

Essential roles for *imuA'*- and *imuB*-encoded accessory factors in DnaE2-dependent mutagenesis in *Mycobacterium tuberculosis*

Digby F. Warner^{a,1}, Duduzile E. Ndwandwe^a, Garth L. Abrahams^a, Bavesh D. Kana^a, Edith E. Machowski^a, Česlovas Venclovas^b, and Valerie Mizrahi^{a,1}

^aMedical Research Council/National Health Laboratory Service/University of the Witwatersrand Molecular Mycobacteriology Research Unit and Department of Science and Technology/National Research Foundation Centre of Excellence for Biomedical Tuberculosis Research, Faculty of Health Sciences, University of the Witwatersrand and the National Health Laboratory Service, Johannesburg 2000, South Africa; and ^bLaboratory of Bioinformatics, Institute of Biotechnology, LT-02241 Vilnius, Lithuania

Edited by Stephen J. Benkovic, The Pennsylvania State University, University Park, PA, and approved June 14, 2010 (received for review March 2, 2010)

In *Mycobacterium tuberculosis* (Mtb), damage-induced mutagenesis is dependent on the C-family DNA polymerase, DnaE2. Included with *dnaE2* in the Mtb SOS regulon is a putative operon comprising *Rv3395c*, which encodes a protein of unknown function restricted primarily to actinomycetes, and *Rv3394c*, which is predicted to encode a Y-family DNA polymerase. These genes were previously identified as components of an *imuA'-imuB-dnaE2*-type mutagenic cassette widespread among bacterial genomes. Here, we confirm that *Rv3395c* (designated *imuA'*) and *Rv3394c* (*imuB*) are individually essential for induced mutagenesis and damage tolerance. Yeast two-hybrid analyses indicate that ImuB interacts with both *ImuA'* and DnaE2, as well as with the β -clamp. Moreover, disruption of the ImuB- β clamp interaction significantly reduces induced mutagenesis and damage tolerance, phenocopying *imuA'*, *imuB*, and *dnaE2* gene deletion mutants. Despite retaining structural features characteristic of Y-family members, ImuB homologs lack conserved active-site amino acids required for polymerase activity. In contrast, replacement of DnaE2 catalytic residues reproduces the *dnaE2* gene deletion phenotype, strongly implying a direct role for the α -subunit in mutagenic lesion bypass. These data implicate differential protein interactions in specialist polymerase function and identify the split *imuA'-imuB/dnaE2* cassette as a compelling target for compounds designed to limit mutagenesis in a pathogen increasingly associated with drug resistance.

drug resistance | induced mutagenesis | Y-family polymerase | *Mycobacterium smegmatis*

DNA damage-induced base substitution mutagenesis in *Mycobacterium tuberculosis* (Mtb) depends on DnaE2 (1), a C-family DNA polymerase implicated in error-prone bypass of DNA lesions. Loss of DnaE2 activity renders Mtb hypersensitive to DNA damage and eliminates induced mutagenesis. Moreover, *dnaE2* deletion attenuates virulence and reduces the frequency of drug resistance in vivo. Mtb contains two DnaE-type polymerases; the other, DnaE1, provides essential, high-fidelity replicative polymerase function (1). However, the basis for the functional specialization of the DnaE subunits remains unclear (2, 3). Although structural determinants such as active-site architecture contribute significantly to inherent fidelity, it is possible that differential interactions with other DNA metabolic proteins modulate polymerase function.

Bacterial genomes containing a DnaE2-type DNA polymerase almost invariably encode a homolog of ImuB (4–6), a putative Y-family polymerase that is usually present in a LexA-regulated *imuA'-imuB-dnaE2* gene cassette (5). In *Caulobacter crescentus*, both ImuB and ImuA are required for induced mutagenesis and damage tolerance (6) whereas plasmid-encoded DnaE2 and ImuB mediate UV-induced mutagenesis in *Deinococcus deserti* (7). Although distributed widely across the bacterial domain, the *imuA'-imuB-dnaE2* cassette is not found in organisms possessing

umuDC homologs (5). This suggests that the encoded proteins perform an analogous function to DNA polymerase V (8), the Y-family member required for damage-induced mutagenesis in *Escherichia coli* (9).

Mtb contains a putative SOS-inducible operon, *Rv3395c-Rv3394c* (1, 10), located ~24.7 kb upstream of *dnaE2*. *Rv3395c* homologs are found in a limited number of organisms and their function is unknown, whereas *Rv3394c* exhibits significant homology to ImuB proteins (6), identifying the *Rv3395c-Rv3394c* operon as part of a split *imuA'-imuB/dnaE2* cassette (5). *Rv3394c* (ImuB) contains a predicted β -clamp-binding motif, which designates the protein as a DinB3-type Y-family polymerase (11). Notably, neither DnaE2 nor ImuA' possesses an identifiable clamp-binding motif (1, 12). The β -clamp modulates the recruitment to the replication machinery of proteins involved in bacterial DNA replication and repair (13), binding replicative and specialist lesion bypass polymerases simultaneously to enable rapid interchange (14, 15). The β -clamp-binding motif therefore suggested that ImuB might be crucial for DnaE2 function.

In this paper, we investigate the operation of the split *imuA'-imuB/dnaE2* cassette in mycobacteria. Our results establish essential roles for ImuA' and ImuB in DnaE2-dependent mutagenesis and damage tolerance. We also confirm the identity of the β -clamp-binding motif in ImuB and identify a role for the extended ImuB C-terminal domain in binding both DnaE2 and ImuA'. Although structurally similar to Y-family polymerases, homology modeling indicates that ImuB proteins lack highly conserved active-site residues required for polymerase activity. Instead, disruption of the ImuB- β -clamp interaction reproduces the *imuB* gene deletion phenotype, suggesting an essential role for ImuB in mediating access of the other cassette components to the replication fork. In contrast, through targeted replacement of active-site residues, we elucidate a direct catalytic role for DnaE2 in mutagenic lesion bypass. These observations confirm the functionality of the split *imuA'-imuB/dnaE2* cassette and identify differential protein interactions as determinants of specialist polymerase function in Mtb.

Results

DNA Damage-Induced Expression of the Split *imuA'-imuB/dnaE2* Cassette. All mycobacteria contain a split *imuA'-imuB/dnaE2*

Author contributions: D.F.W., D.E.N., G.L.A., B.D.K., C.V., and V.M. designed research; D.F.W., D.E.N., G.L.A., B.D.K., and C.V. performed research; D.F.W. and E.E.M. contributed new reagents/analytic tools; D.F.W., D.E.N., G.L.A., B.D.K., C.V., and V.M. analyzed data; and D.F.W., C.V., and V.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence may be addressed. E-mail: valerie.mizrahi@wits.ac.za or digby.warner@nhls.ac.za.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1002614107/-DCSupplemental.

cassette (5), except *Mycobacterium leprae*, which has lost these genes through genome decay (16). The genomic location varies across mycobacterial species: in *Mycobacterium smegmatis* mc²155 (Msm), *imuA'* and *imuB* are situated only 10.8 kb upstream of *dnaE2*. However, the encoded proteins are highly conserved, and a LexA-binding site is located in the promoter region upstream of *imuA'* in Mtb (17) and Msm (Fig. S1B). Previous studies identified all three cassette components in the Mtb SOS regulon (1, 10), and expression of *dnaE2* was shown to parallel *recA* induction (1). We analyzed transcript abundances of the cassette components relative to a *sigA* control (18) during mid log-phase growth of Mtb and following UV exposure (Table S1) and compared with the non-inducible Y-family polymerases, *dinB1* and *dinB2* (10). All cassette genes were significantly up-regulated within 6 h of UV treatment and remained elevated at 24 h. *dnaE1*, which encodes the essential replicative α -subunit (1), was also up-regulated following UV irradiation. Although there is no identifiable SOS box (17), a promoter motif regulating the LexA-independent damage response (10) is situated immediately upstream of *dnaE1* (19), consistent with damage induction.

ImuA' and ImuB Are Individually Essential for DnaE2 Function in Mycobacteria. To investigate the roles of *imuA'* and *imuB* in DNA damage-induced mutagenesis, allelic exchange mutants of Mtb H37Rv were constructed in which *imuB* alone, or both *imuA'* and *imuB*, were deleted and replaced by an antibiotic resistance marker (Fig. S1A). Neither mutant exhibited a phenotype under standard in vitro growth conditions. However, UV-induced mutagenesis was eliminated in both strains (Fig. 1A) as determined by mutation frequencies to rifampicin resistance (Rif^R). The damage-induced mutator phenotype was restored by complementation with a fragment carrying both *imuA'* and *imuB* genes integrated at the *attB* site. Moreover, loss of *imuA'-imuB* reproduced the *dnaE2* deletion phenotype (1), confirming the functionality of the split *imuA'-imuB/dnaE2* cassette. To determine the individual requirements for *imuA'* and *imuB*, the similar genomic arrangement in Msm was exploited (Fig. S1B). Deletion of *imuB* significantly reduced the UV-induced Rif^R frequency in Msm (Fig. 1B), mimicking inactivation of *dnaE2* (1). The same phenotype resulted from deletion of *imuA'* and was complemented by integration of a wild-type copy of *imuA'*. These data established the individual essentiality of ImuA' and ImuB for induced mutagenesis and confirmed the designation of the mycobacterial ImuA' as a distant homolog of ImuA (5, 6). Msm *imuA'* and *imuB* deletion mutants were hypersensitive to mitomycin C (MMC), a genotoxic agent known to induce the SOS response in mycobacteria (1) (Fig. S1C). Deletion of all three components ($\Delta dnaE2 \Delta imuA' \Delta imuB$) did not increase damage sensitivity relative to individual deletion mutants ($\Delta imuA'$, $\Delta imuB$, or $\Delta dnaE2$) (Fig. S1C). Moreover, integration of wild-type copies of the deleted genes at the *attB* (*imuA'-imuB*) and *attL* (*dnaE2*) sites reversed the MMC-hypersensitive phenotype of

the $\Delta dnaE2 \Delta imuA' \Delta imuB$ triple mutant, confirming that all three genes operate in a single pathway.

ImuB Lacks Active Site Residues Required for DNA Polymerase Activity. ImuB occupies a distinct branch of the UmuC subfamily of Y-family polymerases (6) and possesses all structural domains typical of Y-family members, including the defining little finger (20) (Fig. 2A). A putative β -clamp-binding motif (³⁵⁴QLPLWG³⁵⁹) that is located between the Y polymerase-like N-terminal region and the C-terminal extension characteristic of ImuB proteins identified Mtb ImuB as founder of the DinB3-type Y-family polymerases (11) (Fig. S2). Although classified as a Y-family member, comparative sequence analyses and homology modeling indicate that the expected carboxylates (21) are replaced by other residues in the ImuB active site (Fig. 2B). The absence of the complete triad of active-site acidic residues is a feature of all ImuB homologs analyzed (Fig. S2) and strongly implies that these proteins cannot function as DNA polymerases (22). In turn, this observation suggests a model for translesion synthesis in which DnaE2 itself catalyzes lesion bypass (6).

Catalytic Activity of DnaE2 Is Required for Induced Mutagenesis in Mtb. All three DNA polymerase III α active-site acidic residues (23, 24) are present in Mtb DnaE2 (D⁴³⁹, D⁴⁴¹, D⁵⁷⁹; Fig. S3). To determine the requirement for DnaE2 polymerase activity in damage tolerance and induced mutagenesis in mycobacteria, a site-directed Msm mutant was generated in which catalytic function was crippled by replacement of two corresponding active-site aspartates with alanines (⁴⁴¹DID⁴⁴³ \rightarrow ⁴⁴¹AIA⁴⁴³). The chromosomal *dnaE2*^{Ala} mutation eliminated UV-induced mutagenesis (Fig. 3A) and rendered Msm hypersensitive to MMC (Fig. 3B). The fact that the *dnaE2*^{Ala} mutation phenocopied the $\Delta dnaE2$ knockout strongly implied a direct role for DnaE2 in catalyzing translesion synthesis.

QLPLWG Clamp-Binding Motif Mediates the Interaction of ImuB with β . A β -clamp-binding motif (⁹⁴⁶QFDLF⁹⁵⁰) (12) is situated in the β -binding domain of Mtb DnaE1 (Fig. S3). The corresponding region in Mtb DnaE2 (⁹²⁴RPDRLPGVG⁹³²) is invariant in mycobacterial genomes, but does not resemble known β -binding motifs (11, 12). We showed recently (25) that the *dnaN*-encoded β -subunit interacts with DnaE1 and with itself in yeast two-hybrid (Y2H) assays, but does not bind DnaE2 (Fig. 4). Similarly, Mtb ImuA' failed to bind β , whereas an interaction between ImuB and the β -clamp was readily detected. In combination, these data establish ImuB as the only cassette component with β -binding capacity in Mtb. The observed interaction was consistent with the β -binding motif (³⁵⁴QLPLWG³⁵⁹) in ImuB that characterizes the DinB3 family (11) and is highly conserved among mycobacterial ImuBs. Moreover, a Q354A mutation (ImuB^{ALPLWG}) eliminated the ImuB- β interaction in the Y2H system (Fig. 4), confirming a role for this motif in clamp binding.

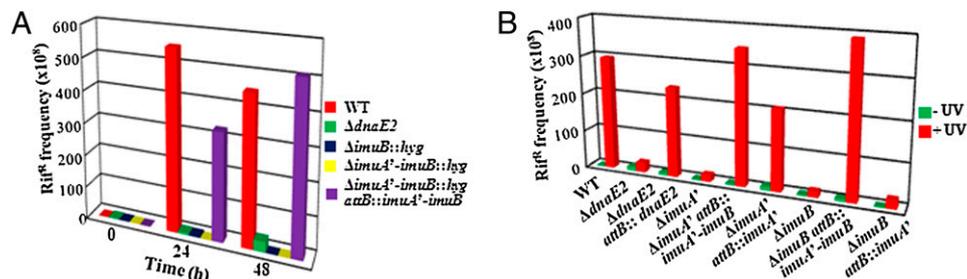


Fig. 1. ImuA' and ImuB are essential for induced mutagenesis. UV-induced mutation frequencies to rifampicin resistance (Rif^R) in (A) Mtb at 24 and 48 h and (B) Msm at 4.5 h post UV treatment. Data represent single experiments performed in triplicate.

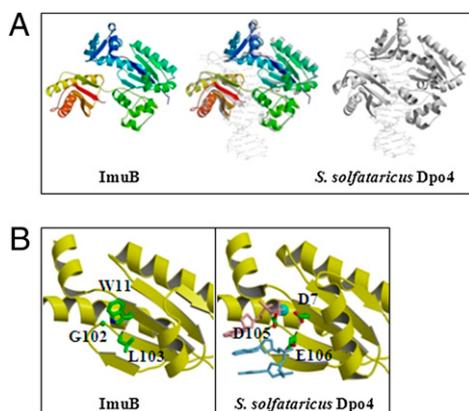


Fig. 2. ImuB is a homolog of Y-family polymerases but lacks active-site acidic residues. (A) Comparison of the Mtb ImuB homology model with the X-ray structure of *Sulfolobus solfataricus* Dpo4 complexed with DNA and incoming nucleotide (PDB id: 1jx4) (21). ImuB is colored from N terminus (blue) to C terminus (red). (Center) The superimposed structures. (B) Close-up of the active site of Dpo4 and corresponding residues in ImuB. Only the palm domain is shown in both structures. The incoming nucleotide in the Dpo4–DNA complex is colored pink, and the magnesium ion is shown as a cyan sphere.

Protein Interactions of the Split *imuA'*-*imuB*/*dnaE2* Cassette Components in Mtb. Y2H analyses identified an ImuA'–ImuB interaction that was retained in the ImuB^{ALPLWG} mutant (Fig. 4), confirming the specificity of the ³⁵⁴QLPLWG³⁵⁹ motif for binding β. In addition, both ImuB and ImuB^{ALPLWG} interacted with DnaE2, whereas DnaE2 and ImuA' failed to interact with any protein analyzed other than ImuB. The active-site mutant form of DnaE2, DnaE2^{ΔIA}, retained the ability to bind ImuB (Fig. 4), confirming that loss of DnaE2 catalytic activity—and not disruption of the inferred ImuB–DnaE2 interaction—was responsible for the phenotypes of the Msm *dnaE2*^{ΔIA} mutant strain (Fig. 3). Interactions of ImuB with DnaE1 and with itself were also elucidated (Fig. 4), identifying multiple partners for ImuB (Fig. S4A).

C-Terminal Region of ImuB Is Required for Interaction with Other Cassette Components. We could not detect a relationship between the C-terminal region of Mtb ImuB and any known protein structure. However, this region contains stretches of predicted structural disorder (Fig. S4B) that are a hallmark of protein–protein interaction sites (26). We generated Y2H constructs in which nonsense codons were introduced into ImuB to preserve N-terminal Y-family structural domains while progressively eliminating C-terminal segments (Fig. 4). In Mtb ImuB^{C168}, the entire 168-amino-acid C-terminal region is eliminated downstream of the intact ³⁵⁴QLPLWG³⁵⁹ motif. The truncated ImuB retained its interaction with β, confirming the identity of the clamp-binding motif. In contrast, the ability to bind β was eliminated in ImuB^{CB} in which both the C-terminal region and the ³⁵⁴QLPLWG³⁵⁹ motif were deleted (Fig. 4). Both truncation mutants, ImuB^{C168} and

ImuB^{CB}, lost the ability to bind DnaE1, DnaE2, and ImuA', confirming the involvement of the C-terminal region of ImuB in multiple protein interactions. To evaluate the requirement for the C-terminal-dependent ImuB interactions in cassette function, we constructed integrating vectors carrying wild-type Msm *imuA'* and mutant *imuB* alleles with equivalent C-terminal truncations (*imuA'*-*imuB*^{C168} and *imuA'*-*imuB*^{CB}). Both constructs restored full function to the Msm *ΔimuA'* mutant, but failed to complement the induced mutagenesis (Fig. S5A) and damage tolerance (Fig. S5B) phenotypes of the single *ΔimuB* and double *ΔimuA'* *ΔimuB* knockout strains, thereby supporting a critical role for ImuB–DnaE2 and/or ImuB–ImuA' interaction(s) in pathway function. By mapping the ImuA' sequence onto the RecA structure (Fig. 5A), we also designed N- and C-terminal truncations of ImuA' and assessed their effects on the interaction with ImuB (Fig. 4). Removal of 31 (ImuA'^{N31}) or 48 (ImuA'^{N48}) amino acids from the N terminus (eliminating the predicted N-terminal subdomain; Fig. 5A and Fig. S6) had no impact on the ability of ImuA' to bind ImuB in Y2H assays (Fig. 4). In contrast, a 44-amino-acid C-terminal truncation (ImuA'^C) eliminated ImuB binding. The corresponding truncation in Msm *imuA'* abrogated the ability of a complementing vector to restore damage tolerance to the *ΔimuA'* deletion mutant (Fig. 5B). However, as the C-terminal truncation might have affected the structural integrity of the protein, it is uncertain whether this region alone or the entire structural domain is required for the ImuA'–ImuB interaction.

ImuB–β-Clamp Interaction Is Required for Induced Mutagenesis. A survey of representative *imuB*/*dnaE2*-containing bacteria revealed clamp-binding motifs in ImuB and DnaE2 proteins (Figs. S2 and S3), suggesting that at least one of these components must bind β for cassette function. To test this hypothesis, we constructed an integration vector carrying wild-type *imuA* and *imuB*^{ALPLWG}, which encodes ImuB with a mutated β-binding motif (³⁵²ALPLWG³⁵⁷), and evaluated its ability to complement the Msm *ΔimuA'* *ΔimuB* double mutant. The equivalent mutation had disrupted the ImuB–β interaction in Y2H assays (Fig. 4). However, a single Q352A substitution had no impact on the ability of the complementing vector to restore damage tolerance and induced mutagenesis to wild-type levels in Msm (Fig. S7), probably reflecting the inability of the yeast system to reproduce all elements contributing to protein–protein interactions in the natural mycobacterial host. Therefore, we constructed additional alleles containing multiple mutations that targeted hydrophobic and/or aromatic residues in the putative ImuB β-binding motif. Two alleles containing double mutations, *imuB*^{AAPLWG} and *imuB*^{ALPLGG}, retained ImuB function (Fig. 6). However, mutation of the first five β-binding motif residues in *imuB*^{AAAAAGG} eliminated the ability of the resulting vector to complement the induced mutagenesis (Fig. 6A) and damage tolerance (Fig. 6B) phenotypes of *ΔimuA'* *ΔimuB*, strongly suggesting that the ImuB–β-clamp interaction is required for cassette operation.

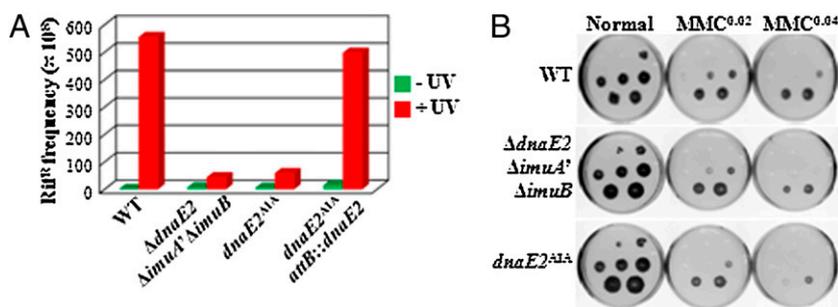


Fig. 3. DnaE2 catalytic activity is required for induced mutagenesis and damage tolerance. (A) UV-induced mutation frequencies to rifampicin resistance (Rif^R). (B) Sensitivity to MMC treatment. Log-fold dilutions were plated on antibiotic-free medium or on medium containing MMC at 0.02 and 0.04 μg·ml⁻¹. Data are from a representative experiment performed in triplicate.

	β clamp	ImuA'	ImuB	ImuB ^{ALPLWG}	DnaE1	DnaE2	DnaE2 ^{AIA}
DnaE2	×	×	○	○	ND	ND	ND
DnaE2 ^{AIA}	×	×	○	○	ND	ND	ND
ImuB	○	○	○	○	○	○	○
ImuB ^{C168}	○	×	○	ND	×	×	×
ImuB ^{CB}	×	×	○	ND	×	×	×
ImuB ^{ALPLWG}	×	○	○	ND	○	○	○
ImuA'	×	×	○	○	×	×	×
ImuA' ^{N91}	ND	ND	○	○	ND	ND	ND
ImuA' ^{N48}	ND	ND	○	○	ND	ND	ND
ImuA' ^C	ND	ND	×	×	ND	ND	ND

Fig. 4. Interactions of cassette components. Summary of Mtb protein interactions detected (●) or absent (×) on highest stringency growth medium as identified in this study and previously (25). ND, not determined. An open circle (○) indicates interactions maintained where bait or prey vector contains full-length ImuB; ImuB self-association is eliminated when both vectors carry the ImuB^{C168} allele. A blue bar indicates the position of the β -clamp-binding motif (³⁵⁴QLPLWG³⁵⁹) deleted in ImuB^{CB} and mutated (XXX) in ImuB^{ALPLWG}. The ⁴³⁹AIA⁴⁴¹ mutation in Mtb DnaE2 is indicated (XXX). Lines are to scale.

Discussion

The phenotypes associated with knockout of Mtb *dnaE2* suggested a functional analogy with *E. coli* DNA polymerase V (1). Here, we extended those observations, demonstrating the essentiality of ImuA' and ImuB for damage tolerance and induced mutagenesis and establishing that these proteins act in the same pathway as DnaE2. We also confirmed the inducible expression of *dnaE2*, *imuA'*, and *imuB* and identified *dnaE1* as a component of the Mtb damage response. Previous studies included DnaE2 and ImuB among three damage-responsive DNA polymerases in Mtb (1, 10), the other being a putative DNA polymerase X (Rv3856c). However, the absence of active-site carboxylates appears incompatible with catalytic function in ImuB, whereas a truncated polymerase domain probably precludes polymerase function in Rv3856c (Fig. S8). Therefore, in contrast to *E. coli* DNA polymerase V, which is part of an SOS regulon that includes the additional specialist DNA polymerases II (B family) and IV (Y family) (reviewed in ref. 27), the polymerases induced as part of the Mtb damage response seem to be limited to two α -subunits, DnaE1 and DnaE2.

Point mutations in the DnaE2 active site eliminated induced mutagenesis and damage tolerance, reproducing the *dnaE2* deletion phenotype. In combination with the prediction that ImuB lacks polymerase activity, this observation implies that DnaE2 catalyzes translesion synthesis. Unlike Y-family polymerases whose structures are adapted to specialist lesion bypass (20), sequence analysis reveals few clues to DnaE2 function. All major DNA polymerase III α structural/functional domains (23, 24) are readily identified in DnaE2, except for the very C-terminal region which in *E. coli* has been implicated in the interaction of α with the clamp-loader subunit, τ (28, 29). A strong α - τ interaction enables simultaneous leading and lagging-strand synthesis by the DNA polymerase III holoenzyme (13), and the absence of this region in DnaE2 and all other nonessential *dnaE*-type α -subunits (Fig. S3) might correlate with an inability to substitute essential replicative function. Evidence implicating a defective α - τ interaction in an *E. coli* mutator phenotype (30) further reinforces the potential

contribution of τ -binding to the functional specialization of DnaE2 subunits.

Mtb DnaE2 does not bind β . However, a highly conserved motif is located immediately downstream of the predicted OB fold domain in DnaE2 proteins (Fig. S3), which suggests the potential for other intermolecular interactions. In *E. coli*, replicative polymerase fidelity is enhanced by the association of the α -subunit with the *dnaQ*-encoded proofreading exonuclease, and disruptions to proofreading activity enable DNA polymerase III-mediated lesion bypass in the absence of DNA polymerases IV and V (31, 32). In combination, these observations suggest that differential interactions with DnaQ-like proteins might determine DnaE subunit function in Mtb. Additional 3'-5' exonuclease activity was recently located to the PHP domain of DnaE-type polymerases (33), although the contribution of this function to polymerase fidelity remains to be determined. The corresponding domain of Mtb DnaE1 retains all amino acids associated with exonuclease activity (34–36), whereas key residues are absent in DnaE2. Notably, the residues present in Mtb DnaE1 are conserved in the sole DnaE polymerase in *M. leprae* (Fig. S3), an organism marked by a complete lack of DnaQ homologs (37). Therefore, although the functional consequence of these differences remains to be elucidated, it is tempting to speculate that loss of conserved residues in the DnaE2 PHP domain impacts relative fidelity.

Our evidence indicates a central role for ImuB in cassette function that is independent of polymerase activity. Structurally, ImuB resembles Y-family polymerases; however, the active-site architecture is inconsistent with catalysis and instead reinforces the likely importance of protein interactions for ImuB function. Consistent with this idea, we identified multiple partners for ImuB, including DnaE1 and DnaE2, as well as ImuA' and the β -clamp. We also established the capacity of ImuB to interact with itself and implicated the C-terminal extension in this self-association and in the interaction of ImuB with DnaE1, DnaE2, and ImuA'. The fact that the C terminus binds multiple partners suggests that this region mediates protein interactions that distinguish ImuB from other Y-family members. Our Y2H experiments also confirmed the role of the DinB3-type QLPLWG motif in the interaction of ImuB with β . Notably, disruption of this motif in the

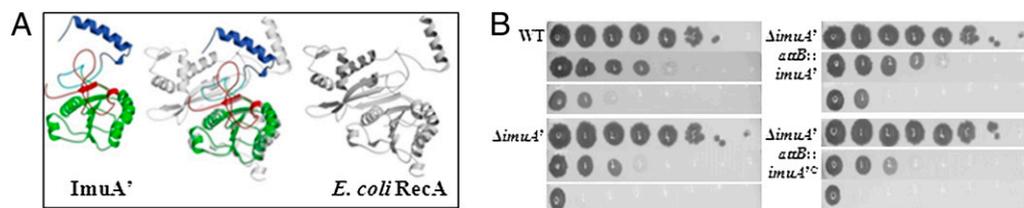


Fig. 5. ImuA' is required for damage tolerance. (A) Comparison of the Mtb ImuA' homology model and the X-ray structure of *E. coli* RecA (PDB id: 1u94) (57). ImuA' is colored according to truncation variants: blue and cyan indicate N-terminal truncations, and red is the C-terminal truncation. (B) Deletion of 44 C-terminal amino acids (*imuA'^C*) renders Msm sensitive to MMC treatment. Log-fold dilutions were plated on antibiotic-free medium (Left and Right, Rows 1 and 4), 0.02 $\mu\text{g}\cdot\text{mL}^{-1}$ (Left and Right, Rows 2 and 5), and 0.04 $\mu\text{g}\cdot\text{mL}^{-1}$ (Left and Right, Rows 3 and 6) MMC. Data are from a representative experiment performed in duplicate.

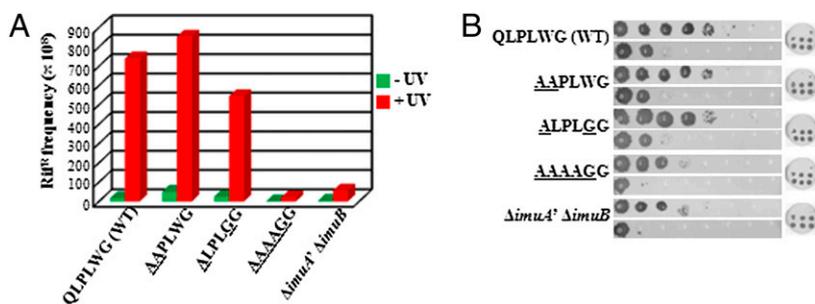


Fig. 6. The ImuB- β -clamp interaction is required for induced mutagenesis and damage tolerance. (A) UV-induced mutation frequencies to rifampicin resistance (Rif^R). (B) Sensitivity to MMC treatment. Log-fold dilutions of Msm $\Delta imuA'$ $\Delta imuB$ complemented with full-length *imuA'-imuB* containing wild-type (WT) and mutant *imuB* alleles were plated on 0.02 $\mu\text{g}\cdot\text{mL}^{-1}$ (Rows 1, 3, 5, 7, and 9) and 0.04 $\mu\text{g}\cdot\text{mL}^{-1}$ (Rows 2, 4, 6, 8, and 10) MMC; at right is the untreated control. Data are from a representative experiment performed in duplicate.

imuB^{AAAAGG} allele eliminated cassette function, strongly implying that ImuB—through β —is required to mediate access of DnaE2 to the replication fork. It is difficult, therefore, to reconcile the highly conserved ImuB (Fig. S2) and DnaE2 (Fig. S3) proteins with conflicting reports of biological function: specifically, the observation that ImuB and DnaE2 fulfill antagonistic roles in *Pseudomonas putida* (22), whereas the homologous *Pseudomonas aeruginosa* proteins are dispensable for damage tolerance (38) but not for induced mutagenesis (39). The results of Sanders et al. (39) are more consistent with the data presented here in that they indicate essential, and interdependent, roles for DnaE2 and ImuB proteins in damage tolerance and induced mutagenesis.

In *E. coli*, RecA constitutes a key component of the DNA polymerase V “mutasome” comprising UmuD₂C-RecA-ATP (8). Mtb ImuA' resembles RecA at a structural level, but lacks the characteristic C-terminal domain. We observed the direct interaction of ImuA' with ImuB involving the C-terminal regions of both proteins; however, no evidence of ImuA' self-association was detected that might indicate RecA-like filamentation (40). Instead, structural peculiarities such as the absence of a RecA-like nucleotide-binding motif, as well as extensive differences in the regions corresponding to RecA DNA-binding loops, suggest the possible functional specialization of ImuA'. Therefore, although broadly consistent with the identification of the *imuA'-imuB-dnaE2* cassette as the nonorthologous replacement of the *umuDC* system (5), our data reveal a “mutasome” comprising distinct components: a damage-inducible α -subunit that does not bind β but catalyzes translesion synthesis, a putative polymerase-inactive Y-family homolog that binds β as well as the other cassette components, and a predicted DNA-binding protein. Fundamental questions therefore remain regarding the relevance of the inferred protein-protein interactions to cassette function, the temporal order of those interactions, the potential for complex formation, and the possibility that the components are subject to additional regulation.

Mtb engages multiple strategies to subvert tuberculosis (TB) chemotherapy (41), which favors the emergence of antibiotic resistance (42). One approach to the development of antimicrobials is to disarm mechanisms of induced mutagenesis (43). The encoded components of the split *imuA'-imuB/dnaE2* cassette are essential for induced mutagenesis and so present compelling targets for the discovery of anti-TB drugs.

Materials and Methods

Bacterial Strains and Growth Conditions. Strains, plasmids, and oligonucleotides are described in Tables S2 and S3. Mtb was grown on Middlebrook 7H10 (Merck) supplemented with 0.5% glycerol and Middlebrook Oleic Acid Dextrose Catalase (OADC) enrichment (Merck) or in Middlebrook 7H9 supple-

mented with 0.2% glycerol, Middlebrook OADC, and 0.05% Tween 80. For Msm, 0.1% Tween 80 was used.

Yeast Two-Hybrid Analyses. Protein-protein interactions were assessed using the Clontech Matchmaker Y2H system. Fusion of Mtb ImuB to the binding domain resulted in autoactivation on low-stringency growth medium; therefore, all interactions involving Mtb ImuB were inferred from experiments using an ImuB-activation domain fusion construct only. All other proteins were cloned as activation domain and binding domain fusions.

Construction of Mutant Strains of Mtb and Msm. Mtb and Msm mutants were constructed by allelic exchange (44) using suicide plasmids described in Table S2. Alleles for site-directed mutagenesis were generated by the megaprimer method (45). Genetically complemented derivatives were generated by integration of L5 or Tweety-based vectors at *attB* (46) or *attL* (47).

DNA-Damaging Treatments and Determination of Mutation Frequencies. UV-induced mutation frequencies were determined as described (1). Survival assays were performed by plating 10-fold serial dilutions of log-phase Msm cultures on solid media containing MMC.

Gene Expression Analysis. Bacteria were grown in 7H9 medium and aliquots sampled during different phases of growth. Real-time quantitative reverse transcription-PCR was carried out with the primer pairs detailed in Table S3, as described (25).

Homology Modeling. Models were generated using previously described methods (48). Briefly, structural templates were selected by PSI-BLAST (49), and the position-dependent reliability of alignments was estimated by PSI-BLAST-ISS (50). In addition, HHsearch (51), COMPASS (52), and COMA (53) were used to reduce uncertain regions in alignments with structural templates. The agreement of all three algorithms was required before regions were considered reliably aligned. Three-dimensional models were refined by iterative steps of construction and assessment of different alignment variants for unreliable regions and/or combinations of structural templates. Models corresponding to different sequence-structure alignment variants were built by Modeler (54), and residue side chains were positioned with SCWRL3 (55). Modeled structures were assessed iteratively with Prosa2003 (56) until no further improvement in energy scores was possible and the visual assessment revealed no significant flaws.

ACKNOWLEDGMENTS. This work was funded by grants from the Medical Research Council of South Africa (to V.M.), the National Research Foundation (V.M.), the Howard Hughes Medical Institute (International Research Scholar's grants to V.M. and C.V.), the University of the Witwatersrand (V.M., D.F.W. and D.E.N.), and the South African Tuberculosis and AIDS Training (SATBAT) program (National Institutes of Health/Fogarty International Center 1U2RTW007370/3) (D.E.N.). We thank Sabine Ehrh, Dirk Schappinger, Graham Hatfull, and Diane Kuhnert for providing vectors and Helena Boshoff and members of the Mizrahi laboratory for valuable advice and assistance.

- Boshoff HI, Reed MB, Barry CE, III, Mizrahi V (2003) DnaE2 polymerase contributes to in vivo survival and the emergence of drug resistance in *Mycobacterium tuberculosis*. *Cell* 113:183–193.
- Bruck I, Goodman MF, O'Donnell M (2003) The essential C family DnaE polymerase is error-prone and efficient at lesion bypass. *J Biol Chem* 278:44361–44368.
- Le Chatelier E, et al. (2004) Involvement of DnaE, the second replicative DNA polymerase from *Bacillus subtilis*, in DNA mutagenesis. *J Biol Chem* 279:1757–1767.

- Abella M, et al. (2004) Widespread distribution of a *lexA*-regulated DNA damage-inducible multiple gene cassette in the *Proteobacteria* phylum. *Mol Microbiol* 54:212–222.
- Erill I, Campoy S, Mazon G, Barbé J (2006) Dispersal and regulation of an adaptive mutagenesis cassette in the bacteria domain. *Nucleic Acids Res* 34:66–77.
- Galhardo RS, Rocha RP, Marques MV, Menck CF (2005) An SOS-regulated operon involved in damage-inducible mutagenesis in *Caulobacter crescentus*. *Nucleic Acids Res* 33:2603–2614.

7. Dulerio R, Fochesato S, Blanchard L, de Groot A (2009) Mutagenic lesion bypass and two functionally different RecA proteins in *Deinococcus deserti*. *Mol Microbiol* 74: 194–208.
8. Jiang Q, Karata K, Woodgate R, Cox MM, Goodman MF (2009) The active form of DNA polymerase V is UmuD' (γ C-RecA-ATP. *Nature* 460:359–363.
9. Tang M, et al. (2000) Roles of *E. coli* DNA polymerases IV and V in lesion-targeted and untargeted SOS mutagenesis. *Nature* 404:1014–1018.
10. Rand L, et al. (2003) The majority of inducible DNA repair genes in *Mycobacterium tuberculosis* are induced independently of RecA. *Mol Microbiol* 50:1031–1042.
11. Dalrymple BP, Wijffels G, Kongsuwan K, Jennings P (2003) Towards an understanding of protein-protein interaction network hierarchies. Analysis of DnaN-binding peptide motifs in members of protein families interacting with the eubacterial processivity clamp, the subunit of DNA Polymerase III. *First Asia-Pacific Bioinformatics Conference (APBC2003)*, ed Chen Y-PP (Australian Computer Society, Sydney, Australia), pp 153–162.
12. Dalrymple BP, Kongsuwan K, Wijffels G, Dixon NE, Jennings PA (2001) A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. *Proc Natl Acad Sci USA* 98:11627–11632.
13. Johnson A, O'Donnell M (2005) Cellular DNA replicases: Components and dynamics at the replication fork. *Annu Rev Biochem* 74:283–315.
14. Heltzel JM, Maul RW, Scouten Ponticelli SK, Sutton MD (2009) A model for DNA polymerase switching involving a single cleft and the rim of the sliding clamp. *Proc Natl Acad Sci USA* 106:12664–12669.
15. Pagès V, Fuchs RP (2002) How DNA lesions are turned into mutations within cells? *Oncogene* 21:8957–8966.
16. Cole ST, et al. (2001) Massive gene decay in the leprosy bacillus. *Nature* 409: 1007–1011.
17. Davis EO, Dullaghan EM, Rand L (2002) Definition of the mycobacterial SOS box and use to identify LexA-regulated genes in *Mycobacterium tuberculosis*. *J Bacteriol* 184: 3287–3295.
18. Manganeli R, Dubnau E, Tyagi S, Kramer FR, Smith I (1999) Differential expression of 10 σ factor genes in *Mycobacterium tuberculosis*. *Mol Microbiol* 31:715–724.
19. Gamulin V, Cetkovic H, Ahel I (2004) Identification of a promoter motif regulating the major DNA damage response mechanism of *Mycobacterium tuberculosis*. *FEMS Microbiol Lett* 238:57–63.
20. Yang W, Woodgate R (2007) What a difference a decade makes: Insights into translesion DNA synthesis. *Proc Natl Acad Sci USA* 104:15591–15598.
21. Ling H, Boudsocq F, Woodgate R, Yang W (2001) Crystal structure of a Y-family DNA polymerase in action: A mechanism for error-prone and lesion-bypass replication. *Cell* 107:91–102.
22. Koorits L, et al. (2007) Study of involvement of ImuB and DnaE2 in stationary-phase mutagenesis in *Pseudomonas putida*. *DNA Repair (Amst)* 6:863–868.
23. Bailey S, Wing RA, Steitz TA (2006) The structure of *T. aquaticus* DNA polymerase III is distinct from eukaryotic replicative DNA polymerases. *Cell* 126:893–904.
24. Lamers MH, Georgescu RE, Lee SG, O'Donnell M, Kuriyan J (2006) Crystal structure of the catalytic α subunit of *E. coli* replicative DNA polymerase III. *Cell* 126:881–892.
25. Kana BD, et al. (2010) Role of the DinB homologs Rv1537 and Rv3056 in *Mycobacterium tuberculosis*. *J Bacteriol* 192:2220–2227.
26. Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6:197–208.
27. Goodman MF (2002) Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu Rev Biochem* 71:17–50.
28. Kim DR, McHenry CS (1996) Biotin tagging deletion analysis of domain limits involved in protein-macromolecular interactions. Mapping the τ binding domain of the DNA polymerase III α subunit. *J Biol Chem* 271:20690–20698.
29. López de Saro FJ, Georgescu RE, O'Donnell M (2003) A peptide switch regulates DNA polymerase processivity. *Proc Natl Acad Sci USA* 100:14689–14694.
30. Gawel D, Pham PT, Fijalkowska IJ, Jonczyk P, Schaaper RM (2008) Role of accessory DNA polymerases in DNA replication in *Escherichia coli*: Analysis of the *dnaX36* mutator mutant. *J Bacteriol* 190:1730–1742.
31. Borden A, et al. (2002) *Escherichia coli* DNA polymerase III can replicate efficiently past a T-T cis-syn cyclobutane dimer if DNA polymerase V and the 3' to 5' exonuclease proofreading function encoded by *dnaQ* are inactivated. *J Bacteriol* 184:2674–2681.
32. Vandewiele D, Borden A, O'Grady PI, Woodgate R, Lawrence CW (1998) Efficient translesion replication in the absence of *Escherichia coli* Umu proteins and 3'-5' exonuclease proofreading function. *Proc Natl Acad Sci USA* 95:15519–15524.
33. Stano NM, Chen J, McHenry CS (2006) A coproofreading Zn(2+)-dependent exonuclease within a bacterial replicase. *Nat Struct Mol Biol* 13:458–459.
34. Baños B, Lázaro JM, Villar L, Salas M, de Vega M (2008) Editing of misaligned 3'-termini by an intrinsic 3'-5' exonuclease activity residing in the PHP domain of a family X DNA polymerase. *Nucleic Acids Res* 36:5736–5749.
35. Nakane S, Nakagawa N, Kuramitsu S, Masui R (2009) Characterization of DNA polymerase X from *Thermus thermophilus* HB8 reveals the POLXc and PHP domains are both required for 3'-5' exonuclease activity. *Nucleic Acids Res* 37:2037–2052.
36. Wieczorek A, McHenry CS (2006) The NH₂-terminal php domain of the α subunit of the *Escherichia coli* replicase binds the ϵ proofreading subunit. *J Biol Chem* 281: 12561–12567.
37. Dawes SS, Mizrahi V (2001) DNA metabolism in *Mycobacterium leprae*. *Lepr Rev* 72: 408–414.
38. Cirz RT, O'Neill BM, Hammond JA, Head SR, Romesberg FE (2006) Defining the *Pseudomonas aeruginosa* SOS response and its role in the global response to the antibiotic ciprofloxacin. *J Bacteriol* 188:7101–7110.
39. Sanders LH, Rockel A, Lu H, Wozniak DJ, Sutton MD (2006) Role of *Pseudomonas aeruginosa* *dinB*-encoded DNA polymerase IV in mutagenesis. *J Bacteriol* 188: 8573–8585.
40. Datta S, et al. (2003) Crystal structures of *Mycobacterium smegmatis* RecA and its nucleotide complexes. *J Bacteriol* 185:4280–4284.
41. Warner DF, Mizrahi V (2006) Tuberculosis chemotherapy: The influence of bacillary stress and damage response pathways on drug efficacy. *Clin Microbiol Rev* 19: 558–570.
42. Kohanski MA, DePristo MA, Collins JJ (2010) Sublethal antibiotic treatment leads to multidrug resistance via radical-induced mutagenesis. *Mol Cell* 37:311–320.
43. Smith PA, Romesberg FE (2007) Combating bacteria and drug resistance by inhibiting mechanisms of persistence and adaptation. *Nat Chem Biol* 3:549–556.
44. Parish T, Stoker NG (2000) Use of a flexible cassette method to generate a double unmarked *Mycobacterium tuberculosis* *tlyA* *plcABC* mutant by gene replacement. *Microbiology* 146:1969–1975.
45. Smith AM and Klugman KP (1997) "Megaprimer" method of PCR-based mutagenesis: The concentration of megaprimer is a critical factor. *Biotechniques* 22:438, 442.
46. Lee MH, Pascopella L, Jacobs WR, Jr, Hatfull GF (1991) Site-specific integration of mycobacteriophage L5: Integration-proficient vectors for *Mycobacterium smegmatis*, *Mycobacterium tuberculosis*, and bacille Calmette-Guérin. *Proc Natl Acad Sci USA* 88: 3111–3115.
47. Pham TT, Jacobs-Sera D, Pedulla ML, Hendrix RW, Hatfull GF (2007) Comparative genomic analysis of mycobacteriophage Tweety: Evolutionary insights and construction of compatible site-specific integration vectors for mycobacteria. *Microbiology* 153:2711–2723.
48. Venclovas Č, Margelevičius M (2009) The use of automatic tools and human expertise in template-based modeling of CASP8 target proteins. *Proteins* 77 (Suppl 9):81–88.
49. Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402.
50. Margelevičius M, Venclovas Č (2005) PSI-BLAST-ISS: An intermediate sequence search tool for estimation of the position-specific alignment reliability. *BMC Bioinformatics* 6:185.
51. Söding J (2005) Protein homology detection by HMM-HMM comparison. *Bioinformatics* 21:951–960.
52. Sadreyev R, Grishin N (2003) COMPASS: A tool for comparison of multiple protein alignments with assessment of statistical significance. *J Mol Biol* 326:317–336.
53. Margelevičius M, Venclovas Č (2010) Detection of distant evolutionary relationships between protein families using theory of sequence profile-profile comparison. *BMC Bioinformatics* 11:89.
54. Šali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234:779–815.
55. Canutescu AA, Shelenkov AA, Dunbrack RL, Jr (2003) A graph-theory algorithm for rapid protein side-chain prediction. *Protein Sci* 12:2001–2014.
56. Sippl MJ (1993) Recognition of errors in three-dimensional structures of proteins. *Proteins* 17:355–362.
57. Xing X, Bell CE (2004) Crystal structures of *Escherichia coli* RecA in a compressed helical filament. *J Mol Biol* 342:1471–1485.