

## The Effectors and Sensory Sites of Formaldehyde-Responsive Regulator FrmR and Metal-Sensing Variant

Deenah Osman<sup>‡1</sup>, Cecilia Piergentili<sup>‡1</sup>, Junjun Chen<sup>§</sup>, Lucy N. Sayer<sup>‡</sup>, Isabel Usón<sup>¶</sup>, Thomas G. Huggins<sup>§</sup>, Nigel J. Robinson<sup>‡†</sup> and Ehmke Pohl<sup>‡</sup>

From the <sup>‡</sup>Durham University, School of Biological and Biomedical Sciences, Department of Chemistry, DH1 3LE, UK, <sup>§</sup>Procter and Gamble, Mason Business Centre, Cincinnati, Ohio 45040, USA, <sup>¶</sup>Instituto de Biología Molecular de Barcelona (IBMB-CSIC) Barcelona Science Park, Barcelona, Spain, and <sup>1</sup>ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain.

**Running title:** *The Effectors of FrmR*

<sup>†</sup>To whom correspondence should be addressed. E-mail: [nigel.robinson@durham.ac.uk](mailto:nigel.robinson@durham.ac.uk)

**Key words:** formaldehyde, metal, zinc, cobalt, glutathione, *Salmonella enterica*, *Escherichia coli*, RcnR/CsoR, FrmR, allosteric regulation

The DUF156 family of DNA-binding, transcriptional-regulators include metal-sensors which respond to cobalt and/or nickel (RcnR, InrS) or copper (CsoR), plus CstR which responds to persulfide, and formaldehyde-responsive FrmR. Unexpectedly, the allosteric mechanism of FrmR from *Salmonella enterica* serovar Typhimurium is triggered by metals *in vitro* and variant FrmR<sup>E64H</sup> gains responsiveness to Zn(II) and cobalt *in vivo*. Here we establish that the allosteric mechanism of FrmR is triggered directly by formaldehyde *in vitro*. Sensitivity to formaldehyde requires a cysteine (Cys<sup>35</sup> in FrmR) conserved in all DUF156 proteins. A crystal structure of metal- and formaldehyde-sensing FrmR<sup>E64H</sup> reveals that an FrmR-specific amino-terminal Pro<sup>2</sup> is proximal to Cys<sup>35</sup> and these residues form the deduced formaldehyde-sensing site. Evidence is presented which implies that residues spatially close to the conserved cysteine tune the sensitivities of DUF156 proteins above or below critical thresholds for different effectors, generating the semblance of specificity within cells. Relative to FrmR, RcnR is less responsive to formaldehyde *in vitro* and RcnR does not sense formaldehyde *in vivo*, but reciprocal mutations FrmR<sup>P25</sup> and RcnR<sup>S2P</sup> respectively impair or enhance formaldehyde-reactivity *in vitro*. Formaldehyde-detoxification by FrmA requires S-(hydroxymethyl)glutathione, yet glutathione inhibits formaldehyde detection by FrmR *in vivo* and *in vitro*. Quantifying the number of FrmR

molecules per cell and modelling formaldehyde modification as a function of [formaldehyde], demonstrates that FrmR-reactivity is optimised such that FrmR is modified, and *frmRA* de-repressed, at lower [formaldehyde] than required to generate S-(hydroxymethyl)glutathione. Expression of FrmA is thereby coordinated with the accumulation of its substrate.

---

Formaldehyde (H<sub>2</sub>C=O), as a strong electrophile, is capable of alkylating and cross-linking the reactive groups (such as thiols and amines) of proteins and DNA (1-5). This reactivity and subsequent damage to biological macromolecules makes formaldehyde a highly cytotoxic compound. In addition to environmental sources, formaldehyde is generated intracellularly by a number of cellular processes. In methylotrophic and methanotrophic bacteria, it is well known that formaldehyde is generated as a bi-product of methanol and methane oxidation (6-9), consistent with the presence of genetically encoded formaldehyde detoxification systems in these organisms (2,8,10-12). Intracellular formaldehyde generation in bacteria that do not use these C<sub>1</sub> substrates as a carbon source has been less well studied. Formaldehyde is produced by the alternative heme degradation pathway (IsdG and IsdI) in *Staphylococcus aureus* to acquire iron (13,14). The recent detection of trimethylamine N-oxide (TMAO) demethylase activity in cell

extracts suggests this activity may be an endogenous source of formaldehyde in *Escherichia coli* (15). Demethylation of nucleic acids and production of methylglyoxal from glyceraldehyde-3-phosphate and dihydroxyacetone-phosphate during glycolysis represent more widespread physiological sources of formaldehyde (16-18). In addition, several mechanisms for the generation of formaldehyde at the host-pathogen interface have recently been proposed (2).

Inducible formaldehyde detoxification mechanisms have now been recognised in most bacteria (2,3,19). A glutathione-dependent pathway represents the most widespread formaldehyde-detoxification system, although the functional proteins and/or genomic arrangement may vary (19-26). In *E. coli*, this pathway is encoded by the *frmRAB* operon which includes *frmA*, encoding a Zn(II)-binding glutathione-dependent formaldehyde dehydrogenase, and *frmB*, encoding *S*-formylglutathione hydrolase (Figs. 1 and 2A) (26-29). Regulation of the *frmRAB* operon upon formaldehyde accumulation is mediated by the first gene product, FrmR, a DNA-binding transcriptional regulator (26).

FrmR is a member of the RcnR/CsoR-family (DUF156) of (predominantly) metal-sensing transcriptional repressors (30-32). This family can be divided into subgroups which have evolved to detect distinct and specific effectors in a cellular context by modification of a relatively conserved protein scaffold, in a manner similar to ArsR-, MerR- and Fur-family regulators (33-36). In addition to FrmR, characterised DUF156-subgroups to date include the metal-sensors RcnR and DmeR which respond to Ni(II)/Co(II), CsoR and RicR which respond to Cu(I), InrS which responds to Ni(II), and the non-metal sensor CstR which undergoes cysteine modification by sodium sulphite, selenite and tellurite (31,32,37-41). Upon binding of an allosteric effector (e.g. metal ion), affinity for DNA is weakened, alleviating repression from the target operator-promoter (30). At the time of writing, CsoR represents the only member of this family for which a structure has been reported (31,42-44). CsoR forms a three-helix bundle which adopts a tetrameric assembly made up of a dimer of dimers. The known effector sensory sites of metal-sensing DUF156 proteins exploit side-chains of conserved residues at a

dimer interface, denoted the 'WXYZ' fingerprint, characteristic of each subgroup but all involving a conserved Cys-thiolate (position 'X') located at the amino-terminal end of helix  $\alpha 2$  (31,45,46). Analogous information is not yet available for the sensory sites of FrmR.

*E. coli* FrmR-mediated transcriptional repression is alleviated following exposure of cells to exogenous formaldehyde, CO-releasing molecules, chloride treatment, and during anaerobic respiration using TMAO as the terminal electron acceptor (15,26,47-49). However, the effector directly detected by FrmR in each case remains unexplored. We recently identified an FrmR homologue in *Salmonella enterica* serovar Typhimurium strain SL1344 (hereafter referred to as *Salmonella*) which, as observed for *E. coli*, responds to exogenous formaldehyde *in vivo* (Fig. 1) (50). Unlike *E. coli* FrmR (containing two), *Salmonella* FrmR possesses three (four including Glu<sup>64</sup>) putative metal-ligands at positions 'WXY(Z)', within the metal-binding fingerprint of metal-sensing DUF156 members (50). Moreover, *Salmonella* FrmR can bind Co(II), Cu(I) and Zn(II). Unexpectedly, Cu(I) and Zn(II) are capable of triggering an allosteric response that weakens FrmR DNA affinity *in vitro* (50). Metal-responsiveness is not observed *in vivo* because FrmR is less sensitive than the endogenous *Salmonella* sensors for these metals. However, generation of a variant FrmR, responsive to cobalt and Zn(II) in addition to formaldehyde *in vivo*, is achieved by single amino-acid substitution at the putative metal-binding site (Glu<sup>64</sup> to His<sup>64</sup>) (Fig. 1). The combined effect of tighter metal affinity and weaker DNA-affinity of the apo-form, relative to wild type FrmR, confers metal-sensing gain of function to FrmR<sup>E64H</sup> (50). Evidence that *Salmonella* FrmR is competent to respond to metals raises the possibility that formaldehyde-sensing could be indirect and mediated by an effect on metal availability to FrmR by formaldehyde (Fig. 1). Notably, FrmA also requires Zn(II) for catalytic activity (47). The extent to which Zn(II) might be required to act as a signal transducer of formaldehyde accumulation in a cell now needs to be addressed.

In addition to FrmR, transcriptional regulators that respond following exposure to exogenous formaldehyde include HxlR (MarR-family) from *Bacillus subtilis* and NmlR/AdhR

(MerR-family) identified in *Neisseria sp.* and other Gram-positive pathogens (51-56). However, the effector(s) detected by any formaldehyde-responsive transcriptional regulator has yet to be biochemically identified. Despite the requirement of glutathione for formaldehyde detoxification by FrmA, the extent to which glutathione plays a role in the regulation of expression of glutathione-dependent formaldehyde dehydrogenase in any organism is unknown (Fig. 1).

We present the first *in vitro* evidence that formaldehyde is a direct allosteric effector of *Salmonella* FrmR. The FrmR sensory site is particularly reactive to formaldehyde such that the related *Salmonella* RcnR-sensor is less-responsive to formaldehyde *in vitro* and *in vivo*. We determine the crystal structure of FrmR<sup>E64H</sup> in order to define the effectors and sensory site(s) of this formaldehyde- and metal-sensing variant. Residues required for Zn(II)/Co(II)- and formaldehyde-sensing are determined and support a mechanism involving an FrmR-specific formaldehyde cross-link between Pro<sup>2</sup> and Cys<sup>35</sup>. An RcnR variant with enhanced sensitivity for formaldehyde *in vitro* was generated based on the deduced FrmR sensory site and mechanism. Implications for the basis of effector-specificity within DUF156-family proteins, and the chemical-species detected by FrmR *in vivo*, are discussed.

## RESULTS

*Salmonella* FrmR and FrmR<sup>E64H</sup> Retain Responsiveness to Formaldehyde and Metals When Expressed in *E. coli*—Despite *E. coli* and *Salmonella* being co-linear for most genes (57), the *Salmonella* *frm* operon occurs at a distinct genomic location compared with *E. coli* and lacks the *frmB* gene for S-formylglutathione hydrolase (Fig. 2A). *E. coli* and *Salmonella* FrmRs share only 52.3 % sequence identity (Fig. 2B) compared to an average ~ 85 % identity for orthologous gene products between these organisms (58,59), and analysis of the DUF156 FrmR subgroup demonstrates they are polyphyletic (Fig. 2C). This is also reflected by the distinct operator-promoter sequences upstream of each *frm* locus (Fig. 2D). The significance of these differences in relation to formaldehyde detoxification remains unknown, but may reflect specific requirements for formaldehyde detoxification in the respective

cellular backgrounds. To investigate the response of *Salmonella* FrmR in an *E. coli* cytosol, P<sub>frmRA</sub>-*frmR* reporter constructs comprising the *Salmonella* *frmRA* promoter (P<sub>frmRA</sub>) and *frmR* coding sequence fused to *lacZ* were expressed in *E. coli* cells which lacked the endogenous *E. coli* *frmR* gene ( $\Delta$ *frmR*) (Fig. 2E). As observed in *Salmonella* (50), expression from P<sub>frmRA</sub>-*frmR* was de-repressed in the heterologous *E. coli* host following exposure of cells to maximum non-inhibitory concentrations (MNIC) of formaldehyde, whereas exposure to MNIC CoCl<sub>2</sub> and ZnCl<sub>2</sub> did not alleviate repression (Fig. 2E). FrmR variant, FrmR<sup>E64H</sup>, which responds to CoCl<sub>2</sub>, ZnCl<sub>2</sub> and formaldehyde in *Salmonella* cells (50), retains the same effector responsiveness when P<sub>frmRA</sub>-*frmR*<sup>E64H</sup> is expressed in *E. coli*  $\Delta$ *frmR* (Fig. 2F). This demonstrates that the ability of FrmR<sup>E64H</sup> to respond to metals (and formaldehyde) *in vivo* is not exclusive to *Salmonella* cells.

*FrmR* Senses Formaldehyde Directly—Repression by FrmR (and FrmR<sup>E64H</sup>) is alleviated by exogenous formaldehyde *in vivo* (Figs. 2E and F), but DNA binding to the target *frmRA* operator-promoter (*frmRAPro*) (Fig. 2D) is weakened by Zn(II) (and Cu(I)) *in vitro* (50). To explore whether the *in vivo* response might be transduced by metals during formaldehyde stress or whether formaldehyde is able to act directly on FrmR, fluorescence anisotropy was used to monitor the interaction of FrmR with fluorescently labelled *frmRAPro* in the presence of formaldehyde (Fig. 3A). FrmR has previously been shown to bind *frmRAPro* with a stoichiometry of two tetramers per DNA molecule and a K<sub>DNA</sub> of 9.9 (± 0.3) × 10<sup>-8</sup> M for each tetramer, in the absence of effector (50) (also confirmed here in Fig. 4C). Consequently, a limiting concentration (10 nM) of *frmRAPro* was used for titration with FrmR in the presence of 10 or 20 μM formaldehyde, concentrations chosen to minimise non-specific formaldehyde cross-linking, which is likely at higher formaldehyde concentrations (60). EDTA was included as a metal-chelator to eliminate any effect that may arise due to the presence of (allosterically-effective) trace metals. The anisotropy data were fit to a model describing the binding of two non-dissociable FrmR-tetramers per DNA molecule and revealed that DNA binding of FrmR to *frmRAPro* was weakened by ~ 6.5-fold and ≥ 70-fold (compared to the published value

(50), Table 2) in the presence of 10  $\mu\text{M}$  and 20  $\mu\text{M}$  formaldehyde, respectively (Fig. 3A). This identifies formaldehyde as a direct allosteric effector of FrmR.

The ability of FrmR to respond directly to formaldehyde opens up the possibility that other aldehydes or related alcohols may also act as allosteric effectors. To test this hypothesis, DNA-binding was monitored in the presence of acetaldehyde ( $\text{CH}_3\text{HC}=\text{O}$ ), which differs from formaldehyde by replacement of a carbonyl-bonded proton with a methyl group. Inclusion of 20  $\mu\text{M}$  acetaldehyde did not weaken the DNA-binding affinity of FrmR (Fig. 3A), and expression from  $P_{frmRA}\text{-frmR}$  was not de-repressed when *Salmonella* cells were exposed to MNIC acetaldehyde (Fig. 3B). Similarly, FrmR did not respond to ethanol, methanol, 1-butanol, 1-propanol and 2-propanol *in vivo* (Fig. 3B). Together, these data show that the response of FrmR demonstrates specificity for formaldehyde over other organic molecules and suggest metal ions are not required to transduce the formaldehyde signal to FrmR *in vivo*.

*Structure of FrmR<sup>E64H</sup> and Visualisation of its Inferred Metal (Zn(II)/Co(II)) Sensory Site*—To identify the sensory site(s) of FrmR<sup>E64H</sup> for metal and formaldehyde, diffraction quality crystals were generated and an X-ray crystal structure was determined to 2.19 Å resolution (Fig. 4A). FrmR<sup>E64H</sup> exists as a homotetrameric assembly composed of a dimer of dimers as observed for the structurally characterised metal-sensing regulator CsoR (31,42-44) (Fig. 4A). FrmR<sup>E64H</sup> has a kink (~45 degrees) in helix  $\alpha_3$ , not observed in (any) CsoR structure, which is enabled by Gly<sup>83</sup>, a residue specific to *Salmonella* FrmR (distinct from Ile<sup>83</sup> in *E. coli* FrmR). The electrostatic surface potential highlights a region of positive potential composed of positively-charged residues from helices  $\alpha_1$  and  $\alpha_2$  within a single monomer subunit (Fig. 4B). This region (as suggested for metal-sensor CsoR (31,43,45)) is anticipated to enable binding of FrmR<sup>E64H</sup> to the *frmRA* operator-promoter, although the precise nature of the protein-DNA interactions for any DUF156 member are unknown. The *Salmonella frmRA* operator-promoter comprises a C<sub>6</sub>-tract flanked by a T/A-rich inverted repeat (Fig. 2D). The requirement of these inverted repeats for FrmR and FrmR<sup>E64H</sup> binding was investigated by

fluorescence anisotropy using a modified *frmRAPro* dsDNA fragment in which one flanking repeat had been mutated (*frmRAPro\**) (Figs. 2D and 4C). No binding of FrmR or FrmR<sup>E64H</sup> to *frmRAPro\** (10 nM) was detected, indicating a considerably weaker DNA-binding affinity ( $> 10^{-5}$  M) than determined for *frmRAPro*. This demonstrates that the T/A-rich inverted repeat is required for tight-affinity (physiologically-relevant) DNA-binding to the *frmRA* operator-promoter. The *frmRA* operator-promoter supports binding of two FrmR (or FrmR<sup>E64H</sup>) tetramers (50), and these data are consistent with obligatory tetramer interaction with the *frmRAPro* inverted repeat.

A candidate metal-binding site of FrmR<sup>E64H</sup> is formed by the side-chains of His<sup>60</sup> and His<sup>64</sup> from one subunit, and Cys<sup>35</sup> from the second subunit within the dimeric assembly (Fig. 4D). These residues match the ‘XYZ’ motif required for metal binding in related metal-sensors RcnR, CsoR and InrS, and this was the rationale behind the Glu<sup>64</sup> to His<sup>64</sup> substitution (31,39,45,46,50). To investigate the role of Cys<sup>35</sup> and His<sup>60</sup> in metal binding, site-directed mutants FrmR<sup>C35A</sup> and FrmR<sup>H60L</sup> were generated and assayed for their ability to bind Zn(II). Following pre-incubation with excess ZnCl<sub>2</sub>, neither variant retained Zn(II) during size-exclusion chromatography in contrast to wild type FrmR and FrmR<sup>E64H</sup>, which each co-migrate with one molar equivalent of Zn(II) (Fig. 4E and (50)). This indicates that the affinities of FrmR<sup>C35A</sup> and FrmR<sup>H60L</sup> for Zn(II) are considerably weaker than wild type FrmR, and implicates these residues in Zn(II) (and by inference Co(II)) coordination. Candidate residues for a fourth ligand required for the tetrahedral coordination geometry observed for Co(II) and inferred for Zn(II) (50), include His<sup>3</sup> (denoted position ‘W’ in RcnR (46,61)), Asp<sup>63</sup>, the amino-terminus, or solvent (Fig. 4D).

*Proposed Formaldehyde Sensory Site and Reaction Mechanism*—To define the functional formaldehyde sensory site, residues specifically conserved within the FrmR-subgroup of DUF156 family of transcriptional regulators were identified. Protein sequences previously ascribed to the FrmR-subgroup (45) were used to generate a multiple-sequence alignment with *Salmonella* FrmR (Fig. 5A). Twelve residues are conserved within the FrmR-subgroup but absent from the

closely-related Ni(II)/Co(II)-sensing RcnR-subgroup. Two-thirds of the conserved residues are clustered in helix  $\alpha 1$  based on the structure of FrmR<sup>E64H</sup> (Figs. 5A-C). Sensing of formaldehyde may proceed via reaction with Cys<sup>35</sup>, also implicated in the FrmR metal-site (Figs. 4D and E) due to its conservation in all characterised DUF156 members. Formation of an *S*-formyl adduct at this Cys-thiol followed by reaction with a primary amine has been suggested as a possible reaction mechanism of FrmR with formaldehyde (30). The pyrrolidine side chain of proline residue 2 ( $\alpha 1$ ) is in close proximity (3.0-3.2 Å in the four independent locations within the tetrameric structure) to Cys<sup>35</sup> from  $\alpha 2'$  (Figs. 5B and C, and supplemental Fig. S1A). A second FrmR-specific proline (Pro<sup>5</sup>) acts to terminate helix  $\alpha 1$  and positions the amino terminus of FrmR<sup>E64H</sup> adjacent to Cys<sup>35</sup> (Fig. 5B). Pro<sup>2</sup> is the first residue identified in the FrmR<sup>E64H</sup> structure and is positioned in a pocket at the dimer interface, leaving no space (and no observed electron density) for the amino terminal methionine predicted by the primary sequence (Fig. 5D and supplemental Fig. S1A). The amino-terminal region has been implicated in the co-ordination of Ni(II)/Co(II) by RcnR and of Ni(II) by InrS (61,62). In the absence of Met<sup>1</sup>, the terminal secondary amine of Pro<sup>2</sup> and a Cys<sup>35</sup>-thiolate are both ideal candidates for nucleophilic addition to formaldehyde (Fig. 5D and E) (63,64). Either reaction with Pro<sup>2</sup> followed by Cys<sup>35</sup> via an *N*-methylol intermediate, or reciprocally via an *S*-hydroxymethyl intermediate, are plausible (Fig. 5E). In both cases, the end product would be a methylene bridge between the two residues, requiring a 1:1 formaldehyde:FrmR monomer (4 possible sites per tetramer) reaction stoichiometry.

The cleavage of FrmR Met<sup>1</sup> was examined by multiple reaction monitoring mass-spectrometry (MRM-MS) using purified FrmR. Amino-terminal peptide PHSPED was detected confirming that FrmR is a substrate for methionine aminopeptidase (Fig. 6A). To investigate the requirement of Cys<sup>35</sup> and Pro<sup>2</sup> for formaldehyde-sensing, transcriptional fusions of P<sub>frmRA</sub>-frmR<sup>C35A</sup> and P<sub>frmRA</sub>-frmR<sup>P2S</sup> with lacZ were generated, introduced into *Salmonella* and compared to wild type (P<sub>frmRA</sub>-frmR). Expression from P<sub>frmRA</sub>-frmR is de-repressed by exposure of cells to formaldehyde in a concentration dependent manner (Fig. 6B).

Repression of P<sub>frmRA</sub> is retained following mutation of either Cys<sup>35</sup> or Pro<sup>2</sup>, but de-repression in response to formaldehyde is completely abolished (Fig. 6B). To confirm that formaldehyde was unable to act as an allosteric effector of these FrmR variants, the DNA binding properties of FrmR<sup>C35A</sup> and FrmR<sup>P2S</sup> were characterised using fluorescence anisotropy (Fig. 6C and D). Titration of frmR<sub>Pro</sub> with apo-FrmR<sup>C35A</sup> or apo-FrmR<sup>P2S</sup> in the presence of excess EDTA revealed  $K_{DNA}$  values comparable with wild type FrmR and consistent with the observed repression by both mutants *in vivo* ( $K_{DNA}^{apo-FrmR^{C35A}} = 1.6 (\pm 0.2) \times 10^{-7}$  M and  $K_{DNA}^{apo-FrmR^{P2S}} = 1.5 (\pm 0.2) \times 10^{-7}$  M) (Figs. 6C-D and Table 2). However, in contrast to wild type FrmR (Fig. 3A), DNA-binding by FrmR<sup>C35A</sup> was unaffected by the presence of 20  $\mu$ M formaldehyde, inferring a loss of formaldehyde reactivity (Fig. 6C). The reactivity of FrmR<sup>P2S</sup> to formaldehyde was significantly decreased compared to wild type FrmR with apparent DNA binding affinity weaker than apo-FrmR<sup>P2S</sup> by only ~ 4-fold in the presence of 20  $\mu$ M formaldehyde (compared with  $\geq 70$ -fold for wild type FrmR (Fig. 3A)) (Fig. 6D). Consequently, the FrmR<sup>C35A</sup> substitution impairs formaldehyde-reactivity more severely than FrmR<sup>P2S</sup> (Fig. 6E). The proposed mechanism (1:1 formaldehyde:FrmR stoichiometry (Fig. 5E)) and observed DNA binding by FrmR infers an affinity at the allosteric site(s) of FrmR for formaldehyde in the 10 – 20  $\mu$ M range, whereas formaldehyde-affinities of FrmR<sup>C35A</sup> and FrmR<sup>P2S</sup> variants are inferred to be  $\gg 100$   $\mu$ M and  $> 50$   $\mu$ M, respectively (Fig. 6E). These data implicate Cys<sup>35</sup> and Pro<sup>2</sup> in formaldehyde-mediated de-repression *in vivo*, and impaired DNA binding *in vitro* (Fig. 6). The Zn(II)/Co(II)-site also requires Cys<sup>35</sup> (Figs. 4D and E), implying overlap between the two effector sensory-sites.

*FrmR is More Sensitive to Formaldehyde Than RcnR*—Although FrmR shares sequence similarities with Ni(II)/Co(II)-sensing RcnR (30,32,45,46), expression from *Salmonella* rcnR-P<sub>rcnRA</sub> fused to lacZ is not de-repressed by formaldehyde *in vivo* (Fig. 7A). Analysis of the rcnR-rcnA intergenic region identified two putative RcnR DNA-binding sequences in the target RcnR operator-promoter (Supplemental table S1). The interaction of RcnR with a fluorescently labelled double-stranded DNA

fragment containing these sequences, *rcnRAPro*, was monitored by fluorescence anisotropy. The stoichiometry of *Salmonella* RcnR binding to *rcnRAPro* was first confirmed by titration of RcnR into a relatively high concentration of DNA (2.5  $\mu$ M) with saturation observed at 8 molar equivalents of RcnR (monomer) consistent with binding of two tetramers (one per site) (Fig. 7B) as observed for *E. coli* RcnR (65). A limiting concentration of DNA (10 nM) and a model describing the binding of two non-dissociable RcnR-tetramers per DNA molecule was subsequently used to determine the  $K_{\text{DNA}}$  of RcnR and *rcnRAPro* as  $1.5 (\pm 0.8) \times 10^{-7}$  M for apo-RcnR (Fig. 7C and Table 2). As predicted, titration of *rcnRAPro* with either Ni(II)-RcnR or Co(II)-RcnR dramatically weakened DNA-binding (Fig. 7C and Table 2);  $K_{\text{DNA}}^{\text{Ni(II)-RcnR}} \geq 5.9 (\pm 1.3) \times 10^{-6}$  M and  $K_{\text{DNA}}^{\text{Co(II)-RcnR}} \geq 1.5 (\pm 0.2) \times 10^{-5}$  M. The allosteric-coupling free energy ( $\Delta G_c$ ) which couples effector-binding to DNA-binding (66-68) is calculated to be  $\geq +2.2 (\pm 0.2)$  and  $\geq +2.7 (\pm 0.2)$  kcal mol $^{-1}$  for Ni(II)- and Co(II)-RcnR, respectively. Conversely, DNA binding by RcnR is unaffected by inclusion of 20  $\mu$ M formaldehyde (Fig. 7D), a concentration which weakens FrmR DNA binding by  $\geq 70$ -fold (Fig. 6E). Importantly, these data establish a degree of specificity of FrmR over RcnR for formaldehyde.

*Generation of an RcnR<sup>S2P</sup> Variant With Enhanced Response to Formaldehyde In Vitro*—Increasing the concentration of formaldehyde during fluorescence anisotropy to 50  $\mu$ M and 100  $\mu$ M does impair binding of RcnR to *rcnRAPro* by  $\sim 2.5$ -fold and  $\sim 5$ -fold (relative to apo-RcnR), respectively (Figs. 7C and D), providing an assay to monitor changes in RcnR formaldehyde reactivity. Introduction of the proposed formaldehyde sensing site of FrmR into *Salmonella* RcnR was achieved by mutation of Ser<sup>2</sup> to Pro<sup>2</sup> (Cys<sup>35</sup> is conserved in both proteins). Titration of RcnR<sup>S2P</sup> into *rcnRAPro* (10 nM) confirms that this variant binds *rcnRAPro* with equal affinity to wild type RcnR ( $K_{\text{DNA}}^{\text{apo-RcnRS2P}} = 1.6 (\pm 0.1) \times 10^{-7}$  M) (Fig. 7E and Table 2). DNA-binding by RcnR<sup>S2P</sup> was subsequently assessed in the presence of formaldehyde: inclusion of 50  $\mu$ M and 100  $\mu$ M formaldehyde weakened the apparent DNA-affinity of RcnR<sup>S2P</sup> by  $\sim 11$ -fold and  $\sim 17$ -fold, respectively (Figs. 7D and E). Thus, the single Ser<sup>2</sup> to Pro<sup>2</sup> point mutation generates an

RcnR variant with increased reactivity to formaldehyde compared with wild type RcnR (Figs. 7D). However, this increase was not sufficient to gain formaldehyde-sensing by *rcnR<sup>S2P</sup>-P<sub>rcnRA</sub>* *in vivo* (Fig. 7A). Repression from *P<sub>rcnRA</sub>* under control conditions (without inducer) combined with an observed cobalt-responsiveness, confirmed that the expressed RcnR<sup>S2P</sup> was functional (Fig. 7A). It is inferred that the threshold for formaldehyde detection is not met by RcnR<sup>S2P</sup> *in vivo*.

*Glutathione Inhibits Formaldehyde Sensing*—The substrates of the FrmR-regulated alcohol dehydrogenase from *Salmonella* (FrmA) are predicted to be the formaldehyde and nitrosylated adducts of reduced glutathione (GSH), *S*-(hydroxymethyl)glutathione and *S*-nitrosoglutathione, respectively, by analogy to *E. coli* (Fig. 2B) (47,69). Despite the evidence that DNA binding by FrmR is directly weakened by formaldehyde *in vitro* (Fig. 3A), glutathione-adducts of formaldehyde might represent the predominant available species during formaldehyde stress conditions. Notably, glutathione has been shown to act positively on metal-detection by FrmR<sup>E64H</sup> *in vivo*, suggesting that the protein may interact with glutathione-adducts (Fig. 1) (50). Deletion of *gshA*, encoding  $\gamma$ -glutamate-cysteine ligase (70), required for the first step in glutathione biosynthesis renders *Salmonella* more sensitive to exogenous formaldehyde compared to the wild type strain (Fig. 8A), as expected if (like in *E. coli*) glutathione is required for formaldehyde detoxification in *Salmonella* by formation of *S*-(hydroxymethyl)glutathione. However, formaldehyde mediated de-repression of *P<sub>frmRA</sub>-frmR* was not impaired in  $\Delta gshA$  cells (Fig. 8B) indicating that formation of formaldehyde glutathione-adducts is not an absolute requirement for FrmR-responsiveness to formaldehyde *in vivo*. Indeed, expression levels from *P<sub>frmRA</sub>-frmR* were higher in  $\Delta gshA$  than wild type, at equivalent exogenous formaldehyde concentrations (Fig. 8B), consistent with FrmR detecting increased formaldehyde accumulation in the cytosol of  $\Delta gshA$  cells, due to reduced FrmA activity and/or due to glutathione acting negatively on the modification of FrmR by formaldehyde.

Binding of FrmR to *frmRAPro* was monitored by fluorescence anisotropy in the

presence of 800  $\mu\text{M}$  GSH. GSH alone has a minimal but detectable effect on DNA binding by apo-FrmR ( $\sim 2.5$  fold tighter) (Fig. 8C). The ability of FrmR to respond to formaldehyde in the presence of GSH was assessed by subsequent titration of FrmR into *frmR*Pro in the presence of both formaldehyde (20  $\mu\text{M}$ ) and excess GSH (800  $\mu\text{M}$ ). The apparent DNA-affinity of FrmR was weaker (relative to FrmR and GSH alone, without formaldehyde) by  $\sim 9$ -fold, but critically the magnitude of the response by FrmR to formaldehyde is diminished by GSH (compare open symbols in Figs. 3A and 8C). These data show that GSH competes with FrmR for formaldehyde, rather than contributing towards its reactivity. How then can FrmR detect free formaldehyde *in vivo*, since glutathione is expected to be in a large molar excess? Under aerobic conditions, the intracellular glutathione concentration in *Salmonella* cells was determined to be 4.2 ( $\pm 0.5$ ) mM (Fig. 8D), whereas the abundance of FrmR was 9.7 ( $\pm 2.6$ ) tetramers per cell (16.1  $\pm 0.2$  nM), as determined by quantitative mass-spectrometry (Fig. 8E and supplemental Table S2). Repression by *E. coli* FrmR is alleviated during TMAO-mediated anaerobic respiration, likely due to TMAO-demethylase activity and intracellular formaldehyde generation (15). The concentration of glutathione in *Salmonella* drops to 1.2 ( $\pm 0.4$ ) mM when cells are grown anaerobically using TMAO as an alternative electron acceptor (Fig. 8D). These data have been used to model formation of *S*-(hydroxymethyl)glutathione and the modification of FrmR as a function of [formaldehyde], with implications for the species detected by FrmR *in vivo* discussed below (Fig. 8F).

## DISCUSSION

Detection of metals and formaldehyde by *Salmonella* FrmR<sup>E64H</sup> is retained when the sensor is expressed in a heterologous *E. coli* host (Fig. 2). Zn(II) is not required to transduce the formaldehyde signal *in vivo* since formaldehyde directly allosterically activates wild type *Salmonella* FrmR *in vitro* (Fig. 3). The allosteric response to organic molecules is specific to formaldehyde and not acetaldehyde *in vitro* and *in vivo* (Fig. 3). Deduced sensory sites for Zn(II)/Co(II) and for formaldehyde overlap with

both effectors requiring Cys<sup>35</sup> (Figs. 4 and 5). Substitution of either Cys<sup>35</sup> or Pro<sup>2</sup> decreases the reactivity of FrmR to formaldehyde *in vitro* and abolishes sensing *in vivo* (Fig. 6). The sensory site of FrmR is more reactive to formaldehyde than the related Ni(II)/Co(II)-sensor RcnR *in vitro*, and RcnR does not respond to formaldehyde *in vivo* (Fig. 7). Introduction of the deduced formaldehyde sensory site to generate RcnR<sup>S2P</sup>, confers increased reactivity to formaldehyde *in vitro* (Fig. 7). Although *S*-(hydroxymethyl)glutathione is a substrate for FrmA, free formaldehyde is the allosteric effector of FrmR, and glutathione competes with FrmR for formaldehyde both *in vitro* and *in vivo* (Fig. 8).

The unexpected ability of Zn(II) and Cu(I) to weaken *Salmonella* FrmR  $K_{\text{DNA}}$  *in vitro* (50) raised the possibility that metals might act as signal transducers of intracellular formaldehyde accumulation (Fig. 1). Moreover, there is precedence for a Zn(II)-dependent alcohol dehydrogenase being regulated in response to Zn(II) by Zap1 (Zinc-responsive Activator Protein) transcription factor as a Zn(II)-sparing mechanism in yeast (71,72). FrmR-regulated glutathione-dependent formaldehyde dehydrogenase, FrmA, similarly requires Zn(II) for activity (47). However, here we eliminate the requirement of Zn(II) during FrmR-mediated derepression of *frmRA*, as formaldehyde is shown by fluorescence anisotropy to be a direct allosteric effector of FrmR (Figs. 2E and F and Fig. 3). The related metal-sensor RcnR (which shares 40 % identity with FrmR), is less reactive to formaldehyde by at least an order of magnitude (Figs. 3A, 6E, and 7C-D). Candidate effector sensory sites for formaldehyde and Zn(II)/Co(II) were identified by structural characterisation of FrmR<sup>E64H</sup> (Figs. 4 and 5) and shown by site-directed mutagenesis to each require Cys<sup>35</sup> (Figs. 4E, 6B-C and 6E). We show that an FrmR-specific amino terminus, Pro<sup>2</sup>, is also required to react with formaldehyde and propose formation of an inter-dimer methylene bridge between the two residues (Figs. 5, 6D-E, and supplemental Fig. S1A). Introduction of such a cross-link would only alter the distance between Cys<sup>35</sup> and Pro<sup>2</sup> by  $\sim +0.5$  Å relative to the crystal structure. Future studies should aim to visualise the proposed methylene bridge and the nature of allosteric coupling between formaldehyde-modification and DNA-

binding. Indeed, such coupling is yet to be characterised for any DUF156 family-member. The unique (to date) Gly<sup>83</sup> in *Salmonella* FrmR, and resulting kink in  $\alpha 3$ , may make this protein distinct.

Glutathione is not required for FrmR to respond to formaldehyde *in vivo* (Fig. 8B): Rather than aid detection, glutathione competes with FrmR for formaldehyde *in vitro* and inhibits the response *in vivo* (Figs. 8B and C). In contrast, although glutathione acts positively towards cobalt detection by FrmR<sup>E64H</sup> *in vivo* (50), the present data argue against a suggestion that FrmR<sup>E64H</sup> preferentially detects cobalt due to its interaction with glutathione conjugates (Fig. 8). Because glutathione is such an abundant biomolecule (Fig. 8D), *S*-(hydroxymethyl)glutathione (the substrate for FrmA) might be expected to predominate over formaldehyde in a cell. However, the affinity of FrmR for formaldehyde is inferred to be  $\sim 10^{-5}$  M from measured DNA-affinities (Fig. 6E), substantially tighter than the affinity for formation of *S*-(hydroxymethyl)glutathione from formaldehyde and GSH ( $1.77 \times 10^{-3}$  M, (73)) (Fig. 8G). Thus, even though GSH is at least five orders of magnitude more abundant than FrmR (determined to be  $16.1 \pm 0.2$  nM, Fig. 8E and supplemental Table S2), FrmR will nonetheless be  $> 85$  % modified by formaldehyde at cellular [formaldehyde] where only 4 % of the GSH pool is in the *S*-(hydroxymethyl)glutathione form (*dashed red line* and *grey line*, Figs. 8F and G). Crucially, this means that expression of *frmA* will be de-repressed as cellular *S*-(hydroxymethyl)glutathione begins to accumulate (Fig. 8F). Importantly, the ability of FrmR to respond to formaldehyde directly prevents [formaldehyde] from rising to levels where crosslinking of other cellular proteins (e.g. RcnR) (*blue dashed line*, Fig. 8F), or significant depletion of the GSH pool would occur.

The *Salmonella* and *E. coli* *frm* operons are distinct (Fig. 2), which could reflect requirements specific to pathogenicity with a suggestion that formaldehyde generation may arise following the macrophage respiratory burst (2). Consistent with this, the *frmRA* locus is known to be up-regulated during intracellular survival of *Salmonella* within macrophages (74,75). Unlike *E. coli*, the *Salmonella* *frm* locus does not present a complete formaldehyde detoxification pathway (or recycling

of GSH) due to the lack of *frmB*. YeiG, capable of catalysing the formation of formate and GSH from *S*-formylglutathione in *E. coli* (27), is also present in *Salmonella* (Fig. 2B) and is an obvious candidate to function in the absence of FrmB (Fig. 8G). Furthermore, *yeiG*, which is not FrmR-regulated, is co-expressed with genes encoded by the *Salmonella* pathogenicity island-2 (SPI-2), notable for being up-regulated and absolutely required during replication within eukaryotic cells (74-76). It is formally possible that differences may emerge between the effectors and sensory sites of *Salmonella* versus *E. coli* FrmR (e.g. residues surrounding sensory sites, specificities to organic molecules, metals as allosteric effectors, nature of allostery) which reflect the demands for survival in the distinctive niches inhabited by each organism.

The FrmR sensory site is more reactive to formaldehyde than, for example, RcnR. The amino terminus of mature FrmR becomes a pyrrolidine secondary amine from Pro<sup>2</sup>, in contrast to the primary amino group of RcnR (from Ser<sup>2</sup>). Consequently, the nucleophilic reactivity of the FrmR amino-terminus is predicted to be greater than RcnR (63,64) and more able to undergo nucleophilic addition to the formaldehyde carbonyl group (Fig. 5). A Cys<sup>35</sup>-thiolate in both proteins also presents a particularly nucleophilic group capable of this reaction (Fig. 5) (63,64). We propose that the presence of this reactive pair in FrmR would allow formation of an inter-dimer cross-link (Fig. 5). Consistent with this, Pro<sup>2</sup> and Cys<sup>35</sup> are required for formaldehyde detection by FrmR (Fig. 6), and creation of the proposed formaldehyde site in RcnR<sup>S2P</sup> increases sensitivity to formaldehyde *in vitro* (Fig. 7). However, mutation of the RcnR amino-terminus alone is not sufficient to confer the same degree of reactivity exhibited by FrmR; therefore it is likely that additional residues optimise formation of a cross-link. Most notable is Pro<sup>5</sup>, another FrmR-specific residue which terminates helix  $\alpha 1$  and may confer a degree of rigidity to the amino terminus, positioning Pro<sup>2</sup> into the sensory site adjacent to Cys<sup>35</sup>.

Effector selectivity of DUF156 family transcriptional repressors can be changed by relatively modest sequence alterations. Conservation of a cysteine at the effector site is now confirmed to be common to a formaldehyde



sensing family member (Figs. 5 and 6), as well as the metal-sensors. Changing single residues proximal to this active cysteine has: Increased the ability of RcnR to sense formaldehyde *in vitro* in RcnR<sup>S2P</sup> (Figs. 7D and E); gained metal-sensing *in vivo* by FrmR<sup>E64H</sup> (Fig. 2F and (50)); and switched metal specificity of RcnR<sup>H3E</sup> *in vivo* (61). Notably, although RcnR<sup>S2P</sup> is more reactive to formaldehyde than RcnR *in vitro*, it still cannot respond *in vivo* (Fig. 7). Furthermore, FrmR can respond to metal *in vitro* but not *in vivo* (Fig. 2E and (50)), the latter being achieved by the FrmR<sup>E64H</sup> variant. In the case of FrmR<sup>E64H</sup>, the threshold for Zn(II)-responsiveness *in vivo* was met by a tighter Zn(II)-affinity and weaker apo-DNA affinity (relative to wild type FrmR), rendering FrmR<sup>E64H</sup> competitive relative to cognate Zn(II) sensors, ZntR and Zur (50). FrmR responds to formaldehyde in a cell placing it above some threshold of reactivity for this effector (Figs. 2E, 3B and 8F). FrmR<sup>C35A</sup> and FrmR<sup>P2S</sup> variants, along with RcnR and RcnR<sup>S2P</sup>, must be below the threshold for formaldehyde sensing (Figs. 6B-E and 7C-E). Presumably, cells do not survive at [effector] sufficient to trigger such sensors. Among the FrmR-RcnR DUF156 proteins (and yet to be tested for CstR and CsoR), subtle quantitative changes to effector-responses tune these sensors above or below different cellular thresholds, and this is sufficient to confer the necessary level of specificity *in vivo*.

## EXPERIMENTAL PROCEDURES

**Bacterial Strains and DNA Manipulations**—*S. enterica* sv. *Typhimurium* strain SL1344 was used as wild type, and strain LB5010a was used as a restriction-deficient modification-proficient host for DNA manipulations (50). Deletion derivatives of SL1344 lacking *frmR* and *gshA* were generated previously (50). *E. coli* strains BW25113Δ*frmR*, in which the *frmR* coding sequence is disrupted by a *kan<sup>R</sup>* cassette (77), was used for β-galactosidase assays. This was a gift from D. Weinkove (Durham University). *E. coli* strain DH5α was used for routine cloning and strain BL21(DE3) for recombinant protein over-expression. Bacteria were cultured aerobically (with shaking) at 37 °C in Luria-Bertani (LB) medium or M9 minimal medium (78), supplemented with thiamine (0.001 % w/v) and either L-histidine (20 μg ml<sup>-1</sup>) for

*Salmonella* or 1 μM C<sub>6</sub>H<sub>5</sub>FeO<sub>7</sub> for *E. coli*. Carbenicillin (100 μg ml<sup>-1</sup>), kanamycin (25 μg ml<sup>-1</sup>) and TMAO (40 mM) were added where appropriate. Cells were transformed to antibiotic resistance as described (78,79). For glutathione quantification under aerobic and anaerobic conditions, glucose was replaced with glycerol as a non-fermentable carbon source. For survival assays, overnight cultures were grown in M9 minimal medium, diluted 1:50 into fresh medium in 14 ml culture tubes containing indicated concentrations of formaldehyde, and grown to mid-logarithmic phase. Growth was assessed by measuring absorbance at 600 nm and calculating the percentage survival compared to the control condition for each strain. Experiments were performed in triplicate on at least three separate occasions. Generated plasmid constructs were checked by sequence analysis. Primers are listed in Supplemental Table S1.

**Bioinformatic analysis**—Fourteen FrmR and nine RcnR non-redundant primary amino acid sequences identified in (45) and still present in UniProtKB (80) were aligned with the *Salmonella enterica* serovar Typhimurium SL1344 FrmR sequence (UniProtKB identifier: A0A0H3NLH8) using PRALINE multiple sequence alignment tool (81). UniProtKB identifiers for the FrmR sequences were: B2SZZ0, *Burkholderia phytofirmans* (strain DSM 17436 / PsJN); Q1IAA5, *Pseudomonas entomophila* (strain L48); B9BFA7, *Burkholderia multivorans* CGD1; F0DZ53, *Pseudomonas* sp. (strain TJI-51); B5JUQ3, *gamma proteobacterium* HTCC5015; H4ZQC4, *Escherichia coli* DEC8C; D8A2B2, *Escherichia coli* (strain MS 21-1); D7ZJL9, *Escherichia coli* MS 69-1; F4VAD6, *Escherichia coli* H591; P0AAP3 (*blue*), *Escherichia coli* (K12); G7LSK1, *Brenneria* sp. EniD312; I0QLA2, *Serratia* sp. M24T3; Q8KKB0, *Proteus vulgaris*; D1P3L2, *Providencia rustigianii* DSM 4541. Residues present in FrmR, but not RcnR, sequences were identified as FrmR specific. For three sequences (D8A2B2, D7ZJL9, F4VAD6) amino terminal residues annotated by UniprotKB were not predicted to be coding residues using the NCBI (National Center for Biotechnology Information) database and were removed. Phylogenetic analysis was performed using ClustalW2 Phylogeny (82); *E.coli* FrmR sequences except *E. coli* K12 were removed.

Distance values relate to the number of substitutions as a proportion of the length of the alignment (excluding gaps). Amino acid sequence identities were determined using Clustal Omega (83).

*Generation of Promoter-lacZ Fusion Constructs and  $\beta$ -Galactosidase Assays*—Promoter-lacZ fusions  $P_{frmRA}$ -*frmR*,  $P_{frmRA}$ -*frmR*<sup>E64H</sup> and *rcnR*- $P_{rcnRA}$  have been described previously (50). Sub-cloning vector pGEM-T containing either the  $P_{frmRA}$ -*frmR* or *rcnR*- $P_{rcnRA}$  DNA fragment (50) was used as a template for site-directed mutagenesis via the Quickchange® protocol (Stratagene) using primers 1 and 2 to generate  $P_{frmRA}$ -*frmR*<sup>P2S</sup>, primers 3 and 4 to generate  $P_{frmRA}$ -*frmR*<sup>C35A</sup> or primers 5 and 6 to generate *rcnR*<sup>S2P</sup>- $P_{rcnRA}$  (Primers listed in Supplemental Table S1). Digested fragments were cloned into the *SmaI/BamHI* site of pRS415 (84). Constructs were introduced into *E. coli* strain BW25113/*AfrmR* as appropriate or *Salmonella* strain LB5010a, prior to SL1344 (and derivatives).  $\beta$ -galactosidase assays were performed as described (50,85), in triplicate, and on at least three separate occasions. Briefly, overnight cultures were grown in M9 minimal medium, diluted 1:50 in fresh medium supplemented with up to maximum non-inhibitory concentrations (MNIC; defined as the maximum concentration which inhibited growth by ~ 10%) of metal, alcohol or aldehyde and grown to mid-logarithmic phase prior to assays. MNICs under these growth conditions were 5  $\mu$ M CoCl<sub>2</sub>, 50  $\mu$ M ZnCl<sub>2</sub>, 300 mM ethanol, 600 mM methanol, 5 mM 1-butanol, 50 mM 1-propanol, 200 mM 2-propanol, 50  $\mu$ M formaldehyde and 3 mM acetaldehyde, with the exception that 1  $\mu$ M CoCl<sub>2</sub> was the MNIC for cells expressing *rcnR*- $P_{rcnRA}$  or *rcnR*<sup>S2P</sup>- $P_{rcnRA}$ . Time-course experiments were performed as described (50) by exposing logarithmic cells to MNIC metal or formaldehyde for 2-h at 25 °C. Where stated,  $\beta$ -galactosidase activity (nmoles *o*-nitrophenol min<sup>-1</sup> mg protein<sup>-1</sup>) was normalised to the control data for cells expressing the wild type protein conducted in parallel.

*Protein Expression and Purification*—Vectors for overexpression of FrmR, FrmR<sup>E64H</sup> and RcnR have been described previously (50). Site-directed mutagenesis was conducted as described above using template pETfrmR and primers 7-12 to generate pETfrmR<sup>P2S</sup>,

pETfrmR<sup>C35A</sup> and pETfrmR<sup>H60L</sup>, or template pETrcnR and primers 13 and 14 to generate pETrcnR<sup>S2P</sup>. Proteins were expressed and purified as described previously (50). Mutant variants were purified exactly as described for the respective wild type protein. Protein purity was assessed by SDS-PAGE. Anaerobic protein stocks (maintained in an anaerobic chamber) were prepared as described and confirmed to be  $\geq 90$  % reduced and  $\geq 95$  % metal-free (50). FrmR and variants were stored in 100 mM NaCl, 400 mM KCl, 10 mM HEPES pH 7.0. RcnR and variants were stored in 200 mM NaCl, 800 mM KCl, 10 mM HEPES pH 7.0. All *in vitro* experiments were carried out under anaerobic conditions using Chelex-treated and N<sub>2</sub>-purged buffers as described previously (50). Due to the absence of any thiol groups, experiments with FrmR<sup>C35A</sup> were carried out under aerobic conditions.

*Inductively coupled plasma-mass spectrometry (ICP-MS)*—Anaerobic protein stocks (10 – 20  $\mu$ M) or size-exclusion chromatography fractions were diluted ten-fold in 2.5 % Suprapur HNO<sub>3</sub> (Merck Millipore). Quantitative analysis of metal content was determined using an XSERIES-2 inductively coupled plasma-mass spectrometer (Thermo Fisher Scientific) following calibration with elemental standards that were matrix-matched to the sample by inclusion of appropriate buffer system.

*Fluorescence Anisotropy*—Fluorescently labelled double-stranded DNA probe, *frmRAPro*, containing the identified FrmR-binding site has been described previously (50). Complementary single-stranded oligonucleotides 15 (hexachlorofluorescein-labelled) and 16 (containing two identified RcnR-binding sites (32,65) and flanking oligonucleotides) or 17 (hexachlorofluorescein-labelled) and 18 (*frmRAPro* but with mutation of one T/A-rich inverted repeat) (Supplemental Table S1) were annealed by heating 10 or 200  $\mu$ M of each strand in 150 mM NaCl, 10 mM HEPES pH 7.0 to 95 °C and cooled slowly overnight, to generate *rcnRAPro* (35 bp) or *frmRAPro\** (33 bp). Fluorescently labelled annealed probes were analysed by native-PAGE (12 % (w/v)) and RcnR:*rcnRAPro* stoichiometry experiments were performed as described (50) by titration of RcnR (prepared in 100 mM NaCl, 400 mM KCl, 10 mM HEPES pH 7.0, and 5 mM EDTA) into 2.5  $\mu$ M

*rcnRAPro* in 60 mM NaCl, 240 mM KCl, 10 mM HEPES pH 7.0, and 5 mM EDTA. For  $K_{\text{DNA}}$  determination in the absence of effector, *frmRAPro* or *rcnRAPro* were diluted to 10 nM in the same buffer. Formaldehyde, acetaldehyde and GSH were included as outlined in figure legends. For metal-loaded experiments, EDTA was replaced with 5  $\mu\text{M}$  NiCl<sub>2</sub> or CoCl<sub>2</sub>. FrmR (and variants) and RcnR (and variants) were prepared as described previously and above (50) or replacing EDTA with 1.2 molar eq. per protein monomer of NiCl<sub>2</sub> or CoCl<sub>2</sub> as appropriate. Formaldehyde was prepared daily from single-use sealed ampules of methanol-free 16% (v/v) formaldehyde (Pierce) and stored under anaerobic conditions for the course of the experiment to prevent oxidation. A concentrated GSH stock (5 mM) was prepared in Chelex-treated, N<sub>2</sub>-purged 100 mM NaCl, 400 mM KCl, 10 mM HEPES pH 7.0, under anaerobic conditions. This stock was confirmed to be  $\geq 90\%$  reduced by reaction with DTNB, stored anaerobically to prevent oxidation and used within three days. Changes in anisotropy ( $\Delta r_{\text{obs}}$ ) were measured using a modified Cary Eclipse fluorescence spectrophotometer (Agilent Technologies) as described (50). Control titrations of apo-FrmR and apo-FrmR<sup>E64H</sup> into *frmRAPro* (Fig. 4C) are new unpublished data sets and are presented here to demonstrate reproducibility and for comparative purposes. Data (for both FrmR and RcnR) were fit to a model describing binding of two non-dissociable tetramers ( $K_{\text{tet}}$  fixed at  $10^{-20}$  M) to a target DNA probe with equal affinity (50,86), using Dynafit (87) (see figure legends and Table 2 for details; sample Dynafit script shown in Supplemental Material). For experiments where DNA-binding did not saturate, the average fitted  $\Delta r_{\text{obs}}$  maximum value from apo-protein experiments was used as a fixed parameter. Coupling free energies ( $\Delta G_{\text{C}}$ ) linking DNA-binding to effector-binding (62,66,67), were determined as described previously (50) calculated from the full set of (equally-weighted) possible pairwise permutations of  $K_{\text{C}}$ .

*Crystallization of FrmR<sup>E64H</sup> and Data Collection*—Concentrated FrmR<sup>E64H</sup> (~ 1 mM) was diluted to 0.5 mM in 400 mM NaCl, 1 mM EDTA, 1 mM DTT, and 10 mM HEPES pH 7.0, and stored aerobically at 4 °C for up to two weeks. Initial crystallisation trials were conducted using the Screenmaker 96+8™ Xtal (Innovadyne

Technologies) and commercially available screening kits (Molecular Dimensions). Subsequent FrmR<sup>E64H</sup> crystals were obtained in 20 mM NaCl, 23 % (w/v) poly(ethylene glycol) 4000, and 10 mM BisTris pH 6.5 by hanging drop vapour diffusion at 20 °C. Crystals were physically fragile and disintegrated rapidly when cryo-protectants were added. Multiple crystals using a wide range of cryo-conditions were frozen and tested. Results were obtained from a crystal soaked in 25 % (v/v) glycerol mounted in cryoloops (88). Overall data quality was compromised by residual ice rings and anisotropic diffraction potentially giving rise to higher than expected R-factors of the final model. FrmR<sup>E64H</sup> diffraction data were collected at the Diamond Light Source on beam line I03 at 77 K with a Pilatus pixel detector (89). Diffraction data were initially processed using Mosfilm (90) to a resolution of 2.1 Å to enable *ab-initio* solution, and reprocessed (to 2.19 Å) with Xia2 (91) for structure refinement. Initial molecular replacement trials using MolRep (92) and Phaser (93) using PDB entry 2HH7 (Cu(I)-CsoR from *Mycobacterium tuberculosis*) (31) were unsuccessful, presumably due to differences in the orientation of the three helices of the monomer, and significant differences in monomer-monomer as well as dimer-dimer orientations in the homotetramer. The structure was solved using Arcimboldo installed on a Condor grid computer (94,95). The initial model was completed by iterative cycles of model building and refinement using COOT (96) and REFMAC (97). The final model contained one homotetramer in the asymmetric unit with each chain containing residues 2-89, and residue 90 in chains B and C, and 103 water molecules. The structure was refined against intensities with local non-crystallographic symmetry restraints (98), using Phenix (99). Applying local non-crystallographic symmetry restraints enabled the tracing of all four chains despite relatively weak density, particularly in  $\alpha 3$ . This confers higher than expected overall real-space R-value Z-score and R-factors. A number of polar surface residues where no electron density was observed for the side chains were refined as alanines (Chain A: His<sup>3</sup>, Lys<sup>8</sup>, Lys<sup>9</sup>, Glu<sup>69</sup>, Ile<sup>82</sup>; Chain B: Lys<sup>62</sup>, Glu<sup>69</sup>, Ile<sup>82</sup>, Leu<sup>90</sup>; Chain C: Glu<sup>30</sup>, Glu<sup>69</sup>, Leu<sup>90</sup>; Chain D: His<sup>3</sup>, Lys<sup>8</sup>, Lys<sup>9</sup>, Glu<sup>55</sup>, Lys<sup>62</sup>, Glu<sup>69</sup>, Ile<sup>82</sup>).

Ramachandran plot analysis using RAMPAGE Ramachandran Plot Assessment (100) of FrmR<sup>E64H</sup> demonstrates that 98.8 % of residues are in the favored region (supplemental Fig. S1B). The final data and refinement statistics are provided in Table 1 with the structure deposited in the PDB under accession code 5LCY.

*Protein-Metal Migration by Size-exclusion Chromatography*—Experiments were carried out as described previously (50). FrmR, FrmR<sup>E64H</sup>, FrmR<sup>C35A</sup>, or FrmR<sup>H60L</sup> were incubated (120 min) with an excess of ZnCl<sub>2</sub> in 100 mM NaCl, 400 mM KCl, 10 mM HEPES pH 7.0, and an aliquot (0.5 ml) was resolved by size-exclusion chromatography (PD10 Sephadex G25, GE Healthcare) in the same buffer conditions. Fractions were analysed for zinc by ICP-MS and protein by Bradford assay as described (50). The control experiments with FrmR and FrmR<sup>E64H</sup> (Fig. 4E) are new unpublished data sets and are presented here to demonstrate reproducibility and for comparative purposes.

*FrmR In Vivo Quantification and Detection of Met<sup>I</sup> Cleavage by Liquid Chromatography-Tandem Mass Spectrometry*—Quantification of FrmR in cellular lysates of SL1344 was performed exactly as described previously using aerobically grown logarithmic cells (50). To detect FrmR Met<sup>I</sup> cleavage, a tryptic digest was performed with 5 µg FrmR and 14 µg trypsin in 50 mM NH<sub>4</sub>HCO<sub>3</sub> with shaking (1000 rpm) at 37 °C for 16 h and stopped by addition of 15 % (v/v) formic acid (5 µl). The digested samples were separated by gradient elution at 0.3 ml min<sup>-1</sup> using a Zorbax Eclipse Plus C18 column (2.1 x 150 mm, 3.5 µm particles, Agilent Technologies) at room temperature. Mobile phase A and B consisted of 0.1 % (v/v) formic acid in water and 0.1 % (v/v) formic acid in acetonitrile, respectively. Detection of FrmR amino terminal peptide PHSPEDK was achieved by applying an aliquot (10 µl) to a 6500 triple quadrupole mass spectrometer (AB Sciex) operating in positive ionization mode. Acquisition methods used the following parameters: 5500 V ion spray voltage; 25 psi curtain gas; 60 psi source gas; 550 °C interface heating temperature; 40 V declustering potential; 26 V collision energy; and 27 V collision cell exit potential. Scheduled multiple reaction monitoring (MRM) of ion transition 405.19/488.24 was performed with a 90

s MRM detection window and 1.00 s target scan time.

*Quantification of intracellular glutathione*—Intracellular glutathione was measured as described (50). Lysates from logarithmically growing cells were prepared from overnight cultures grown in M9 minimal medium with glycerol as a carbon source, diluted in 1:50 fresh medium, and grown at 37°C either in round-bottom flasks with shaking to maintain aerobic conditions, or with addition of TMAO (40 mM) and static incubation of completely filled parafilm-sealed 1.5 ml tubes to maintain anaerobic conditions. No growth was observed under anaerobic conditions when TMAO was not included as an electron acceptor. Viable cells were enumerated on LB agar, and cell volume was estimated as 1 fl.

*Fractional Occupancy Model to Describe Formaldehyde Modification In Vivo*—Fractional modification of FrmR, RcnR and GSH with formaldehyde as a function of formaldehyde concentration was determined using Dynafit (87) with the following values as fixed parameters: affinities of FrmR and RcnR for formaldehyde were estimated following fluorescence anisotropy to be 10<sup>-5</sup> M and 10<sup>-4</sup> M, respectively; the dissociation constant for GSH and S-(hydroxymethyl)glutathione was 1.77 × 10<sup>-3</sup> M (73); the total intracellular concentration of glutathione during anaerobic growth with TMAO as an electron acceptor was 1.2 × 10<sup>-3</sup> M and assumed to be in the reduced form (90-99 % of the glutathione pool is GSH in resting *Salmonella* cells (101, 102)); the intracellular concentration of FrmR tetramer under aerobic conditions was calculated to be 1.61 × 10<sup>-8</sup> M and used as an estimate for the intracellular concentrations of FrmR and RcnR tetramers under anaerobic growth conditions. Cell volume was 1 fl. A sample Dynafit script is shown in Supplemental Material.

## ACKNOWLEDGEMENTS

This work was supported by a Biotechnology and Biological Sciences Research Council Industrial Partnership Award BB/J017787/1. The authors would like to thank Diamond Light Source for access to beamline I03. We thank DBS Genomics (Durham University) for sequencing of plasmid constructs, David Weinkove (Durham University) for the *E. coli frmR* deletion strain, and Jeff Green

(University of Sheffield), Peter T. Chivers and Andrew W. Foster (both Durham University) for insightful discussions. The support and advice of Elena Lurie-Luke and Charlie Saunders (both Procter and Gamble) is gratefully acknowledged.

### CONFLICT OF INTEREST

This work was supported by a Biotechnology and Biological Sciences Research Council Industrial Partnership Award with support from Procter and Gamble.

### AUTHOR CONTRIBUTIONS

D.O. carried out the *in vivo* survival and gene expression experiments, bioinformatic analysis,

and developed the fractional occupancy model for formaldehyde modification. C.P. did the *in vitro* metal-binding and fluorescence anisotropy experiments, and generated FrmR<sup>E64H</sup> crystals. E.P. and I.U. solved the X-ray crystal structure of FrmR<sup>E64H</sup>. J.C. and T.G.H. performed the LC-MS/MS. L.N.S. and C.P. determined the intracellular glutathione concentration. D.O. and C.P. made equivalent contributions to data preparation, and with N.J.R. interpreted the significance of the data. D.O. and N.J.R. drafted the manuscript. N.J.R. had overall responsibility for the design and coordination of the programme. All authors reviewed the results, edited, and approved the final version of the manuscript.

### REFERENCES

1. Bolt, H. M. (1987) Experimental toxicology of formaldehyde. *J. Cancer Res. Clin. Oncol.* **113**, 305-309
2. Chen, N. H., Djoko, K. Y., Veyrier, F. J., and McEwan, A. G. (2016) Formaldehyde stress responses in bacterial pathogens. *Front. Microbiol.* **7**, 257
3. Conaway, C. C., Whysner, J., Verna, L. K., and Williams, G. M. (1996) Formaldehyde mechanistic data and risk assessment: endogenous protection from DNA adduct formation. *Pharmacol. Ther.* **71**, 29-55
4. Feldman, M. Y. (1973) Reactions of nucleic acids and nucleoproteins with formaldehyde. *Prog. Nucleic Acid Res. Mol. Biol.* **13**, 1-49
5. Ma, T. H., and Harris, M. M. (1988) Review of the genotoxicity of formaldehyde. *Mutat. Res.* **196**, 37-59
6. Chistoserdova, L., Kalyuzhnaya, M. G., and Lidstrom, M. E. (2009) The expanding world of methylotrophic metabolism. *Annu. Rev. Microbiol.* **63**, 477-499
7. Hanson, R. S., and Hanson, T. E. (1996) Methanotrophic bacteria. *Microbiol. Rev.* **60**, 439-471
8. Vorholt, J. A. (2002) Cofactor-dependent pathways of formaldehyde oxidation in methylotrophic bacteria. *Arch. Microbiol.* **178**, 239-249
9. Yurimoto, H., Kato, N., and Sakai, Y. (2005) Assimilation, dissimilation, and detoxification of formaldehyde, a central metabolic intermediate of methylotrophic metabolism. *Chem. Rec.* **5**, 367-375
10. Ferenci, T., Strom, T., and Quayle, J. R. (1974) Purification and properties of 3-hexulose phosphate synthase and phospho-3-hexuloisomerase from *Methylococcus capsulatus*. *Biochem. J.* **144**, 477-486
11. Marx, C. J., Chistoserdova, L., and Lidstrom, M. E. (2003) Formaldehyde-detoxifying role of the tetrahydromethanopterin-linked pathway in *Methylobacterium extorquens* AM1. *J. Bacteriol.* **185**, 7160-7168
12. Marx, C. J., Miller, J. A., Chistoserdova, L., and Lidstrom, M. E. (2004) Multiple formaldehyde oxidation/detoxification pathways in *Burkholderia fungorum* LB400. *J. Bacteriol.* **186**, 2173-2178
13. Matsui, T., Nambu, S., Ono, Y., Goulding, C. W., Tsumoto, K., and Ikeda-Saito, M. (2013) Heme degradation by *Staphylococcus aureus* IsdG and IsdI liberates formaldehyde rather than carbon monoxide. *Biochemistry* **52**, 3025-3027

14. Streit, B. R., Kant, R., Tokmina-Lukaszewska, M., Celis, A. I., Machovina, M. M., Skaar, E. P., Bothner, B., and DuBois, J. L. (2016) Time-resolved studies of IsdG protein identify molecular signposts along the non-canonical heme oxygenase pathway. *J. Biol. Chem.* **291**, 862-871
15. Denby, K. J., Rolfe, M. D., Crick, E., Sanguinetti, G., Poole, R. K., and Green, J. (2015) Adaptation of anaerobic cultures of *Escherichia coli* K-12 in response to environmental trimethylamine-N-oxide. *Environ. Microbiol.* **17**, 2477-2491
16. Thornalley, P., Wolff, S., Crabbe, J., and Stern, A. (1984) The autoxidation of glyceraldehyde and other simple monosaccharides under physiological conditions catalysed by buffer ions. *Biochim. Biophys. Acta* **797**, 276-287
17. Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T., and Sedgwick, B. (2002) Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage. *Nature* **419**, 174-178
18. Thornalley, P. J. (1993) The glyoxalase system in health and disease. *Mol. Aspects Med.* **14**, 287-371
19. Harms, N., Ras, J., Reijnders, W. N., van Spanning, R. J., and Stouthamer, A. H. (1996) S-formylglutathione hydrolase of *Paracoccus denitrificans* is homologous to human esterase D: a universal pathway for formaldehyde detoxification? *J. Bacteriol.* **178**, 6296-6299
20. Goenrich, M., Bartoschek, S., Hagemeyer, C. H., Griesinger, C., and Vorholt, J. A. (2002) A glutathione-dependent formaldehyde-activating enzyme (Gfa) from *Paracoccus denitrificans* detected and purified via two-dimensional proton exchange NMR spectroscopy. *J. Biol. Chem.* **277**, 3069-3072
21. Wilson, S. M., Gleisten, M. P., and Donohue, T. J. (2008) Identification of proteins involved in formaldehyde metabolism by *Rhodobacter sphaeroides*. *Microbiology* **154**, 296-305
22. Uotila, L., and Koivusalo, M. (1974) Purification and properties of S-formylglutathione hydrolase from human liver. *J. Biol. Chem.* **249**, 7664-7672
23. Kaulfers, P. M., and Marquardt, A. (1991) Demonstration of formaldehyde dehydrogenase activity in formaldehyde-resistant Enterobacteriaceae. *FEMS Microbiol. Lett.* **63**, 335-338
24. Rosario Fernandez, M., Jornvall, H., Moreno, A., Kaiser, R., and Pares, X. (1993) Cephalopod alcohol dehydrogenase: purification and enzymatic characterization. *FEBS Lett.* **328**, 235-238
25. Gutheil, W. G., Kasimoglu, E., and Nicholson, P. C. (1997) Induction of glutathione-dependent formaldehyde dehydrogenase activity in *Escherichia coli* and *Hemophilus influenzae*. *Biochem. Biophys. Res. Commun.* **238**, 693-696
26. Herring, C. D., and Blattner, F. R. (2004) Global transcriptional effects of a suppressor tRNA and the inactivation of the regulator *frmR*. *J. Bacteriol.* **186**, 6714-6720
27. Gonzalez, C. F., Proudfoot, M., Brown, G., Korniyenko, Y., Mori, H., Savchenko, A. V., and Yakunin, A. F. (2006) Molecular basis of formaldehyde detoxification. Characterization of two S-formylglutathione hydrolases from *Escherichia coli*, FrmB and YeiG. *J. Biol. Chem.* **281**, 14514-14522
28. Mason, R. P., Sanders, J. K. M., Crawford, A., and Hunter, B. K. (1986) Formaldehyde metabolism by *Escherichia coli*. Detection by *in vivo* carbon-13 NMR spectroscopy of S-(hydroxymethyl)glutathione as a transient intracellular intermediate. *Biochemistry* **25**, 4504-4507
29. Uotila, L., and Koivusalo, M. (1974) Formaldehyde dehydrogenase from human liver. Purification, properties, and evidence for the formation of glutathione thiol esters by the enzyme. *J. Biol. Chem.* **249**, 7653-7663
30. Higgins, K. A., and Giedroc, D. (2014) Insights into protein allostery in the CsoR/RcnR family of transcriptional repressors. *Chem. Lett.* **43**, 20-25
31. Liu, T., Ramesh, A., Ma, Z., Ward, S. K., Zhang, L., George, G. N., Talaat, A. M., Sacchettini, J. C., and Giedroc, D. P. (2007) CsoR is a novel *Mycobacterium tuberculosis* copper-sensing transcriptional regulator. *Nat. Chem. Biol.* **3**, 60-68
32. Iwig, J. S., Rowe, J. L., and Chivers, P. T. (2006) Nickel homeostasis in *Escherichia coli* - the *rcnR-rcnA* efflux pathway and its linkage to NikR function. *Mol. Microbiol.* **62**, 252-262

33. Brown, N. L., Stoyanov, J. V., Kidd, S. P., and Hobman, J. L. (2003) The MerR family of transcriptional regulators. *FEMS Microbiol. Rev.* **27**, 145-163
34. Lee, J. W., and Helmann, J. D. (2007) Functional specialization within the Fur family of metalloregulators. *Biometals* **20**, 485-499
35. Osman, D., and Cavet, J. S. (2010) Bacterial metal-sensing proteins exemplified by ArsR-SmtB family repressors. *Nat. Prod. Rep.* **27**, 668-680
36. Pennella, A. M., and Giedroc, P. D. (2005) Structural determinants of metal selectivity in prokaryotic metal-responsive transcriptional regulators. *Biometals* **18**, 413-428
37. Rubio-Sanz, L., Prieto, R. I., Imperial, J., Palacios, J. M., and Brito, B. (2013) Functional and expression analysis of the metal-inducible *dmeRF* system from *Rhizobium leguminosarum* *bv. viciae*. *Appl. Environ. Microbiol.* **79**, 6414-6422
38. Festa, R. A., Jones, M. B., Butler-Wu, S., Sinsimer, D., Gerads, R., Bishai, W. R., Peterson, S. N., and Darwin, K. H. (2011) A novel copper-responsive regulon in *Mycobacterium tuberculosis*. *Mol. Microbiol.* **79**, 133-148
39. Foster, A. W., Patterson, C. J., Pernil, R., Hess, C. R., and Robinson, N. J. (2012) Cytosolic Ni(II) sensor in cyanobacterium: nickel detection follows nickel affinity across four families of metal sensors. *J. Biol. Chem.* **287**, 12142-12151
40. Grossoehme, N., Kehl-Fie, T. E., Ma, Z., Adams, K. W., Cowart, D. M., Scott, R. A., Skaar, E. P., and Giedroc, D. P. (2011) Control of copper resistance and inorganic sulfur metabolism by paralogous regulators in *Staphylococcus aureus*. *J. Biol. Chem.* **286**, 13522-13531
41. Luebke, J. L., Arnold, R. J., and Giedroc, D. P. (2013) Selenite and tellurite form mixed seleno- and telluro-trisulfides with CstR from *Staphylococcus aureus*. *Metallomics* **5**, 335-342
42. Coyne, H. J., 3rd, and Giedroc, D. P. (2013) Backbone resonance assignments of the homotetrameric (48 kD) copper sensor CsoR from *Geobacillus thermodenitrificans* in the apo- and Cu(I)-bound states: insights into copper-mediated allostery. *Biomol. NMR Assign.* **7**, 279-283
43. Dwarakanath, S., Chaplin, A. K., Hough, M. A., Rigali, S., Vijgenboom, E., and Worrall, J. A. (2012) Response to copper stress in *Streptomyces lividans* extends beyond genes under direct control of a copper-sensitive operon repressor protein (CsoR). *J. Biol. Chem.* **287**, 17833-17847
44. Sakamoto, K., Agari, Y., Agari, K., Kuramitsu, S., and Shinkai, A. (2010) Structural and functional characterization of the transcriptional repressor CsoR from *Thermus thermophilus* HB8. *Microbiology* **156**, 1993-2005
45. Chang, F. M., Coyne, H. J., Cubillas, C., Vinuesa, P., Fang, X., Ma, Z., Ma, D., Helmann, J. D., Garcia-de los Santos, A., Wang, Y. X., Dann, C. E., 3rd, and Giedroc, D. P. (2014) Cu(I)-mediated allosteric switching in a copper-sensing operon repressor (CsoR). *J. Biol. Chem.* **289**, 19204-19217
46. Iwig, J. S., Leitch, S., Herbst, R. W., Maroney, M. J., and Chivers, P. T. (2008) Ni(II) and Co(II) sensing by *Escherichia coli* RcnR. *J. Am. Chem. Soc.* **130**, 7592-7606
47. Gutheil, W. G., Holmquist, B., and Vallee, B. L. (1992) Purification, characterization, and partial sequence of the glutathione-dependent formaldehyde dehydrogenase from *Escherichia coli*: a class III alcohol dehydrogenase. *Biochemistry* **31**, 475-481
48. Nobre, L. S., Al-Shahrour, F., Dopazo, J., and Saraiva, L. M. (2009) Exploring the antimicrobial action of a carbon monoxide-releasing compound through whole-genome transcription profiling of *Escherichia coli*. *Microbiology* **155**, 813-824
49. Wang, S., Deng, K., Zaremba, S., Deng, X., Lin, C., Wang, Q., Tortorello, M. L., and Zhang, W. (2009) Transcriptomic response of *Escherichia coli* O157:H7 to oxidative stress. *Appl. Environ. Microbiol.* **75**, 6110-6123
50. Osman, D., Piergentili, C., Chen, J., Chakrabarti, B., Foster, A. W., Lurie-Luke, E., Huggins, T. G., and Robinson, N. J. (2015) Generating a metal-responsive transcriptional regulator to test what confers metal sensing in cells. *J. Biol. Chem.* **290**, 19806-19822

51. Yurimoto, H., Hirai, R., Matsuno, N., Yasueda, H., Kato, N., and Sakai, Y. (2005) HxlR, a member of the DUF24 protein family, is a DNA-binding protein that acts as a positive regulator of the formaldehyde-inducible *hxlAB* operon in *Bacillus subtilis*. *Mol. Microbiol.* **57**, 511-519
52. Kidd, S. P., Potter, A. J., Apicella, M. A., Jennings, M. P., and McEwan, A. G. (2005) NmlR of *Neisseria gonorrhoeae*: a novel redox responsive transcription factor from the MerR family. *Mol. Microbiol.* **57**, 1676-1689
53. Kidd, S. P., Jiang, D., Tikhomirova, A., Jennings, M. P., and McEwan, A. G. (2012) A glutathione-based system for defense against carbonyl stress in *Haemophilus influenzae*. *BMC Microbiol.* **12**, 159
54. Stroehler, U. H., Kidd, S. P., Stafford, S. L., Jennings, M. P., Paton, J. C., and McEwan, A. G. (2007) A pneumococcal MerR-like regulator and S-nitrosoglutathione reductase are required for systemic virulence. *J. Infect. Dis.* **196**, 1820-1826
55. McEwan, A. G., Djoko, K. Y., Chen, N. H., Counago, R. L., Kidd, S. P., Potter, A. J., and Jennings, M. P. (2011) Novel bacterial MerR-like regulators their role in the response to carbonyl and nitrosative stress. *Adv. Microb. Physiol.* **58**, 1-22
56. Chen, N. H., Counago, R. M., Djoko, K. Y., Jennings, M. P., Apicella, M. A., Kobe, B., and McEwan, A. G. (2013) A glutathione-dependent detoxification system is required for formaldehyde resistance and optimal survival of *Neisseria meningitidis* in biofilms. *Antioxid. Redox. Signal* **18**, 743-755
57. Sanderson, K. E. (1976) Genetic relatedness in the family Enterobacteriaceae. *Annu. Rev. Microbiol.* **30**, 327-349
58. McClelland, M., Florea, L., Sanderson, K., Clifton, S. W., Parkhill, J., Churcher, C., Dougan, G., Wilson, R. K., and Miller, W. (2000) Comparison of the *Escherichia coli* K-12 genome with sampled genomes of a *Klebsiella pneumoniae* and three *Salmonella enterica* serovars, Typhimurium, Typhi and Paratyphi. *Nucleic Acids Res.* **28**, 4974-4986
59. Sharp, P. M. (1991) Determinants of DNA sequence divergence between *Escherichia coli* and *Salmonella typhimurium*: codon usage, map position, and concerted evolution. *J. Mol. Evol.* **33**, 23-33
60. Brodolin, K. (2010) *Protein-DNA crosslinking with formaldehyde in vitro*, Oxford University Press Inc., New York
61. Higgins, K. A., Chivers, P. T., and Maroney, M. J. (2012) Role of the N-terminus in determining metal-specific responses in the *Escherichia coli* Ni- and Co-responsive metalloregulator, RcnR. *J. Am. Chem. Soc.* **134**, 7081-7093
62. Foster, A. W., Pernil, R., Patterson, C. J., and Robinson, N. J. (2014) Metal specificity of cyanobacterial nickel-responsive repressor InrS: cells maintain zinc and copper below the detection threshold for InrS. *Mol. Microbiol.* **92**, 797-812
63. Brotzel, F., Chu, Y. C., and Mayr, H. (2007) Nucleophilicities of primary and secondary amines in water. *J. Org. Chem.* **72**, 3679-3688
64. Brotzel, F., and Mayr, H. (2007) Nucleophilicities of amino acids and peptides. *Org. Biomol. Chem.* **5**, 3814-3820
65. Iwig, J. S., and Chivers, P. T. (2009) DNA recognition and wrapping by *Escherichia coli* RcnR. *J. Mol. Biol.* **393**, 514-526
66. Grosseohme, N. E., and Giedroc, D. P. (2012) Illuminating allostery in metal sensing transcriptional regulators. *Methods. Mol. Biol.* **875**, 165-192
67. Guerra, A. J., and Giedroc, D. P. (2012) Metal site occupancy and allosteric switching in bacterial metal sensor proteins. *Arch. Biochem. Biophys.* **519**, 210-222
68. Ma, Z., Jacobsen, F. E., and Giedroc, D. P. (2009) Coordination chemistry of bacterial metal transport and sensing. *Chem. Rev.* **109**, 4644-4681
69. Liu, L., Hausladen, A., Zeng, M., Que, L., Heitman, J., and Stamler, J. S. (2001) A metabolic enzyme for S-nitrosothiol conserved from bacteria to humans. *Nature* **410**, 490-494



70. Hibi, T., Nii, H., Nakatsu, T., Kimura, A., Kato, H., Hiratake, J., and Oda, J. (2004) Crystal structure of  $\gamma$ -glutamylcysteine synthetase: Insights into the mechanism of catalysis by a key enzyme for glutathione homeostasis. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 15052-15057
71. Bird, A. J., Gordon, M., Eide, D. J., and Winge, D. R. (2006) Repression of ADH1 and ADH3 during zinc deficiency by Zap1-induced intergenic RNA transcripts. *EMBO J.* **25**, 5726-5734
72. Lyons, T. J., Gasch, A. P., Gaither, L. A., Botstein, D., Brown, P. O., and Eide, D. J. (2000) Genome-wide characterization of the Zap1p zinc-responsive regulon in yeast. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7957-7962
73. Sanghani, P. C., Stone, C. L., Ray, B. D., Pindel, E. V., Hurley, T. D., and Bosron, W. F. (2000) Kinetic mechanism of human glutathione-dependent formaldehyde dehydrogenase. *Biochemistry* **39**, 10720-10729
74. Eriksson, S., Lucchini, S., Thompson, A., Rhen, M., and Hinton, J. C. D. (2003) Unravelling the biology of macrophage infection by gene expression profiling of intracellular *Salmonella enterica*. *Mol. Microbiol.* **47**, 103-118
75. Srikumar, S., Kröger, C., Hébrard, M., Colgan, A., Owen, S. V., Sivasankaran, S. K., Cameron, A. D. S., Hokamp, K., and Hinton, J. C. D. (2015) RNA-seq brings new insights to the intramacrophage transcriptome of *Salmonella* Typhimurium. *PLoS Pathog.* **11**, e1005262
76. Shea, J. E., Hensel, M., Gleeson, C., and Holden, D. W. (1996) Identification of a virulence locus encoding a second type III secretion system in *Salmonella typhimurium*. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 2593-2597
77. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008
78. Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
79. Datsenko, K. A., and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6640-6645
80. UniProt Consortium. (2015) UniProt: a hub for protein information. *Nucleic Acids Res.* **43**, D204-212
81. Simossis, V. A., and Heringa, J. (2005) PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. *Nucleic Acids Res.* **33**, W289-294
82. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948
83. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539
84. Simons, R. W., Houman, F., and Kleckner, N. (1987) Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. *Gene* **53**, 85-96
85. Osman, D., Waldron, K. J., Denton, H., Taylor, C. M., Grant, A. J., Mastroeni, P., Robinson, N. J., and Cavet, J. S. (2010) Copper homeostasis in *Salmonella* is atypical and copper-CueP is a major periplasmic metal complex. *J. Biol. Chem.* **285**, 25259-25268
86. Liu, T., Chen, X., Ma, Z., Shokes, J., Hemmingsen, L., Scott, R. A., and Giedroc, D. P. (2008) A Cu(I)-sensing ArsR family metal sensor protein with a relaxed metal selectivity profile. *Biochemistry* **47**, 10564-10575
87. Kuzmic, P. (1996) Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal. Biochem.* **237**, 260-273
88. Teng, T.-Y. (1990) Mounting of crystals for macromolecular crystallography in a free-standing thin film. *J. Appl. Cryst.* **23**, 387-391

89. Broennimann, C., Eikenberry, E. F., Henrich, B., Horisberger, R., Huelsen, G., Pohl, E., Schmitt, B., Schulze-Briese, C., Suzuki, M., Tomizaki, T., Toyokawa, H., and Wagner, A. (2006) The PILATUS 1M detector. *J. Synchrotron Radiat.* **13**, 120-130
90. Powell, H. R., Johnson, O., and Leslie, A. G. W. (2013) Autoindexing diffraction images with iMosflm. *Acta. Crystallogr. D Biol. Crystallogr.* **69**, 1195-1203
91. Winter, G. (2010) xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Cryst.* **43**, 186-190
92. Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP. *Acta. Crystallogr. D Biol. Crystallogr.* **66**, 22-25
93. Bunkóczi, G., Echols, N., McCoy, A. J., Oeffner, R. D., Adams, P. D., and Read, R. J. (2013) Phaser.MRage: automated molecular replacement. *Acta. Crystallogr. D Biol. Crystallogr.* **69**, 2276-2286
94. Millan, C., Sammito, M., and Uson, I. (2015) Macromolecular *ab initio* phasing enforcing secondary and tertiary structure. *IUCrJ* **2**, 95-105
95. Proper, K., Meindl, K., Sammito, M., Dittrich, B., Sheldrick, G. M., Pohl, E., and Uson, I. (2014) Structure solution of DNA-binding proteins and complexes with ARCIMBOLDO libraries. *Acta. Crystallogr. D Biol. Crystallogr.* **70**, 1743-1757
96. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta. Crystallogr. D Biol. Crystallogr.* **60**, 2126-2132
97. Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta. Crystallogr. D Biol. Crystallogr.* **67**, 355-367
98. Uson, I., Pohl, E., Schneider, T. R., Dauter, Z., Schmidt, A., Fritz, H.-J., and Sheldrick, G. M. (1999) 1.7 Å structure of the stabilized REIv mutant T39K. Application of local NCS restraints. *Acta. Crystallogr. D Biol. Crystallogr.* **55**, 1158-1167
99. Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., and Zwart, P.H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta. Crystallogr. D Biol. Crystallogr.* **66**, 213-221
100. Lovell, S.C., Davis, I.W., Arendall, W.B. 3rd, de Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S., Richardson, D.C. (2003) Structure validation by C $\alpha$  geometry: phi,psi and C $\beta$  deviation. *Proteins* **50**, 437-450
101. Henard, C. A., Bourret, T. J., Song, M., and Vázquez-Torres, A. (2010) Control of redox balance by the stringent response regulatory protein promotes antioxidant defenses of *Salmonella*. *J. Biol. Chem.* **285**, 36785-36793
102. Song, M., Husain, M., Jones-Carson, J., Liu, L., Henard, C. A., and Vázquez-Torres, A. (2013) Low-molecular-weight thiol-dependent antioxidant and antinitrosative defences in *Salmonella* pathogenesis. *Mol. Microbiol.* **87**, 609-622
103. Pettersen E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605-1612

## FOOTNOTES

<sup>1</sup>Both authors contributed equally to the experimentation.

<sup>2</sup>The abbreviations used are: MNIC, maximum non-inhibitory concentrations; GSH, reduced glutathione; GSSG, oxidised glutathione; ICP-MS, inductively-coupled plasma mass spectrometry; TMAO, trimethylamine-*N*-oxide; MRM-MS, multiple reaction monitoring-mass spectrometry.

<sup>3</sup>The atomic coordinates and structural factors for the FrmR<sup>E64H</sup> crystal structure have been deposited in the Worldwide Protein Data Bank (<http://www.wwpdb.org/>) under PDB # 5LCY.

## FIGURE LEGENDS

**FIGURE 1. Formaldehyde detoxification and sensing**

Spontaneous reaction of formaldehyde with the reduced form of glutathione (GSH) generates S-(hydroxymethyl)glutathione (S-HMG), the substrate oxidised by FrmA to S-formylglutathione (28). Following hydrolysis of S-formylglutathione by FrmB, formate is produced as the final product and GSH is regenerated (22,29). An additional enzyme, YeiG, is also implicated in formaldehyde detoxification as it demonstrates hydrolytic activity against S-formylglutathione, although *yeiG* is not FrmR-regulated (27). Notably, FrmB is present in the *E. coli* but not the *Salmonella frm* operon. *Salmonella* FrmR (or FrmR<sup>E64H</sup>) represses expression from the *frm* promoter which is alleviated by exogenous formaldehyde. The intracellular effector of (any) FrmR is unknown and possibilities include formaldehyde alone (1) or S-(hydroxymethyl)glutathione (2), in which case GSH could act negatively (3) or positively (4) on FrmR-mediated de-repression. Alternatively, FrmR de-repression may be transduced by a metal intermediate (5), or require activation of formaldehyde by metal (6). FrmR<sup>E64H</sup> additionally responds to Zn(II) and cobalt, however the response to metals is lost in cells lacking glutathione (7) (50).

**FIGURE 2. *E. coli* and *Salmonella* FrmRs have distinct origins but *Salmonella* FrmR and FrmR<sup>E64H</sup> retain their effector-sensitivities in *E. coli***

A, Schematic representation of the *frm* operon (to scale) from *E. coli* K12 (blue) and *Salmonella* (strain SL1344, red), with the nucleotide position at the start and end of each gene cluster indicated. B, Percent identity of *Salmonella* proteins required for formaldehyde detoxification compared to their *E. coli* orthologue. C, Rooted phylogenetic tree of eleven sequences from the DUF156 FrmR subgroup identified previously (45). Organism details and UniProtKB identifiers are outlined in *Experimental Procedures*. D, Alignment of the *frm* promoter from *E. coli* and *Salmonella*. The position relative to the translational start site is labelled. G/C tracts are underlined. T/A-rich inverted repeats are highlighted in grey. The sequence corresponding to one strand of *frmRAPro*, used for fluorescence anisotropy, is in bold. Mutations to generate *frmRAPro*\* are highlighted in red. E and F,  $\beta$ -galactosidase activity as a function of time in *E. coli* BW25113  $\Delta$ *frmR* containing P<sub>*frmRA*</sub>-*frmR* (E) or P<sub>*frmRA*</sub>-*frmR*<sup>E64H</sup> (F), fused to *lacZ* following exposure of logarithmic cells to MNIC formaldehyde (50  $\mu$ M; diamonds), Zn(II) (50  $\mu$ M; triangles), Co(II) (5  $\mu$ M; squares), or untreated control (circles). Values are means of at least three biological replicates (each performed in triplicate) with standard deviation.

**FIGURE 3. FrmR responds specifically to formaldehyde *in vitro* and *in vivo***

A, Anisotropy change upon titration of a limiting concentration of *frmRAPro* (10 nM) with FrmR in the presence of 5 mM EDTA and either 20  $\mu$ M acetaldehyde (black symbols), 10  $\mu$ M formaldehyde (grey symbols), or 20  $\mu$ M formaldehyde (open symbols). Symbol shapes represent individual experiments. Data were fit to a model describing a 2:1 protein tetramer (nondissociable):DNA stoichiometry (binding with equal affinity) (50,86), and lines represent simulated curves produced from the average (apparent)  $K_{DNA}$  determined across the experimental replicates shown. The dashed red line (largely obscured) is a simulated curve based on the published  $K_{DNA}$  of apo-FrmR (50), presented here for comparative purposes. B,  $\beta$ -galactosidase activity in SL1344 containing P<sub>*frmRA*</sub>-*frmR* fused to *lacZ* grown to mid-exponential phase in M9 minimal medium in the absence (control) or presence of MNIC indicated alcohol or aldehyde (see *Experimental Procedures* for concentrations). Values are means of three biological replicates (each performed in triplicate) with standard deviation.

**FIGURE 4. Structure of FrmR<sup>E64H</sup> and inferred Zn(II)/Co(II)-sensory site**

A, Ribbon representation of the 2.19 Å resolution crystal structure of FrmR<sup>E64H</sup> tetramer (PDB code: 5LCY; see Table 1 for a summary of the crystallographic data). Each monomer is colored differently and secondary structural units are labeled on the cyan monomer. B, Electrostatic surface potential of FrmR<sup>E64H</sup> tetramer using Chimera (103). The color scale is from -10 (negative potential; red) to +10 (positive

potential; *blue*) kcal/mol·e. *C*, Anisotropy change upon titration of a limiting concentration (10 nM) of *frmRAPro* (*solid symbols*) or *frmRAPro\** (half site defined in Fig. 2D; *open symbols*,) with FrmR (*circles*) or FrmR<sup>E64H</sup> (*triangles*) in the presence of 5 mM EDTA. The *lines* are fits of the data to a model describing a 2:1 protein tetramer (nondissociable):DNA stoichiometry (binding with equal affinity) (50,86). *D*, Expansion of the dimeric interface with backbone helices from two different monomers shaded *green* and *cyan* (the same colors as used in *A*). The inferred Zn(II)/Co(II)-binding site comprises Cys<sup>35</sup> from  $\alpha 2'$ , and His<sup>60</sup> and His<sup>64</sup> from  $\alpha 2$  (belonging to the XYZ motif required for metal-binding in DUF156 members CsoR, RcnR and InrS (39,46,68), with His<sup>3</sup> from  $\alpha 1$  (position W (46, 61)) and Asp<sup>63</sup> presenting candidate fourth ligands. *E*, Analysis of fractions (0.5 ml) for protein by Bradford assay (*open circles*) and metal by ICP-MS (*filled circles*) following size-exclusion chromatography of FrmR, FrmR<sup>E64H</sup>, FrmR<sup>C35A</sup> (50  $\mu$ M, monomer) or FrmR<sup>H60L</sup> (in this case [monomer] = 32.5  $\mu$ M), preincubated with 150  $\mu$ M ZnCl<sub>2</sub>.

### FIGURE 5. Conservation of residues in the DUF156 FrmR subgroup and proposed formaldehyde-sensory site

*A*, Alignment of *Salmonella* FrmR with nonredundant UniProtKB DUF156 sequences previously attributed to the FrmR subgroup (45). Organism details and UniProtKB identifiers are outlined in *Experimental Procedures*. Highlighted in *grey* are residues conserved in both FrmR and RcnR subgroups. Highlighted in *red* are residues conserved in the FrmR, but not RcnR subgroup. Highlighted in *yellow* is the invariant cysteine present in all DUF156 proteins. The secondary structure elements of the FrmR<sup>E64H</sup> crystal structure are shown below (*black bars*). The inferred Zn(II)/Co(II)-sensory site is identified by *orange arrows*. The proposed formaldehyde sensing site is identified by *green arrows*. *B* and *C*, Dimeric representation of FrmR<sup>E64H</sup> (rotated relative Fig. 4A) with the side-chains for Cys<sup>35</sup> and FrmR-subgroup specific residues *labelled*. Each monomer is *colored* differently (using the same colors as in 3A) with secondary structure units *labelled* on the *cyan* subunit. *D*, Solvent-accessible surface representation of the proposed formaldehyde-binding site which comprises Pro<sup>2</sup> (subunit 1, *cyan*) and Cys<sup>35</sup> (subunit 2, *green*). *E*, Proposed reaction of formaldehyde with FrmR Cys<sup>35</sup> (*green*) followed by Pro<sup>2</sup> (*cyan*) (both deprotonated ultimately to water) via an *S*-hydroxymethyl intermediate. The reciprocal reaction with Pro<sup>2</sup> followed by Cys<sup>35</sup> via an *N*-methylol intermediate is also possible. In both cases, a methylene bridge (*black*) between the two residues is the final product. The nucleophile(s) responsible for deprotonation of Cys<sup>35</sup> and Pro<sup>2</sup> remain unknown.

### FIGURE 6. Pro<sup>2</sup> and Cys<sup>35</sup> are required for formaldehyde sensing by FrmR

*A*, LC-MS chromatogram following multiple reaction monitoring of purified FrmR. Ion transition 405.19/488.24 is for analyte PHSPED. *B*,  $\beta$ -galactosidase activity in SL1344 containing P<sub>*frmRA*</sub>-*frmR* (*circles*), P<sub>*frmRA*</sub>-*frmR*<sup>C35A</sup> (*triangles*) or P<sub>*frmRA*</sub>-*frmR*<sup>P2S</sup> (*squares*) fused to *lacZ* grown to mid-exponential phase in M9 minimal medium in the presence of formaldehyde (MNIC = 50  $\mu$ M). Values are means of three biological replicates (each performed in triplicate) with standard deviation. *C* and *D*, Anisotropy change upon titration of a limiting concentration of *frmRAPro* (10 nM) with FrmR<sup>C35A</sup> (*C*) or FrmR<sup>P2S</sup> (*D*) in the presence of 5 mM EDTA (*closed symbols*) and with addition of 20  $\mu$ M formaldehyde (*open symbols*). Data were fit to a model describing a 2:1 protein tetramer (nondissociable):DNA stoichiometry (binding with equal affinity) (50,86), and *lines* represent simulated curves produced from the average (apparent)  $K_{DNA}$  determined across the experimental replicates shown. *Symbol shapes* represent individual experiments. *E*, Apparent  $K_{DNA}$  values of FrmR (*black symbols*), FrmR<sup>P2S</sup> (*grey symbols*), and FrmR<sup>C35A</sup> (*open symbols*) with increasing formaldehyde concentration. Values are means of three replicates with standard deviation.

### FIGURE 7. RcnR is less formaldehyde responsive but RcnR<sup>S2P</sup> gains reactivity

*A*,  $\beta$ -galactosidase activity in SL1344 containing *rcnR*-P<sub>*rcnRA*</sub> (*solid bars*) or *rcnR*<sup>S2P</sup>-P<sub>*rcnRA*</sub> (*open bars*) fused to *lacZ* grown to mid-exponential phase in M9 minimal medium in the absence or presence of MNIC formaldehyde (50  $\mu$ M) or CoCl<sub>2</sub> (1  $\mu$ M). Values are means of three biological replicates (each

performed in triplicate) with standard deviation. *B*, Anisotropy change upon titration of a high concentration of *rcnRAPro* (2.5  $\mu$ M) with RcnR in the presence of 5 mM EDTA, or *C*, a limiting concentration of *rcnRAPro* (10 nM) with RcnR in the presence of 5 mM EDTA (*black symbols*) and with addition of 50  $\mu$ M formaldehyde (*open symbols*), or Ni(II)-RcnR (*light grey symbols*) and Co(II)-RcnR (*dark grey symbols*) in the presence of 5  $\mu$ M NiCl<sub>2</sub> or 5  $\mu$ M CoCl<sub>2</sub>, respectively. *Symbol shapes* represent individual experiments. Data were fit to a model describing a 2:1 protein tetramer (nondissociable):DNA stoichiometry (binding with equal affinity) (50,86), and *lines* represent simulated curves produced from the average (apparent)  $K_{DNA}$  determined across the experimental replicates shown. *D*, Apparent  $K_{DNA}$  values of RcnR (*black symbols*) and RcnR<sup>S2P</sup> (*open symbols*) with increasing formaldehyde concentration. Values are means of three replicates with standard deviation. *E*, as described in ‘*C*’ except with RcnR<sup>S2P</sup> in the presence of 5 mM EDTA (*black symbols*) and with addition of 50  $\mu$ M formaldehyde (*open symbols*).

### FIGURE 8. The relationship between glutathione and formaldehyde-sensing by FrmR

*A*, Survival of wild type *Salmonella* SL1344 (*solid circles*) or  $\Delta$ *gshA* (*open circles*) grown to mid-exponential phase in M9 minimal medium in the presence of formaldehyde. Values are means of three biological replicates (each performed in triplicate) with standard deviation. *B*,  $\beta$ -galactosidase activity of SL1344 (*solid symbols*) or  $\Delta$ *gshA* (*open symbols*) containing P<sub>*frmRA*</sub>-*frmR* fused to *lacZ* grown to mid-exponential phase in M9 minimal medium in the presence of formaldehyde (MNIC = 50 and 20  $\mu$ M for wildtype and  $\Delta$ *gshA*, respectively; see supplemental Fig. S2 for corresponding growth data). Values are means of at least three biological replicates (each performed in triplicate) with standard deviation. *C*, Anisotropy change upon titration of a limiting concentration of *frmRAPro* (10 nM) with FrmR in the presence of 5 mM EDTA and 800  $\mu$ M reduced glutathione (GSH), in the absence (*grey symbols*) or presence (*open symbols*) of 20  $\mu$ M formaldehyde. *Symbol shapes* represent individual experiments. Data were fit to a model describing a 2:1 protein tetramer (nondissociable):DNA stoichiometry (binding with equal affinity) (50,86), and *lines* represent simulated curves produced from the average (apparent)  $K_{DNA}$  determined across the experimental replicates shown. *D*, Intracellular glutathione concentration in *Salmonella* cells following growth to exponential phase in M9 minimal medium aerobically (O<sub>2</sub>) or anaerobically with TMAO as an alternative electron acceptor. Values are means of three biological replicates with standard deviation. *E*, Representative (n=3) LC-MS chromatograms of ion transitions detected in mid-logarithmic *Salmonella* SL1344 cells under aerobic growth conditions. Transitions are for analyte GQVEALER (*solid black line*) or labelled internal standard (IS) GQVEALER[<sup>13</sup>C<sub>6</sub>, <sup>15</sup>N<sub>4</sub>] (<sup>13</sup>C, <sup>15</sup>N-Arg residue) (*dashed grey line*). *F*, Fractional modification by formaldehyde of FrmR (*solid black line*), GSH (*solid grey line*), or FrmR (*dashed red line*) and RcnR (*dashed blue line*; tighter-limit as indicated by the *blue arrow*) in the presence of GSH in *Salmonella* cells grown anaerobically with TMAO. Formaldehyde affinities of 10<sup>-5</sup>, 10<sup>-4</sup> (tighter-limit) and 1.77  $\times$  10<sup>-3</sup> M (73) were used for FrmR, RcnR and GSH, respectively. Intracellular abundance was determined for FrmR (16.1  $\pm$  0.2 nM) and GSH (1.2  $\pm$  0.4 mM) and estimated for RcnR as described in *Experimental Procedures*. *G*, The role of glutathione in formaldehyde detoxification and sensing in *Salmonella*: In the absence of effector, *Salmonella* FrmR represses the *frm* promoter. Formaldehyde directly modifies FrmR (reaction ① in Fig. 1) via a deduced inter-subunit methylene bridge between Pro<sup>2</sup> and Cys<sup>35</sup> (Fig. 5, up to four per tetramer) de-repressing *frm* expression. GSH inhibits formaldehyde detection (③ in Fig. 1) and despite the high [glutathione], the affinity of FrmR for formaldehyde is sufficiently tight relative to GSH to enable expression of FrmA to coincide with the appearance of its substrate. The *Salmonella frm* operon lacks *frmB*, and YeiG may catalyse the final detoxification step. S-HMG, S-(hydroxymethyl)glutathione; GSH, reduced glutathione; GSSG, oxidised glutathione.

**TABLE 1**  
**Crystallographic data collection and refinement**  
**statistics**

	<b>FrmR<sup>E64H</sup></b>
<b>Data collection</b>	
Beam line	I03
Wavelength (Å)	0.9762
Space group	P2 <sub>1</sub>
Cell dimensions	
<i>a, b, c</i> (Å)	68.79, 25.68, 100.50
$\alpha, \beta, \gamma$ (°)	90, 103.1, 90
Resolution (Å)	29.7-2.19
<i>R</i> <sub>merge</sub>	0.081 (0.495) <sup>a</sup>
<i>I</i> / $\sigma$ <i>I</i>	12.1 (3.2) <sup>a</sup>
Multiplicity	6.3 (6.4) <sup>a</sup>
Completeness (%)	99.3
Wilson B-factor <sup>b</sup>	33
<b>Refinement</b>	
Resolution (Å)	29.7-2.19
No. reflections	17990
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.24/0.31
No. of atoms	
Protein	2767
Water	103
<i>B</i> -factors (Å <sup>2</sup> )	
Chain A	51
Chain B	47
Chain C	47
Chain D	53
Protein	30
Water	48
Root mean square deviations	
Bond lengths (Å)	0.01
Bond angles (°)	1.1

<sup>a</sup>The values in parentheses refer to the last resolution shell (2.25-2.19 Å).

<sup>b</sup>Calculated using phenix.xtriage (99)

**TABLE 2**  
**DNA binding affinities and allosteric coupling free energies for FrmR and RcnR**

Values were determined from fluorescence anisotropy experiments. The conditions used were as follows: 25 °C, 10 mM HEPES, pH 7.0, 60 mM NaCl, 240 mM KCl with addition of 5 mM EDTA for titrations with apo-protein, or 5  $\mu$ M NiCl<sub>2</sub> or CoCl<sub>2</sub> for metal-loaded titrations. RcnR was incubated with 1.2 molar eq of NiCl<sub>2</sub> or CoCl<sub>2</sub> per monomer as indicated.

Protein	Effector	$K_{DNA}^a$
		<i>M</i>
<b>FrmR</b>	Apo <sup>b</sup>	$9.9 \pm 0.3 \times 10^{-8}$
<b>FrmR<sup>C35A</sup></b>	Apo	$1.6 \pm 0.2 \times 10^{-7}$
<b>FrmR<sup>P2S</sup></b>	Apo	$1.5 \pm 0.2 \times 10^{-7}$
<b>RcnR</b>	Apo	$1.5 \pm 0.8 \times 10^{-7}$
	Ni(II)	$\geq 5.9 \pm 1.3 \times 10^{-6}$
	Co(II)	$\geq 1.5 \pm 0.2 \times 10^{-5}$
<b>RcnR<sup>S2P</sup></b>	Apo	$1.6 \pm 0.1 \times 10^{-7}$

<sup>a</sup>Data were fit to a model describing two nondissociable tetramers of FrmR or RcnR binding with equal affinity to *frmRAPro* or *rcnRAPro*, respectively.

<sup>b</sup>Determined previously under the same conditions (50).

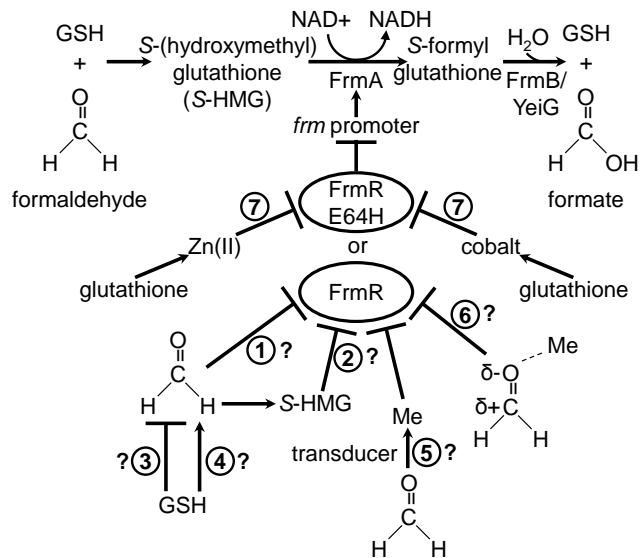


FIGURE 1



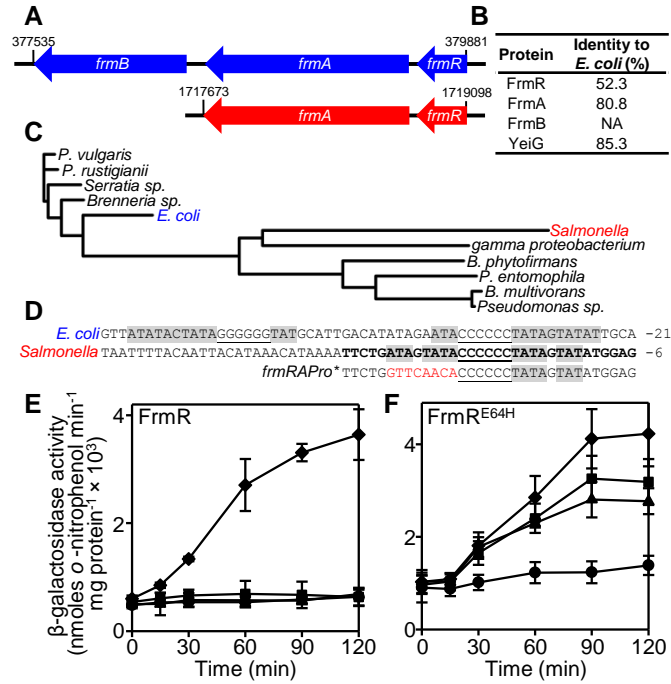


FIGURE 2

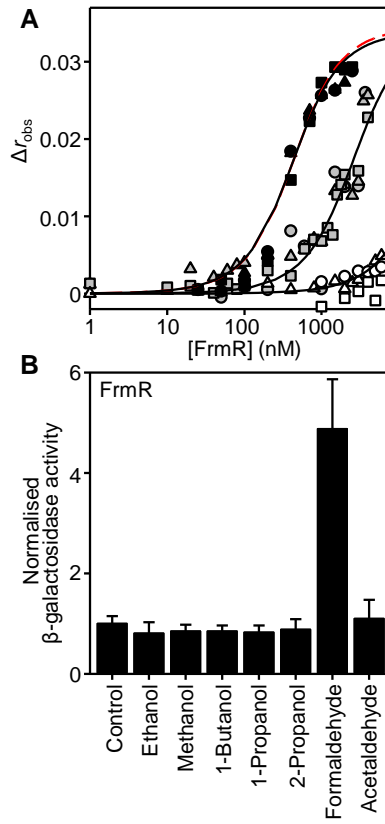


FIGURE 3

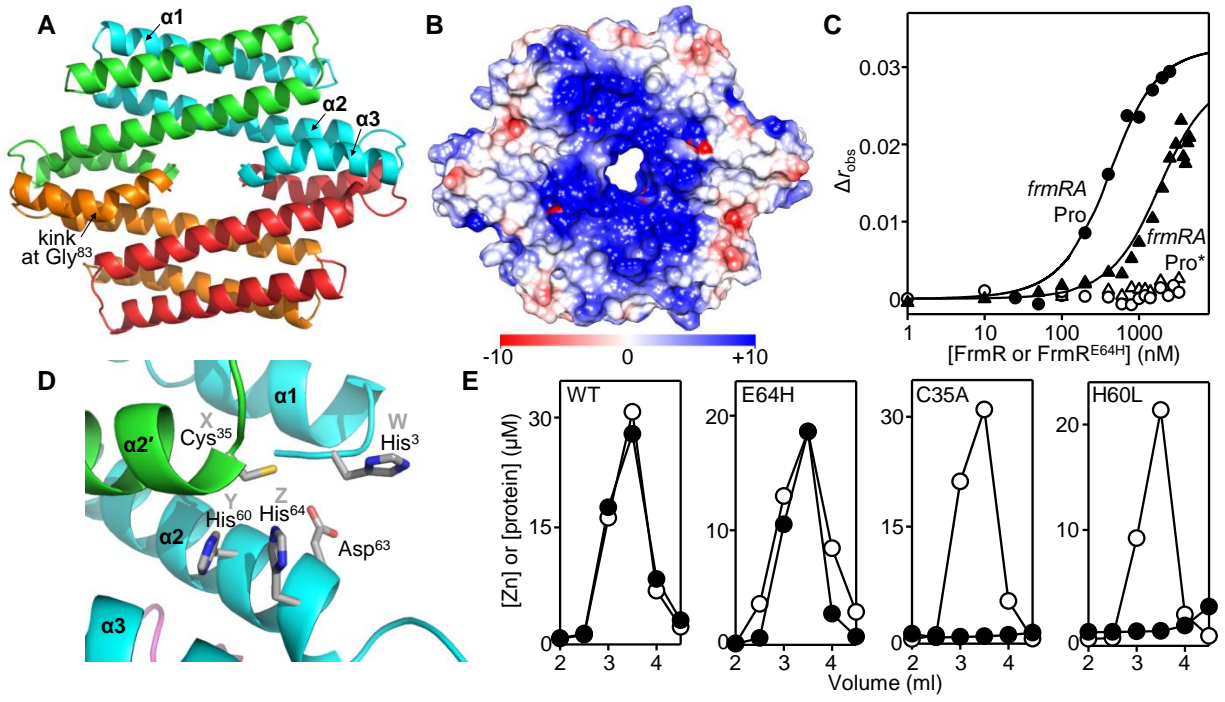


FIGURE 4

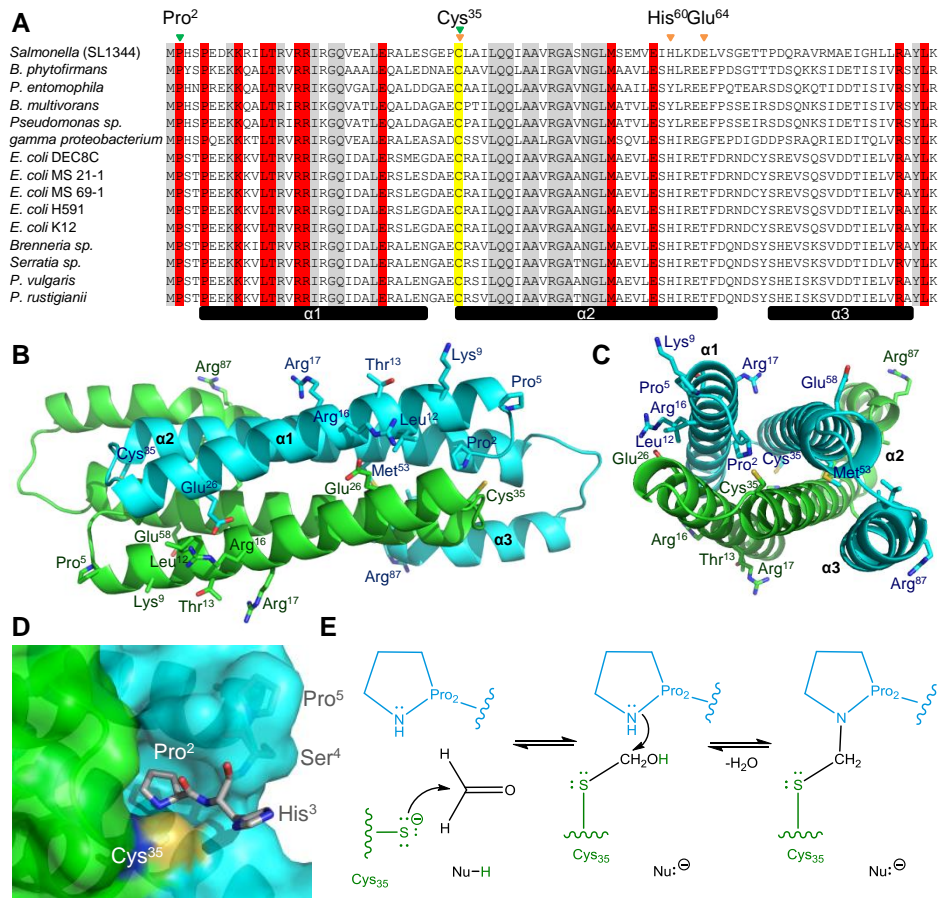


FIGURE 5

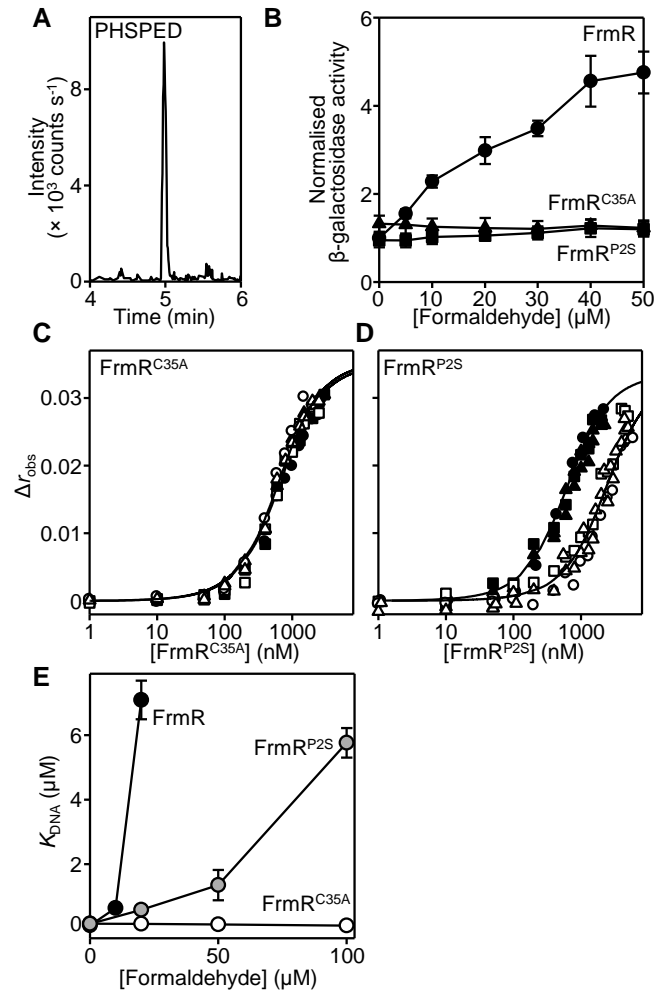


FIGURE 6

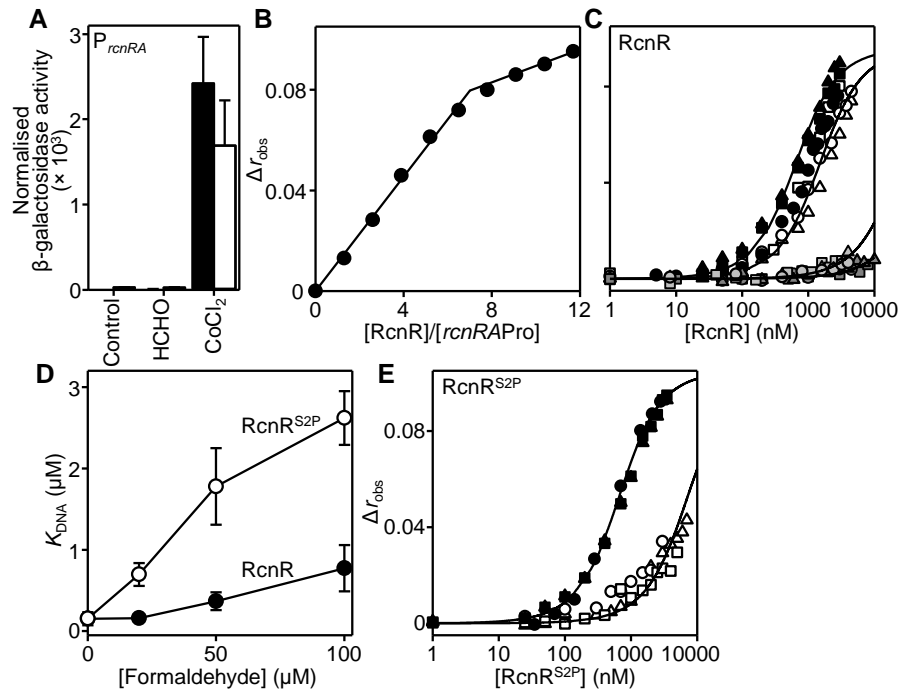


FIGURE 7

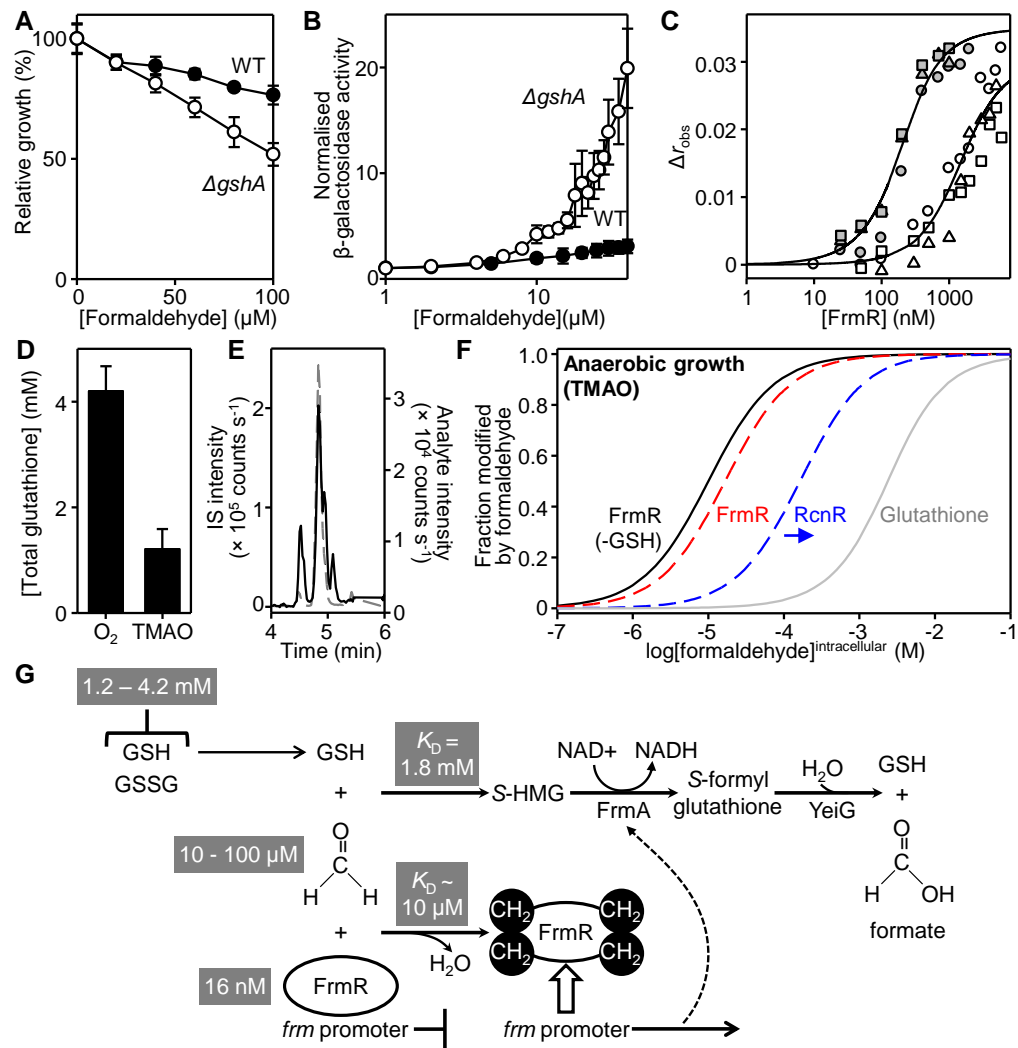


FIGURE 8