Supplementary Material

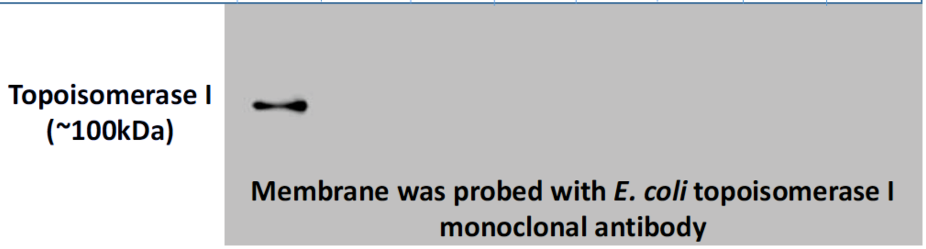
Novel Mechanism Evolved for Mycobacteria RNA polymerase and Topoisomerase I Protein-Protein Interaction

Srikanth Banda, Nan Cao, Yuk-Ching Tse-Dinh*

*** Correspondence:** Yuk-Ching Tse-Dinh: ytsedinh@fiu.edu

Figure S1. Rabbit MtbTopoI polyclonal antibodies can cross-react with MsmTopoI. Rabbit Polyclonal antibodies that were generated against MtbTopoI can cross-react with MsmTopoI due to a high sequence homology between these proteins. Western blotting of the soluble lysates from *M. smegmatis* mc^2 155 was carried out to verify the cross-reactivity of the antibody. Lane 1: Soluble lysate (10 µg) from *M. tuberculosis* H37Rv. Lane 2-4: Purified MtbTopoI (25 ng, 50 ng, 75 ng). Lane 5: Purified MsmTopoI (25 ng). Lane 6-8: Soluble lysates (10 µg) of *M. smegmatis* mc^2 155 over the course of its growth (Lane 6: OD$_{600}$- 0.8, Lane 7: OD$_{600}$- 1.5, Lane 8: OD$_{600}$- 3.0).

	1	2	3	4	5	6	7	8
M. smegmatis **RNAP β' (15nM)**	---	---	+	+	+	---	---	---
M. smegmatis **RNAP β (15nM)**	---	---	---	---	---	+	+	+
E. coli **Topoisomerase I (nM)**	---	20	5	10	20	5	10	20
E. coli **Topoisomerase I (25ng)**	+	---	---	---	---	---	---	---
HisPur cobalt resin	---	+	+	+	+	+	+	+



Membrane was probed with *E. coli* topoisomerase I monoclonal antibody

Figure S2. Lack of inter-species cross-interactions between purified topoisomerase I, and RNA polymerase β, β' subunits. An assay was carried out to verify the lack of cross-interaction of the purified RNA polymerase β, β' subunits of *M. smegmatis* with *E. coli* topoisomerase I. Recombinant (6xHis) RNA polymerase subunit of *M. smegmatis* was first incubated with the *E. coli* topoisomerase I, and later captured on the HisPur cobat resin. The recombinant protein bound resin was washed in pull-down wash buffer (10 mM HEPES, pH: 7.5, 10 mM Imidazole, 0.005% Tween -20), and finally elutions were made with pull-down elution buffer (10 mM HEPES, pH: 7.5, 350 mM Imidazole). The eluates were analyzed by SDS-PAGE and Western blot with a monoclonal antibody against *E. coli* topoisomerase I. Lane 1: Purified *E. coli* topoisomerase I was loaded directly. Lane 2: Eluate from pull-down reaction of *E. coli* topoisomerase I alone. Lanes 3, 4, 5 are loaded with eluates from the pull-down reactions of recombinant His-tagged RNA polymerase β' subunit. Lanes 6, 7, 8 represent the eluates from the pull-down with recombinant His-tagged RNA polymerase β.

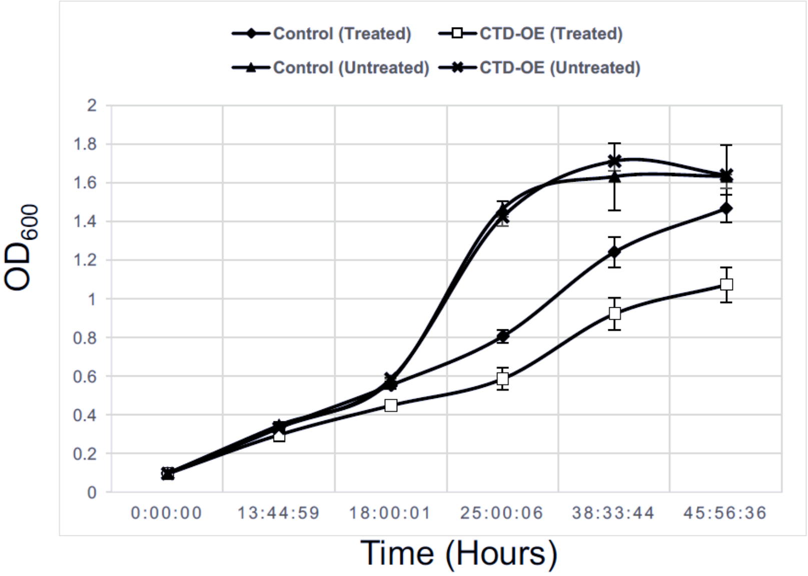
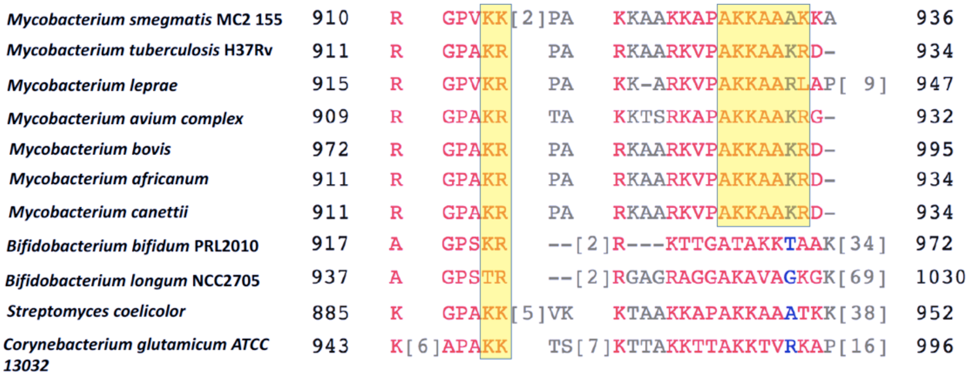


Figure S3. Effect of MsmTopoI-CTD overexpression on growth: The growth curves of the MsmTopoI-CTD overexpression strain and the control strain, induced with 25 ng/ml of tetracycline in 7H9 media, were monitored over a period of 48 hours by reading the absorbance (595 nm) at different time points. The overexpression of MsmTopoI-CTD did not influence the growth rate (Untreated curves). However, a slower growth rate of the MsmTopoI-CTD overexpression strain was observed in the presence of 0.15μM Moxifloxacin (Treated curves). Error bars represent the standard deviation (n=3).



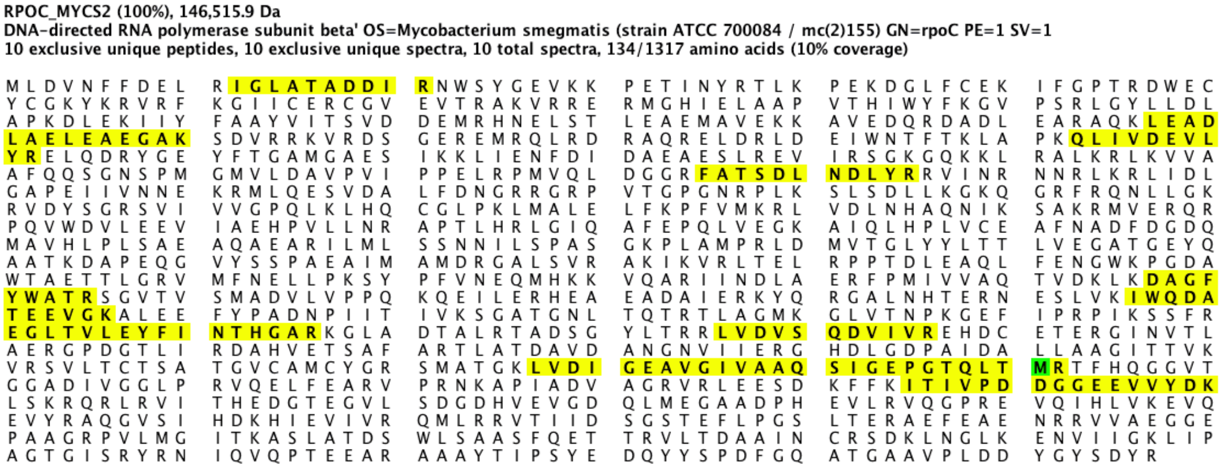
Mycobacterium smegmatis MC2 155	910	R	GPVKK[2]PA	KKAAKKAPAKKAAAKKA	936
Mycobacterium tuberculosis H37Rv	911	R	GPAKR PA	RKAARKVPAKKAAKRD-	934
Mycobacterium leprae	915	R	GPVKR PA	KK-ARKVPAKKAARLAP[9]	947
Mycobacterium avium complex	909	R	GPAKR TA	KKTSRKAPAKKAAKRG-	932
Mycobacterium bovis	972	R	GPAKR PA	RKAARKVPAKKAAKRD-	995
Mycobacterium africanum	911	R	GPAKR PA	RKAARKVPAKKAAKRD-	934
Mycobacterium canettii	911	R	GPAKR PA	RKAARKVPAKKAAKRD-	934
Bifidobacterium bifidum PRL2010	917	A	GPSKR --[2]	R---KTTGATAKKTAAK[34]	972
Bifidobacterium longum NCC2705	937	A	GPSTR --[2]	RGAGRAGGAKAVAGKGK[69]	1030
Streptomyces coelicolor	885	K	GPAKK[5]VK	KTAAKKAPAKKAAATKK[38]	952
Corynebacterium glutamicum ATCC 13032	943	K[6]APAKK	TS[7]	KTTAKKTTAKKTVRKAP[16]	996

Figure S4. The C-terminal tail of the topoisomerase I from Actinobacteria is rich in basic amino acids. Amino acid sequences of topoisomerase I from different Actinobacteria were aligned using Constraint-based multiple alignment tool, COBALT[1]. The regions of high conservation are highlighted. Mycobacterial topoisomerase I has a highly-conserved sequence (AKKAAAK) in the tail region.

Reference

1. Papadopoulos J. S., Agarwala R. (2007) COBALT: constraint-based alignment tool for multiple protein sequences. Bioinformatics 23, 1073-1079. btm076 [pii].

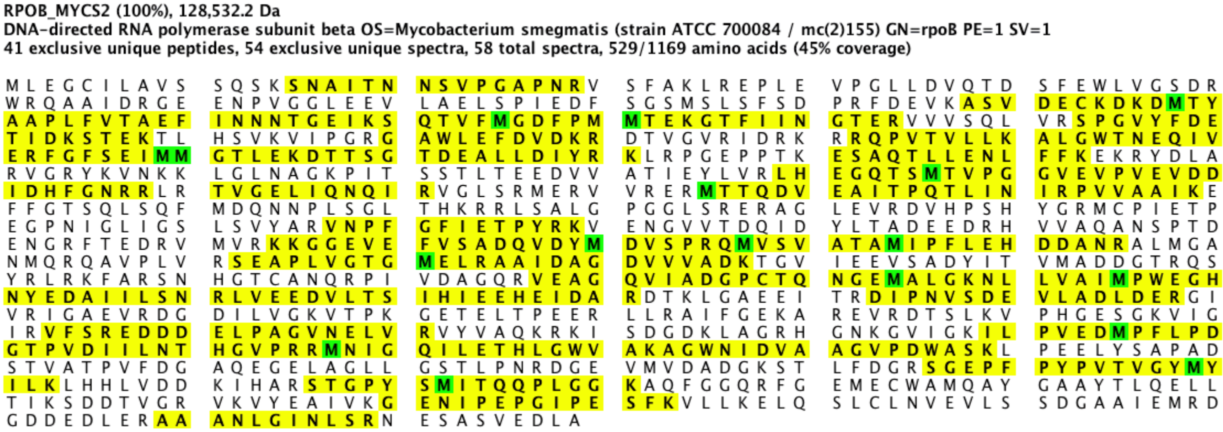
A



RPOC_MYCS2 (100%), 146,515.9 Da
DNA-directed RNA polymerase subunit beta' OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpoC PE=1 SV=1
10 exclusive unique peptides, 10 exclusive unique spectra, 10 total spectra, 134/1317 amino acids (10% coverage)

```
MLDVNFFDEL RIGLATADDI RNWSYGEVKK PETINYRTLK PEKDGLFCEK IFGPTRDWEC
YCGKYKRVRF KGIICERCGV EVTRAKVRRE RMGHIELAAP VTHIWYFKGV PSRLGYLLDL
APKDLEKIIY FAAYVITSVD DEMRHNELST LEAEMAVEKK AVEDQRDADL EARAQKLEAD
LAELEAEGAK SDVRRKVRDS GEREMRQLRD RAQRELDRLD EIWNTFTKLA PKQLIVDEVL
YRELQDRYGE YFTGAMGAES IKKLIENFDI DAEAESLREV IRSGKGQKKL RALKRLKVVA
AFQQSGNSPM GMVLDAVPVI PPELRPMVQL DGGRFATSDL NDLYRRVINR NNRLKRLIDL
GAPEIIVNNE KRMLQESVDA LFDNGRRGRP VTGPGNRPLK SLSDLLKGKQ GRFRQNLLGK
RVDYSGRSVI VVGPQLKLHQ CGLPKLMALE LFKPFVMKRL VDLNHAQNIK SAKRMVERQR
PQVWDVLEEV IAEHPVLLNR APTLHRLGIQ AFEPQLVEGK AIQLHPLVCE AFNADFDGDQ
MAVHLPLSAE AQAEARILML SSNNILSPAS GKPLAMPRLD MVTGLYYLTT LVEGATGEYQ
AATKDAPEQG VYSSPAEAIM AMDRGALSVR AKIKVRLTEL RPPTDLEAQL FENGWKPGDA
WTAETTLGRV MFNELLPKSY PFVNEQMHKK VQARIINDLA ERFPMIVVAQ TVDKLKDAGF
YWATRSGVTV SMADVLVPPQ KQEILERHEA EADAIERKYQ RGALNHTERN ESLVKIWQDA
TEEVGKALEE FYPADNPIIT IVKSGATGNL TQTRTLAGMK GLVTNPKGEF IPRPIKSSFR
EGLTVLEYFI NTHGARKGLA DTALRTADSG YLTRRLVDVS QDVIVREHDC ETERGINVTL
AERGPDGTLI RDAHVETSAF ARTLATDAVD ANGNVIIERG HDLGDPAIDA LLAAGITTVK
VRSVLTCTSA TGVCAMCYGR SMATGKLVDI GEAVGIVAAQ SIGEPGTQLT MRTFHQGGVT
GGADIVGGLP RVQELFEARV PRNKAPIADV AGRVRLEESD KFFKITIVPD DGGEEVVYDK
LSKRQRLRVI THEDGTEGVL SDGDHVEVGD QLMEGAADPH EVLRVQGPRE VQIHLVKEVQ
EVYRAQGVSI HDKHIEVIVR QMLRRVTIID SGSTEFLPGS LTERAEFEAE NRRVVAEGGE
PAAGRPVLMG ITKASLATDS WLSAASFQET TRVLTDAAIN CRSDKLNGLK ENVIIGKLIP
AGTGISRYRN IQVQPTEEAR AAAYTIPSYE DQYYSPDFGQ ATGAAVPLDD YGYSDYR
```

B



RPOB_MYCS2 (100%), 128,532.2 Da
DNA-directed RNA polymerase subunit beta OS=Mycobacterium smegmatis (strain ATCC 700084 / mc(2)155) GN=rpoB PE=1 SV=1
41 exclusive unique peptides, 54 exclusive unique spectra, 58 total spectra, 529/1169 amino acids (45% coverage)

```
MLEGCILAVS SQSKSNAITN NSVPGAPNRV SFAKLREPLE VPGLLDVQTD SFEWLVGSDR
WRQAAIDRGE ENPVGGLEEV LAELSPIEDF SGSMSLSFSD PRFDEVKASV DECKDKDMTY
AAPLFVTAEF INNNTGEIKS QTVFMGDFPM MTEKGTFIIN GTERVVVSQL VRSPGVYFDE
TIDKSTEKTL HSVKVIPGRG AWLEFDVDKR DTVGVRIDRK RRQPVTVLLK ALGWTNEQIV
ERFGFSEIMM GTLEKDTTSG TDEALLDIYR KLRPGEPPTK ESAQTLLENL FFKEKRYDLA
RVGRYKVNKK LGLNAGKPIT SSTLTEEDVV ATIEYLVRLH EGQTSMTVPG GVEVPVEVDD
IDHFGNRRLR TVGELIQNQI RVGLSRMERV VRERMTTQDV EAITPQTLIN IRPVVAAIKE
FFGTSQLSQF MDQNNPLSGL THKRRLSALG PGGLSRERAG LEVRDVHPSH YGRMCPIETP
EGPNIGLIGS LSVYARVNPF GFIETPYRKV ENGVVTDQID YLTADEEDRH VVAQANSPTD
ENGRFTEDRV MVRKKGGEVE FVSADQVDYM DVSPRQMVSV ATAMIPFLEH DDANRALMGA
NMQRQAVPLV RSEAPLVGTG MELRAAIDAG DVVVADKTGV IEEVSADYIT VMADDGTRQS
YRLRKFARSN HGTCANQRPI VDAGQRVEAG QVIADGPCTQ NGEMALGKNL LVAIMPWEGH
NYEDAIILSN RLVEEDVLTS IHIEEHEIDA RDTKLGAEEI TRDIPNVSDE VLADLDERGI
VRIGAEVRDG DILVGKVTPK GETELTPEER LLRAIFGEKA REVRDTSLKV PHGESGKVIG
IRVFSREDDD ELPAGVNELV RVYVAQKRKI SDGDKLAGRH GNKGVIGKIL PVEDMPFLPD
GTPVDIILNT HGVPRRMNIG QILETHLGWV AKAGWNIDVA AGVPDWASKL PEELYSAPAD
STVATPVFDG AQEGELAGLL GSTLPNRDGE VMVDADGKST LFDGRSGEPF PYPVTVGYMY
ILKLHHLVDD KIHARSTGPY SMITQQPLGG KAQFGGQRFG EMECWAMQAY GAAYTLQELL
TIKSDDTVGR VKVYEAIVKG ENIPEPGIPE SFKVLLKELQ SLCLNVEVLS SDGAAIEMRD
GDDEDLERAA ANLGINLSRN ESASVEDLA
```

Figure S5. Unique peptides of *M. smegmatis* RNAP subunits identified by mass spectrometry of the eluates from the Co-IP assay (A) RNAP beta’ subunit (B) RNAP beta subunit.

Table S1. PCR primers for Gibson cloning

Primer	Gene	Sequence (5’-3’)
RNA Polymerase beta_LIC_FP	rpoB	CTGTACTTCCAATCCAATGTGCTGGAAGGATGCA
RNA Polymerase beta_LIC_RP	rpoB	ATCCGTTATCCACTTCCAATCTACGCGAGATCCTCGAC
RNA Polymerase beta’_LIC_FP	rpoC	CTGTACTTCCAATCCAATGTGCTAGACGTCAACTTC
RNA Polymerase beta’_LIC_RP	rpoC	CGTTATCCACTTCCAATTTAGCGGTAATCCGAGTAG
MsmTopoI_pKW08_FP	MsmtopA	TTCGCGGATCCTTGGCTGGCGGCGACCGCGG
MsmTopoI_pKW08_RP	MsmtopA	TTCTCAAGCTTCTAGGCCTTCTTGGCGGCGG
MsmTopoI_2OT_FP	MsmtopA	GGGATCGAGGAAAACCTGTACTTCCAAATGGCTGGCGGCGACCG
MsmTopoI_2OT_RP	MsmtopA	GCGGATCCGTTATCCACTTCCAATATTGTTCGGCGGAAACCTAGGCCTTCTT
D1-D8_MsmTopoI_2OT_FP	D1-D8 MsmtopA	GGGATCGAGGAAAACCTGTACTTCCAAATGGCTGGCGGCGACCG
D1-D8_MsmTopoI_2OT_RP	D1-D8 MsmtopA	GCGGATCCGTTATCCACTTCCAATATTGTTAGGCACGGCGGTCGG

Table S2. Primers used for Site-Directed Mutagenesis

Primer	Sequence (5'-3')	Description
CTD-MsmTopoI_ pKW08_FP	GGATCCTGTCAGGATTCCACGATGAGAG	Deletion of the N-terminal domains from the full-length *topA* gene that was previously cloned into pKW08.
CTD-MsmTopoI_ pKW08_RP	GGCGTCGAGGGTTCGATCGCG	
NTD-MsmTopoI_2OT_FP	TAGAACCTCGAAGGCATCGACGC	Insertion of a stop codon in the *topA* gene for early termination at the end of the N-terminus domains (D1-D4).
NTD-MsmTopoI_2OT_RP	GCCGCCGACGAGCTGCTT	
D1-D5_MsmTopoI_2OT_FP	GCGTCCCTCTTGCTATGTGGCGAAGAGCTTTTCGG	Pro702 of MsmTopoI was substituted with a stop codon. The resulting protein (1-701) is termed as MsmTopoI-701t (D1-D5).
D1-D5_MsmTopoI_2OT_RP	CCGAAAAGCTCTTCGCCACATAGCAAGAGGGACGC	
D1-D6_MsmTopoI_2OT_FP	GCCCACGACGCGCTACAGCGACAGCAGC	Pro786 of MsmTopoI was substituted with a stop codon. The resulting protein (1-785) is termed as MsmTopoI-785t (D1-D6).
D1-D6_MsmTopoI_2OT_RP	GCTGCTGTCGCTGTAGCGCGTCGTGGGC	
D1-D7_MsmTopoI_2OT_FP	CGGCCACGGCGTTTCTACTCGGCGTAGATCTT	Pro840 of MsmTopoI was substituted with a stop codon. The resulting protein (1-839) is termed as MsmTopoI-839t (D1-D7).
D1-D7_MsmTopoI_2OT_RP	AAGATCTACGCCGAGTAGAAACGCCGTGGCCG	