikh, S. S., C. D. Mol, D. J. Hosfield and J. A. Tainer (1999). "Envisioning the molecular choreography of DNA base excision repair." Curr Opin Struct Biol 9(1): 37-47.

Pascal, J. M., P. J. O'Brien, A. E. Tomkinson and T. Ellenberger (2004). "Human DNA ligase I completely encircles and partially unwinds nicked DNA." Nature 432(7016): 473-478.

Pascal, J. M., P. J. O'Brien, A. E. Tomkinson and T. Ellenberger (2004). "Human DNA ligase I completely encircles and partially unwinds nicked DNA." Nature 432(7016): 473-478.

Perkins, A. and P. Robershotte (1992). "Dancing without the elephants. How local managed care and national insurer contracts match up." Contemp Longterm Care 15(5): 28-30, 78.

Perkins, A. and P. Robershotte (1992). "Dancing without the elephants. How local managed care and national insurer contracts match up." Contemp Longterm Care 15(5): 28-30, 78.

Plocinski, P., N. C. Brissett, J. Bianchi, A. Brzostek, M. Korycka-Machala, A. Dziembowski, J. Dziadek and A. J. Doherty (2017). "DNA Ligase C and Prim-PolC participate in base excision repair in mycobacteria." Nat Commun 8(1): 1251.

Plocinski, P., N. C. Brissett, J. Bianchi, A. Brzostek, M. Korycka-Machala, A. Dziembowski, J. Dziadek and A. J. Doherty (2017). "DNA Ligase C and Prim-PolC participate in base excision repair in mycobacteria." Nat Commun 8(1): 1251.

Poidevin, L. and S. A. MacNeill (2006). "Biochemical characterisation of LigN, an NAD+-dependent DNA ligase from the halophilic euryarchaeon Haloferax volcanii that displays maximal in vitro activity at high salt concentrations." BMC Mol Biol 7: 44.

Poidevin, L. and S. A. MacNeill (2006). "Biochemical characterisation of LigN, an NAD+-dependent DNA ligase from the halophilic euryarchaeon Haloferax volcanii that displays maximal in vitro activity at high salt concentrations." BMC Mol Biol 7: 44.

Poletto, M., A. J. Legrand and G. L. Dianov (2017). "DNA Base Excision Repair: The Achilles' Heel of Tumour Cells and their Microenvironment?" Curr Pharm Des 23(32): 4758-4772.

Poletto, M., A. J. Legrand and G. L. Dianov (2017). "DNA Base Excision Repair: The Achilles' Heel of Tumour Cells and their Microenvironment?" Curr Pharm Des 23(32): 4758-4772.

Prasad, R., R. K. Singhal, D. K. Srivastava, J. T. Molina, A. E. Tomkinson and S. H. Wilson (1996). "Specific interaction of DNA polymerase beta and DNA ligase I in a multiprotein base excision repair complex from bovine testis." J Biol Chem 271(27): 16000-16007.

Prasad, R., R. K. Singhal, D. K. Srivastava, J. T. Molina, A. E. Tomkinson and S. H. Wilson (1996). "Specific interaction of DNA polymerase beta and DNA ligase I in a multiprotein base excision repair complex from bovine testis." J Biol Chem 271(27): 16000-16007.

Puri, R. V., P. V. Reddy and A. K. Tyagi (2014). "Apurinic/apyrimidinic endonucleases of Mycobacterium tuberculosis protect against DNA damage but are dispensable for the growth of the pathogen in guinea pigs." PLoS One 9(5): e92035.

Puri, R. V., P. V. Reddy and A. K. Tyagi (2014). "Apurinic/apyrimidinic endonucleases of Mycobacterium tuberculosis protect against DNA damage but are dispensable for the growth of the pathogen in guinea pigs." PLoS One 9(5): e92035.

Ranalli, T. A., S. Tom and R. A. Bambara (2002). "AP endonuclease 1 coordinates flap endonuclease 1 and DNA ligase I activity in long patch base excision repair." J Biol Chem 277(44): 41715-41724.

Ranalli, T. A., S. Tom and R. A. Bambara (2002). "AP endonuclease 1 coordinates flap endonuclease 1 and DNA ligase I activity in long patch base excision repair." J Biol Chem 277(44): 41715-41724.

Richardson, A. R., K. C. Soliven, M. E. Castor, P. D. Barnes, S. J. Libby and F. C. Fang (2009). "The Base Excision Repair system of Salmonella enterica serovar

typhimurium counteracts DNA damage by host nitric oxide." PLoS Pathog 5(5): e1000451.

Richardson, A. R., K. C. Soliven, M. E. Castor, P. D. Barnes, S. J. Libby and F. C. Fang (2009). "The Base Excision Repair system of Salmonella enterica serovar typhimurium counteracts DNA damage by host nitric oxide." PLoS Pathog 5(5): e1000451.

Robertson, A. B., A. Klungland, T. Rognes and I. Leiros (2009). "DNA repair in mammalian cells: Base excision repair: the long and short of it." Cell Mol Life Sci 66(6): 981-993.

Robertson, A. B., A. Klungland, T. Rognes and I. Leiros (2009). "DNA repair in mammalian cells: Base excision repair: the long and short of it." Cell Mol Life Sci 66(6): 981-993.

Roldan-Arjona, T., R. R. Ariza and D. Cordoba-Canero (2019). "DNA Base Excision Repair in Plants: An Unfolding Story With Familiar and Novel Characters." Front Plant Sci 10: 1055.

Roldan-Arjona, T., R. R. Ariza and D. Cordoba-Canero (2019). "DNA Base Excision Repair in Plants: An Unfolding Story With Familiar and Novel Characters." Front Plant Sci 10: 1055.

Sancar, A. (1994). "Mechanisms of DNA excision repair." Science 266(5193): 1954-1956.

Savery, N. J. (2007). "The molecular mechanism of transcription-coupled DNA repair." Trends Microbiol 15(7): 326-333.

Savery, N. J. (2007). "The molecular mechanism of transcription-coupled DNA repair." Trends Microbiol 15(7): 326-333.

Shapiro, A. B. (2014). "Complete steady-state rate equation for DNA ligase and its use for measuring product kinetic parameters of NAD(+)-dependent DNA ligase from Haemophilus influenzae." BMC Res Notes 7: 287.

Sheng, Z. Z., Y. Q. Zhao and J. F. Huang (2011). "Functional Evolution of BRCT Domains from Binding DNA to Protein." Evol Bioinform Online 7: 87-97.

Shuman, S. (1995). "Vaccinia virus DNA ligase: specificity, fidelity, and inhibition." Biochemistry 34(49): 16138-16147.

Shuman, S. (1996). "Closing the gap on DNA ligase." Structure 4(6): 653-656.

Shuman, S. (1996). "Closing the gap on DNA ligase." Structure 4(6): 653-656.

Shuman, S. (2009). "DNA ligases: progress and prospects." J Biol Chem 284(26): 17365-17369.

Shuman, S. and C. D. Lima (2004). "The polynucleotide ligase and RNA capping enzyme superfamily of covalent nucleotidyltransferases." Curr Opin Struct Biol 14(6): 757-764.

Shuman, S. and C. D. Lima (2004). "The polynucleotide ligase and RNA capping enzyme superfamily of covalent nucleotidyltransferases." Curr Opin Struct Biol 14(6): 757-764.

Shuman, S. and X. M. Ru (1995). "Mutational analysis of vaccinia DNA ligase defines residues essential for covalent catalysis." Virology 211(1): 73-83.

Sidorenko, V. S., M. A. Rot, M. L. Filipenko, G. A. Nevinsky and D. O. Zharkov (2008). "Novel DNA glycosylases from Mycobacterium tuberculosis." Biochemistry (Mosc) 73(4): 442-450.

Sidorenko, V. S., M. A. Rot, M. L. Filipenko, G. A. Nevinsky and D. O. Zharkov (2008). "Novel DNA glycosylases from Mycobacterium tuberculosis." Biochemistry (Mosc) 73(4): 442-450.

Singleton, M. R., K. Hakansson, D. J. Timson and D. B. Wigley (1999). "Structure of the adenylation domain of an NAD+-dependent DNA ligase." Structure 7(1): 35-42.

Soderhall, S. (1976). "DNA ligases during rat liver regeneration." Nature 260(5552): 640-642.

Soderhall, S. and T. Lindahl (1976). "DNA ligases of eukaryotes." FEBS Lett 67(1): 1-8.

Soderhall, S. and T. Lindahl (1976). "DNA ligases of eukaryotes." FEBS Lett 67(1): 1-8.

Spampinato, C. and P. Modrich (2000). "The MutL ATPase is required for mismatch repair." J Biol Chem 275(13): 9863-9869.

Spampinato, C. and P. Modrich (2000). "The MutL ATPase is required for mismatch repair." J Biol Chem 275(13): 9863-9869.

Springer, B., P. Sander, L. Sedlacek, W. D. Hardt, V. Mizrahi, P. Schar and E. C. Bottger (2004). "Lack of mismatch correction facilitates genome evolution in mycobacteria." Mol Microbiol 53(6): 1601-1609.

Springer, B., P. Sander, L. Sedlacek, W. D. Hardt, V. Mizrahi, P. Schar and E. C. Bottger (2004). "Lack of mismatch correction facilitates genome evolution in mycobacteria." Mol Microbiol 53(6): 1601-1609.

Sriskanda, V. and S. Shuman (1998). "Chlorella virus DNA ligase: nick recognition and mutational analysis." Nucleic Acids Res 26(2): 525-531.

Sriskanda, V. and S. Shuman (1998). "Chlorella virus DNA ligase: nick recognition and mutational analysis." Nucleic Acids Res 26(2): 525-531.

Sriskanda, V. and S. Shuman (1998). "Mutational analysis of Chlorella virus DNA ligase: catalytic roles of domain I and motif VI." Nucleic Acids Res 26(20): 4618-4625.

Sriskanda, V. and S. Shuman (2002). "Conserved residues in domain Ia are required for the reaction of Escherichia coli DNA ligase with NAD+." J Biol Chem 277(12): 9695-9700.

Sriskanda, V. and S. Shuman (2002). "Conserved residues in domain Ia are required for the reaction of Escherichia coli DNA ligase with NAD+." J Biol Chem 277(12): 9695-9700.

Sriskanda, V. and S. Shuman (2002). "Role of nucleotidyltransferase motifs I, III and IV in the catalysis of phosphodiester bond formation by Chlorella virus DNA ligase." Nucleic Acids Res 30(4): 903-911.

Sriskanda, V. and S. Shuman (2002). "Role of nucleotidyltransferase motifs I, III and IV in the catalysis of phosphodiester bond formation by Chlorella virus DNA ligase." Nucleic Acids Res 30(4): 903-911.

Sriskanda, V., B. Schwer, C. K. Ho and S. Shuman (1999). "Mutational analysis of Escherichia coli DNA ligase identifies amino acids required for nick-ligation in vitro and for in vivo complementation of the growth of yeast cells deleted for CDC9 and LIG4." Nucleic Acids Res 27(20): 3953-3963.

Sriskanda, V., R. W. Moyer and S. Shuman (2001). "NAD+-dependent DNA ligase encoded by a eukaryotic virus." J Biol Chem 276(39): 36100-36109.

Sriskanda, V., R. W. Moyer and S. Shuman (2001). "NAD+-dependent DNA ligase encoded by a eukaryotic virus." J Biol Chem 276(39): 36100-36109.

Srivastava, S. K., D. Dube, N. Tewari, N. Dwivedi, R. P. Tripathi and R. Ramachandran (2005). "Mycobacterium tuberculosis NAD+-dependent DNA ligase is selectively inhibited by glycosylamines compared with human DNA ligase I." Nucleic Acids Res 33(22): 7090-7101.

Srivastava, S. K., D. Dube, N. Tewari, N. Dwivedi, R. P. Tripathi and R. Ramachandran (2005). "Mycobacterium tuberculosis NAD+-dependent DNA ligase is selectively inhibited by glycosylamines compared with human DNA ligase I." Nucleic Acids Res 33(22): 7090-7101.

Srivastava, S. K., D. Dube, N. Tewari, N. Dwivedi, R. P. Tripathi and R. Ramachandran (2005). "Mycobacterium tuberculosis NAD+-dependent DNA ligase is selectively inhibited by glycosylamines compared with human DNA ligase I." Nucleic Acids Res 33(22): 7090-7101.

Srivastava, S. K., D. Dube, V. Kukshal, A. K. Jha, K. Hajela and R. Ramachandran (2007). "NAD+-dependent DNA ligase (Rv3014c) from Mycobacterium tuberculosis: novel structure-function relationship and identification of a specific inhibitor." Proteins 69(1): 97-111.

Srivastava, S. K., D. Dube, V. Kukshal, A. K. Jha, K. Hajela and R. Ramachandran (2007). "NAD+-dependent DNA ligase (Rv3014c) from Mycobacterium tuberculosis: novel structure-function relationship and identification of a specific inhibitor." Proteins 69(1): 97-111.

Srivastava, S. K., D. Dube, V. Kukshal, A. K. Jha, K. Hajela and R. Ramachandran (2007). "NAD+-dependent DNA ligase (Rv3014c) from Mycobacterium tuberculosis: novel structure-function relationship and identification of a specific inhibitor." Proteins 69(1): 97-111.

Srivastava, S. K., D. Dube, V. Kukshal, A. K. Jha, K. Hajela and R. Ramachandran (2007). "NAD+-dependent DNA ligase (Rv3014c) from Mycobacterium tuberculosis: novel structure-function relationship and identification of a specific inhibitor." Proteins 69(1): 97-111.

Srivastava, S. K., R. P. Tripathi and R. Ramachandran (2005). "NAD+-dependent DNA Ligase (Rv3014c) from Mycobacterium tuberculosis. Crystal structure of the adenylation domain and identification of novel inhibitors." J Biol Chem 280(34): 30273-30281.

Srivastava, S. K., R. P. Tripathi and R. Ramachandran (2005). "NAD+-dependent DNA Ligase (Rv3014c) from Mycobacterium tuberculosis. Crystal structure of the adenylation domain and identification of novel inhibitors." J Biol Chem 280(34): 30273-30281.

Srivastava, S. K., R. P. Tripathi and R. Ramachandran (2005). "NAD+-dependent DNA Ligase (Rv3014c) from Mycobacterium tuberculosis. Crystal structure of the adenylation domain and identification of novel inhibitors." J Biol Chem 280(34): 30273-30281.

Stallings, C. L. and M. S. Glickman (2010). "Is Mycobacterium tuberculosis stressed out? A critical assessment of the genetic evidence." Microbes Infect 12(14-15): 1091-1101.

Stiborova, M., J. Sejbal, L. Borek-Dohalska, D. Aimova, J. Poljakova, K. Forsterova, M. Rupertova, J. Wiesner, J. Hudecek, M. Wiessler and E. Frei (2004). "The anticancer drug ellipticine forms covalent DNA adducts, mediated by human cytochromes P450, through metabolism to 13-hydroxyellipticine and ellipticine N2-oxide." Cancer Res 64(22): 8374-8380.

Stiborova, M., J. Sejbal, L. Borek-Dohalska, D. Aimova, J. Poljakova, K. Forsterova, M. Rupertova, J. Wiesner, J. Hudecek, M. Wiessler and E. Frei (2004). "The anticancer drug ellipticine forms covalent DNA adducts, mediated by human cytochromes P450, through metabolism to 13-hydroxyellipticine and ellipticine N2-oxide." Cancer Res 64(22): 8374-8380.

Sukhanova, M. V., S. N. Khodyreva, N. A. Lebedeva, R. Prasad, S. H. Wilson and O. I. Lavrik (2005). "Human base excision repair enzymes apurinic/apyrimidinic endonuclease1 (APE1), DNA polymerase beta and poly(ADP-ribose) polymerase 1:

interplay between strand-displacement DNA synthesis and proofreading exonuclease activity." Nucleic Acids Res 33(4): 1222-1229.

Tainer, J. A. and E. C. Friedberg (2000). "Dancing with the elephants: envisioning the structural biology of DNA repair pathways." Mutat Res 460(3-4): 139-141.

Tainer, J. A. and E. C. Friedberg (2000). "Dancing with the elephants: envisioning the structural biology of DNA repair pathways." Mutat Res 460(3-4): 139-141.

Taylor, R. M., B. Wickstead, S. Cronin and K. W. Caldecott (1998). "Role of a BRCT domain in the interaction of DNA ligase III-alpha with the DNA repair protein XRCC1." Curr Biol 8(15): 877-880.

Tell, G., D. Fantini and F. Quadrifoglio (2010). "Understanding different functions of mammalian AP endonuclease (APE1) as a promising tool for cancer treatment." Cell Mol Life Sci 67(21): 3589-3608.

Tell, G., D. Fantini and F. Quadrifoglio (2010). "Understanding different functions of mammalian AP endonuclease (APE1) as a promising tool for cancer treatment." Cell Mol Life Sci 67(21): 3589-3608.

Thakur, S., M. Dhiman, G. Tell and A. K. Mantha (2015). "A review on protein-protein interaction network of APE1/Ref-1 and its associated biological functions." Cell Biochem Funct 33(3): 101-112.

Thakur, S., M. Dhiman, G. Tell and A. K. Mantha (2015). "A review on protein-protein interaction network of APE1/Ref-1 and its associated biological functions." Cell Biochem Funct 33(3): 101-112.

Timson, D. J. and D. B. Wigley (1999). "Functional domains of an NAD+-dependent DNA ligase." J Mol Zhu, H. and S. Shuman (2005)

Timson, D. J. and D. B. Wigley (1999). "Functional domains of an NAD+-dependent DNA ligase." J Mol Biol 285(1): 73-83.

Tomkinson, A. E., S. Vijayakumar, J. M. Pascal and T. Ellenberger (2006). "DNA ligases: structure, reaction mechanism, and function." Chem Rev 106(2): 687-699.

Tomkinson, A. E., T. R. Howes and N. E. Wiest (2013). "DNA ligases as therapeutic targets." Transl Cancer Res 2(3).

Tomkinson, A. E., T. R. Howes and N. E. Wiest (2013). "DNA ligases as therapeutic targets." Transl Cancer Res 2(3).

Tyedmers, J., M. L. Madariaga and S. Lindquist (2008). "Prion switching in response to environmental stress." PLoS Biol 6(11): e294.

Tyedmers, J., M. L. Madariaga and S. Lindquist (2008). "Prion switching in response to environmental stress." PLoS Biol 6(11): e294.

Udwadia, Z. and D. Vendoti (2013). "Totally drug-resistant tuberculosis (TDR-TB) in India: every dark cloud has a silver lining." J Epidemiol Community Health 67(6): 471-472.

Udwadia, Z. and D. Vendoti (2013). "Totally drug-resistant tuberculosis (TDR-TB) in India: every dark cloud has a silver lining." J Epidemiol Community Health 67(6): 471-472.

Unciuleac, M. C., Y. Goldgur and S. Shuman (2017). "Two-metal versus one-metal mechanisms of lysine adenylylation by ATP-dependent and NAD(+)-dependent polynucleotide ligases." Proc Natl Acad Sci U S A 114(10): 2592-2597.

Unciuleac, M. C., Y. Goldgur and S. Shuman (2017). "Two-metal versus one-metal mechanisms of lysine adenylylation by ATP-dependent and NAD(+)-dependent polynucleotide ligases." Proc Natl Acad Sci U S A 114(10): 2592-2597.

Verhoeven, E. E., M. van Kesteren, J. J. Turner, G. A. van der Marel, J. H. van Boom, G. F. Moolenaar and N. Goosen (2002). "The C-terminal region of Escherichia coli UvrC contributes to the flexibility of the UvrABC nucleotide excision repair system." Nucleic Acids Res 30(11): 2492-2500.

Verhoeven, E. E., M. van Kesteren, J. J. Turner, G. A. van der Marel, J. H. van Boom, G. F. Moolenaar and N. Goosen (2002). "The C-terminal region of Escherichia coli UvrC contributes to the flexibility of the UvrABC nucleotide excision repair system." Nucleic Acids Res 30(11): 2492-2500.

Vidal, A. E., S. Boiteux, I. D. Hickson and J. P. Radicella (2001). "XRCC1 coordinates the initial and late stages of DNA abasic site repair through protein-protein interactions." EMBO J 20(22): 6530-6539.

Vidal, A. E., S. Boiteux, I. D. Hickson and J. P. Radicella (2001). "XRCC1 coordinates the initial and late stages of DNA abasic site repair through protein-protein interactions." EMBO J 20(22): 6530-6539.

Wallace, S. S. (2002). "Biological consequences of free radical-damaged DNA bases." Free Radic Biol Med 33(1): 1-14.

Wallace, S. S. (2002). "Biological consequences of free radical-damaged DNA bases." Free Radic Biol Med 33(1): 1-14.

Wallace, S. S. (2013). "DNA glycosylases search for and remove oxidized DNA bases." Environ Mol Mutagen 54(9): 691-704.

Wallace, S. S. (2013). "DNA glycosylases search for and remove oxidized DNA bases." Environ Mol Mutagen 54(9): 691-704.

Wallace, S. S. (2014). "Base excision repair: a critical player in many games." DNA Repair (Amst) 19: 14-26.

Wang, L. K., H. Zhu and S. Shuman (2009). "Structure-guided Mutational Analysis of the Nucleotidyltransferase Domain of Escherichia coli DNA Ligase (LigA)." J Biol Chem 284(13): 8486-8494.

Wang, L. K., H. Zhu and S. Shuman (2009). "Structure-guided Mutational Analysis of the Nucleotidyltransferase Domain of Escherichia coli DNA Ligase (LigA)." J Biol Chem 284(13): 8486-8494.

Wang, L. K., P. A. Nair and S. Shuman (2008). "Structure-guided mutational analysis of the OB, HhH, and BRCT domains of Escherichia coli DNA ligase." J Biol Chem 283(34): 23343-23352.

Wang, L. K., P. A. Nair and S. Shuman (2008). "Structure-guided mutational analysis of the OB, HhH, and BRCT domains of Escherichia coli DNA ligase." J Biol Chem 283(34): 23343-23352.

Ward, I., J. E. Kim, K. Minn, C. C. Chini, G. Mer and J. Chen (2006). "The tandem BRCT domain of 53BP1 is not required for its repair function." J Biol Chem 281(50): 38472-38477.

Ward, I., J. E. Kim, K. Minn, C. C. Chini, G. Mer and J. Chen (2006). "The tandem BRCT domain of 53BP1 is not required for its repair function." J Biol Chem 281(50): 38472-38477.

Waters, T. R., P. Gallinari, J. Jiricny and P. F. Swann (1999). "Human thymine DNA glycosylase binds to apurinic sites in DNA but is displaced by human apurinic endonuclease 1." J Biol Chem 274(1): 67-74.

Weller, G. R., B. Kysela, R. Roy, L. M. Tonkin, E. Scanlan, M. Della, S. K. Devine, J. P. Day, A. Wilkinson, F. d'Adda di Fagagna, K. M. Devine, R. P. Bowater, P. A. Jeggo, S. P. Jackson and A. J. Doherty (2002). "Identification of a DNA nonhomologous end-joining complex in bacteria." Science 297(5587): 1686-1689.

Weller, G. R., B. Kysela, R. Roy, L. M. Tonkin, E. Scanlan, M. Della, S. K. Devine, J. P. Day, A. Wilkinson, F. d'Adda di Fagagna, K. M. Devine, R. P. Bowater, P. A. Jeggo, S. P. Jackson and A. J. Doherty (2002). "Identification of a DNA nonhomologous end-joining complex in bacteria." Science 297(5587): 1686-1689.

Whitaker, A. M. and B. D. Freudenthal (2018). "APE1: A skilled nucleic acid surgeon." DNA Repair (Amst) 71: 93-100.

Wilkinson, A., A. Smith, D. Bullard, M. Lavesa-Curto, H. Sayer, A. Bonner, A. Hemmings and R. Bowater (2005). "Analysis of ligation and DNA binding by Escherichia coli DNA ligase (LigA)." Biochim Biophys Acta 1749(1): 113-122.

Wilkinson, A., A. Smith, D. Bullard, M. Lavesa-Curto, H. Sayer, A. Bonner, A. Hemmings and R. Bowater (2005). "Analysis of ligation and DNA binding by Escherichia coli DNA ligase (LigA)." Biochim Biophys Acta 1749(1): 113-122.

Williamson, A. and H. S. Leiros (2019). "Structural intermediates of a DNA-ligase complex illuminate the role of the catalytic metal ion and mechanism of phosphodiester bond formation." Nucleic Acids Res 47(14): 7147-7162.

Williamson, A. and H. S. Leiros (2019). "Structural intermediates of a DNA-ligase complex illuminate the role of the catalytic metal ion and mechanism of phosphodiester bond formation." Nucleic Acids Res 47(14): 7147-7162.

Wilson, D. M., 3rd, M. Takeshita and B. Demple (1997). "Abasic site binding by the human apurinic endonuclease, Ape, and determination of the DNA contact sites." Nucleic Acids Res 25(5): 933-939.

Wilson, D. M., 3rd, M. Takeshita and B. Demple (1997). "Abasic site binding by the human apurinic endonuclease, Ape, and determination of the DNA contact sites." Nucleic Acids Res 25(5): 933-939.

Wilson, S. H. and T. A. Kunkel (2000). "Passing the baton in base excision repair." Nat Struct Biol 7(3): 176-178.

Wilson, S. H. and T. A. Kunkel (2000). "Passing the baton in base excision repair." Nat Struct Biol 7(3): 176-178.

Wood, R. D. (1996). "DNA repair in eukaryotes." Annu Rev Biochem 65: 135-167.

Wood, R. D. (1996). "DNA repair in eukaryotes." Annu Rev Biochem 65: 135-167.

Xia, L., L. Zheng, H. W. Lee, S. E. Bates, L. Federico, B. Shen and T. R. O'Connor (2005). "Human 3-methyladenine-DNA glycosylase: effect of sequence context on excision, association with PCNA, and stimulation by AP endonuclease." J Mol Biol 346(5): 1259-1274.

Xia, L., L. Zheng, H. W. Lee, S. E. Bates, L. Federico, B. Shen and T. R. O'Connor (2005). "Human 3-methyladenine-DNA glycosylase: effect of sequence context on excision, association with PCNA, and stimulation by AP endonuclease." J Mol Biol 346(5): 1259-1274.

Yadav, N., T. Khanam, A. Shukla, N. Rai, K. Hajela and R. Ramachandran (2015). "Tricyclic dihydrobenzoxazepine and tetracyclic indole derivatives can specifically target bacterial DNA ligases and can distinguish them from human DNA ligase I." Org Biomol Chem 13(19): 5475-5487.

Yadav, N., T. Khanam, A. Shukla, N. Rai, K. Hajela and R. Ramachandran (2015). "Tricyclic dihydrobenzoxazepine and tetracyclic indole derivatives can specifically target bacterial DNA ligases and can distinguish them from human DNA ligase I." Org Biomol Chem 13(19): 5475-5487.

Yadav, N., T. Khanam, A. Shukla, N. Rai, K. Hajela and R. Ramachandran (2015). "Tricyclic dihydrobenzoxazepine and tetracyclic indole derivatives can specifically target

bacterial DNA ligases and can distinguish them from human DNA ligase I." Org Biomol Chem 13(19): 5475-5487.

Zhu, H. and S. Shuman (2005). "Structure-guided mutational analysis of the nucleotidyltransferase domain of Escherichia coli NAD+-dependent DNA ligase (LigA)." J Biol Chem 280(13): 12137-12144.

Zhu, H. and S. Shuman (2005). "Structure-guided mutational analysis of the nucleotidyltransferase domain of Escherichia coli NAD+-dependent DNA ligase (LigA)." J Biol Chem 280(13): 12137-12144.

Zhu, H. and S. Shuman (2010). "Gap filling activities of Pseudomonas DNA ligase D (LigD) polymerase and functional interactions of LigD with the DNA end-binding Ku protein." J Biol Chem 285(7): 4815-4825.

Zhu, H. and S. Shuman (2010). "Gap filling activities of Pseudomonas DNA ligase D (LigD) polymerase and functional interactions of LigD with the DNA end-binding Ku protein." J Biol Chem 285(7): 4815-4825.

Appendix

Composition for buffers, reagents and medium:

1. 50X TAE Buffer (1 L) - 242 g Tris Base 57.1 mL glacial acetic Acid 100 mL 0.5 M EDTA Make up to 1 liter with dH_2O

2. 10X TBE (1 L) - 108 g Tris 55 g boric acid 40 ml 0.5M EDTA Make up to 1 liter with dH_2O

3. 10x PBS (1 L)- 5.75 g disodium hydrogen phosphate 1.48 g sodium dihydrogen phosphate 9 g sodium Chloride Make up to 1 liter with dH_2O

4. Coomassie brilliant blue R250 stain (500 ml)- 0.5 g Coomassie Brilliant Blue R250 50 ml glacial acetic acid 450 ml methanol and water (1:1 v/v)

5. Urea PAGE loading dye- 0.1% w/v bromophenol blue 0.1% w/v xylene cyanol 95% formamide deionized 25 mM EDTA pH 8.0

6. 5X Laemmli Buffer (20 ml)- 5 ml 1 M Tris (pH 6.8) 2 g SDS 10 ml glycerol 700 µl β-mercaptoethanol Make up to 20 ml by adding dH_2O

7. Terrific Broth medium- 12 g tryptone 24 g yeast extract 4 ml glycerol Make volume up to 900 ml. Autoclave and add 100 ml sterile phosphate buffer.

8. Phosphate buffer for TB medium (100 ml) -2.31 g KH_2PO_4 12.54 g K_2HPO_4

9. Transfer buffer (1 L) -2.42 g Tris 11.26 g glycine Make up to 800 ml with dH_2O and add 200 ml methanol

10. Tris-glycine-SDS running buffer (1 L) -3 g Tris 14.4 g glycine 10 ml of 10% SDS Make up to 1 liter with dH_2O

11. Tris-EDTA Buffer (TE buffer) - 10mM Tris pH 8.0, 0.1 mM EDTA (pH 8.0). For 100 ml, 1 ml 1M Tris pH 8.0 and 20 µl 0.5 M EDTA added and volume made up to 100 ml with dH_2O

12. 10 X Orange G loading dye (10 ml) -20 mg Orange G 3 ml glycerol 7 ml dH_2O

13. Ponceau S stain - 0.1% w/v in 5% glacial acetic acid For 50 ml, add 50 mg Ponceaus S to 2.5 ml GAA + 48 ml dH_2O

Composition of SDS polyacrylamide gel:

Resolving gel (10 ml)

Reagent	12%
1.1 M Tris pH=8.8	2ml
30% acrylamide :bis acryalmide (29:1)	4.0ml
10% SDS	100µl
10% APS	100µl
TEMED	20µl
dH2O	3.33ml

Stacking gel (5 ml)

Reagent	5%
1 M Tris pH= 6.8	1.25ml
30% acrylamide :bis acryalmide (29:1)	0.6ml
10% SDS	50µl
10% APS	50µl
TEMED	10µl
dH2O	3.0ml

Composition of Urea polyacrylamide gel:

15% UREA PAGE (10 ml) - 1 ml 10 X TBE 7.5 ml 20% Acrylamide/bisacrylamide (19:1) 4.8 g Urea Components were mixed and warm at 37°C until warm. Added 100 µl 10% APS and 20 µl TEMED for polymerization.

6% UREA PAGE (10 ml) - 1 ml 10 X TBE 3 ml 20% Acrylamide/bisacrylamide (19:1) 4.8 g Urea Components were mixed and warm at 37°C until warm. Added 100 µl 10% APS and 20 µl TEMED for polymerization.

Composition of Native polyacrylamide gel (10ml) :

3 ml 20% Acrylamide/bisacrylamide (19:1) 2.5ml of 1M Tris pH 8.8 Components were mixed and warm at 37°C until warm. Added 100 µl 10% APS and 20 µl TEMED for polymerization.

List of Publications

1) Tricyclic dihydrobenzoxazepine and tetracyclic indole derivatives can specifically target bacterial DNA ligases and can distuinguish them from human DNA ligase I, N Yadav, T Khanam, **A Shukla,** N Rai, K Hajela, R Ramachandran. Organic & biomolecular chemistry 13 (19), 5475-5487, 2015.
2) Crticial determinants for substrate recognition and catalysis in the *M. tuberculosis* class II AP-endonuclease/3'-5' exonuclease III, T Khanam, **A Shukla**, N Rai, R Ramachandran. Biochemica et Biophysica Acta (BBA)- Proteins and Proteomics 1854 (5), 505-516, 2015.
3) M. tuberculosis class II apurinic/apyrimidinic-endonuclease/3'-5' exonuclease (XthA) engages with NAD^+ -dependent DNA ligase A (LigA) to counter futile cleaavge and ligation cycles in Base Excision Repair, T Khanam*, M Afsar*, **A Shukla***, F. Alam, S Kumar, H Soyar, K Dolma, Ashish, M Pashupileti, K K Srivastava, R S Ampapathi, R Ramachandran. Nucleic Acid Research (Under revision) *Authors contibuted equally.