iScience, Volume 25

Supplemental information

The RNA repair proteins RtcAB regulate

transcription activator RtcR via its

CRISPR-associated Rossmann fold domain

Ioly Kotta-Loizou, Maria Grazia Giuliano, Milija Jovanovic, Jorrit Schaefer, Fuzhou Ye, Nan Zhang, Danai Athina Irakleidi, Xiaojiao Liu, Xiaodong Zhang, Martin Buck, and Christoph Engl

Table S1. Genetic lesions affecting P*rtcBA*-*lacZ* promoter activity and *rtc* mRNA expression levels, Related to Table 1.

↑ induction (P-value<0.05)

↑↑ induction (P-value<0.001)

↓ repression (P-value<0.05)

↓↓ repression (P-value<0.001)

= no change (compared to wild-type)

Table S2. Genes affecting P*rtcBA*-*lacZ* promoter activity and *rtc* mRNA expression, Related to Table 1.

↑ induction (P-value<0.05)

↓ repression (P-value<0.05)

= no change (compared to wild-type)

Figure S1. Regulation of the Rtc system, Related to Table 1. (A) The *rtcBA* promoter activity is induced (*N*=15) and **(B)** the *rtc* mRNA levels are increased (*N*=3) in the gene deletion mutants Δ*gor*, Δ*mazF*, Δ*srmB*, Δ*ybaK* and Δ*yobF*, compared to the wild-type strain. Both RtcR and RtcB expression from the pBAD18cm vector are required for *rtcBA* promoter activity in the gene deletion mutants **(C)** Δ*gor* (*N*=6) and **(D)** Δ*yobF* (*N*=8). RtcR, RtcB and RtcA expression

from the pBAD18cm vector is required for *rtcBA* promoter activity in the gene deletion mutants **(E)** Δ*mazF* (*N*=9) and **(G)** Δ*ybaK* (*N*=9). The effects of the *rtc* gene deletion mutants on the *rtcBA* promoter activity in cells already lacking **(E)** Δ*mazF* and **(G)** Δ*ybaK* cannot be complemented by overexpressed Rtc proteins. **(F)** The effect of the gene deletion mutant Δ*mazF* on the *rtcBA* promoter activity can be complemented with overexpressed MazF protein from the pBAD18cm vector (*N*=9). The *rtcBA* promoter activity is increased in an RtcR-independent manner by the gene deletion mutants **(H)** Δ*rppH* (*N*=4), **(I)** Δ*rnhA* (*N*=9) and **(J)** Δ*hfq* in complete medium (*N*=6), while the *rtcBA* promoter activity is repressed by **(K)** Δ*hfq* in minimal medium (*N*=4). The latter is the only identified gene deletion mutant acting as both an inducer of the Rtc system in rich media, and as a repressor in minimal media. Overexpression of the *rof* (*N*=9) and *yedV* (*N*=10) genes from the pCA24N vector results in an induction of the *rtcBA* promoter activity in **(L,N)** minimal and **(M,O)** complete medium. In all panels, beta-galactosidase activity or mRNA levels of the wild-type strain is set as 100%. Data are shown as mean and error bars represent standard deviation from the mean. *N* represents total number of independent biological replicates, with 3 technical replicates each. ANOVA * P-value < 0.05; ** P-value < 0.01; *** Pvalue < 0.001; **** P-value < 0.0001.

Figure S2. Schematic representation of constructs, Related to Figure 2. (A) MBP- or His-tagged purified Rtc proteins used for gel filtration chromatography and crosslinking; **(B)** T18/T25 N-terminal and C-terminal fusions of full-length Rtc proteins produced by the bacterial two-hybrid vectors; **(C)** T18/T25 N-terminal and C-terminal fusions of truncated RtcR domains produced by the bacterial two-hybrid vectors; **(D)** T18/T25 fusions of Hrp proteins produced by the bacterial two-hybrid vectors.

Figure S3. Interactions between Rtc proteins revealed by bacterial 2-hybrid, Related to Figure 2. (A) RtcR interacts with itself. **(B)** HrpS interacts with itself and HrpV, but not with RtcA or RtcB. **(C)** RtcR interacts with RtcA and RtcB affects their interaction. **(D)** RtcA affects the interaction between RtcR and RtcB. **(E)** The interaction between RtcR proteins is mediated via the HTH DNA binding domain at the C-terminus. **(F)** The interaction between RtcR and RtcA proteins is mediated via the CARF signalling domain at the N-terminus. **(G)** TheRtcR CARF domain interacts with RtcA and RtcB affects their interaction. **(H)** The RtcR CARF domain interacts with RtcB and RtcA affects their interaction.

Data are shown as mean and error bars represent standard deviation from the mean. *N*=3 and represents total number of independent biological replicates, with 3 technical replicates each. Black columns but not dark grey columns are significantly different (ANOVA P-value < 0.0001) as compared to the negative control (T18/T25). Different letters indicate statistically significant differences (ANOVA P-value < 0.01 at least).

Data S1. *In vitro* **and** *in vivo* **protein crosslinking, Related to Figure 2.**

In vitro glutaraldehyde protein crosslinking experiments support the pairwise interactions of RtcA and RtcB with RtcR: as estimated from image analysis of stained gels and immunoblots, the amount of high molecular weight complexes in the presence of MBP-RtcR together with His-RtcA or His-RtcB is 2-3 fold higher as compared to samples of His-RtcA, His-RtcB or MBP-RtcR alone treated with the crosslinker (Fig. S4A). This increase was not evident in the presence of MBP-RtcR_{ΔCARF} together with His-RtcA or His-RtcB (Fig. S4B). Since the interaction between RtcR and RtcB was not detected *in vivo* using the bacterial two-hybrid assay, although it was observed in all the *in vitro* experiments, a different approach was adopted: *in vivo* DSP crosslinking was performed on cells expressing MBP-RtcR and RtcB-His, separately and together, and the outcome was assessed by SDS-PAGE and immunoblotting (Fig. S4D). The levels of both proteins were lower in the samples in which they were co-expressed as compared to when they were expressed separately (Fig. S4D). When DSP was added, the quantity of both MBP-RtcR and RtcB-His detected was reduced in the presence as compared to the absence of the other Rtc protein. This reduction was reversed by the addition of DTT (Fig. S4D), suggesting that a complex may be formed between MBP-RtcR and RtcB-His. Since a single discrete new crosslinked complex was not evident by immunoblotting, it may be present as multiple dispersed crosslinked species distributed across the gel lane. Alternatively, or in addition, a high molecular weight complex unable to enter the gel or transfer may be formed. Such complexes may include cell components such as those of the ribosome that is known to interact with RtcB and RtcA (Fig. S4E; Temmel *et al.*, 2017).

DTT-mediated crosslinking reversal. Quantification using ImageJ of His-RtcB and MBP-RtR following SDS-PAGE and immunoblotting. **(E)** Sucrose gradient centrifugation of *E. coli* extracts under associative ribosome profiling conditions shows that His-RtcA and His-RtcB each associate with high molecular weight complexes characteristic of ribosome assemblies. In all panels, data are shown as mean and error bars represent standard deviation from the mean of triplicates.

Table S3. List of potential RtcR RNA ligands tested for transcriptional activation, Related to Figure 3.

Figure S5. Potential RtcR ligands and activators, Related to Figure 3. UpGpGpG spRNAs were electrophoresed on 20% (w/v) urea-PAGE following production from the super-coiled P*rtcBA-nifH* hybrid promoter template in the presence of **(A)** RtcA and/or RtcB; **(B)** Mg²⁺ and Mn²⁺, on their own or with cyclic tetra/hexa (4/6) adenylates (cOA). % A represents percentage of activity. (C) Fragments of Cy3-tRNA^{Glu(UUC}) were electrophoresed on 4% (^w/_v) PAGE in the presence of the RtcR CARF domain and full length RtcR.

 (A)

Table S4. SNPs in Hpx- compared to wild-type MG1655, Related to Figure 4.

Table S5. Gene deletion mutants associated with oxidative stress, Related to Figure 4.

¹ responsible for detoxifying the majority of H_2O_2 in the cells (Winterbourn, 2008)

Table S6. Predicted effects of cysteine mutations on the RtcR CARF domain, Related to Figure 4.

* as compared to the *in silico* structural model of the RtcR CARF domain (Fig. S6)

Figure S6. The Rtc system and oxidative stress, Related to Figure 4. (A) The negative control test *hrp* promoter activity is not induced in the Hpx- strain after 24 h (*N*=3) and **(B)** growth of the Hpx- strain compared to the wild-type is not inhibited by the presence of the P*hrp-lacZ* reporter plasmid (*N*=3). **(C)** The *rtcBA* promoter activity is not induced by the gene deletion mutants Δ*fur*, Δ*iscR* and Δ*oxyR* in minimal medium after 24 h (*N*=3) and **(D)** growth of the Δ*fur*, Δ*iscR* and Δ*oxyR* strains is not inhibited by deletion of the *rtcB* gene (*N*=3). **(E)** The 314 CARF domain sequences in Pfam contain a number of cysteine residues ranging from zero to four. **(F)** Two of the cysteine residues are conserved among the aligned 314 CARF domain sequences as illustrated by the weblogo; arrows indicate the positions of the four cysteines in the *E. coli* RtcR CARF domain. **(G)** The *E. coli* RtcR CARF domain cysteine residues could potentially form disulphide bridges based on the *in silico* structural model. **(H)** Site-directed mutagenesis of the RtcR CARF domain cysteine residues into alanine residues does not affect the repression of the *rtcBA* promoter activity under non inducing conditions (*N*=4). **(I)** Site-directed mutagenesis of the RtcR CARF domain cysteine residues into alanine

residues does not affect the activation of the *rtcBA* promoter activity under inducing conditions (*N*=3). In all panels, beta-galactosidase activity of the wild-type strain is set as 100%. Data are shown as mean and error bars represent standard deviation from the mean. *N* represents total number of independent biological replicates, with 3 technical replicates each. ANOVA **** P-value < 0.0001.

Data S2. Transcriptome profiling of Rtc inducing conditions, Related to Figure 5.

In all cases, quality assessment revealed that the majority of reads (> 80%) correspond to known features of the annotated *E. coli* genome (Fig. S7A), while the gene deletion mutants and the induction of the Rtc system as judged by *rtcBA* up regulation were also verified (Fig. S7B).

In *E. coli* cells lacking *gor* or *mazF*, 30 and 33 genes, respectively were differentially expressed as compared to the wildtype. In total, 24 of these transcripts were common to both gene deletion mutants (Fig. 5A) and the vast majority were protein-coding. Decreases in the expression levels of genes forming the glycerol-3-phosphate dehydrogenase complex were observed (Fig. S8B), potentially due to increases in the expression levels of *glpR*, the glycerol-3-phosphate regulon repressor located adjacent to the Rtc system on the *E. coli* chromosome. This link was also evident following GO enrichment analysis (Tables S7,S8). Furthermore, the Psp membrane stress response protein transcripts were increased in abundance and the system responsible for arabinose catabolism was affected (Fig. S8D).

Only 3 transcripts were significantly more abundant in *E. coli* cells lacking *srmB*: the outer membrane autotransporter *flu* and the small RNA *isrC*, whose expression is inhibited by OxyR, together with the tRNA for aspartic acid. Valine tRNA was also more abundant but this was not statistically significant potentially due to discrepancies between the replicates.

A very large number, up to 40% of the total *E. coli* genes, were differentially expressed in Hpx-[RtcON] and Hpx- [RtcOFF] as compared to the wild-type: 699 and 1777 respectively 8 h post-inoculation, and 1144 and 460 respectively 24 h post-inoculation. In all cases the vast majority (over 85%) of the differentially expressed transcripts encode proteins (Fig. S7C). Principal component analysis revealed that the major source of variation (PC1) among the samples is time of harvest, while the second major source or variation (PC2) is the genetic background of the strains used (Fig. 5B). All wild-type strains group together, while Hpx-[RtcON] and Hpx-[RtcOFF] were distinct both from the wild-type and from each other (Fig. 5B), potentially due to the lack of *rtcBA* expression in the latter.

GO enrichment analysis of the differentially expressed transcripts revealed overrepresentation of genes forming the glycerol-3-phosphate dehydrogenase complex in Hpx-[RtcON] both at 8 (Table S9) and 24 h post-inoculation (Cellular Component 'glycerol-3-phosphate dehydrogenase complex'; fold enrichment 6.28; p-value 2.34·10-02), in agreement with that was observed in the cells lacking *gor* and *mazF* (Tables S7,S8), and Hpx-[RtcOFF] (Fig. S8B). KEGG pathway analysis also showed an effect on carbohydrate and lipid metabolism (Tables S11,12). Similarly, and possibly linked to changes in lipid metabolism, the Psp membrane stress response protein transcripts were more abundant in both Hpx- [RtcON] and Hpx-[RtcOFF] (Fig. S8D).

Additionally, both GO enrichment and KEGG pathway analysis illustrated an effect on ribosomal proteins, most prominent in the genes down regulated in Hpx-[RtcON] at 8 h post-inoculation (Tables S9,S11). Indeed, 25/31 and 14/21 proteins of the large and small ribosomal subunits, respectively, were significantly regulated in Hpx- at 8 h postinoculation (Fig. S9). A similar trend was noted in Hpx-[RtcON] (Cellular Component 'cytosolic large ribosomal subunit'; fold enrichment 4.02; p-value 2.23·10-02) and Hpx-[RtcOFF] (Cellular Component 'ribosomal subunit'; fold enrichment 2.33; p-value 4.28·10-02 and 'cytosolic ribosome'; fold enrichment 2.33; p-value 3.98·10-02) at 24 h post-inoculation (Table S11). In contrast, GO enrichment and KEGG pathway analysis did not reveal any significant overall effect on ribosomal proteins in Hpx-[RtcOFF] at 8 h post-inoculation, even though individual genes are up regulated or down regulated (Fig. S9). The down regulation of ribosomal proteins and the constituents of the bacterial-type flagellum

hook at 8 h post-inoculation was notable exclusively in Hpx-[RtcON] but not Hpx-[RtcOFF] (Table S10). However, an effect on nucleotide metabolism (Table S11) and a global down regulation of ribosomal RNAs (Fig. 5D) was observed in Hpx-[RtcOFF] at 8 h post-inoculation but not in any other condition.

Finally, the expression levels of genes whose deletion or overexpression modulates the RtcR system was not significantly affected in *E. coli* cells lacking *rtcR*, but some were differentially regulated in *E. coli* cells lacking *rtcA* or *rtcB* (Fig. S11) and their numbers were significantly higher than those expected by chance (P < 0.0001 and P < 0.01 for cells lacking *rtcA* or *rtcB*, respectively). Although evidence for feedback loops affecting RtcAB expression arose from examining the relationships between *rtc* expression and particular gene deletions and over expressions, but these relationships are complex and not fully explained and there was no discernible pattern associating the expression levels of the genes with their regulatory effect on *rtc* or their function; e.g. there are (1) genes, such as those related to oxidative stress including the alkyl hydroperoxide reductase component *ahpC*, whose deletion induces Rtc and *rtc* deletion increases their expression; (2) genes, including the ribotoxin *mazF* and the ribonuclease *rnhA*, whose deletion induces Rtc and *rtc* deletion decreases their expression; (3) genes, such as the transcription antiterminator *rof*, whose overexpression induces Rtc and *rtc* deletion increases their expression; (4) genes, such as the histidine kinase *yedV*, whose overexpression induces Rtc and *rtc* deletion decreases their expression; and (5) genes, such as the energydependent translational throttle protein *ettA* (formerly *yjjK*), whose overexpression represses Rtc and *rtc* deletion increases their expression.

Figure S7. NGS quality assessment and initial analysis, Related to Figure 5. (A) The majority of reads (> 80%) mapped to the annotated *E. coli* MG1655 genome, as calculated by HTSeq. **(B)** The gene deletion mutants *gor*, *mazF* and *srmB* have lower RPKM values compared to the wild-type. The *lacZ* mRNA, produced by the P_{rtcBA}-lacZ reporter plasmid and therefore corresponding to *rtcBA* induction, has a higher RPKM value in the gene deletion mutants compared to the

wild-type. Data are shown as mean and error bars represent standard deviation from the mean. **(C)** Pie charts illustrating the general functional roles of genes differentially expressed in Hpx-[RtcON] and Hpx-[RtcOFF] compared to the wild-type *E. coli* and to each other at 8 and 24 hours post inoculation.

Table S7. GO enrichment analysis of genes differentially expressed in *E. coli* cells lacking *gor*, Related to Figure 5.

Table S8. GO enrichment analysis of genes differentially expressed in *E. coli* cells lacking *mazF*, Related to Figure 5.

Table S9. GO enrichment analysis of genes expressed in Hpx-[RtcON] cells 8 hpi, Related to Figure 5.

Table S10. GO enrichment analysis of genes differentially expressed in Hpx-[RtcON] but not Hpx-[RtcOFF] cells 8 hpi, Related to Figure 5.

Figure S9. Expression of ribosomal proteins, Related to Figure 5. Heat map of expression levels of genes encoding ribosomal proteins for **(A)** the large and **(B)** the small ribosomal subunit in *E. coli* cells lacking *rtc* and in the Hpx-strain, as shown by NGS. Rows and columns have been grouped based on a hierarchical clustering algorithm. Limma/DESeq2 * adjusted P-value < 0.05; ** adjusted P-value < 0.01; *** adjusted P-value < 0.001; **** adjusted P-value < 0.0001.

Figure S10. Expression of cyclases under Rtc-inducing conditions, Related to Figure 5. Heat map of expression levels of genes encoding cyclases in *E. coli* cells lacking *gor* or *mazF* and in the Hpx- strain, as shown by NGS. Rows and columns have been grouped based on a hierarchical clustering algorithm. DESeq2 * adjusted P-value < 0.05; ** adjusted P-value < 0.01; *** adjusted P-value < 0.001; **** adjusted P-value < 0.0001.

Figure S11. Expression of selected genes in *rtc* **gene deletion mutants, Related to Figure 5.** Heat map of expression levels of various genes in *E. coli* cells lacking *rtc* compared to the wild-type, as shown by NGS. Deletion or overexpression of these genes is known to induce or repress *rtc* expression levels. Rows and columns have been grouped based on a hierarchical clustering algorithm. Limma * adjusted P-value < 0.05; ** adjusted P-value < 0.01.

Data S3. Identification of potentially cleaved RNAs, Related to Figure 6.

In order to identify potentially cleaved RNAs in transcriptomes, Pearson's correlation coefficient (PCC) of NGS coverage was calculated for each gene individually between datasets derived from both the same and different *E. coli* strains. The method was initially validated on NGS datasets from *E. coli* cells expressing the ribotoxin VapC (Engl *et al.*, 2016): PCC of initiator tRNAs (tRNAfMet) is > 0.9 within VapC-expressing and VapC non-expressing conditions, while it is < 0 when comparing VapC-expressing to VapC non-expressing conditions (Fig. S12A). In contrast, PCC is > 0.9 for elongator tRNAs (tRNAMet) which are not cleaved (Winther & Gerdes, 2011) regardless of VapC expression (Fig. S12A). The majority of genes (over 80%) have a PCC > 0.5 when comparing datasets derived from the same *E. coli* strains: wildtype, Hpx-[RtcON] and Hpx-[RtcOFF] at 8 h post-inoculation (Fig. 6A). Comparison between Hpx-[RtcON] and wild-type revealed a marginally lower correlation but the trend becomes more evident when comparing Hpx-[RtcOFF] to wildtype (Fig. 6A), while a similar distribution was noted at 24 h post-inoculation (Fig. S12D). Transcripts with PCC > 0.5 when comparing datasets of the same strain but < 0.5 when comparing both Hpx-[RtcON] and Hpx-[RtcOFF] to wildtype were considered potentially damaged. Transcripts with PCC > 0.5 when comparing datasets of the same strain and Hpx-[RtcON] to wild-type but < 0.5 when comparing Hpx-[RtcOFF] to wild-type were considered potentially damaged and then repaired by the Rtc system. The cutoff was set to PCC = 0.75 when comparing Δ*gor*, Δ*mazF* or Δ*srmB* to wild-type.

Figure S12. Analysis of damaged RNAs, Related to Figure 6. (A) Initiator but not elognator tRNAs are cleaved by VapC (top) and this is evident using Pearson's Correlation Coefficient (PCC). Distribution of all reads mapped to *metV* in VapC expressing (bottom left) and VapC non-expressing (bottom right) cells reveals the VapC cleavage site. **(B)** Selected tRNAs are damaged in cells lacking *gor*, *mazF* and *srmB* as compared to wild-type. **(C)** The general functional roles of transcripts damaged in Hpx-[RtcON] compared to wild-type (left) and Hpx-[RtcOFF] compared to wild-type (middle) and Hpx-[RtcOFF] compared to Hpx-[RtcON] (right) cells 8 h post-inoculation. **(D)** Transcripts damaged in Hpx-[RtcON] compared to wild-type (left), Hpx-[RtcOFF] compared to wild-type (middle) and Hpx-[RtcOFF] compared to Hpx- [RtcON] (right) cells 24 h post-inoculation, as illustrated by Pearson's Correlation Coefficient (PCC).

Figure S13. RtcB mediated ligation, Related to Figure 7. (A) Weblogo of 2709 3' termini of cellular RNAs and 5' termini of adapters following RtcB-mediated ligation. Maximum value of the Y-axis for nucleotide sequences is 2.0. **(B)** The tmRNA *ssrA* is targeted more frequently by RtcB ligase in Hpx- compared to the wild-type. **(C)** Expression levels of *ssrA* are upregulated in Δ*rtcA* and Δ*rtcB* compared to the wild-type, as shown by RT-qPCR; mRNA levels of the wild-type strain is set as 100%. Data are shown as mean and error bars represent standard deviation from the mean. *N*=3 and represents total number of independent biological replicates, with 3 technical replicates each. ANOVA ** P-value < 0.01; **** P-value < 0.0001. **(D)** Heat map of *ssrA* expression levels in a range of mutants, as shown by NGS. Rows and columns have been grouped based on a hierarchical clustering algorithm. DESeq2/Limma ** adjusted P-value < 0.01; **** P-value < 0.0001.

Figure S14. Repression of the Rtc system by gene overexpression, Related to Figure 8. Overexpression of the *yheS* gene and its paralogues, *uup*, *ybiT* and *yjjK* results in repression of the *rtcBA* promoter activity in **(A)** minimal and **(B)** complete medium. **(C)** Overexpression of the YheS protein has an epistatic effect on the gene deletion mutants, resulting in repression the the *rtcBA* promoter activity. Beta-galactosidase activity of the wild-type strain is set as 100%. Data are shown as mean and error bars represent standard deviation from the mean. *N*=3 and represents total number of independent biological replicates, with 3 technical replicates each. ANOVA * P-value < 0.05; ** P-value < 0.01; *** P-value < 0.001; **** P-value < 0.0001.

Table S13. Plasmids, Related to STAR Methods.

Table S14. Bacterial strains, Related to STAR Methods.

