SUPPLEMENTARY INFORMATION

M. tuberculosis class II apurinic/apyrimidinic-endonuclease/3′-5′ exonuclease (XthA) engages with NAD+ -dependent DNA ligase A (LigA) to counter futile cleavage and ligation cycles in Base Excision Repair

Taran Khanam[1‡†], Mohammad Afsar[1‡], Ankita Shukla[1‡], Faiyaz Alam[2], Sanjay Kumar[1], Horam Soyar[3], Kunzes Dolma[4], Ashish[4], Mukesh Pasupuleti[3], Kishore Kumar Srivastava[3], Ravi Sankar Ampapathi[2] & Ravishankar Ramachandran[1*]

Table S1. List of primers used to generate mutants:

MtbLigA^{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′
MtbLigA^{G621I}	Forward	5′- GTC ACC ATA TCG CTG ACC GGT TTC TCC-3′
	Reverse	5′- GGT TAG CGABTAT GGT GAC CAC GAT GG-3′
MtbLigA^{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′
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′

Table S2. SAXS data fitting and validation

Sample	Dammif		Supcomb	Crysol	EOM
	χ^2	NSD	NSD	χ^2	χ^2
MtbXthA	0.5765	0.514±0.06	1.94	-	-
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
MtbXthA+MtbLigA	1.02	0.645±0.041	2.94	-	-
MtbXthA+MtbLigA+DNA	0.8048	0.695±0.02	2.8	1.951	-

NSD = normalized spatial discrepancy

Table S3. DNA substrates used in the study:

Substrate	Oligonucleotides
DNA substrate used in SAXS and SEC analysis **S1**	**27mer B3 lower** 5'- GGT AGA TCA GTG TCT AAT GTA TGT CAG-3' **17mer U6 upper** 5'-CTG ACA TAC ATT AGA CA-3' **10mer U7 upper** 5'-CTGATCTACC-3'
DNA substrate used in AP site incision assay **N1**	**75mer B1 lower** 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' **75mer U1 upper** 5'-6FAM-CAT TAC AGC GCC TGC AAG CTC AAT TAA TCG TA**F** TGC TAC ACT GGA GTA CCG GAG CGT CGT CAT GAC AAC GAA TGG-3' **F**= Tetrahyrofuran (abasic analog)
DNA substrate used in nick ligation activity assay **L1**	**40mer B2 lower** 5'-ATG TCC AGT GAT CCA GCT AAG GTA CGA GTC TAT GTC CAG G-3' **18mer U2 upper** 5'-pAGC TGG ATC ACT GGA CAT-6-FAM3' **22mer U3 upper** 5'-CCT GGA CAT AGA CTC GTA CCT T-3'
L2	**40mer B2 lower** 5'-ATG TCC AGT GAT CCA GCT AAG GTA CGA GTC TAT GTC CAG G-3' **18mer U4 upper** 5'-p**F**GC TGG ATC ACT GGA CAT-3' **22mer U5 lower** 5'-6FAM-CCT GGA CAT AGA CTC GTA CCT T-3'
L3	**40mer B2 lower** 5'-ATG TCC AGT GAT CCA GCT AAG GTA CGA GTC TAT GTC CAG G-3' **18mer U4 upper** 5'-p**U**GC TGG ATC ACT GGA CAT-3' **U**= Uracil base **22mer U5 upper** 5'-6FAM-CCT GGA CAT AGA CTC GTA CCT T-3'

DNA substrates were prepared by mixing 1 µM of the upper and lower oligonucleotides (labelled as U and B, respectively) in annealing buffer (10 mM Tris-HCl, pH 7.8, 50 mM KCl and 1mM EDTA). The mixture was heated to 95˚C for 5 min and gradually cooled to room temperature. All the oligos were purchased from M/S Integrated DNA technologies Inc.

Table S4: The Maximum response (R_{max}), Chi^2 and KD values obtained by SPR

Interactions	R_{max}	Chi^2	K_D value (M)
MtbXthA-MtbLigA	49.1	3.19	$6.77e^{-8}$
MtbXthA-BRCT domain	61.1	3.58	$6.77e^{-8}$
$MtbXthA^{ID}$-MtbLigA	65.1	1.32	$6.13e^{-7}$
MtbXthA- $MtbLigA^{G614I}$	24.9	0.217	$7.81e^{-7}$
MtbXthA- $MtbLigA^{G621I}$	16	0.268	$1.54e^{-7}$
MtbXthA-$MtbLigA^{G639V}$	27.6	0.152	$2.77e^{-7}$
$MtbXthA^{ID}$-$MtbLigA^{G639V}$	4.64e-7	0.173	9.01e-6

Table S5. Thermodynamic parameters of MtbXthA derived peptide binding to BRCT domain of MtbLigA obtained from ITC experiments

Ligand / Concentration (mM)	BRCT (µM)	N	K_d (µM)	ΔG (kcal mol^{-1})	ΔH (kcal mol^{-1})	-TΔS (kcal mol^{-1})
DGQ (0.4)	5	2.44	2.4752	-4.98729	-9.958	-14.94529
YDV (0.4)	5	4.19	88.491	-5.63434	0.9137	-6.54804

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

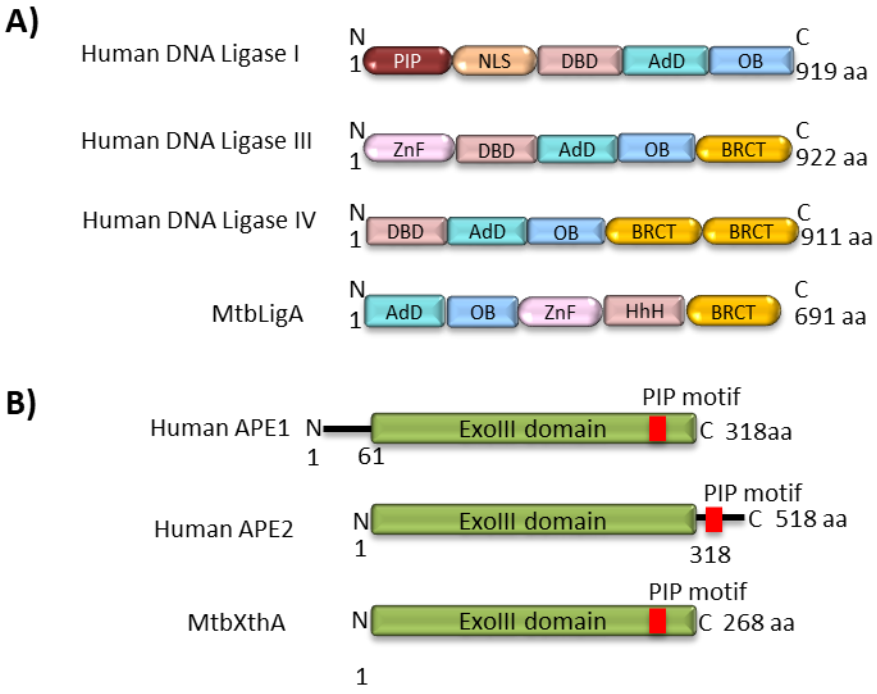


FIGURE S1. Architectural differences in eukaryotic and prokaryotic DNA ligase and AP endonuclease homologs . **(A)** Eukaryotic and mycobacterial DNA ligases have functionally similar domains, but they are jumbled in their linear sequences with other minor differences. AdD is adenylation domain, OB fold is oligomer binding fold, ZnF stands for Zinc finger, HhH represents helix hairpin helix, BRCT denotes BRCA1 C terminal like domain. **(B)** 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.

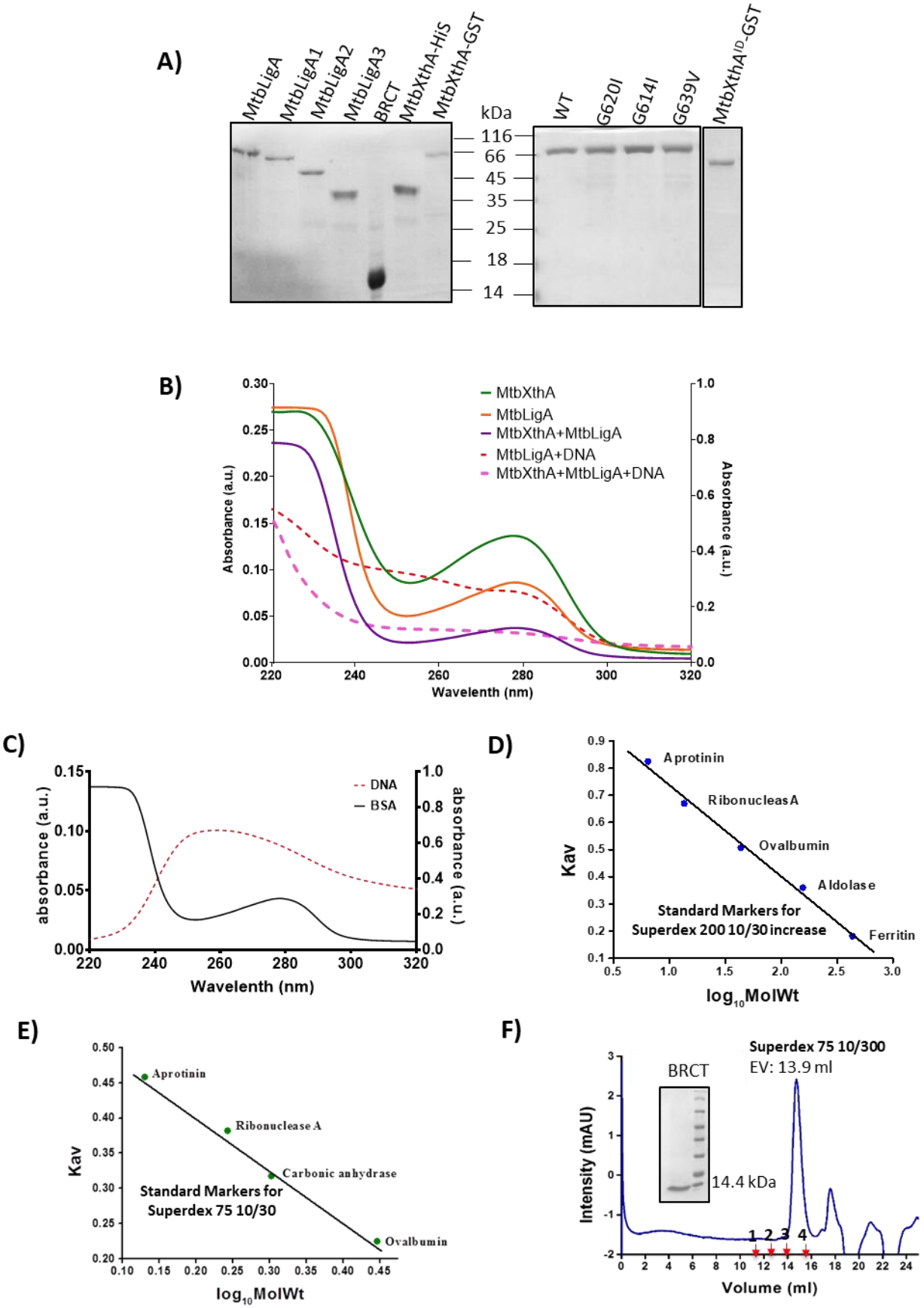

A)
MtbLigA
MtbLigA1
MtbLigA2
MtbLigA3
BRCT
MtbXthA-HiS
MtbXthA-GST
kDa
116
66
45
35
25
18
14
WT
G620I
G614I
G639V
MtbXthA^ID-GST
B)
MtbXthA
MtbLigA
MtbXthA+MtbLigA
MtbLigA+DNA
MtbXthA+MtbLigA+DNA
Absorbance (a.u.)
Wavelenth (nm)
C)
DNA
BSA
absorbance (a.u.)
Wavelenth (nm)
D)
Aprotinin
RibonucleasA
Ovalbumin
Aldolase
Ferritin
Standard Markers for
Superdex 200 10/30 increase
Kav
log10MolWt
E)
Aprotinin
Ribonuclease A
Carbonic anhydrase
Ovalbumin
Standard Markers for
Superdex 75 10/30
Kav
log10MolWt
F)
Superdex 75 10/300
EV: 13.9 ml
BRCT
14.4 kDa
Intensity (mAU)
Volume (ml)
1 2 3 4

FIGURE S2. Purification of recombinant protein components. (A) SDS-PAGE analysis of purified recombinant proteins used in the present study. Electrophoresis was performed using 10% polyacrylamide gel, followed by Coomassie Blue staining. **(B)** UV-Visible spectral analysis of eluted peak fractions of proteins/complexes in SEC. The corresponding SEC profiles are shown in main text in Fig. 1E. (**C**) UV/Vis spectra of BSA (black) showing peak maxima at 280 nm and nicked ds-DNA showing peak maxima at 260nm (dotted red), used as controls. **(D)** Standard curve for calibration of S200 10/30 increase size exclusion column. (**E**) 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, 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. **(F)** 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).

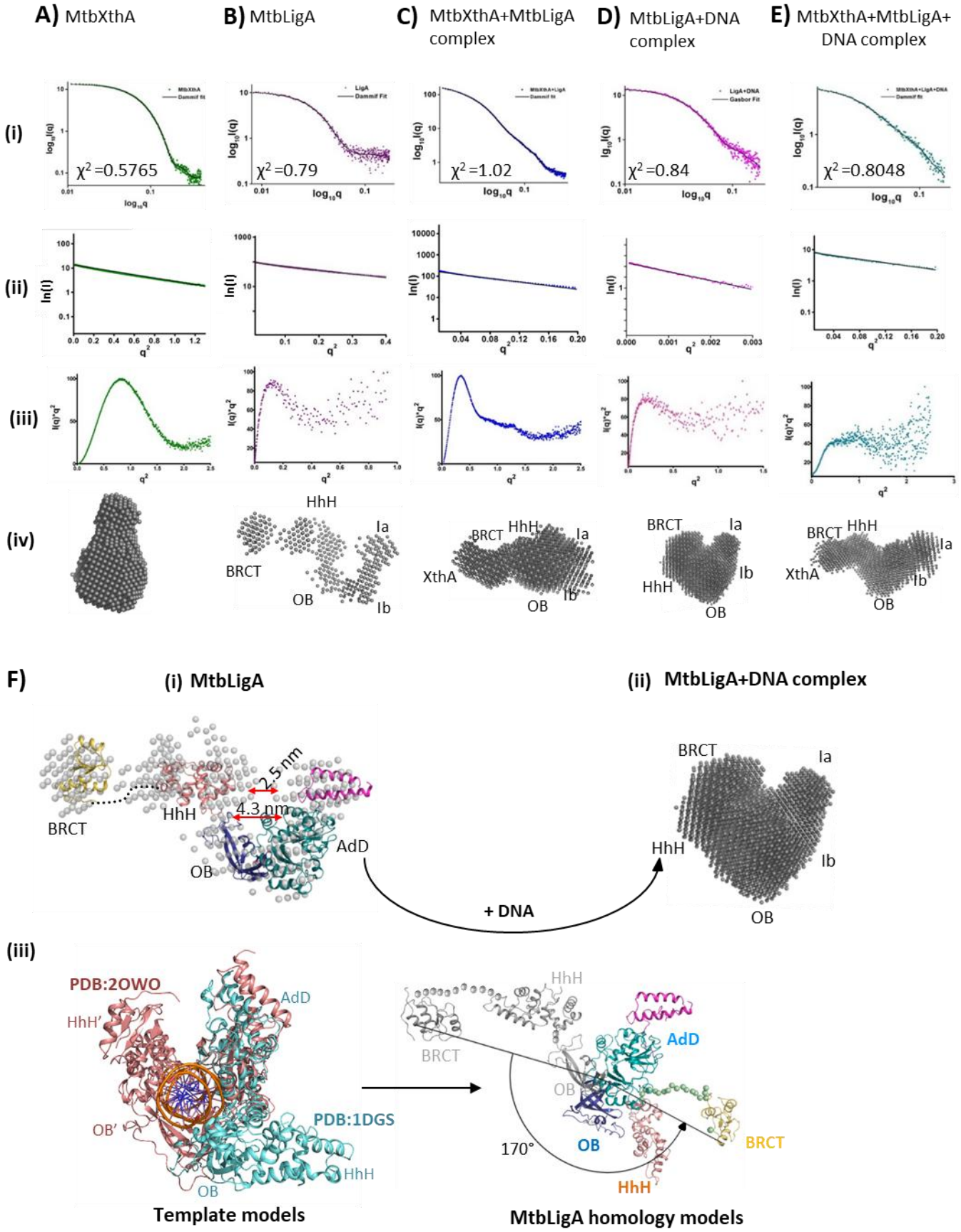

A) MtbXthA
B) MtbLigA
C) MtbXthA+MtbLigA complex
D) MtbLigA+DNA complex
E) MtbXthA+MtbLigA+ DNA complex
(i)
(ii)
(iii)
(iv)
χ2 =0.5765
χ2 =0.79
χ2 =1.02
χ2 =0.84
χ2 =0.8048
HhH
Ia
BRCT
OB
Ib
XthA
F)
(i) MtbLigA
(ii) MtbLigA+DNA complex
2.5 nm
4.3 nm
AdD
+ DNA
(iii)
PDB:2OWO
HhH'
OB'
PDB:1DGS
Template models
170°
MtbLigA homology models

FIGURE S3. Small angle X-Ray scattering. (**i and ii**) Guinier plots (**iii**) Normalised Kratky plots indicating degree of disorder of respective protein and complexes (**iv**) Bead models build by DAMMIF using SAXS data, for **(A)** MtbXthA **(B)** MtbLigA **(C)** MtbXthA-MtbLigA binary complex **(D)** MtbLigA-DNA complex **(E)** MtbXthA-MtbLigA-DNA ternary complex. **(F)** Energy minimized homology model of MtbLigA is superimposed within the envelope encompassing the averaged ab initio model in (**i**) the absence and (**ii**) presence of nicked DNA substrate. The homology models for apo and DNA bound MtbLigA were built using crystal structures of *Thermus filiformis* LigA (PDB:1DGS) and *E.coli* LigA bound to nicked DNA (PDB:2OWO) as templates, respectively. (**iii**) **Left**: The template structures were superimposed at Ib domain and Y324 residue (present at the linker region between the AdD domain and OB domain) was used as reference to measure conformational changes in MtbLigA in the presence of nicked DNA. Gray and colored model depicts apo MtbLigA and DNA bound MtbLigA, respectively. Rotation of nearly 170° was seen in MtbLigA Ia subdomain in the presence of nicked DNA.

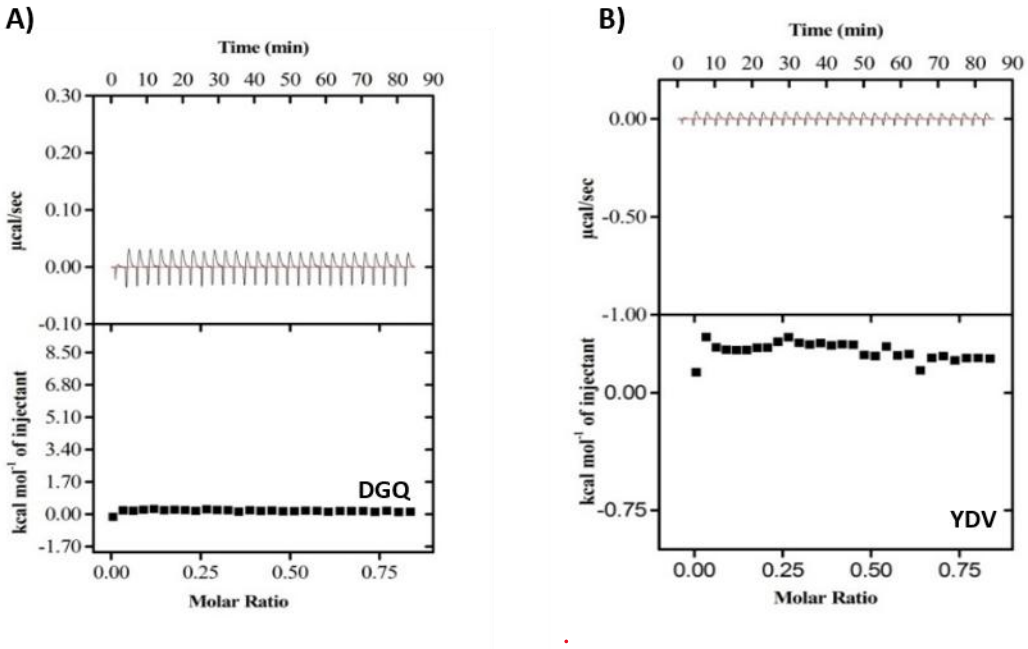


FIGURE S4. The control experiments involving interaction between **(A)** DGQ and **(B)** YDV peptide with reaction buffer

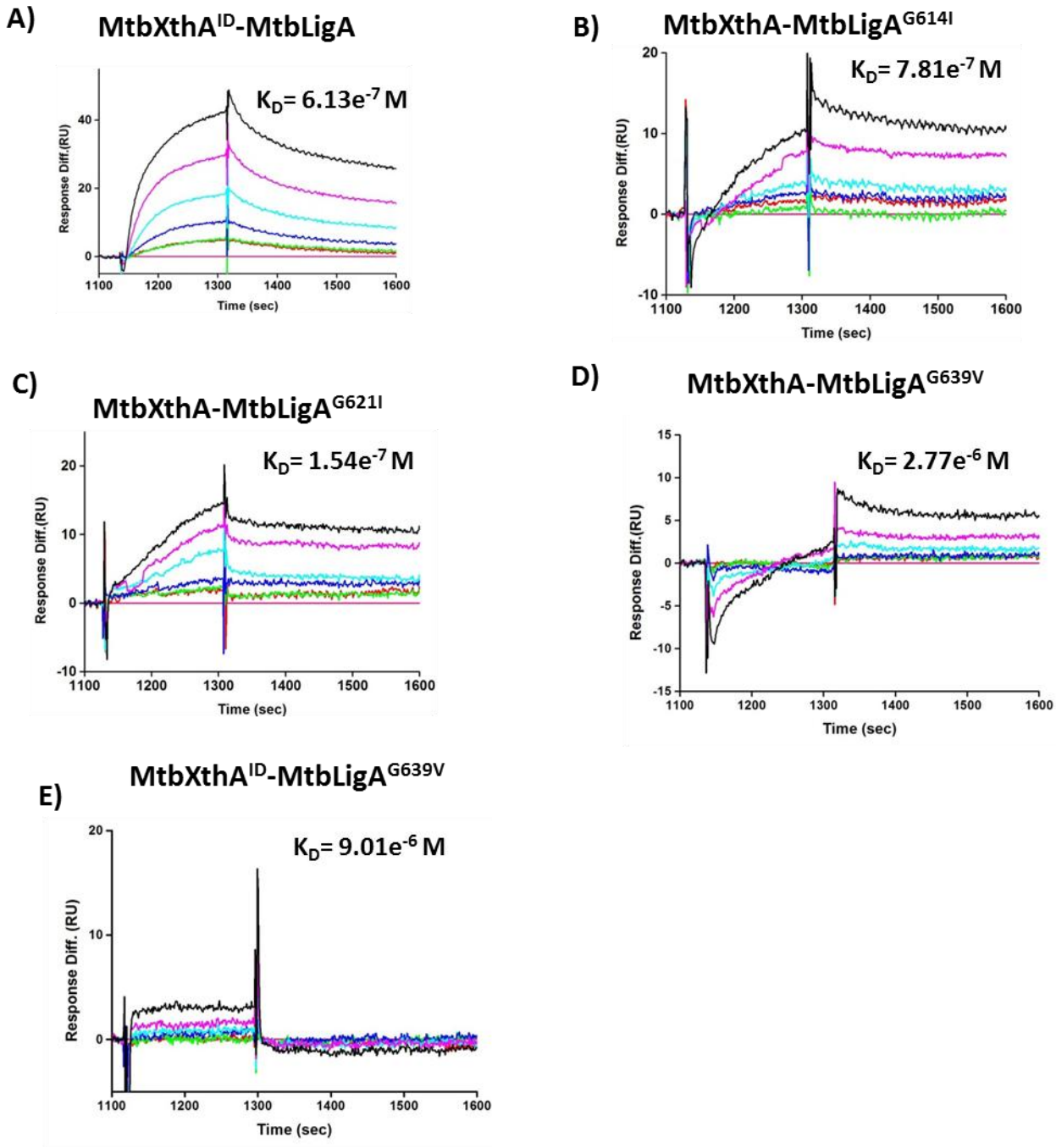
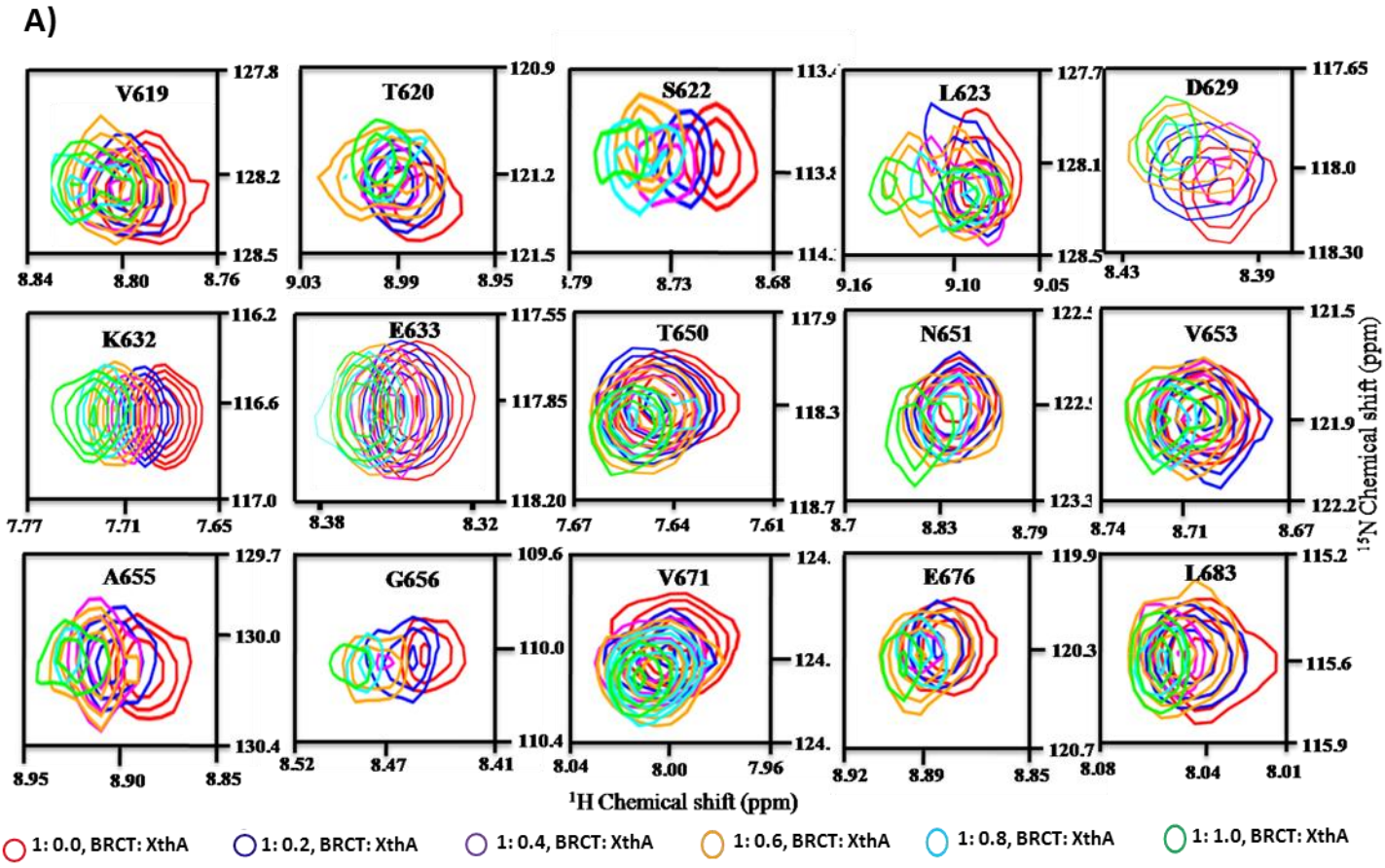
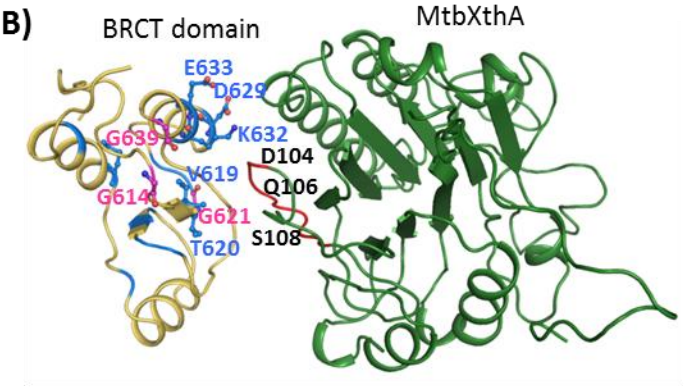
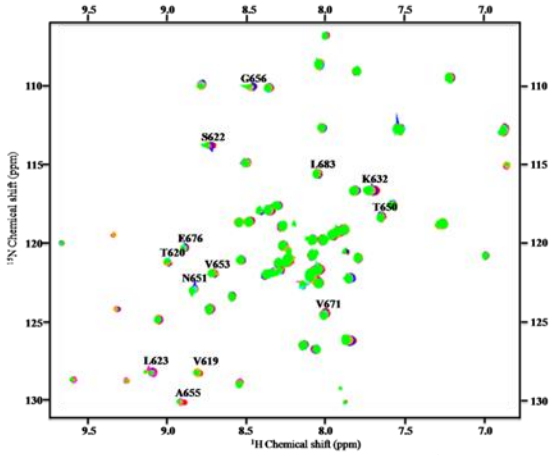


FIGURE S5. Surface plasma resonance (SPR): The SPR assay to characterize the binding kinetics of MtbXthA and MtbLigA mutants. The binding kinetics of **(A)** MtbXthA^{ID} and MtbLigA **(B)** MtbXthA and MtbLigA^{G614I} **(C)** MtbXthA and MtbLigA^{G621I} **(D)** MtbXthA and MtbLigA^{G639V} and **(E)** MtbXthA^{ID} and MtbLigA^{G639V}. The respective K_D values are mentioned in the figure.







D)

```
                590       600       610       620       630       640
M.tb            VDKWRAAGVRMVDER.....DESVPRTLAGLTIVVTGSLTGFSRDDAKEAIVARGGKAAG
M.smeg          VDKWRAAGVRMADER.....DASIERTLEGLSIVVTGSLAGFSRDQAKEAIIARGGKAAG
T.filiformis    VRRLKEAGVSMES.......KEEVSDLLSGLTFVLTGELSR.PREEVKALLQRLGAKVTD
E.coli          ISELLAEGV..HWPAPIVINAEEIDSPFAGKTVVLTGSLSQMSRDDAKARLVELGAKVAG
B.subtilis      LNELQELGVNTLYKGPKKVKAEDSDSYFAGKTIVLTGKLEELSRNEAKAQIEALGGKLTG
S.aurese        IQKLKDKHVNMIYKGIKTSDIE.GHPEFSGKTIVLTGKLHQMTRNEASKWLASQGAKVTS

                650       660       670       680       690
M.tb            SVSKKTNYVVAGDSPGSKYDKAVELGVPILDEDGFRRLLADGPASRT..
M.smeg          SVSKKTAYVVAGDAPGSKYDKAVELGVPVLDEDGFRRLLEQGPPVEPAE
T.filiformis    SVSRKTSYLVVGENPGSKLEKARALGVAVLTEEEFWRFLKEKGAPVPA.
E.coli          SVSKKTDLVIAGEAAGSKLAKAQELGIEVIDEAEMLRLLGS........
B.subtilis      SVSKNTDLVIAGEAAGSKLTKAQELNIEVWNEEQLMGELKK........
S.aurese        SVTKNTDVVIAGEDAGSKLTKAQSLGIEIWTEQQFVDKQNELNS.....
```

FIGURE S6. (A) Expansion of $^{15}N/^{1}H$-HSQC spectrum of BRCT domain showing chemical shift perturbation in the residues shown here. The red, blue, purple, orange, cyan and green color represents 1:0, 1:0.2, 1:0.4, 1:0.8 and 1:1 BRCT:XthA titration ratio. (**B)** The docked model of BRCT domain and XthA showing the binding of $_{104}$DGQ$_{112}$ motif of XthA near the V619, T620, G621, D629, K633 and G639 residues. **(C)** $^{15}N/^{1}H$-HSQC spectrum of BRCT domain titrated with unlabelled XthA protein from 1: 0 to 1:1 titration ratio. Residues which shows significant changes in average chemical shift are labeled. **(D)** Multiple sequence alignment of BRCT domain of LigA homologs. The highly conserved residues are boxed in red. The amino acid sequence of LigA homologs were aligned from *M. tuberculosis*, *M. smegmatis*, *M. filiformiss*, *Escherichia coli, Bacillus subtilis, Staphylococcus aurese.*

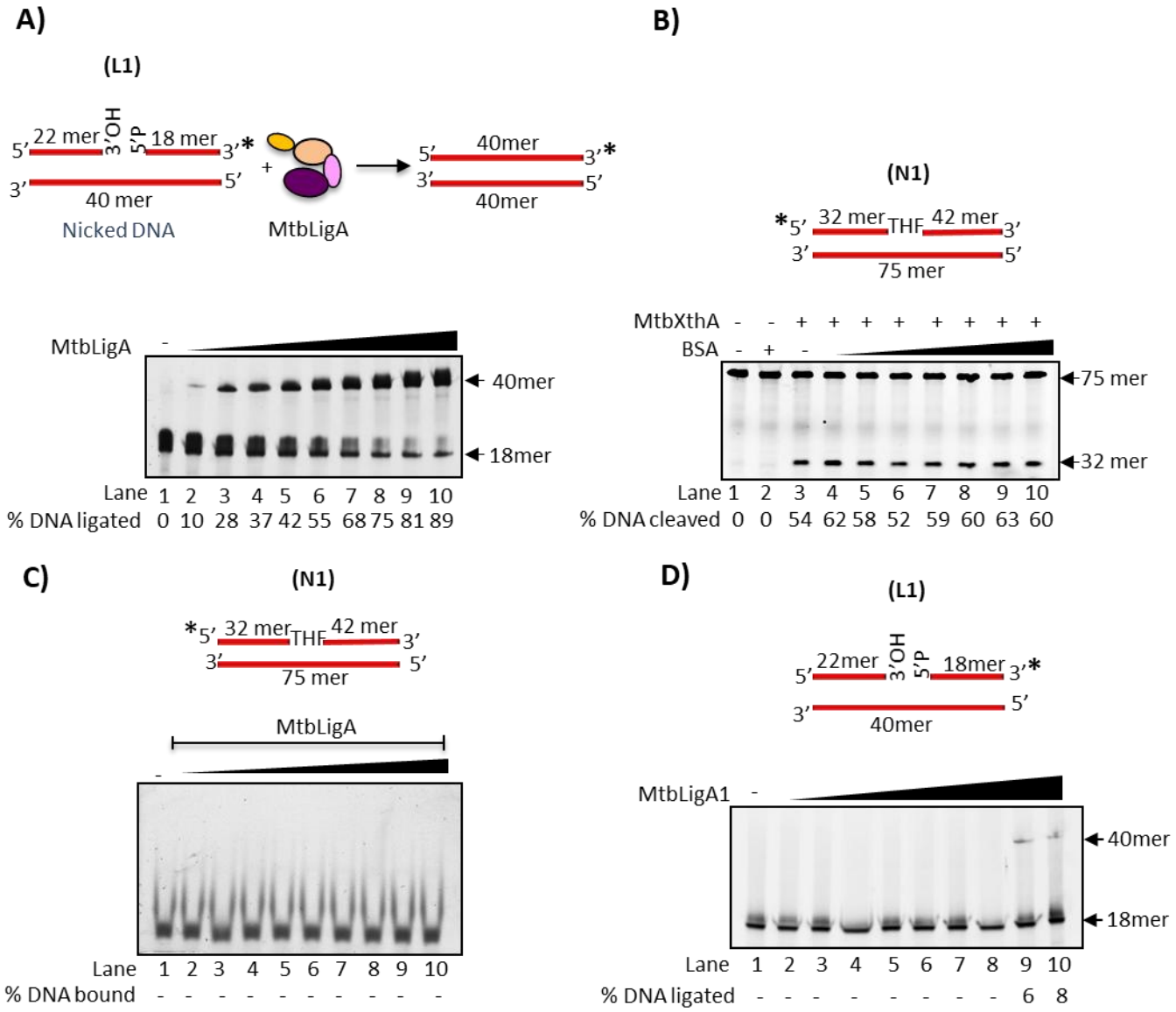


FIGURE S7. (A) Top; Diagram to show the action of MtbLigA on the nicked duplex DNA substrate (**L1**) used in the ligation assays. (Bottom) Ligation activity with respect to increasing concentration of MtbLigA. **(B)** Effect of BSA on AP endonuclease activity of MtbXthA. **(C)** Gel shift assay to monitor the affinity of MtbLigA for abasic site (THF) containing DNA substrate (N1). 10 nM abasic DNA (N1) was incubated with increasing concentration (0, 0.01, 0.03, 0.1, 0.3, 1, 3, 10, 20, 40 µM) of MtbLigA (lanes 1-10). **(D)** Ligation activity with respect to increasing concentration of BRCT deleted mutant MtbLigA1. In **A**, **B**, **D** and **E**, DNA substrate (10 nM) was incubated without protein (lane 1) or with increasing concentrations of MtbLigA/MtbLigA mutants (0.01, 0.05, 0.1, 0.5, 1, 2, 3, 5, 10, 20 nM) (lane 2 onwards). Reaction products were analyzed on 8 M urea-12% polyacrylamide gels for **A**, **B**, **D** and **E**, while 6% native polyacrylamide gel was used for product analysis in **C**, gel shift assay. Intensity of the fluorescent bands corresponding to the products of respective activities were scanned and quantified by using ImageQuant LAS 4000 and ImageQuantTL 8.1 software (GE Healthcare). The images are single representative image of experiments carried out in duplicate. A standard deviation of ± 3 was

obtained for the % DNA ligated. Asterisks denote the 6-FAM labelled termini. The positions of the ligated and incised products are indicated at 40mer and 75mer, respectively.

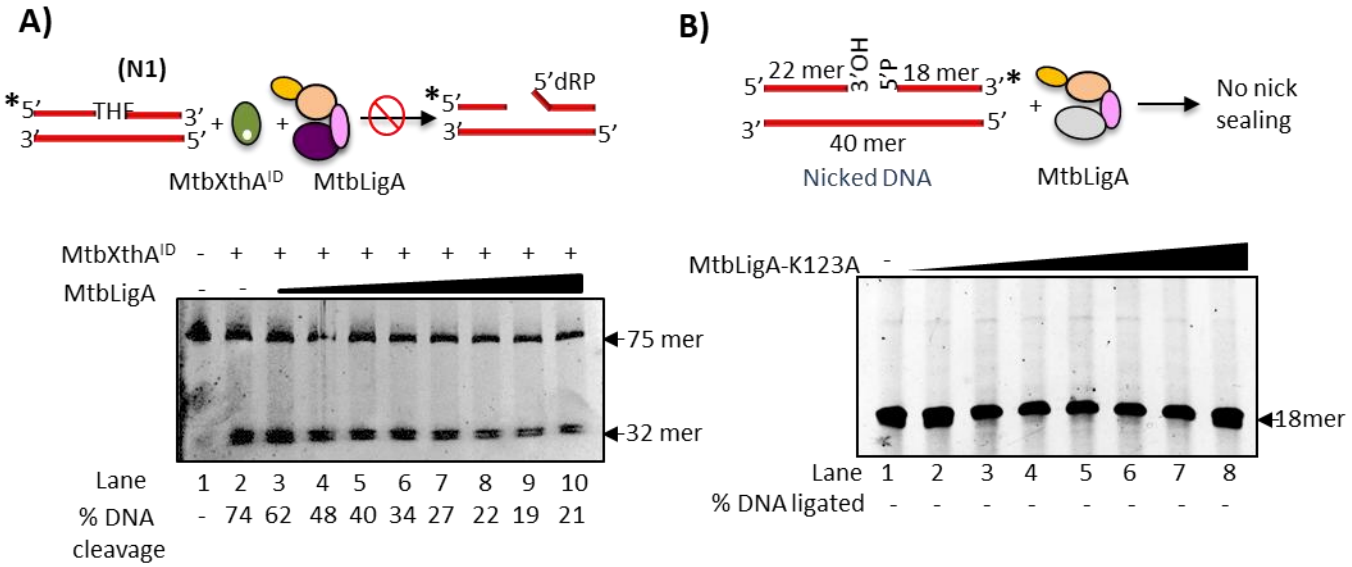
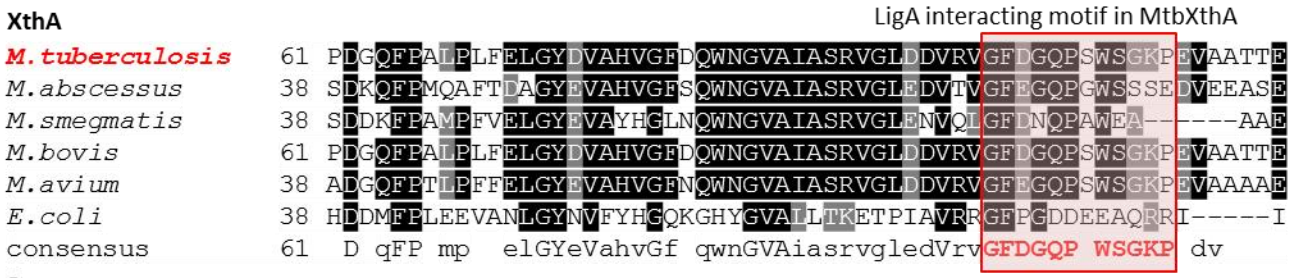
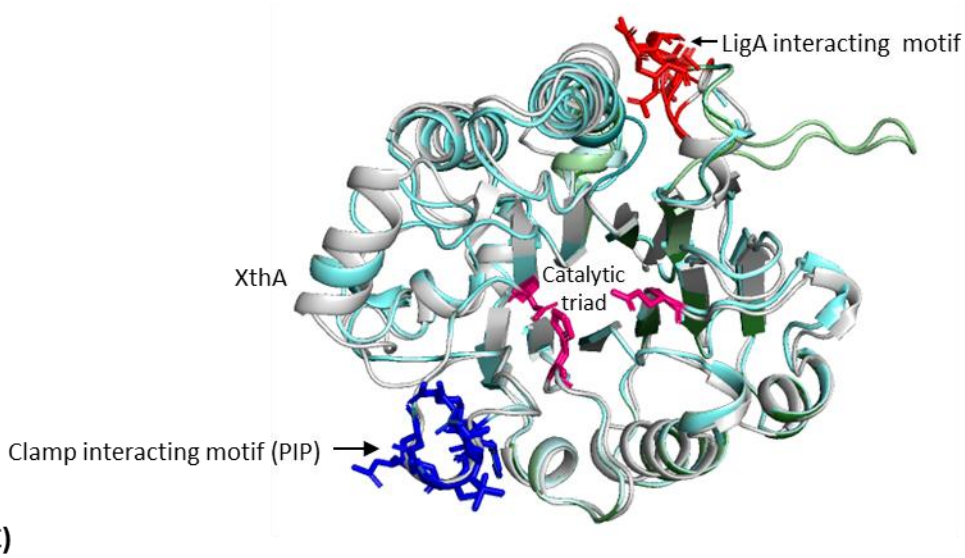


FIGURE S8. (A) Effect of LigA on the endonuclease activity of MtbXthA^{ID} mutant. The DNA substrate (20 nM) was incubated with MtbXthA^{ID} in the absence (Lane 2; 10 nM) and presence of increasing concentration of MtbLigA (lane 3-10; 5, 10, 20, 40, 60, 80, 100 nM) **(B)** The catalytic site mutant MtbLigA-K123A is inactive in sealing the nick. The DNA substrate (10 nM) was incubated with increasing concentration of MtbLigA-K123A (lane 2-8; 0.1, 0.5, 1, 2, 3, 5, 10, nM). Reaction products were analyzed on 8 M urea-12% polyacrylamide gels for **A** and **B.** Intensity of the fluorescent bands corresponding to the products of respective activities were scanned and quantified by using ImageQuant LAS 4000 and ImageQuantTL 8.1 software (GE Healthcare).

A)


XthA
LigA interacting motif in MtbXthA
M.tuberculosis 61 PDGQFPALPLFELGYDVAHVGEDQWNGVAIASRVGLDDVRVGFDGQPSWSGKPEVAATTE
M.abscessus 38 SDKQFPMQAFTDAGYEVAHVGESQWNGVAIASRVGLEDVTVGFEGQPGWSSSEDVEEASE
M.smegmatis 38 SDDKFPAMPFVELGYEVAYHGLNQWNGVAIASRVGLENVQLGFDNQPAWEA------AAE
M.bovis 61 PDGQFPALPLFELGYDVAHVGEDQWNGVAIASRVGLDDVRVGFDGQPSWSGKPEVAATTE
M.avium 38 ADGQFPTLPFFELGYEVAHVGENQWNGVAIASRVGLDDVRVGFEGQPSWSGKPEVAAAAE
E.coli 38 HDDMFPLEEVANLGYNVFYHGQKGHYGVALLTKETPIAVRRGFPGDDEEAQRRI-----I
consensus 61 D qFP mp elGYeVahvGf qwnGVAiasrvgledVrvGFDGQP WSGKP dv
e

B)


LigA interacting motif
XthA
Catalytic triad
Clamp interacting motif (PIP)

C)

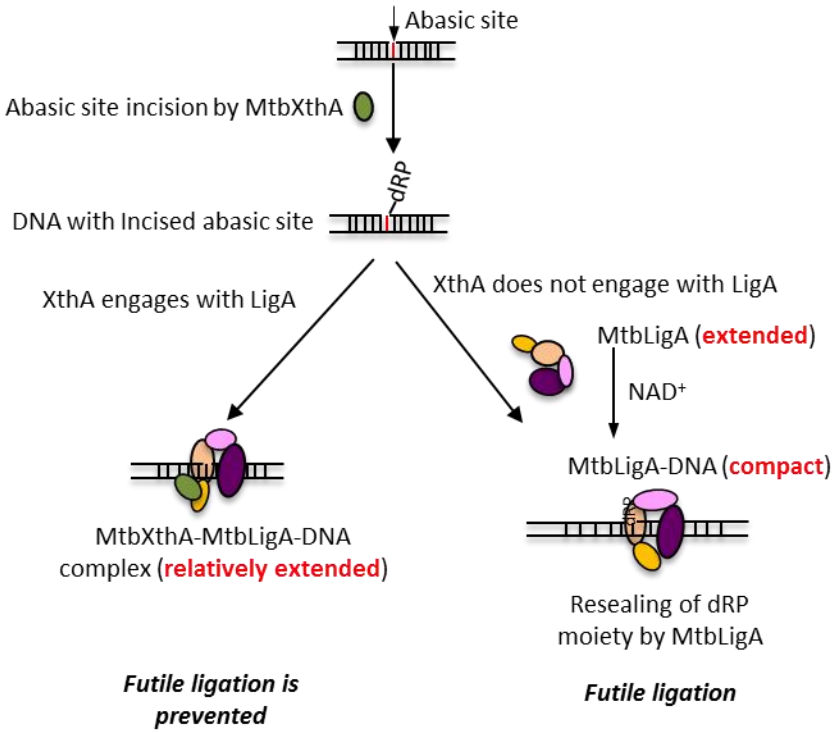

Abasic site
Abasic site incision by MtbXthA
dRP
DNA with Incised abasic site
XthA engages with LigA
XthA does not engage with LigA
MtbLigA (extended)
NAD+
MtbLigA-DNA (compact)
MtbXthA-MtbLigA-DNA complex (relatively extended)
Resealing of dRP moiety by MtbLigA
Futile ligation is prevented
Futile ligation

FIGURE S9. (**A**) Multiple sequence alignment of XthA homologs corresponding to the interaction interface region between MtbXthA and MtbLigA as identified by SAXS analysis, validated by ITC and complex disruption assays (main text). The conserved motif is boxed in red and consensus is shown below. The amino acid sequence of XthA from *M. tuberculosis* was aligned with homologs from *M. abscessus*, *M. smegmatis*, *M. bovis*, *M. avium* and *E. coli.* (**B**) Homology models of XthA from *M. tuberculosis* (green), *M. abscessus* (pale green), *M. smegmatis* (cyan), *M. bovis* (teal), *M. avium* (pale cyan) and *E. coli.* (grey) were structurally aligned using Pymol. The alignment shows conservation of LigA interacting (DGQ-red) and clamp interacting (QLRFPKK-blue) motifs in different XthA homologs. The residues making the catalytic triad (E57, D251, H281 in MtbXthA) are marked in hot pink. (**C**) Diagrammatic representation to show coordinating role of MtbXthA in orchestrating earlier steps of BER, to prevent futile ligation.

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 manager™ 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) and shown in **Table 1.** The HPLC purified nicked ds-DNA (**S1**) and BSA were used as control.