The Mycobacterial LexA/RecA-Independent DNA Damage Response Is Controlled by PafBC and the Pup-Proteasome System

Müller, Andreas U; Imkamp, Frank; Weber-Ban, Eilika

Abstract: Mycobacteria exhibit two DNA damage response pathways: the LexA/RecA-dependent SOS response and a LexA/RecA-independent pathway. Using a combination of transcriptomics and genome-wide binding site analysis, we demonstrate that PafBC (proteasome accessory factor B and C), encoded in the Pup-proteasome system (PPS) gene locus, is the transcriptional regulator of the predominant LexA/RecA-independent pathway. Comparison of the resulting PafBC regulon with the DNA damage response of Mycobacterium smegmatis reveals that the majority of induced DNA repair genes are up-regulated by PafBC. We further demonstrate that RecA, a member of the PafBC regulon and principal regulator of the SOS response, is degraded by the PPS when DNA damage stress has been overcome. Our results suggest a model for the regulation of the mycobacterial DNA damage response that employs the concerted action of PafBC as master transcriptional activator and the PPS for removal of DNA repair proteins to maintain a temporally controlled stress response.

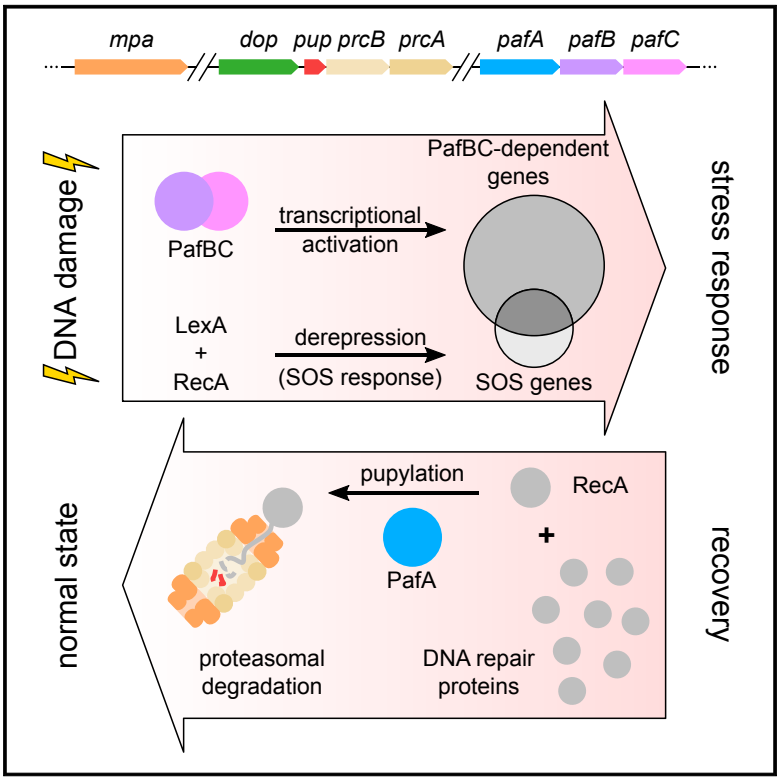
DOI: https://doi.org/10.1016/j.celrep.2018.05.073

Article

Cell Reports

The Mycobacterial LexA/RecA-Independent DNA Damage Response Is Controlled by PafBC and the Pup-Proteasome System

Graphical Abstract



Authors

Andreas U. Müller, Frank Imkamp, Eilika Weber-Ban

Correspondence

eilika@mol.biol.ethz.ch

In Brief

Müller et al. demonstrate that the Pup-proteasome gene locus harbors a transcriptional activator, PafBC, acting as the principal regulator of the mycobacterial DNA damage response, upregulating the majority of DNA repair genes. During stress recovery, the Pup-proteasome system removes DNA repair proteins to return the cell to its normal state.

Highlights

- PafBC is the key transcriptional activator of the mycobacterial DNA damage response
- PafBC upregulates 150 genes via the RecA-NDp motif
- Induced genes are involved in DNA damage repair and the oxidative stress response
- The Pup-proteasome system removes DNA repair proteins during stress recovery

Data and Software Availability

E-MTAB-6497
E-MTAB-6503

Müller et al., 2018, Cell Reports *23*, 3551–3564
June 19, 2018
https://doi.org/10.1016/j.celrep.2018.05.073

CellPress

Cell Reports
Article

The Mycobacterial LexA/RecA-Independent DNA Damage Response Is Controlled by PafBC and the Pup-Proteasome System

Andreas U. Müller,[1] Frank Imkamp,[2] and Eilika Weber-Ban[1,3,*]
[1]ETH Zurich, Institute of Molecular Biology and Biophysics, 8093 Zurich, Switzerland
[2]University of Zurich, Institute of Medical Microbiology, 8006 Zurich, Switzerland
[3]Lead Contact
*Correspondence: eilika@mol.biol.ethz.ch
https://doi.org/10.1016/j.celrep.2018.05.073

SUMMARY

Mycobacteria exhibit two DNA damage response pathways: the LexA/RecA-dependent SOS response and a LexA/RecA-independent pathway. Using a combination of transcriptomics and genome-wide binding site analysis, we demonstrate that PafBC (proteasome accessory factor B and C), encoded in the Pup-proteasome system (PPS) gene locus, is the transcriptional regulator of the predominant LexA/RecA-independent pathway. Comparison of the resulting PafBC regulon with the DNA damage response of *Mycobacterium smegmatis* reveals that the majority of induced DNA repair genes are upregulated by PafBC. We further demonstrate that RecA, a member of the PafBC regulon and principal regulator of the SOS response, is degraded by the PPS when DNA damage stress has been overcome. Our results suggest a model for the regulation of the mycobacterial DNA damage response that employs the concerted action of PafBC as master transcriptional activator and the PPS for removal of DNA repair proteins to maintain a temporally controlled stress response.

INTRODUCTION

The integrity of a cell's genome is essential for survival and stable proliferation. Therefore, all living organisms respond to DNA damage by expressing genes that help repair or circumvent the damage. Mycobacteria are frequently exposed to conditions that can cause DNA damage (Namouchi et al., 2016; Singh, 2017). *Mycobacterium tuberculosis* (Mtb), one of the most successful human pathogens, persists inside host macrophages, the very cells forming the body's first line of defense, withstanding their defensive onslaught of reactive oxygen species and nitrogen intermediates (Adams et al., 1997; MacMicking et al., 1997).

Mycobacteria harbor multiple, partially redundant DNA repair pathways, including base and nucleotide excision repair pathways (BER and NER), homologous recombination (HR), non-homologous end joining (NHEJ), and single strand annealing (SSA) (Singh, 2017). Mitomycin C (MMC), an agent commonly used to establish DNA-damaging conditions in bacterial cultures *in vitro*, is a natural antibiotic derived from *Streptomyces caespitosus*. MMC is reduced by bacterial flavoreductases to a potent bifunctional DNA-alkylating agent that forms monoadducts as well as crosslinked adducts in complementary CpG sequences by N-alkylating guanine nucleosides at the 2-amino group of the base (Tomasz, 1995). The DNA modifications caused by MMC lead to inhibition of DNA synthesis, double-strand breaks, and mutagenesis (Dronkert and Kanaar, 2001). In response to MMC-induced DNA damage, Mtb induces over 100 genes, many of which are involved in DNA repair pathways (Rand et al., 2003).

In *E. coli* and many other bacteria, DNA repair is regulated by the LexA/RecA-dependent SOS response, where DNA repair genes are repressed by binding of the repressor LexA to a sequence preceding these genes referred to as "SOS box" (Shinagawa, 1996). Upon association with recombinase A (RecA) filaments, formed in presence of single stranded DNA (Galletto et al., 2006; Joo et al., 2006), LexA is autocatalytically cleaved (Little et al., 1980), thereby derepressing DNA repair genes. However, the vast majority of inducible DNA repair genes in Mtb are induced independently of LexA/RecA, as among 112 genes consistently induced upon exposure to MMC, only 21 were regulated solely by LexA/RecA and could no longer be upregulated in a *recA*-deficient strain (Rand et al., 2003). Notably, *recA* itself is preceded by two promoters, the LexA/RecA-dependent promoter and a RecA-independent promoter termed RecA-NDp (Ahel et al., 2002; Brooks et al., 2001; Davis et al., 2002; Gamulin et al., 2004; Movahedzadeh et al., 1997). RecA-NDp precedes many of the DNA repair genes induced independent of LexA/RecA and hence the DNA damage response under control of RecA-NDp is referred to as "LexA/RecA-independent pathway." However, the responsible transcriptional regulator and mode of regulation for this pathway remained unclear. A study in mycobacteria suggests that the Clp-gene regulator ClgR is involved in regulating the LexA/RecA-independent pathway (Wang et al., 2011). However, this notion is opposed by the observation that deletion of *clgR* in *Corynebacterium glutamicum* does not affect the transcriptional response to UV exposure or MMC stress (Engels et al., 2005). Interestingly, we recently identified PafBC (proteasome

accessory factor B and C) as a transcriptional regulator that induces *recA* via the LexA/RecA-independent promoter (Fudrini Olivencia et al., 2017).

In mycobacteria and most other actinobacteria, PafBC is associated with the Pup-proteasome system (PPS) gene locus that encodes a ubiquitin-like tagging pathway and the proteasomal degradation machinery (Festa et al., 2007; Fudrini Olivencia et al., 2017). Mycobacterial PafBC is expressed from a single operon together with PafA (proteasome accessory factor A), the ligase that covalently attaches prokaryotic ubiquitin-like protein (Pup) to lysine side chains of proteins in order to render them as substrates for the bacterial proteasome (Pearce et al., 2008; Striebel et al., 2009, 2014). The PPS was shown to support persistence of Mtb inside macrophages (Darwin, 2009; Gandotra et al., 2007).

Here, we define the genome-wide regulon of PafBC in *Mycobacterium smegmatis* by a combination of transcriptomics using next-generation sequencing (RNA-seq) and co-immunoprecipitation with PafBC-specific antibodies followed by massively parallel DNA-sequencing (ChIP-seq). Our results establish PafBC as the regulator of the LexA/RecA-independent DNA damage response pathway. PafBC acts as a global activator, directly upregulating transcription of more than 140 genes in response to MMC exposure, demonstrating that transcriptional activation is a key DNA damage response mechanism in mycobacteria. Furthermore, we provide a functional link to the PPS gene locus by demonstrating that the PPS contributes to the recovery from DNA stress by degradation of key DNA repair proteins as shown on the example of RecA.

RESULTS

PafBC Acts as a Transcriptional Activator of a Large Number of Genes under MMC Stress

We have recently identified PafBC as a transcriptional regulator of the genes encoding RecA and UvrB and showed that PafBC controls *recA* expression via the LexA/RecA-independent promoter (Fudrini Olivencia et al., 2017). In order to investigate if PafBC plays a wider role in the mycobacterial DNA damage response, we used a combined transcriptomics and genomic binding site-mapping approach. To identify the regions of the Msm genome where PafBC binds under DNA stress, Msm cultures were exposed to MMC. Subsequent crosslinking covalently arrested PafBC at its DNA interaction sites. Excision from genomic DNA followed by immunoprecipitation with anti-PafBC antibodies allowed isolation of the genomic ensemble of PafBC-DNA fragments. Only DNA specifically bound to PafBC was enriched, because no DNA was obtained in ChIP-seq experiments using either an unspecific rabbit IgG antibody on wild-type Msm lysate or using the PafBC antibody on Msm Δ*pafBC* lysate (Figure S1). Analysis of the co-immunoprecipitated DNA fragments by next-generation sequencing (ChIP-seq) revealed enrichment of those regions in the Msm genome representing the PafBC binding sites (Figure 1, blue bars; Table S1). A total of 127 peaks were called, most of which were located in intergenic regions. The identified number of binding sites supports a global role of PafBC as transcriptional regulator under DNA stress conditions. In order to compare our finding to a naturally occurring stressor, an additional ChIP-seq analysis was carried out with Msm cultures that were exposed to oxidative stress (H_2O_2). Notably, the resulting peak pattern is very similar to the one obtained under DNA damaging conditions with 107 overlapping peaks (Figure S2). A likely explanation is that MMC and H_2O_2 both modify nucleic acid bases (Imlay and Linn, 1988; Tomasz, 1995), thus generating a similar stimulus within the cell.

To assess the level and the activating or repressing nature of the regulation by PafBC, we carried out whole transcriptome next-generation sequencing (RNA-seq) on the Msm parent and *pafBC* knockout strains grown under standard conditions or MMC stress. Differentially expressed genes are shown as red (decreased in Δ*pafBC*) or gray (increased in Δ*pafBC*) bars in Figure 1 (middle two circles, Table S2). Two observations can be made: first, there are vastly more differentially expressed genes under DNA stress conditions (165) than under standard conditions (27) (see also Figure 2A), indicating PafBC is acting under DNA damage stress; second, the overwhelming majority of differentially expressed genes exhibit lower transcript levels in absence of PafBC, clearly demonstrating that PafBC functions as transcriptional activator.

Genes Upregulated by PafBC Are Mainly Involved in DNA Recombination and Repair

To assess the contribution of PafBC to the DNA damage response of Msm, we compared the genes whose mRNA levels were changed upon MMC stress exposure in the wild-type versus the Δ*pafBC* strain. In the wild-type strain, 243 genes exhibited changed mRNA levels in response to DNA-damaging conditions, while the transcript levels of only 63 genes changed in the *pafBC* knockout (Figure 2B; Table S2). These data point to a central role of PafBC in the DNA damage response of Msm.

To analyze the context in which these genes function, we compared Msm wild-type and Msm Δ*pafBC* transcriptomes under MMC stress and functionally classified the differentially expressed genes according to the cluster of orthologous genes (COG) data base (Figures 2C and 2D; Table S2) (Huerta-Cepas et al., 2016; Tatusov et al., 1997). Occasionally, a gene was assigned to more than one functional category (COG class), accounting for minor differences in the number of genes and class assignments displayed in our graphs. Among the 165 genes affected by deletion of *pafBC*, the largest group with a functional connotation is represented by COG class L (35 genes identified), coding for proteins involved in replication, recombination, and repair. Normalizing the number of identified members of a COG class with the total number of this class in the genome revealed that functional class V is strongly enriched, followed by class L (Figure 2D). Class V codes for proteins involved in defense mechanisms such as restriction endonucleases or toxin-antitoxin systems.

PafBC Regulates Its Target Genes through the RecA-Independent Promoter Motif RecA-NDp

To determine the sequence specificity of PafBC binding, the ChIP-seq enriched genomic regions were analyzed using the MEME motif discovery algorithm (Bailey et al., 2009). Of 127 sequences submitted for motif analysis, 117 were used by the algorithm to build a 22 nt long motif (MEME logo 1) closely

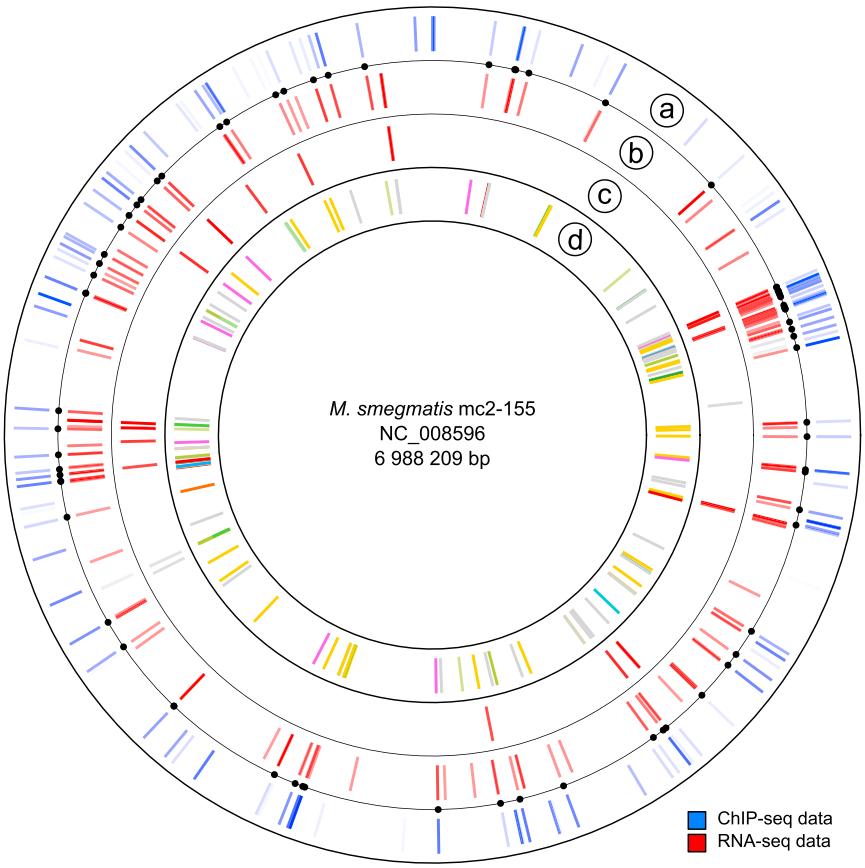


Figure 1. Combined Genome-wide Transcriptomics (RNA-Seq) and Co-immunoprecipitation of PafBC-Crosslinked DNA Fragments (ChIP-Seq) Reveal a Role for PafBC as a Transcriptional Activator under MMC Stress

(A–C) Identified PafBC binding sites plotted on the circular reference genome of Msm (blue bars, outermost circle, shaded according to peak shape score) with a p value < 0.005 (A), and differentially expressed genes (decreased red/increased gray, inner two circles) in the Msm Δ*pafBC* strain exhibiting an absolute log2-fold change >1.5 and an FDR-corrected p value < 0.001 under MMC stress (B) or standard conditions (C). Color shades correspond to the intensity of differential expression. Dots between circles (A) and (B) mark an assignment of a PafBC binding site to a differentially expressed gene within 300 bp distance.

(D) Functional classification of the dataset in (B) according to COG. Color legend can be found in Figures 2 and 4.

See also Figures S1 and S2 and Tables S1 and S2.

resembling the RecA-NDp motif (Gamulin et al., 2004) (Figures 3A, upper panel, and 3B). DREME, an algorithm scanning for very short motifs, determined a 6 nt long motif (TGTCGG; Figure 3A, lower panel) corresponding to the beginning of MEME logo 1. The identified motifs were tested for central enrichment in the input sequences (Figure 3C). Assuming each binding site of a transcriptional regulator targeted by ChIP represents the single point of crosslinking, high clustering of motif positions toward the center of the peak regions would be expected (Bailey and Machanick, 2012). Interestingly, the positioning of the PafBC binding motif (MEME logo 1) shows the highest enrichment 10 nt away from the center while the DREME logo is enriched at a distance of ~20 nt from the center.

Additionally, the motif discovery algorithm identified the LexA binding motif in a minority of the input sequences (15 of 127), but without significant central enrichment (Figure 3D). This agrees with previous studies, reporting that several genes of the DNA damage response exhibit both a LexA motif (SOS box) and the RecA-NDp motif in their promoter region (Gamulin et al., 2004), and only a minority of the DNA repair genes in mycobacteria are inducible via the canonical LexA/RecA pathway (Rand et al., 2003).

Definition of the PafBC Regulon

Correlating the PafBC binding sites with the genes whose transcript levels differ between the *pafBC* knockout and wild-type strains, we can define the regulon for PafBC under MMC stress. Because two independent sources of information (ChIP-seq and RNA-seq) were used, the threshold of log2-fold changes for the transcript levels can be lowered while retaining high confidence. We therefore assigned genes exhibiting an absolute log2-fold change >0.8 (>1.74 fold change) and a false discovery rate (FDR)-adjusted p value < 0.001 in the Δ*pafBC*^{MMC} versus wild-type^{MMC} comparison to the binding sites identified in the ChIP-seq (p value < 0.005) within a distance of 500 bp. Of 127 binding sites identified by ChIP-seq, 87 sites could be assigned to a gene from the transcriptomics data (8 sites received two assignments; 95 assignments in total; 40 sites unassigned). Unassigned sites are mostly located next to genes not differentially expressed according to the applied thresholds. Additionally, every gene located adjacent to an assigned gene in the same orientation and displaying differential expression (>1.74 fold change; FDR-adjusted p value < 0.001) was added to the putative operon, resulting in the addition of 56 putative operon members. The PafBC regulon therefore comprises all genes and putative operons that are differentially expressed between Δ*pafBC*^{MMC} and wild-type^{MMC} and at the same time feature an upstream PafBC binding site (Figure 4A, purple circle; Table S3). Of the 152 genes (42 putative operons, 52 individually regulated genes) in the PafBC regulon, 150 genes showed decreased transcript levels in the *pafBC* knockout under MMC stress (compared to wild-type).

PafBC Activates More Than Half of All Genes Upregulated in the MMC Stress Response

In order to assess the contribution of PafBC to the overall MMC stress response in Msm, we compared the PafBC regulon to the differentially expressed genes under MMC stress in the wild-type

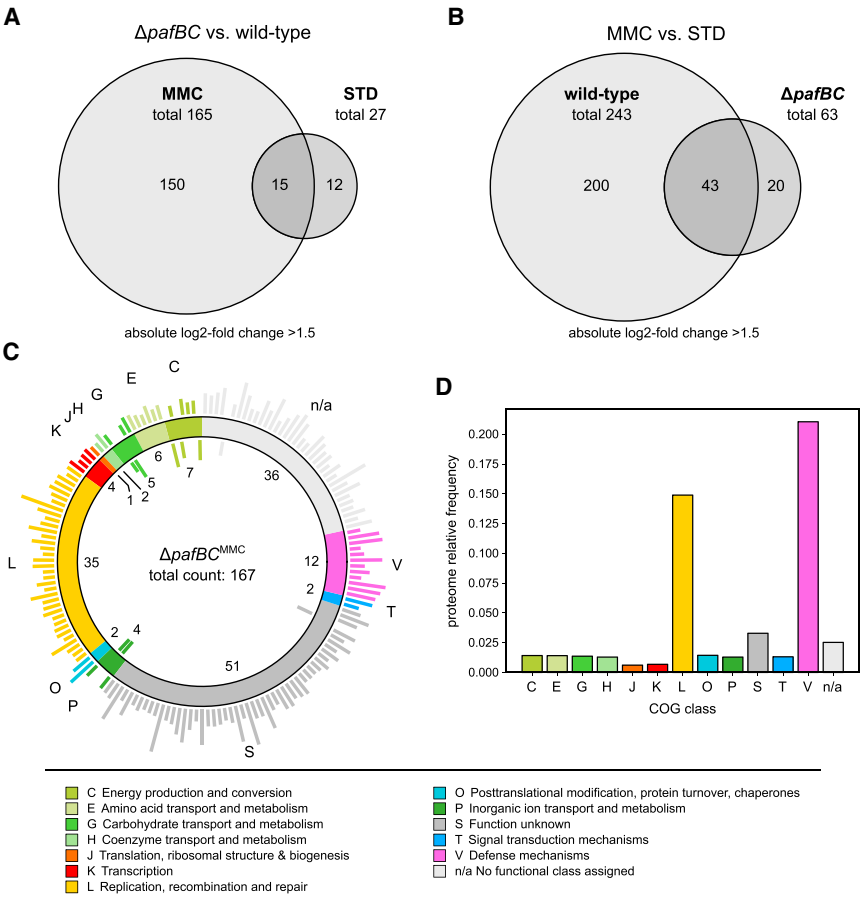


Figure 2. Differentially Expressed Genes Fall Mainly into the Functional Group of "Replication, Recombination, and Repair" (COG Class L)

(A and B) Only a small number of genes is differentially expressed in the Msm Δ*pafBC* strain under standard conditions (STD) compared to MMC stress (A), while upregulation of DNA damage response genes is greatly impaired in the Msm Δ*pafBC* strain (B).

(C) Functional classes of the differentially expressed genes in the Msm Δ*pafBC* strain under MMC stress (dataset Δ*pafBC*^{MMC} in [A]) according to COG comprise mainly genes classified as L ("replication, recombination, and repair"). Log2-fold change values of the individual members of each functional class are depicted around the ring chart (decreased outward, increased inward).

(D) Normalization of differentially expressed genes to the total number of genes in the Msm genome within each functional class reveals an over-representation of classes L and V in the Δ*pafBC*^{MMC} dataset. Total counts in the panels may differ because in a few cases a gene is assigned to more than one functional class in the COG database.

See also Table S2.

strain (>1.74 fold change; FDR-adjusted p value < 0.001). The mRNA levels of 262 genes were increased and the levels of 169 genes were decreased in the wild-type strain upon MMC exposure (431 total; Figure 4A, gray circle). Notably, more than half of the genes with increased levels in the MMC stress response are regulated by PafBC (142 of 262 genes).

Functional categorization of the MMC stress response genes according to the COG classification showed that of classes with known function, class L, coding for proteins involved in DNA replication, recombination, and repair, constitutes the largest group (58 genes) with almost all members upregulated (Figure 4B). On the other hand, the MMC stress response was also found to decrease the transcript levels of a significant number of genes in other classes, for example class G (carbohydrate transport and metabolism), class C (energy production and conversion) and class E (amino acid transport and metabolism).

The PafBC regulon shows a clear enrichment in class L (40 genes; Figure 4C) equivalent to 26% of the entire regulon. Considering that only approximately half of the genes in the regulon have a genuinely assigned function, this is equivalent to 50% of the functionally assigned genes in the regulon. Comparison of Figures 4B and 4C demonstrates that PafBC regulates the majority of the DNA repair genes induced under MMC stress (see Table 1 for a selection). Most known mycobacterial repair pathways are represented with one or more genes in the PafBC regulon (Tables 1 and S3). Proteins involved in all three stages of homologous recombination are upregulated: end resectioning enzymes RecBCD and AdnAB (Sinha et al., 2009) that generate the 3′ single-stranded overhang, recombinase RecA that mediates strand exchange, and the enzymes involved in resolving Holliday junctions, helicase RuvA and RecG as well as resolvase RuvC. PafBC also upregulates base excision repair DNA glycosylase Tag1 predicted to excise 3-methyl adenine from DNA and the type VIII endonuclease Nei2, usually responsible for removing oxidized bases and introducing a single-strand break at the site of the removed base (AP-lyase activity). Endonuclease III Nth and endonuclease IV End that cleave the backbone at abasic sites are also upregulated. The nucleotide excision repair pathway is represented by the three enzymes of the UvrABC excision machinery, which excises a 12 nt stretch around a DNA lesion, along with helicase UvrD1 that removes the excised piece by unpairing from the opposite strand, and DNA polymerase I (PolA) that fills in the gap. Helicase UvrD2 also exhibits PafBC-dependent increased transcript levels (Table S2), but no corresponding peak was detected in the ChIP-seq experiment and hence *uvrD2* was not included in the PafBC regulon.

In addition to DNA repair pathways, the PafBC regulon also encompasses the ergothioneine (EGT) biosynthesis gene cluster (EgtA-E, Table 1). EGT is a redox-active low molecular weight thiol that protects Mtb against oxidative and nitrosative stress in host macrophages and is essential for survival of the bacterium in mice (Richard-Greenblatt et al., 2015; Saini et al., 2016; Salinas and Comini, 2018; Sao Emani et al., 2018).

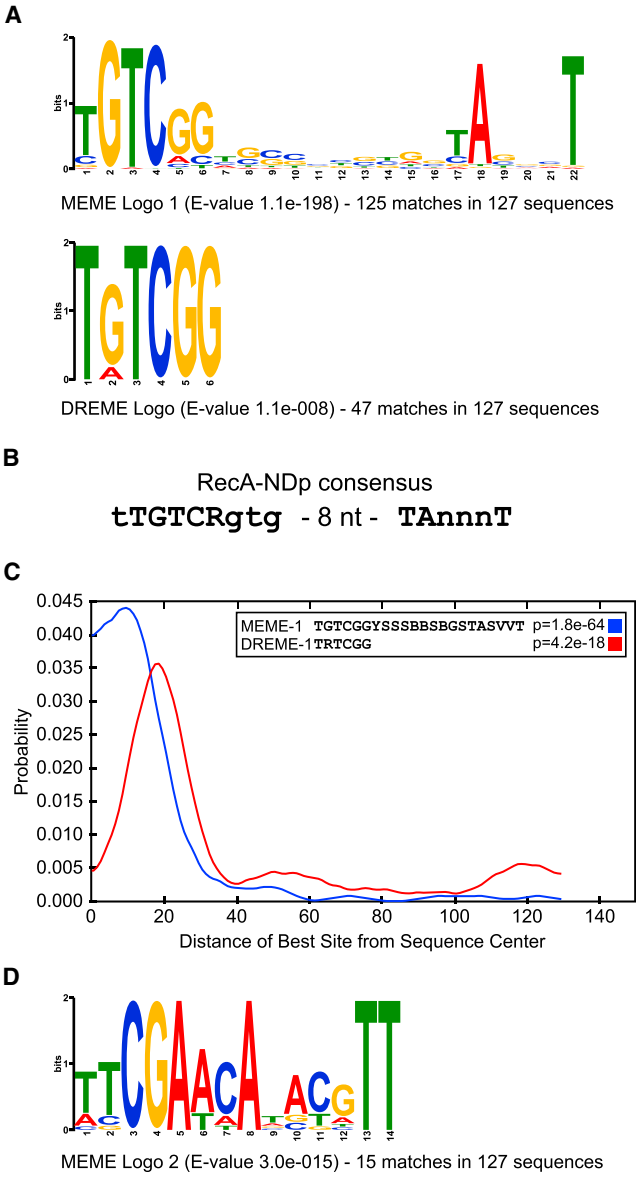


Figure 3. MEME-ChIP Motif Discovery Based on the Underlying Sequences of the ChIP-Seq Enriched Regions for PafBC

(A and B) Logos of the primary motifs identified with the MEME and DREME algorithms (A) closely match the RecA-NDp consensus sequence (B).

(C) The motifs show central enrichment within the ChIP-seq peak regions according to CentriMo analysis.

(D) A second motif identified by the MEME algorithm matches the LexA motif. The input for the analyses comprised ChIP-seq peak sequences of 300 nt length with a p value < 0.005. Motif matches in the input sequences were determined by using FIMO.

See also Table S1.

Furthermore, two sigma factors involved in gene expression under stress conditions are regulated by PafBC: SigH along with its anti-sigma factor RshA, and SigG. In Mtb SigH is activated by heat, oxidative, and nitric oxide stress (Sharp et al., 2016). Its regulon includes members of the oxidative stress response (e.g., thioredoxin, thioredoxin reductase, methionine sulfoxide reductase) and chaperones (Hsp70/40). Only few targets of SigG have been found (Gaudion et al., 2013; Lee et al., 2008). However, SigG is significantly upregulated in Mtb during infection of the macrophage (Cappelli et al., 2006). Additionally, PafBC activates the expression of the transcriptional repressor of the ribonucleotide reductase genes, NrdR.

PafBC, but Not ClgR, Acts via the RecA-NDp Promoter under MMC Stress

Controversial results were reported on the role of the Clp-gene regulator, ClgR, regarding regulation of DNA repair genes (Engels et al., 2005; Wang et al., 2011). To assess how ClgR affects the DNA damage response in mycobacteria, we generated a *clgR* knockout strain of Msm (Figure S3). First, viability of Msm Δ*clgR* in presence of MMC was analyzed using a resazurin-based assay (Figure S4A). There was no difference of Msm Δ*clgR* compared to the wild-type strain, while Msm Δ*pafBC* displayed greatly reduced viability as observed before (Fudrini Olivencia et al., 2017). We then compared the induction of *recA* from different promoter variants upon MMC stress in the wild-type, Δ*pafBC*, and Δ*clgR* strains using a GFP reporter assay (Figure S4B). When GFP was expressed under control of the wild-type *recA* promoter sequence, fluorescence in the wild-type and Δ*clgR* strains increased ~8-fold under MMC stress, while the signal increased only ~3-fold in the Δ*pafBC* strain. We also tested GFP expression from *recA* promoter variants in which either only the LexA/RecA promoter (P2) or the RecA-NDp promoter (P1) was intact. In neither case, the Δ*clgR* strain showed significant differences in GFP fluorescence to the wild-type strain. In accordance with our recent observations (Fudrini Olivencia et al., 2017), no signal increase upon MMC treatment was detected in the Δ*pafBC* strain when only P1 was intact. The unaltered cellular RecA levels in absence of *clgR*—neither under MMC stress nor under standard conditions—strongly suggest that PafBC is so far the only regulator unequivocally shown to induce genes via the RecA-NDp promoter.

The Pup-Proteasome System Plays a Role in Recovery from the DNA Damage Stress Response

Across the group of actinobacteria, *pafB* and *pafC* are tightly associated with the locus encoding the PPS (Figure S5). In mycobacteria, *pafB* and *pafC* are even organized in an operon together with *pafA* (Festa et al., 2007; Fudrini Olivencia et al., 2017). Preservation of operon members during evolution is often indicative of a functional relationship (Moreno-Hagelsieb and Jokic, 2012; Overbeek et al., 1999). Yet, deletion of *pafBC* in Msm has no influence on the protein levels of important PPS members (Fudrini Olivencia et al., 2017). Although we observed a PafBC binding site upstream of *dop*, our transcriptomics analysis did not reveal differences between wild-type and Δ*pafBC* strains in the transcript level of this gene or any of the PPS genes (Table S2). To test whether the PPS might play a role in the DNA damage response, we first assessed the survival of different Msm knockout strains of members of the PPS (Δ*dop*, Δ*pup*, and Δ*prcB*) in comparison to the wild-type and Δ*pafBC* strains in presence of increasing MMC concentrations

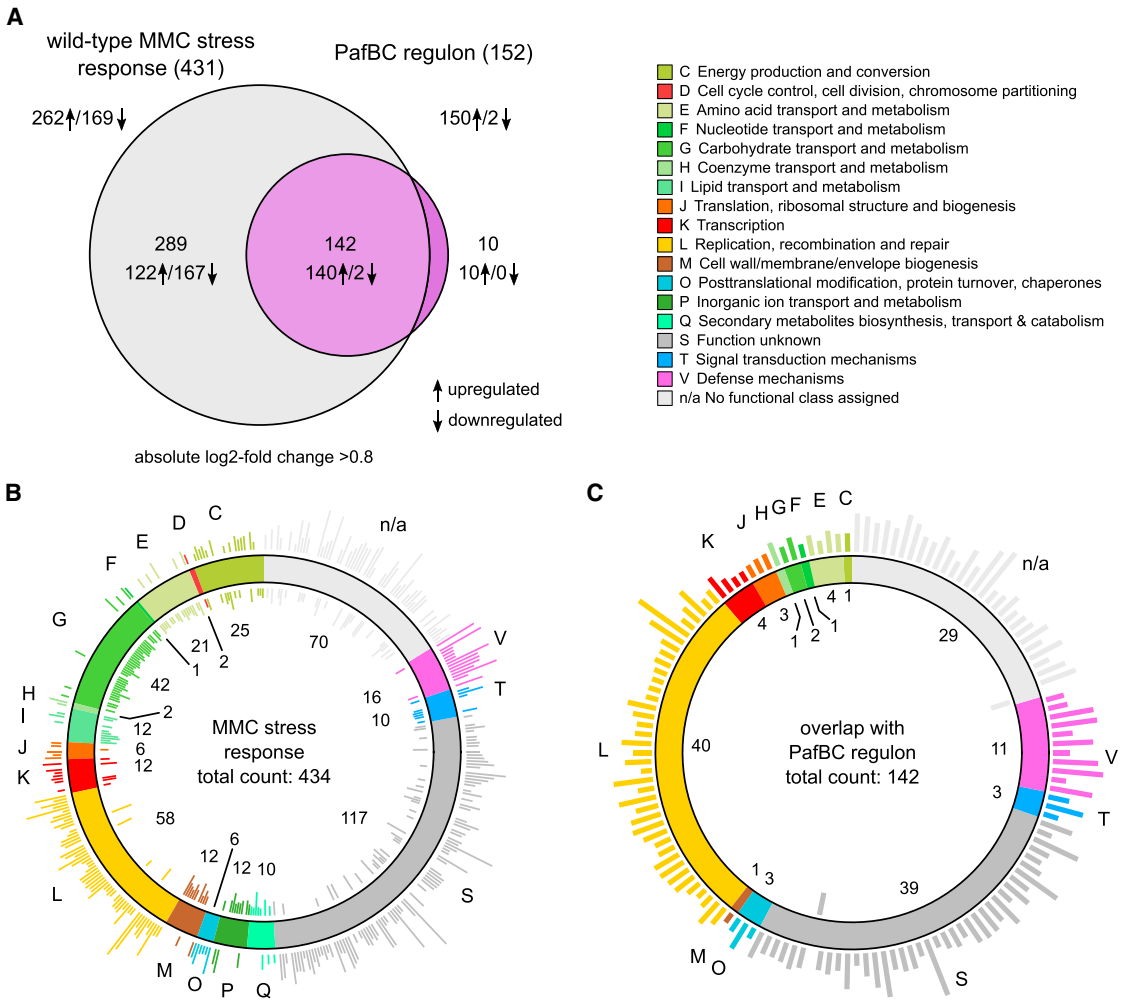


Figure 4. PafBC Regulates the Vast Majority of the DNA Repair Genes under MMC Stress

(A) One-third of the genes differentially expressed during MMC stress (gray) are regulated by PafBC, and more than half of the upregulated genes are contained within the PafBC regulon (purple).

(B and C) The distribution of COG functional classes in the Msm wild-type stress response to MMC is diverse (B), while the PafBC regulon contains more than two thirds of the class L members ("replication, recombination, and repair") of the MMC stress response (C). Differentially expressed genes for the wild-type MMC stress response were chosen with an absolute log2-fold change >0.8 (≈1.74 fold change) and an FDR-adjusted p value < 0.001. Log2-fold change values of the individual members of each functional class are depicted around each ring chart (upregulated outward, downregulated inward). Total counts in the panels may differ because in a few cases a gene is assigned to more than one functional class in the COG database.

See also Tables 1, S2 and S3.

(Figure 5A). Indeed, all PPS knockout strains showed a consistently lower ability to grow under MMC stress compared to wild-type. The growth of the deletion mutants remained unimpaired under standard conditions, exhibiting generation times that were identical within error (Figure S6).

Interestingly, the PafBC regulon includes 17 previously identified pupylation targets (Tables 1 and S3), among them RecA, TopoN, and EgtD (Festa et al., 2010; Poulsen et al., 2010; Watrous et al., 2010; Yun et al., 2012). Alignment of actinobacterial RecA protein sequences revealed high conservation of the Pup-modified lysine residue in actinobacteria harboring a proteasome (Figure 5B). Furthermore, the mid-log phase RecA level in the PPS knockout strains is significantly increased compared to the wild-type (Figure 5C). Thus, the PPS could be responsible for removing upregulated stress response proteins after the stress has been overcome. To test this hypothesis, we followed RecA levels by immunoblotting over the course of DNA damage stress exposure and recovery from the stress. After exposing cultures of the PPS knockout strains, the wild-type strain and Δ*pafBC* strain to MMC for 2 hr, the stress-inducing agent was removed to initiate the recovery phase. The level of RecA increased in all strains during the MMC exposure phase and continued to increase during the recovery phase, reaching its maximum level ~6–9 hr after the removal of MMC (Figure 5D). RecA returned nearly to pre-MMC stress levels in the wild-type and Δ*pafBC* strain by hour 24 of the recovery

Table 1. Overview of Selected Genes Contained in the PafBC Regulon

Locus Tag	Mtb Rv[a]	Name	COG Class	PafBC Regulon[b]	MMC Stress[c]	Pupylated[d]	Description
DNA Recombination and Repair							
MSMEG_2723	Rv2737c	RecA	L	−1.65	3.55	This study*, Festa et al., 2010*, Yun et al., 2012*	Catalyzes strand exchange, LexA co-protease activity
MSMEG_2724	Rv2736c	RecX	S	−1.68	3.88	–	Inhibits RecA co-protease/ strand exchange activity and filament elongation
MSMEG_3808	Rv1638	UvrA	L	−1.62	1.75	This study, Festa et al., 2010	⎤
MSMEG_3816	Rv1633	UvrB	L	−3.63	2.77	–	Nucleotide excision repair complex UvrABC
MSMEG_3078	Rv1420	UvrC	L	−1.11	1.09	–	⎦
MSMEG_1325	Rv0629c	RecD	L	−2.85	2.57	–	⎤
MSMEG_1327	Rv0630c	RecB	L	−3.25	2.85	–	Helicase-nuclease complex RecBCD/exonuclease V0
MSMEG_1328	Rv0631c	RecC	L	−3.35	2.95	–	⎦
MSMEG_1941	Rv3202c	AdnA	L	−8.17	4.46	This study	Helicase-nuclease complex AdnAB
MSMEG_1943	Rv3201c	AdnB	L	−5.08	3.77	–	⎦
MSMEG_4491	Rv2362c	RecO	L	−1.69	0.92	This study	Annealing of ssDNA
MSMEG_2943	Rv2594c	RuvC	L	−1.03	2.43	This study	Holliday junction resolvase/ crossover junction endodeoxyribonuclease
MSMEG_2944	Rv2593c	RuvA	L	−1.09	2.87	This study*	Holliday junction binding protein
MSMEG_2403	Rv2973c	RecG	L	−1.60	1.40	–	ATP-dependent DNA helicase, supports Holliday junction branch migration
MSMEG_2362	Rv3014c	LigA	L	−2.21	1.86	–	NAD(+)-dependent DNA ligase A
MSMEG_2277	Rv3062	LigB	L	−4.04	2.99	–	ATP-dependent DNA ligase B
MSMEG_6079	Rv3585	RadA	O	−1.41	1.98	–	Branch migration protein
MSMEG_6080	Rv3586	DisA	L	−1.46	1.54	Compton et al., 2015*	DNA integrity scanning protein with cyclic-di-AMP synthase activity
MSMEG_0456	NA	TopoM	L	−1.75	1.48	–	Type II topoisomerase
MSMEG_0457	NA	TopoN	L	−1.79	1.83	This study, Watrous et al., 2010*	⎦
MSMEG_1383	Rv0670	End	L	−2.20	1.78	This study, Festa et al., 2010	Endonuclease IV (End/Nfo)
MSMEG_6187	Rv3674c	Nth	L	−1.16	0.82	This study	DNA AP lyase/endonuclease III
MSMEG_1756	Rv3297	Nei2	L	−4.28	3.49	This study	DNA N-glycosylase with an AP lyase activity/endonuclease VIII
MSMEG_3839	Rv1629	PolA	L	−1.94	1.74	Compton et al., 2015*	DNA polymerase I
MSMEG_6275	Rv3711c	DnaQ	L	−3.28	1.36	–	DNA polymerase III epsilon subunit
MSMEG_4259	Rv2191	Cho	L	−5.07	3.16	–	DNA polymerase III epsilon subunit-like/DnaQ-like
Oxidative Stress Response							
MSMEG_6246	Rv3700c	EgtE	E	−2.51	2.42	–	⎤
MSMEG_6247	Rv3701c	EgtD	S	−2.68	2.37	This study, Poulsen et al., 2010*	⎥

(*Continued on next page*)

Table 1. *Continued*

Locus Tag	Mtb Rv[a]	Name	COG Class	PafBC Regulon[b]	MMC Stress[c]	Pupylated[d]	Description
MSMEG_6248	Rv3702c	EgtC	S	−2.78	2.37	This study	Ergothioneine biosynthesis pathway
MSMEG_6249	Rv3703c	EgtB	S	−2.52	2.48	–	|
MSMEG_6250	Rv3704c	EgtA	H	−2.84	2.47	–	⌋
MSMEG_1242[e]	–	IscS	E	−2.26	1.37	This study*, Watrous et al., 2010	Cysteine desulfurase
Transcriptional Control							
MSMEG_1914	Rv3223c	SigH	K	−0.84	0.91	–	ECF RNA polymerase sigma factor
MSMEG_1915	Rv3221A	RshA	S	−1.03	0.94	–	Anti-sigma factor of SigH
MSMEG_0219	Rv0182c	SigG	K	−2.87	2.09	This study	RNA polymerase sigma factor
MSMEG_2743	Rv2718c	NrdR	K	−0.88	2.52	–	Ribonucleotide reductase repressor

[a]Homologous proteins in Mtb with >50% sequence identity.
[b]Log2-fold changes comparing Δ*pafBC*^{MMC} versus wild-type^{MMC} (FDR-adjusted p value < 0.001).
[c]Log2-fold changes comparing wild-type^{MMC} versus wild-type^{STD} (FDR-adjusted p value < 0.001).
[d]Referenced proteomics analyses marked with an asterisk (*) detected a Pup-modified peptide.
[e]Msm exhibits two IscS-like cysteine desulfurases (MSMEG_1242 and MSMEG_2357) of which only MSMEG_2357 has a homolog in Mtb (Rv3025c).

phase. In contrast, there was no decrease of RecA levels observed for the PPS knockout mutants within the same time frame.

Because stress response proteins are usually present transiently during the stress condition, we isolated pupylated proteins from cultures of Msm Δ*pup* complemented with His_{10}-Strep-Pup exposed to MMC using tandem-affinity purification and analyzed them by liquid chromatography-tandem mass spectrometry (LC-MS/MS). A total of 832 pupylated proteins were identified, and for 64 of those a Pup-modified peptide was detected (Table S4). Comparison with the PafBC regulon shows that 26 proteins are pupylated (Table S3), including important DNA repair proteins such as Nth, Xth, Nei2, the Ruv complex, and AdnA (Table 1).

DISCUSSION

Mycobacteria inhabit various biological niches where they face conditions of oxidative, nutritional, and genotoxic stress (Gupta and Chatterji, 2005). The most prominent example is the obligate pathogen Mtb residing inside phagocytic host immune cells. Persistence in the phagosome of host macrophages is only possible because Mtb has developed strategies to counteract the anti-microbial warfare in the form of reactive oxygen and nitrogen species it is exposed to. In our study, we identify PafBC as a global transcriptional activator responsible for directly upregulating 150 different proteins involved in DNA damage repair pathways and the oxidative stress response, and we establish a functional link to the Pup-proteasome system.

In order to survive genotoxic conditions, bacteria mount a stress response that includes the upregulation of enzymes involved in protecting the proteome and genome and in removing or circumventing incurred DNA damage. With such a response the bacteria acquiesce to slowed or arrested growth, error-prone translesion synthesis mechanisms, and increased energy expense (Namouchi et al., 2016). It is evident that prolongation of this state beyond the conditional requirement would be counterproductive for the organism. In this context, the question of a functional connection between the PafBC-mediated DNA damage response and the Pup-proteasome system gene locus (PPS locus) arises. In all mycobacteria, the genes encoding PafBC are located together in an operon with the Pup ligase *pafA* and in other actinobacteria the *pafBC* operon is to the very least loosely connected to the PPS locus (Figure S5). The PPS locus supports persistence of Mtb in host macrophages (Darwin et al., 2003; Gandotra et al., 2007), an environment where the DNA stress response might also play a role. We show that pupylation-mediated proteasomal degradation counteracts PafBC-mediated upregulation of RecA (Figure 5). Exerting tight control over RecA levels seems to be vital for the cell, as overexpression of RecA without co-expression of the RecA-inhibitor RecX (Drees et al., 2004; Venkatesh et al., 2002) is toxic in Msm (Papavinasasundaram et al., 1998), *Streptomyces lividans* (Vierling et al., 2000), *Pseudomonas aeruginosa* (Sano, 1993), and *Xanthomonas oryzae* (Sukchawalit et al., 2001). RecX is encoded in an operon together with RecA in these organisms. Notably, Mtb *recX* is both co-transcribed with *recA* as well as additionally transcribed from a *recA* internal promoter (Forse et al., 2011). Our results not only confirm the previous identification of RecA as a genuine pupylation target, but also demonstrate its degradation by the proteasome. Furthermore, these data establish that the PPS is indeed involved in the removal of proteins after DNA damage. In case of RecA, it appears to be the main degradation pathway, since PPS impairment fully stabilizes RecA within the observed time frame (Figure 5D). Moreover, the PafBC regulon includes 33 genes whose gene products were found to be pupylation targets (Tables 1, S3, and S4) by our analysis or in recent studies

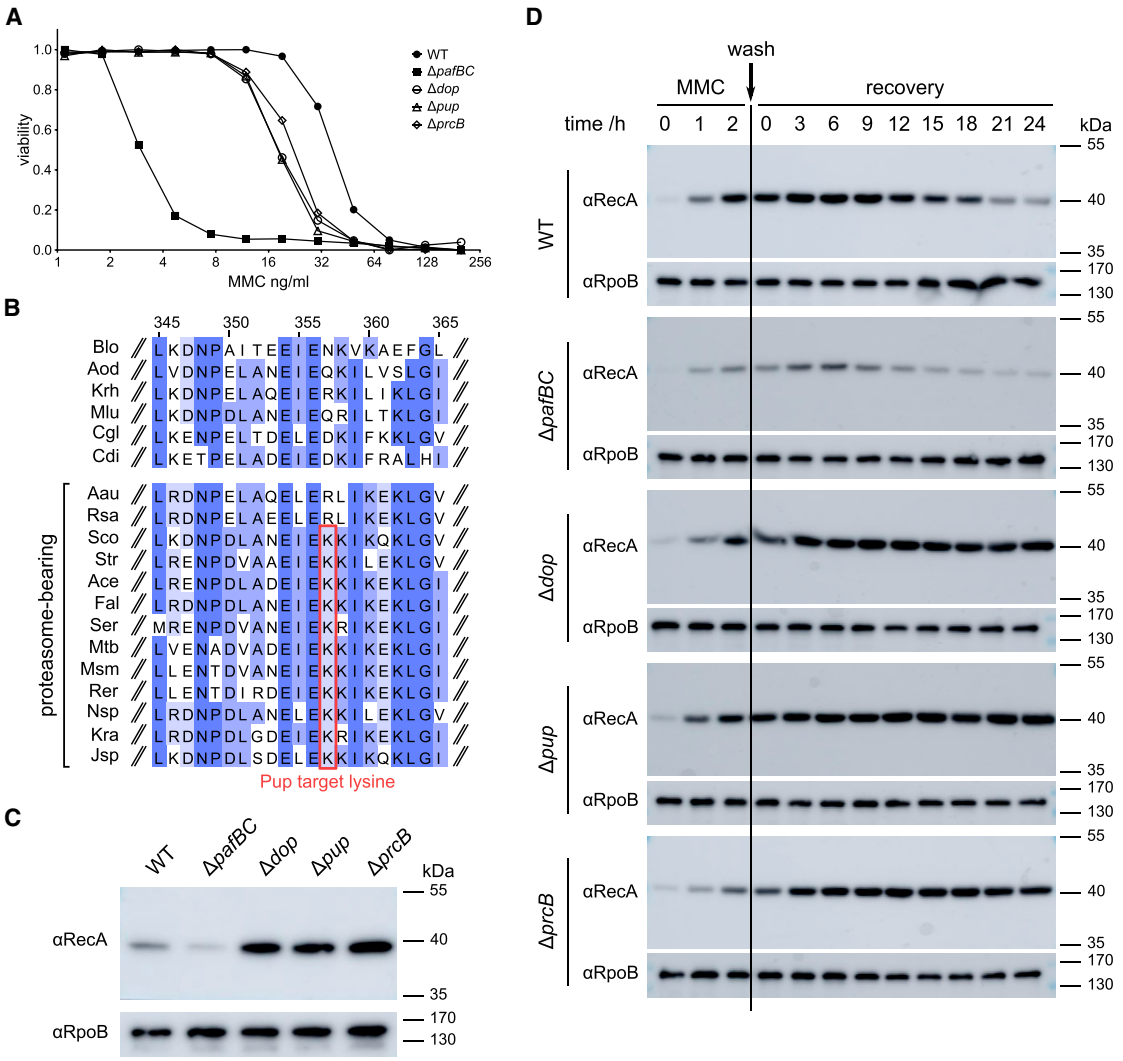


Figure 5. The Pup-Proteasome System Is Required for RecA Degradation during Recovery from the DNA Damage Stress Response

(A) Impaired pupylation sensitizes Msm to MMC. Deletion strains of members of the Pup-proteasome system (PPS) (Msm Δ*dop*, Δ*pup*, and Δ*prcB*) show significantly reduced viability in presence of MMC compared to the wild-type (WT) strain. A representative of three individual experiments is shown.

(B) Alignment of RecA protein sequences from various actinobacteria (alignment positions 345–365 displayed) shows conservation of the pupylated lysine (as determined in Mtb and Rer) in proteasome-bearing actinobacteria. Coloring shows percent identity. Blo, *Bifidobacterium longum*; Aod, *Actinomyces odontolyticus*; Krh, *Kocuria rhizophila*; Mlu, *Micrococcus luteus*; Cgl, *Corynebacterium glutamicum*; Cdi, *Corynebacterium diphtheriae*; Aau, *Paenarthrobacter aurescens*; Rsa, *Renibacterium salmonarium*; Sco, *Streptomyces coelicolor*; Str, *Salinispora tropica*; Ace, *Acidothermus cellulolyticus*; Fal, *Frankia alni*; Ser, *Saccharopolyspora erythrea*; Mtb, *Mycobacterium tuberculosis*; Msm, *Mycobacterium smegmatis*; Rer, *Rhodococcus erythropolis*; Nsp, *Nocardioides sp.*; Kra, *Kineococcus radiotolerans*; Jsp, *Janibacter sp.*

(C) Comparison of RecA levels in Msm wild-type (WT), Δ*pafBC*, Δ*dop*, Δ*pup*, and Δ*prcB* under standard conditions (OD_{600} of 1).

(D) RecA levels monitored over the course of MMC stress and recovery phase in the given Msm strains. Cells were grown to an OD_{600} of 1.0–1.2 and stressed using MMC (80 ng/mL). To remove MMC, cells were washed ("wash") and incubated further. RecA levels were analyzed by immunoblotting at the indicated time points. RpoB was probed in parallel as loading control. A representative result of two independent experiments is shown.

See also Figure S6.

defining the pupylated proteome in various actinobacteria (Compton et al., 2015; Festa et al., 2010; Poulsen et al., 2010; Watrous et al., 2010; Yun et al., 2012). It is likely that proteasomal degradation plays a role in their return to baseline levels as well.

Although mycobacteria possess two major pathways for inducing gene expression upon DNA damage (Brooks et al., 2001; Davis et al., 2002; Movahedzadeh et al., 1997; Rand et al., 2003), only the elements of the LexA/RecA-dependent pathway (SOS response) had been identified so far and it remained unknown how the target genes of the LexA/RecA-independent pathway are induced. Our study shows that PafBC accounts for more than half of the genes upregulated under MMC stress. A common trait of all genes addressed by the LexA/RecA-independent pathway is an upstream promoter

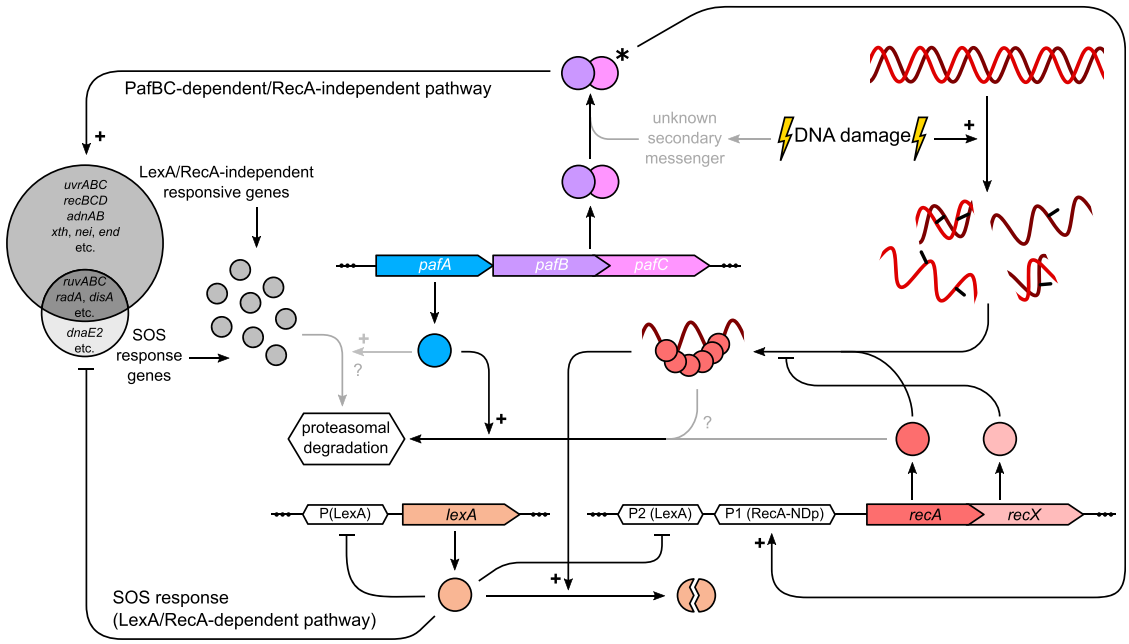


Figure 6. Regulatory Network of the DNA Damage Response in Mycobacteria Involving PafBC as a Key Regulator

In the canonical pathway of the DNA stress response (SOS response), RecA forms filaments on single-stranded DNA generated during DNA damage leading to cleavage and inactivation of the LexA repressor by stimulating its co-protease activity. Relieved repression by LexA enhances expression of LexA, the RecA/RecX operon, and SOS response genes via the LexA/RecA-dependent promoter. In a negative feedback loop, RecX inhibits RecA filament formation, while increased LexA levels lead to shutdown of the SOS response. In the PafBC-dependent/RecA-independent pathway, PafBC is kept at constant levels awaiting the occurrence of an unknown second messenger triggered by DNA damage. Once bound to the unknown ligand (denoted by the asterisk), PafBC is competent to activate expression from the RecA-NDp promoter and controls expression of *recA* and numerous DNA repair genes. In addition, RecA and many other DNA repair proteins are pupylated and removed by the Pup-proteasome system to revert the cell to its initial state.

motif, referred to as RecA-NDp (Gamulin et al., 2004; Rand et al., 2003). Our motif search using the PafBC binding sites resulted in a motif highly similar to the RecA-NDp consensus sequence, arguing that at least in mycobacteria, the RecA/LexA-independent pathway is better referred to as the PafBC-mediated pathway. Importantly, we demonstrate that ClgR is not required for DNA repair induction via the RecA-NDp promoter (Figure S4), showing that the previous data by Wang et al. (2011) are incorrect.

Curiously, the motif was enriched with a small offset to the ChIP-seq peak center (Figure 3C). The relatively large size of PafBC compared to other transcription factors could enable crosslinking of its C-terminal domain to the adjacent DNA region, resulting in a shift of the binding center. Alternatively, the offset might have resulted from formation of a larger complex including PafBC. Transcriptional activation can involve direct binding of the transcription factor to RNA polymerase thereby modulating its activity (Browning and Busby, 2016). If PafBC requires binding to RNA polymerase, the entire complex may be crosslinked to the DNA, again shifting the center of the binding site. It is also possible that yet unknown binders are involved in the transcriptional activation mechanism of PafBC. The latter has been proposed for other transcription factors whose motifs show an offset from the peak center (Bailey and Machanick, 2012; Worsley Hunt et al., 2014). In fact, because PafBC protein levels do not change upon exposure to DNA stress (Fudrini Olivencia et al., 2017), activation of its regulon is likely induced by a response-producing ligand associating with PafBC, rendering it competent for binding to the RecA-NDp motif. This could be a small molecule binder or a macromolecular ligand.

In the LexA/RecA-dependent pathway, transcription of DNA repair genes is suppressed by LexA during non-stress conditions. Derepression upon DNA damage is triggered by single-stranded DNA and upregulation of *recA*, which itself is under control of LexA (Figure 6). With the identification of PafBC as a transcriptional activator and direct regulator of the predominant LexA/RecA-independent (PafBC-mediated) DNA damage response pathway, we demonstrate that mycobacteria employ an active induction mechanism in response to DNA damage via a yet unknown signal. Indeed, PafBC is present in the majority of actinobacteria and the RecA-NDp promoter was shown to precede DNA repair genes in all actinomycetales (Gamulin et al., 2004). Furthermore, LexA/RecA-independent regulation of DNA repair genes has been reported also for other bacterial phyla. In *Pseudomonas aeruginosa* many genes upregulated in response to the DNA gyrase inhibitor ciprofloxacin or the DNA methylating agent methyl-methanesulfonate are not preceded by a LexA promotor (Cirz et al., 2006). A study in *Bacillus subtilis* reported that the levels of over 100 genes were changed independently of the RecA-dependent SOS response upon inhibition of DNA replication (Goranov et al., 2005). In *E. coli*, the important SOS-response gene *uvrB* is under the control of a

LexA-independent promoter in addition to the RecA/LexA promoter (Sancar et al., 1982). Taken together, our study and the observations reported for other bacteria might hint at a widespread existence of a second DNA damage response pathway, mediated by transcriptional activation in addition to the derepression pathway via RecA/LexA.

PafBC appears to be important for the survival of a range of stress conditions ultimately involving DNA repair, because it affects genes involved in most major DNA repair pathways and also genes important for oxidative stress protection, highlighting its general role in protection of the mycobacterial genome. The PafBC binding sites identified under oxidative stress (H_2O_2) using ChIP-seq exhibit a large overlap with those obtained under MMC stress. Likewise, reactive nitrogen species produced by activated macrophages cause deamination and oxidation of bases in DNA, in particular of the guanine base (Burney et al., 1999). This can lead to spontaneous depurination leaving behind abasic (AP) sites. It has been shown that an *E. coli* strain lacking AP endonuclease activity is highly sensitive to NO stress (Spek et al., 2001). PafBC upregulates AP endonucleases End and Nth, indicating that it also supports survival under NO stress. Another mechanism that appears to be important for coping with NO stress is homologous recombination. *E. coli* strains deficient in *recBCD* or the resolvase RuvC show strongly impaired survival when treated with NO gas (Spek et al., 2001). The two main end resectioning complexes in mycobacteria, RecBCD and AdnAB are both part of the PafBC regulon and display considerably lower transcript levels in the *pafBC* knockout strain compared to wild-type under DNA stress conditions. In fact, the transcript level of *adnA* exhibits the biggest difference we observed for any member of the PafBC regulon under MMC stress with an almost 300-fold higher transcript level in the wild-type strain compared to the knockout.

With the discovery of PafBC as the principal activator of the mycobacterial DNA stress response, the complete picture of this important regulatory network begins to emerge (Figure 6): the DNA stress response is launched by a combination of transcriptional derepression and activation involving the repressor LexA and the activator PafBC. In the activation branch, PafBC upregulates the members of its regulon (150 genes) by binding to the RecA-NDp promoter located upstream of each gene or operon in the regulon. As *recA* is a member of the PafBC regulon, PafBC indirectly also influences the LexA/RecA-derepression pathway via induction of *recA* (Figure 6). In fact, RecA levels in the Δ*pafBC* strain are more than one order of magnitude lower (Figure 5) (Fudrini Olivencia et al., 2017), resulting in an impaired response also of the LexA/RecA-dependent pathway as seen on the example of *dnaE2* (Table S2) (Rand et al., 2003; Smollett et al., 2012). At the same time, upregulation of RecA by PafBC also provides RecA activity as part of the double-strand break repair, a pathway strongly induced by MMC as seen on the examples of AdnAB and RecBCD. In the shutdown phase of the response, the functional link between PafBC and the Pup-proteasome gene locus becomes apparent, because RecA and other members of the DNA stress response are pupylation targets and are subject to degradation by the bacterial proteasome complex. Thus, the Pup-proteasome system assists the return to baseline protein levels once DNA stress conditions are overcome.

Our study has established PafBC as the regulator of the LexA/RecA-independent DNA damage response in actinobacteria with a transcriptional activation mode of action. Future work is needed to determine if an analogous "activation branch" in DNA repair gene control might also exist in other bacterial phyla.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Culture Conditions

Strains used were *Mycobacterium smegmatis* (Msm) mc2-155 SMR5 (Sander et al., 1995), Msm Δ*pafBC* (Fudrini Olivencia et al., 2017), Msm Δ*dop* (Imkamp et al., 2010), and Msm Δ*pup* (Fascellaro et al., 2016). The Msm Δ*prcB* strain was generated previously by allelic exchange using the *rpsL* counterselection strategy (Fudrini Olivencia et al., 2017; Sander et al., 1995). Bacteria were routinely grown in Middlebrook 7H9 (supplemented with 0.2% v/v glycerol and 0.05% w/v Tween-80) or on Middlebrook 7H10 agar plates. If applicable, hygromycin B (Carl Roth) or kanamycin (AppliChem) were added to a final concentration of 50 μg/mL or 25 μg/mL, respectively. Liquid cultures were routinely grown at 37°C in a shaking incubator (Infors) and agar plates were incubated at 37°C for growth.

Transcriptomics of Msm Wild-Type and Δ*pafBC* Strains under Standard Growth Conditions and DNA Damage Stress (RNA-Seq)

Msm wild-type and Msm Δ*pafBC* were grown in triplicates in 50 mL shaking cultures to an optical density (OD) 600 of 0.4–0.7. Cultures were incubated for another 1.5 hr in presence or absence of 80 ng/mL mitomycin C (MMC; Roche) to induce DNA damage stress. Briefly, RNA was isolated from 10 mL culture by acidic phenol extraction followed by DNase treatment. Ribosomal RNA was depleted and sequencing libraries were prepared and sequenced. For details, see the Supplemental Experimental Procedures.

The obtained raw sequencing reads were further processed using the CLC Genomics Workbench v10.0.1 software. After trimming and mapping the reads to the reference genome of Msm (NCBI #NC_008596), a statistical analysis for differential expression was performed (for details, see the Supplemental Experimental Procedures). The analysis tool for differential expression of genes models each transcript according to a generalized linear model (GLM) assuming a negative binomial distribution of the read counts similar to EdgeR/DESeq (Anders and Huber, 2010; Robinson et al., 2010). p values of expression level changes were computed using the Wald test and were corrected for FDR (Benjamini and Hochberg, 1995).

Genome-wide PafBC Binding Site Analysis by Chromatin Immunoprecipitation (ChIP-Seq)

Msm wild-type cells were grown as shaking culture to an OD_{600} of 0.5–0.7. The culture was split into 50 mL aliquots from which cells were collected and resuspended in the same volume of fresh medium. DNA damage stress or oxidative stress was induced by addition of MMC (100 ng/mL final) or H_2O_2 (5 mM final) in duplicate cultures and cultures were incubated further. Briefly, formaldehyde crosslinking was performed after 4 hr of stress exposure. Cells were lysed, DNA was sheared by sonication, and protein-DNA complexes were immunoprecipitated from the cleared lysate using an anti-PafBC antibody. After crosslink reversal, protein digestion, and purification of the enriched DNA fragments, sequencing libraries were prepared and sequenced. Enrichment of PafBC-specific DNA fragments was verified by chromatin immunoprecipitation from the *pafBC* deletion strain or using an unspecific antibody (Figure S1). For details, see the Supplemental Experimental Procedures.

The obtained sequencing reads were trimmed and mapped to the reference genome of Msm (NCBI #NC_008596) using CLC Genomics Workbench v8.5.1 (for details, see the Supplemental Experimental Procedures). ChIP-seq peaks with a p value < 0.005 were called using the TF ChIP-seq tool of the software. From the read mapping, the tool learns a filter modeled after a Hotelling observer based on the rationale that highly enriched regions represent positive

examples for ChIP-seq peaks. Such an approach has been previously employed for ChIP-seq peak calling and allows to call peaks based on the characteristic pattern generated by the targeted transcription factor in the immunoprecipitation (Kumar et al., 2013; Stanton et al., 2013).

Of the oxidative stress condition, sample H_2O_2 replicate A exhibited a heavy GC-bias in the reads and was discarded. Both replicates of the MMC stress condition yielded essentially the same binding sites (118 out of 127 for replicate A and 127 for replicate B; Figure S2), and thus, only the data of replicate B were used in all further analyses due to its higher coverage.

Motif Discovery Based on ChIP-Seq Data

Nucleotide sequences underlying the peaks from the ChIP-seq data analysis (p value < 0.005) were extracted from the reference genome with 150 nt up- and downstream of the center of each peak. Sequences were submitted to the MEME-ChIP analysis tool (MEME suite v4.12.0) (Bailey et al., 2009) for motif discovery. For details, see the Supplemental Experimental Procedures.

Functional Classification

The reference proteome of Msm (UniProt #UP000000757) was mapped to COG functional classes (Clusters of Orthologous Groups; revision 4.5) (Huerta-Cepas et al., 2016) using the eggNOG mapper (Huerta-Cepas et al., 2017) with the taxonomic scope set to actinobacteria. Of 6,938 annotated genes in the Msm reference genome (NCBI #NC_008596), 6,717 (96.8%) are listed in UniProt of which 5,512 (82%) could be mapped to a COG class.

Protein Sequence Alignment

Alignment of protein sequences was performed using Clustal Omega (Sievers et al., 2011) with default settings.

Resazurin-Based Viability Assay

Middlebrook 7H9 medium (supplemented with 50 μg/mL streptomycin) containing MMC at the given concentration was inoculated with Msm cells to an OD_{600} of 0.005 and incubated at 37°C for 24 hr. Resazurin was then added to a final concentration of 20 μg/mL and incubation was continued for 3 days at room temperature. Resazurin is reduced by oxidoreductases of living cells to resorufin. Absorbance of resorufin at 605 nm was measured and raw values were normalized within each dataset.

Time-Resolved *In Vivo* RecA Degradation Assay

Msm strains were grown in 300 mL shaking cultures at 37°C in 7H9 medium. Cells were stressed by addition of MMC (80 ng/mL) at an OD_{600} of 1.0 to 1.2. After 2 hr of MMC stress, cells were collected by centrifugation (4,000 × *g*, 10 min, 25°C), washed once in 300 mL fresh 7H9 medium, resuspended in 280 mL fresh 7H9 medium, and transferred to a new, sterile flask to ensure complete removal of MMC. Cultures were incubated for another 24 hr. OD_{600} was intermittently measured to ensure cells from the same growth phase were harvested. Samples of 5 mL culture were withdrawn at the indicated time points and cells were collected by centrifugation (3,200 × *g*, 5 min, 4°C). Pellets were immediately flash-frozen in liquid N_2 and stored at −20°C for later analysis. Immunoblotting for RecA was carried out as previously described (Fudrini Olivencia et al., 2017).

Identification of Pupylated Proteins under DNA Damage Stress

Pupylated proteins were isolated from Msm Δ*pup* expressing His_{10}-Strep-Pup under control of the native *pup* promoter from an integrative plasmid. Cultures (3 L) were grown in 7H9 medium to an OD_{600} of 1.2, MMC was added to 80 ng/mL, and cultures were incubated for another 4 hr before harvesting. Cells were lysed under denaturing conditions, and pupylated proteins were purified from the cleared lysate on NiNTA resin while changing to non-denaturing buffer conditions. After a second purification step using StrepTactin resin, pupylated proteins were concentrated by TCA precipitation and separated by SDS-PAGE. The upper part of the gel above 17 kDa was excised to exclude free Pup and submitted for LC-MS/MS analysis at the Functional Genomics Center Zurich (FGCZ). For details, see the Supplemental Experimental Procedures.

DATA AND SOFTWARE AVAILABILITY

The accession numbers for the RNA-seq and ChIP-seq data reported in this paper refer to the ArrayExpress database (Kolesnikov et al., 2015) at EMBL-EBI ArrayExpress: E-MTAB-6497 and E-MTAB-6503, respectively (https://www.ebi.ac.uk/arrayexpress).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and four tables and can be found with this article online at https://doi.org/10.1016/j.celrep.2018.05.073.

ACKNOWLEDGMENTS

We thank Lucy Poveda of the Functional Genomics Center Zurich (FGCZ) for support with library preparation and next-generation sequencing. The research was supported by the Swiss National Science Foundation (31003A_163314).

AUTHOR CONTRIBUTIONS

A.U.M., F.I., and E.W.-B. designed research and analyzed data. A.U.M. performed experiments. A.U.M., F.I., and E.W.-B. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 9, 2018
Revised: April 16, 2018
Accepted: May 22, 2018
Published: June 19, 2018

REFERENCES

Adams, L.B., Dinauer, M.C., Morgenstern, D.E., and Krahenbuhl, J.L. (1997). Comparison of the roles of reactive oxygen and nitrogen intermediates in the host response to *Mycobacterium tuberculosis* using transgenic mice. Tuber. Lung Dis. *78*, 237–246.

Ahel, I., Vujaklija, D., Mikoč, A., and Gamulin, V. (2002). Transcriptional analysis of the *recA* gene in *Streptomyces rimosus*: identification of the new type of promoter. FEMS Microbiol. Lett. *209*, 133–137.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Bailey, T.L., and Machanick, P. (2012). Inferring direct DNA binding from ChIP-seq. Nucleic Acids Res. *40*, e128.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME Suite: tools for motif discovery and searching. Nucleic Acids Res. *37* (*Suppl 2*), W202–W208.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Royal Statist. Soc. Ser. B. *57*, 289–300.

Brooks, P.C., Movahedzadeh, F., and Davis, E.O. (2001). Identification of some DNA damage-inducible genes of *Mycobacterium tuberculosis*: apparent lack of correlation with LexA binding. J. Bacteriol. *183*, 4459–4467.

Browning, D.F., and Busby, S.J.W. (2016). Local and global regulation of transcription initiation in bacteria. Nat. Rev. Microbiol. *14*, 638–650.

Burney, S., Caulfield, J.L., Niles, J.C., Wishnok, J.S., and Tannenbaum, S.R. (1999). The chemistry of DNA damage from nitric oxide and peroxynitrite. Mutat. Res. *424*, 37–49.

Cappelli, G., Volpe, E., Grassi, M., Liseo, B., Colizzi, V., and Mariani, F. (2006). Profiling of *Mycobacterium tuberculosis* gene expression during human macrophage infection: upregulation of the alternative sigma factor G, a group

of transcriptional regulators, and proteins with unknown function. Res. Microbiol. *157*, 445–455.

Cirz, R.T., O'Neill, B.M., Hammond, J.A., Head, S.R., and Romesberg, F.E. (2006). Defining the *Pseudomonas aeruginosa* SOS response and its role in the global response to the antibiotic ciprofloxacin. J. Bacteriol. *188*, 7101–7110.

Compton, C.L., Fernandopulle, M.S., Nagari, R.T., and Sello, J.K. (2015). Genetic and proteomic analyses of pupylation in *Streptomyces coelicolor*. J. Bacteriol. *197*, 2747–2753.

Darwin, K.H. (2009). Prokaryotic ubiquitin-like protein (Pup), proteasomes and pathogenesis. Nat. Rev. Microbiol. *7*, 485–491.

Darwin, K.H., Ehrt, S., Gutierrez-Ramos, J.-C., Weich, N., and Nathan, C.F. (2003). The proteasome of *Mycobacterium tuberculosis* is required for resistance to nitric oxide. Science *302*, 1963–1966.

Davis, E.O., Springer, B., Gopaul, K.K., Papavinasasundaram, K.G., Sander, P., and Böttger, E.C. (2002). DNA damage induction of *recA* in *Mycobacterium tuberculosis* independently of RecA and LexA. Mol. Microbiol. *46*, 791–800.

Drees, J.C., Lusetti, S.L., Chitteni-Pattu, S., Inman, R.B., and Cox, M.M. (2004). A RecA filament capping mechanism for RecX protein. Mol. Cell *15*, 789–798.

Dronkert, M.L., and Kanaar, R. (2001). Repair of DNA interstrand cross-links. Mutat. Res. *486*, 217–247.

Engels, S., Ludwig, C., Schweitzer, J.-E., Mack, C., Bott, M., and Schaffer, S. (2005). The transcriptional activator ClgR controls transcription of genes involved in proteolysis and DNA repair in *Corynebacterium glutamicum*. Mol. Microbiol. *57*, 576–591.

Fascellaro, G., Petrera, A., Lai, Z.W., Nanni, P., Grossmann, J., Burger, S., Biniossek, M.L., Gomez-Auli, A., Schilling, O., and Imkamp, F. (2016). Comprehensive proteomic analysis of nitrogen-starved *Mycobacterium smegmatis* Δ*pup* reveals the impact of pupylation on nitrogen stress response. J. Proteome Res. *15*, 2812–2825.

Festa, R.A., Pearce, M.J., and Darwin, K.H. (2007). Characterization of the proteasome accessory factor (*paf*) operon in *Mycobacterium tuberculosis*. J. Bacteriol. *189*, 3044–3050.

Festa, R.A., McAllister, F., Pearce, M.J., Mintseris, J., Burns, K.E., Gygi, S.P., and Darwin, K.H. (2010). Prokaryotic ubiquitin-like protein (Pup) proteome of Mycobacterium tuberculosis [corrected]. PLoS ONE *5*, e8589.

Forse, L.N., Houghton, J., and Davis, E.O. (2011). Enhanced expression of *recX* in *Mycobacterium tuberculosis* owing to a promoter internal to *recA*. Tuberculosis (Edinb.) *91*, 127–135.

Fudrini Olivencia, B., Müller, A.U., Roschitzki, B., Burger, S., Weber-Ban, E., and Imkamp, F. (2017). *Mycobacterium smegmatis* PafBC is involved in regulation of DNA damage response. Sci. Rep. *7*, 13987.

Galletto, R., Amitani, I., Baskin, R.J., and Kowalczykowski, S.C. (2006). Direct observation of individual RecA filaments assembling on single DNA molecules. Nature *443*, 875–878.

Gamulin, V., Cetkovic, H., and Ahel, I. (2004). Identification of a promoter motif regulating the major DNA damage response mechanism of *Mycobacterium tuberculosis*. FEMS Microbiol. Lett. *238*, 57–63.

Gandotra, S., Schnappinger, D., Monteleone, M., Hillen, W., and Ehrt, S. (2007). *In vivo* gene silencing identifies the *Mycobacterium tuberculosis* proteasome as essential for the bacteria to persist in mice. Nat. Med. *13*, 1515–1520.

Gaudion, A., Dawson, L., Davis, E., and Smollett, K. (2013). Characterisation of the *Mycobacterium tuberculosis* alternative sigma factor SigG: its operon and regulon. Tuberculosis (Edinb.) *93*, 482–491.

Goranov, A.I., Katz, L., Breier, A.M., Burge, C.B., and Grossman, A.D. (2005). A transcriptional response to replication status mediated by the conserved bacterial replication protein DnaA. Proc. Natl. Acad. Sci. USA *102*, 12932–12937.

Gupta, S., and Chatterji, D. (2005). Stress responses in mycobacteria. IUBMB Life *57*, 149–159.

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., Rattei, T., Mende, D.R., Sunagawa, S., Kuhn, M., et al. (2016). eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res. *44* (*D1*), D286–D293.

Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., von Mering, C., and Bork, P. (2017). Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. Mol. Biol. Evol. *34*, 2115–2122.

Imkamp, F., Rosenberger, T., Striebel, F., Keller, P.M., Amstutz, B., Sander, P., and Weber-Ban, E. (2010). Deletion of *dop* in *Mycobacterium smegmatis* abolishes pupylation of protein substrates *in vivo*. Mol. Microbiol. *75*, 744–754.

Imlay, J.A., and Linn, S. (1988). DNA damage and oxygen radical toxicity. Science *240*, 1302–1309.

Joo, C., McKinney, S.A., Nakamura, M., Rasnik, I., Myong, S., and Ha, T. (2006). Real-time observation of RecA filament dynamics with single monomer resolution. Cell *126*, 515–527.

Kolesnikov, N., Hastings, E., Keays, M., Melnichuk, O., Tang, Y.A., Williams, E., Dylag, M., Kurbatova, N., Brandizi, M., Burdett, T., et al. (2015). ArrayExpress update–simplifying data submissions. Nucleic Acids Res. *43*, D1113–D1116.

Kumar, V., Muratani, M., Rayan, N.A., Kraus, P., Lufkin, T., Ng, H.H., and Prabhakar, S. (2013). Uniform, optimal signal processing of mapped deep-sequencing data. Nat. Biotechnol. *31*, 615–622.

Lee, J.H., Geiman, D.E., and Bishai, W.R. (2008). Role of stress response sigma factor SigG in *Mycobacterium tuberculosis*. J. Bacteriol. *190*, 1128–1133.

Little, J.W., Edmiston, S.H., Pacelli, L.Z., and Mount, D.W. (1980). Cleavage of the *Escherichia coli lexA* protein by the *recA* protease. Proc. Natl. Acad. Sci. USA *77*, 3225–3229.

MacMicking, J.D., North, R.J., LaCourse, R., Mudgett, J.S., Shah, S.K., and Nathan, C.F. (1997). Identification of nitric oxide synthase as a protective locus against tuberculosis. Proc. Natl. Acad. Sci. USA *94*, 5243–5248.

Moreno-Hagelsieb, G., and Jokic, P. (2012). The evolutionary dynamics of functional modules and the extraordinary plasticity of regulons: the *Escherichia coli* perspective. Nucleic Acids Res. *40*, 7104–7112.

Movahedzadeh, F., Colston, M.J., and Davis, E.O. (1997). Determination of DNA sequences required for regulated *Mycobacterium tuberculosis* RecA expression in response to DNA-damaging agents suggests that two modes of regulation exist. J. Bacteriol. *179*, 3509–3518.

Namouchi, A., Gómez-Muñoz, M., Frye, S.A., Moen, L.V., Rognes, T., Tønjum, T., and Balasingham, S.V. (2016). The *Mycobacterium tuberculosis* transcriptional landscape under genotoxic stress. BMC Genomics *17*, 791.

Overbeek, R., Fonstein, M., D'Souza, M., Pusch, G.D., and Maltsev, N. (1999). The use of gene clusters to infer functional coupling. Proc. Natl. Acad. Sci. USA *96*, 2896–2901.

Papavinasasundaram, K.G., Colston, M.J., and Davis, E.O. (1998). Construction and complementation of a *recA* deletion mutant of *Mycobacterium smegmatis* reveals that the intein in *Mycobacterium tuberculosis recA* does not affect RecA function. Mol. Microbiol. *30*, 525–534.

Pearce, M.J., Mintseris, J., Ferreyra, J., Gygi, S.P., and Darwin, K.H. (2008). Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. Science *322*, 1104–1107.

Poulsen, C., Akhter, Y., Jeon, A.H.-W., Schmitt-Ulms, G., Meyer, H.E., Stefanski, A., Stühler, K., Wilmanns, M., and Song, Y.-H. (2010). Proteome-wide identification of mycobacterial pupylation targets. Mol. Syst. Biol. *6*, 386.

Rand, L., Hinds, J., Springer, B., Sander, P., Buxton, R.S., and Davis, E.O. (2003). The majority of inducible DNA repair genes in *Mycobacterium tuberculosis* are induced independently of RecA. Mol. Microbiol. *50*, 1031–1042.

Richard-Greenblatt, M., Bach, H., Adamson, J., Peña-Diaz, S., Li, W., Steyn, A.J.C., and Av-Gay, Y. (2015). Regulation of ergothioneine biosynthesis and

its effect on *Mycobacterium tuberculosis* growth and infectivity. J. Biol. Chem. *290*, 23064–23076.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Saini, V., Cumming, B.M., Guidry, L., Lamprecht, D.A., Adamson, J.H., Reddy, V.P., Chinta, K.C., Mazorodze, J.H., Glasgow, J.N., Richard-Greenblatt, M., et al. (2016). Ergothioneine maintains redox and bioenergetic homeostasis essential for drug susceptibility and virulence of *Mycobacterium tuberculosis*. Cell Rep. *14*, 572–585.

Salinas, G., and Comini, M.A. (2018). Alternative thiol-based redox systems. Antioxid. Redox Signal. *28*, 407–409.

Sancar, G.B., Sancar, A., Little, J.W., and Rupp, W.D. (1982). The *uvrB* gene of *Escherichia coli* has both *lexA*-repressed and *lexA*-independent promoters. Cell *28*, 523–530.

Sander, P., Meier, A., and Böttger, E.C. (1995). *rpsL+*: a dominant selectable marker for gene replacement in mycobacteria. Mol. Microbiol. *16*, 991–1000.

Sano, Y. (1993). Role of the *recA*-related gene adjacent to the *recA* gene in *Pseudomonas aeruginosa*. J. Bacteriol. *175*, 2451–2454.

Sao Emani, C., Williams, M.J., Van Helden, P.D., Taylor, M.J.C., Wiid, I.J., and Baker, B. (2018). Gamma-glutamylcysteine protects ergothioneine-deficient *Mycobacterium tuberculosis* mutants against oxidative and nitrosative stress. Biochem. Biophys. Res. Commun. *495*, 174–178.

Sharp, J.D., Singh, A.K., Park, S.T., Lyubetskaya, A., Peterson, M.W., Gomes, A.L.C., Potluri, L.P., Raman, S., Galagan, J.E., and Husson, R.N. (2016). Comprehensive definition of the SigH regulon of *Mycobacterium tuberculosis* reveals transcriptional control of diverse stress responses. PLoS ONE *11*, e0152145.

Shinagawa, H. (1996). SOS response as an adaptive response to DNA damage in prokaryotes. EXS *77*, 221–235.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. *7*, 539.

Singh, A. (2017). Guardians of the mycobacterial genome: A review on DNA repair systems in *Mycobacterium tuberculosis*. Microbiology *163*, 1740–1758.

Sinha, K.M., Unciuleac, M.C., Glickman, M.S., and Shuman, S. (2009). AdnAB: a new DSB-resecting motor-nuclease from mycobacteria. Genes Dev. *23*, 1423–1437.

Smollett, K.L., Smith, K.M., Kahramanoglou, C., Arnvig, K.B., Buxton, R.S., and Davis, E.O. (2012). Global analysis of the regulon of the transcriptional repressor LexA, a key component of SOS response in *Mycobacterium tuberculosis*. J. Biol. Chem. *287*, 22004–22014.

Spek, E.J., Wright, T.L., Stitt, M.S., Taghizadeh, N.R., Tannenbaum, S.R., Marinus, M.G., and Engelward, B.P. (2001). Recombinational repair is critical for survival of *Escherichia coli* exposed to nitric oxide. J. Bacteriol. *183*, 131–138.

Stanton, K.P., Parisi, F., Strino, F., Rabin, N., Asp, P., and Kluger, Y. (2013). Arpeggio: harmonic compression of ChIP-seq data reveals protein-chromatin interaction signatures. Nucleic Acids Res. *41*, e161.

Striebel, F., Imkamp, F., Sutter, M., Steiner, M., Mamedov, A., and Weber-Ban, E. (2009). Bacterial ubiquitin-like modifier Pup is deamidated and conjugated to substrates by distinct but homologous enzymes. Nat. Struct. Mol. Biol. *16*, 647–651.

Striebel, F., Imkamp, F., Özcelik, D., and Weber-Ban, E. (2014). Pupylation as a signal for proteasomal degradation in bacteria. Biochim. Biophys. Acta *1843*, 103–113.

Sukchawalit, R., Vattanaviboon, P., Utamapongchai, S., Vaughn, G., and Mongkolsuk, S. (2001). Characterization of *Xanthomonas oryzae pv. oryzae recX*, a gene that is required for high-level expression of *recA*. FEMS Microbiol. Lett. *205*, 83–89.

Tatusov, R.L., Koonin, E.V., and Lipman, D.J. (1997). A genomic perspective on protein families. Science *278*, 631–637.

Tomasz, M. (1995). Mitomycin C: small, fast and deadly (but very selective). Chem. Biol. *2*, 575–579.

Venkatesh, R., Ganesh, N., Guhan, N., Reddy, M.S., Chandrasekhar, T., and Muniyappa, K. (2002). RecX protein abrogates ATP hydrolysis and strand exchange promoted by RecA: insights into negative regulation of homologous recombination. Proc. Natl. Acad. Sci. USA *99*, 12091–12096.

Vierling, S., Weber, T., Wohlleben, W., and Muth, G. (2000). Transcriptional and mutational analyses of the *Streptomyces lividans recX* gene and its interference with RecA activity. J. Bacteriol. *182*, 4005–4011.

Wang, Y., Huang, Y., Xue, C., He, Y., and He, Z.-G. (2011). ClpR protein-like regulator specifically recognizes RecA protein-independent promoter motif and broadly regulates expression of DNA damage-inducible genes in mycobacteria. J. Biol. Chem. *286*, 31159–31167.

Watrous, J., Burns, K., Liu, W.-T., Patel, A., Hook, V., Bafna, V., Barry, C.E., 3rd, Bark, S., and Dorrestein, P.C. (2010). Expansion of the mycobacterial "PUPylome". Mol. Biosyst. *6*, 376–385.

Worsley Hunt, R., Mathelier, A., Del Peso, L., and Wasserman, W.W. (2014). Improving analysis of transcription factor binding sites within ChIP-seq data based on topological motif enrichment. BMC Genomics *15*, 472.

Yun, H.-Y., Tamura, N., and Tamura, T. (2012). Rhodococcus prokaryotic ubiquitin-like protein (Pup) is degraded by deaminase of pup (Dop). Biosci. Biotechnol. Biochem. *76*, 1959–1966.