Similar solutions to a common challenge: Regulation of genes encoding *Ralstonia solanacearum* xanthine dehydrogenase

Smitha Sivapragasam,[1] Arpita Ghosh, Sanjay Kumar, Danté T. Johnson,[2] Anne Grove*

Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803, USA

*To whom correspondence should be addressed: Anne Grove, Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803. Telephone: 225-578-5148; Fax: 225-578-8790; E-mail: agrove@lsu.edu

[1]Current Address: School of Molecular Biosciences, Washington State University, Pullman, WA 99164

[2]Current Address: Department of Pharmaceutical Sciences, University of Maryland Baltimore, School of Pharmacy, Baltimore, MD 21201

Key words: *Ralstonia solanacearum*; xanthine dehydrogenase; purine salvage; (p)ppGpp; gene regulation; LTTR

Abstract

The stringent response involves accumulation of (p)ppGpp, and it ensures that survival is prioritized. Production of (p)ppGpp requires purine synthesis, and upregulation of an operon that encodes the purine salvage enzyme xanthine dehydrogenase (Xdh) has been observed during stringent response in some bacterial species, where direct binding of ppGpp to a TetR-family transcription factor is responsible for increased *xdh* gene expression. We show here that the plant pathogen *Ralstonia solanacearum* has a regulatory system in which the LysR-family transcription factor XanR controls expression of the *xan* operon; this operon encodes Xdh as well as other enzymes involved in purine salvage, which favor accumulation of xanthine. XanR bound upstream of the *xan* operon, a binding that was attenuated on addition of either ppGpp or cyclic di-guanosine monophosphate (c-di-GMP). Using a reporter in which enhanced green fluorescent protein (EGFP) is expressed under control of a modified *xan* promoter, XanR was shown to repress EGFP production. Our data suggest that *R. solanacearum* features a regulatory mechanism in which expression of genes encoding purine salvage enzymes is controlled by a transcription factor that belongs to a different protein family, yet performs similar regulatory functions.

Introduction

Ralstonia solanacearum is a soil-borne β-proteobacterium, which causes wilt disease in important crop plants (Genin and Denny 2012, Mansfield, et al. 2012). It typically enters susceptible hosts through a wound and colonizes the water-conducting xylem tissue, which is a nutrient-poor environment. A number of virulence genes are expressed in response to plant-derived signals, such as the reactive oxygen species that are produced as part of the plant defense (Flores-Cruz and Allen 2009, Jacobs, et al. 2012, Mandal, et al. 2011). Several bacterial genes that are required for the oxidative stress response are upregulated under such conditions, along with genes encoding proteins with less obvious roles in stress responses or virulence. One of the genes that were induced both by addition of hydrogen peroxide to *R. solanacearum* grown in liquid culture and during growth *in planta* is *xdhA*, encoding a subunit of xanthine dehydrogenase (Xdh; Figure 1A) (Flores-Cruz and Allen 2009). A screening of a promoter-fusion library also identified the *xdhA* promoter as being active in a host environment (Brown and Allen 2004).

Xdh participates in purine salvage where it converts hypoxanthine to xanthine, thereby favoring the conversion of adenine to guanine (Figure 1B). Xdh also initiates purine degradation by converting xanthine to urate, which is ultimately degraded to ammonia and CO_2 (Xi, et al. 2000). *R. solanacearum* encodes the requisite enzymes for complete purine degradation and it can use xanthine as a nitrogen source (Lundgren, et al. 2015). The Xdh components are encoded in a predicted operon that also includes genes coding for adenine deaminase (Ada) and guanine deaminase (Gda).

In several bacteria, enzymes involved in purine biosynthesis or salvage have been shown to play a vital role in virulence (Eastgate, et al. 1997, Jewett, et al. 2009, Liechti and

Goldberg 2012). The reason appears to be the requirement for synthesis of phosphorylated guanosine-derived signaling molecules. During stress or nutrient limitation, guanosine tetra- and pentaphosphate, (p)ppGpp, is synthesized from GTP and ATP (Gaca, et al. 2015). (p)ppGpp is a global regulator, and it inhibits cellular functions associated with growth to conserve resources and promote survival (Gaca, et al. 2015, Liu, et al. 2015, Ross, et al. 2013, Traxler, et al. 2011). That purine salvage is important for synthesis of (p)ppGpp is also reflected in the observation that a positive feed-back loop exists in both *Streptomyces coelicolor* and *Agrobacterium fabrum*. In these species, the *xdh* operons encoding components of Xdh are repressed by the TetR family transcription factor, XdhR, during balanced growth. During stringent response, expression of the respective *xdh* operons is induced by direct binding of (p)ppGpp to XdhR (Sivapragasam, et al. 2017, Sivapragasam and Grove 2016, Sivapragasam and Grove 2019).

The second messenger c-di-GMP is synthesized from two molecules of GTP, and it is involved in multiple cellular functions, including biofilm formation, motility, and virulence (Jenal, et al. 2017, Martinez-Gil and Ramos 2018). Synthesis of c-di-GMP has also been shown to depend on purine biosynthesis (Kim, et al. 2014).

R. solanacearum Xdh is encoded as part of an operon, which is divergently oriented from a gene encoding a member of the LysR family of transcriptional regulators (Figure 1A). For reference, *S. coelicolor* and *A. fabrum* encode the TetR-family transcription factor XdhR divergently from an operon that encodes four proteins (Figure 1C): The chaperone XdhC, which is required for assembly and maturation of Xdh, and the three Xdh subunits (denoted XdhSML for Small, Medium, and Large) that are assembled into a heterohexameric $(\alpha\beta\gamma)_2$ complex. The *R. solanacearum*-encoded transcription factor is

structurally unrelated to *S. coelicolor* and *A. fabrum* XdhR, and the operon is predicted to encode five proteins: The chaperone XdhC, two Xdh subunits that are assembled into a heterotetrameric $(\alpha\beta)_2$ complex, adenine deaminase (Ada), and guanine deaminase (Gda). The combined action of these enzymes leads to formation of xanthine, which may be salvaged for synthesis of phosphorylated guanosine derivatives, including (p)ppGpp and c-di-GMP. Since the operon includes additional genes encoding proteins involved in purine salvage (favoring production of xanthine), we have named this transcription factor XanR and we refer to the operon as the *xan* operon. This locus is conserved in other members of the genus *Ralstonia* as well as in *Cupriavidus* species. Using a construct comprising the gene encoding *C. necator* XanR divergent from a reporter under control of the *xan* promoter, derepression of the *xan* promoter was observed on addition of xanthine, suggesting that XanR represses the *xan* promoter and that xanthine (or a metabolite thereof) relieves this repression (Hanko, et al. 2020). We show here that DNA binding by XanR is attenuated by ppGpp and c-di-GMP *in vitro*, and that XanR functions to repress gene expression *in vivo*.

Materials and Methods

Protein modeling and ligand docking

A model of XanR was generated using Phyre2 in intensive mode (Kelley, et al. 2015). Three templates were used to generate the final model representing an extended conformation, 2ESNC (27% overall template identity for residues 4-301; 100% confidence; unpublished structure of LTTR from *Pseudomonas aeruginosa*), 5Y2VA (17% identity for residues 2-307; 100% confidence; NdhR from *Synechocystis* PCC6803 (Jiang, et al. 2018)), and 5AE5A (31% identity for residues 3-303; 100% confidence; DntR from *Burkholderia*

cepacia (Lerche, et al. 2016)). Overall, 97% of residues were modeled, with only 8 residues modeled ab initio (residues at the C-terminus). A model based on 2ESNC was used to represent the compact conformation. A tetrameric assembly was created in PyMol based on 2ESN. The structure of ligands was retrieved from PubChem and docked to the final XanR model using EDock (Zhang, et al. 2020). The models with the highest XSCORE, which is an empirical scoring function reflecting binding affinity, were selected.

Cloning, purification, and characterization of XanR

Primers RalxanR_FP and RalxanR_RP (Table S1) were used to amplify the *R. solanacearum xanR* (*Rsc2094*; new annotation *RS_RS10505*) gene using genomic DNA from *R. solanacearum* GM1000 as template. The PCR product was cloned into pET100, and the recombinant plasmid (pET_xanR) was sequenced and transformed into *E. coli* BL21(DE3). Cells were grown in LB containing ampicillin (50 μg mL^{-1}) to OD_{600} 0.4-0.6 and induced with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) for 5 hours. Cells were resuspended in lysis buffer (20 mM sodium phosphate pH 8.0, 150 mM sodium chloride, 5% (v/v) glycerol, 0.15 mM phenylmethylsulfonyl fluoride (PMSF) and 10 mM β-mercaptoethanol) to which lysozyme was added to 200 μg mL^{-1} and the cell suspension was incubated on ice for 30 minutes. Cells were lysed further by sonication and the lysate was centrifuged at 13,225 ×g for 1 hour. The supernatant was mixed with HIS-select Nickel affinity beads (MilliporeSigma, Burlington, MA) for 30 minutes at 4 °C. XanR was purified according to the manufacturer's protocol. The purified protein was stored in storage buffer (20 mM sodium phosphate pH 8.0, 150 mM NaCl, 20% (v/v) glycerol, 0.15 mM PMSF and

10 mM β-mercaptoethanol). Purity was verified by SDS-PAGE followed by staining with Coomassie Brilliant Blue. Protein concentration was determined by MicroBCA kit (BioRad).

Purified His_6-tagged XanR (8.8 μM) was incubated with 1 μL of 0.005%, 0.01% or 0.05% (v/v) glutaraldehyde on ice for 20 minutes. After incubation, the reaction was stopped with Laemmli sample buffer without β-mercaptoethanol and loaded onto a denaturing SDS-PAGE gel. Gels were stained with Coomassie Brilliant Blue.

Thermal stability of XanR was determined by differential scanning fluorimetry, as described (Grove, et al. 2015) using 5 μM protein in assay buffer (50 mM Tris, pH 7.4, 100 mM NaCl, 5× SYPRO Orange dye (Invitrogen)). The Tm is reported as the mean ± SD of triplicate measurements.

Electrophoretic mobility shift assay (EMSA)

The intergenic DNA between *xanR* and the *xan* operon (242 bp amplicon, which also includes 38 bp of the *xanR* coding sequence) was amplified using interFP 5'-CGCAGGAGGTAGGTATCGAG-3' and interRP 5'-CGCCTGTCTCGTGTCTGTAA-3' and *R. solanacearum* GM1000 genomic DNA as template. The DNA was labeled at the 5'-ends with ^{32}P using T4 polynucleotide kinase. Labeled DNA (0.05 nM) was incubated with XanR in binding buffer (25 mM Tris (pH 8.0), 100 mM NaCl, 1 mM EDTA, 0.05% (v/v) Brij58, 5 mM DTT and 2% (v/v) glycerol) for 30 minutes and loaded onto a 4% or 6% native PAGE gel that was prerun for 30 minutes in 0.5X TBE (45 mM Tris-borate, 1 mM EDTA). After electrophoresis, gels were dried, exposed to a phosphor screen, and scanned using a Storm scanner (GE Healthcare). The bands were quantified using ImageQuant; all complexes were included as bound DNA, along with the region on the gel between complex and free DNA.

The data were plotted using KaleidaGraph. K_d was calculated using the Hill equation $f = f_{max}*[X]^{nH} / (K_d^{nH} + [X]^{nH})$ where f is fractional saturation, n_H is the Hill coefficient, K_d refers to apparent equilibrium dissociation constant reflecting half-maximal saturation of the DNA, and X is the protein concentration. Experiments were repeated twice, and the results are reported as mean ± SD. Specificity of XanR binding was assessed by titrating binding reactions with increasing concentrations of polydGdC (10 pg – 1 ng).

To identify the effect of oxidants on DNA binding, hydrogen peroxide, cumene hydroperoxide, or tertiary butyl hydroperoxide were incubated with XanR for 15 minutes. The oxidized protein was then incubated with labeled DNA for 15 minutes and processed as described.

GMP, GTP (MilliporeSigma, Burlington, MA), ppGpp (TriLink Bio Technologies, San Diego, CA) and c-di-GMP (Kerafast, Boston, MA) were used as ligands in EMSA. Ligands were added to DNA, and XanR was added last. After 15 minutes, reaction mixtures were electrophoresed, and fractional binding was quantified as described. IC_{50} was calculated using $f = A + B \times e^{-kL}$ where f is fractional saturation, k is the decay constant, L is ligand concentration, A is the saturation plateau and B is the decay amplitude. The corresponding inhibition constant, Ki, was calculated according to $Ki = IC_{50}/([DNA]_{50}/Kd+[X]_0/Kd+1)$ where $[DNA]_{50}$ is the concentration of DNA at 50% inhibition and $[X]_0$ is the protein concentration at 0% inhibition (Cer, et al. 2009). At least two independent experiments were performed.

DNase I footprinting

A 247 bp DNA fragment encompassing the *xanR-xan* intergenic region and 43 bp of *xanR* coding sequence was obtained by PCR amplification using a forward primer (5'-GGACGCGCAGGAGGTAGGTATCG-3') that was labeled at its 5'-end with carboxyfluorescein (6-FAM) and reverse primer interRP and *R. solanacearum* GM1000 genomic DNA as template. The DNA (50 ng) was incubated with XanR for 10 minutes at room temperature in EMSA binding buffer, at which time the reaction was supplemented with 1× DNase I reaction buffer (New England BioLabs). DNase I (0.02 units) was added, reactions were incubated at room temperature for 3.5 minutes and then stopped by addition of 8 mM Na_2EDTA. The digested DNA was purified and dissolved in 10 µL Hi-Di formamide and subjected to fragment analysis using an ABI 3130 analyzer after addition of LIZ 500 ladder (ABI – Life Technologies). Microsatellite Analysis Software (ThermoFisher) was used for data analysis. Sequences were identified by comparison with a sequencing reaction, as described (Sivapragasam, et al. 2015). Footprints are representative of at least three replicates.

Regulation of reporter construct

The gene encoding a destabilized version of enhanced green fluorescent protein (d1EGFP) was cloned into plasmid pACYC184 (Deochand, et al. 2016). A DNA construct was synthesized, which contained a consensus *E. coli* promoter (TTGACAtatctacctaatttgtgTATAAT; -35 and -10 in capital letters (O'Neill 1989)) followed by the intergenic sequence spanning *xanR* and the *xan* operon. This DNA fragment was flanked by NruI and BamHI sites and cloned upstream of *d1egfp*, generating pACYC_xEGFP.

Plasmids pACYC_xEGFP and pET_xanR, harboring the *xanR* gene, were co-transformed into *E. coli* BL21(DE3) cells. An overnight culture of each strain was sub-cultured after 1:100 dilution in LB containing 50 μg mL^{-1} ampicillin, 25 μg mL^{-1} chloramphenicol, and 0.15 mM IPTG. The cells were grown to OD_{600} of 0.4-0.6, and 1 mL of each culture was concentrated ~10-fold. Four μL of this suspension was spread onto a slide and examined using a Leica DM6 microscope with 63X and 1.4 numerical aperture oil immersion objective lens. Images were captured using Hamamatsu-Flash4-CL-100110 camera with 2 seconds exposure time. Images, analyzed using LASX software and processed and quantified (~200 cells per condition) using ImageJ, are representative of three independent experiments.

For measurement of mRNA abundance, *E. coli* BL21(DE3) was transformed with pACYC_xEGFP and either pET_xanR or pET_pecS (expressing *Pectobacterium atrosepticum pecS* (Deochand, et al. 2016)). Cells were grown in LB containing ampicillin (100 μg mL^{-1}) and chloramphenicol (25 μg mL^{-1}) to OD_{600} of 0.4-0.6 at 37°C. Production of (p)ppGpp was induced by addition of 30 mM serine hydroxamate (MilliporeSigma, Burlington, MA) when OD_{600} reached ~0.6, and cells were grown for an additional 30 min.

A total of 5 mL culture was harvested, and cells were pelleted and washed with ice-cold diethyl pyrocarbonate (DEPC)-treated water and pelleted again. Pellets were frozen at -80°C after discarding the supernatant. Total RNA was extracted using Monarch Total RNA Miniprep Kit (New England BioLabs). Total RNA extracted was rendered free of any genomic DNA contamination using Turbo DNase (Invitrogen) and absence of DNA was verified by PCR. RNA quantification was done using NanoDrop (Thermo Scientific), and all sample concentrations were normalized to 50 ng μl^{-1} by adding the required volume of nuclease free water. The mRNA abundance levels were determined using 200 ng total RNA for each sample

in technical replicates using Luna Universal One-Step RT-qPCR kit (New England BioLabs) and SYBR Green for detection and gene-specific primers (Table S1). All quantitative PCR reactions were performed using 7900HT Fast Real-Time PCR System (Applied Biosystems) using 96-well plates. *EGFP* mRNA abundance was normalized using two reference genes, the genomic *IlvY* and the plasmid-encoded *CmR*, which was included to control for potential changes in plasmid copy number. The geometric mean of C_T values for the two reference genes was used for normalization (Vandesompele, et al. 2002), and mRNA abundance is represented relative to levels in cells expressing PecS. All mRNA abundance levels were calculated using the $\Delta\Delta C_T$ method from three biological replicates. Statistical significance was evaluated using a Student's t-test.

Results

Model of XanR

R. solanacearum XanR is a LysR-type transcriptional regulator (LTTR), and these proteins typically exist as homotetramers (Maddocks and Oyston 2008). Modeling of XanR was performed using Phyre2 in intensive mode, using three templates to generate the final model, which represented an extended conformation. This model was combined with a model of a compact subunit, which was based on template 2ESNC, and a tetrameric assembly was created containing two compact and two extended subunits, representing the typical association of LTTR tetramers (Jiang, et al. 2018). Modeling revealed that each XanR subunit consists of the characteristic N-terminal winged helix-turn-helix domain (wHTH; Figure 1DE, cyan) and a C-terminal domain that is predicted to bind to inducers (red and orange); these two domains are connected by a mostly helical linker (gray). Two

subunits, one in the extended conformation and one compact, associate such that their wHTH domains are positioned for interaction of a pair of recognition helices in consecutive DNA major grooves (Figure 1E; one subunit colored according to the domain designation in Figure 1D and the other is blue; recognition helices are marked by open arrows). The other pair of subunits is arranged with their recognition helices on the opposite face of the tetrameric assembly (light brown). LTTRs typically bind a so-called LTTR-box for which a consensus sequence T-N_{11}-A has been reported (Maddocks and Oyston 2008). Binding to a long 50-60 bp site is predicted to occur by inducing a significant DNA bend to accommodate interaction with the two paired wHTH domains (Muraoka, et al. 2003). The inducer-binding region is composed of two Rossman-like domains (RD1/RD2), with inducer typically binding in the interface between these domains.

DNA binding by XanR

R. solanacearum XanR (Mr~35.6 kDa) was overexpressed in *E. coli* with an N-terminal His_6-tag (Figure 2A). Glutaraldehyde cross-linking revealed the formation of dimers and tetramers as well as multimers that fail to enter the gel. That XanR tetrameric assemblies may be detected after glutaraldehyde crosslinking is consistent with its predicted existence as a tetramer in solution. A determination of thermal stability revealed that XanR exhibits an apparent two-state melting transition with Tm = 63.6 ± 0.2 °C, indicating that it is quite stable at physiological temperatures (Figure 2B).

DNA binding by XanR was assessed using electrophoretic mobility shift assays (EMSA) using a DNA fragment comprising the *xanR-xan* intergenic DNA (and 38 bp of *xanR* coding sequence). XanR bound to the intergenic region forming more than one complex

with an apparent Kd of 45 ± 2 nM (tetramer); the Hill co-efficient of 1.2 ± 0.05 indicates that binding was not cooperative (Figure 2CD). Since LysR family proteins generally bind AT-rich DNA, XanR-DNA complexes were challenged with poly-dGdC, which only modestly affected binding. Addition of poly-dGdC did reveal the existence of a faster-migrating complex (Figure 2E). Differences in electrophoretic mobility of complexes may either be due to the stoichiometry of protein to DNA or the extent of DNA bending, with complexes containing more severely bent DNA migrating slower. The appearance of the faster-migrating complex on addition of poly-dGdC may therefore reflect either displacement of non-specifically bound XanR or the disruption of binding to a lower-affinity site by the second pair of wHTH domains within a XanR tetramer, resulting in loss of DNA bending.

DNA binding by XanR is unaffected by oxidant

Since *xdhA* expression is increased both *in planta* and *in vitro* upon exposure to H_2O_2 (Flores-Cruz and Allen 2009), we investigated the effect of oxidants on DNA binding. The inorganic oxidant H_2O_2 was used as well as the organic oxidants cumene hydroperoxide and *tert*-butyl hydroperoxide; XanR contains two cysteine residues per monomer, both in the RD2 ligand-binding domain. EMSAs with oxidized protein revealed that XanR binding was essentially unaffected by oxidant, except at very high oxidant concentration (Figure 3F and data not shown). Since H_2O_2 causes an increase in *xdhA* expression, a possible interpretation would be that XanR oxidation leads to an altered binding mode that promotes *xan* expression, but cannot be distinguished by EMSA. Alternatively, induction by H_2O_2 may be indirect or unrelated to XanR.

297 **Identification of XanR binding site**

298 DNaseI footprinting was used to map XanR binding. A concentration-dependent
299 protection was observed in the DNA region encompassing the start of the *xanR* open
300 reading frame and extending upstream of the start codon Figure 3). Increasing the
301 concentration of XanR from 320 nM (Figure 3A) to 1,000 nM (Figure 3B) resulted in better
302 protection of the identified sequence, but not in an extended footprint. Examination of this
303 sequence revealed a signature T-N_{11}-A motif twice within the region of protection, raising
304 the possibility that either site represents the cognate XanR site. No protection was
305 observed downstream, towards the *xan* operon, and this pattern of protection was
306 consistent, regardless of protein concentrations (up to 1 μM). That only a partial footprint
307 was observed likely speaks to the rate of complex dissociation during the time of
308 incubation with DNaseI, with a half-life of <3.5 min (the time of incubation with DNaseI)
309 indicating an apparent first order dissociation rate constant of >0.003 s^{-1}.
310
311 **Ligands disrupt DNA binding by XanR**

312 The TetR-family transcription factor XdhR, which controls expression of the *xdh*
313 operon in both *S. coelicolor* and *A. fabrum* (Figure 1C), binds phosphorylated guanosine
314 derivatives directly as a result of which DNA binding is attenuated (Sivapragasam, et al.
315 2017, Sivapragasam and Grove 2016). We therefore examined the ability of GMP, GTP,
316 ppGpp, and c-di-GMP to modulate DNA-XanR complex formation. EMSA revealed that GMP
317 disrupted complex formation only at a high concentration, with a Ki of 6.6 ± 0.3 mM (Figure
318 4A; IC_{50} = 15.8 ± 0.8 mM). GTP was a more potent ligand, with Ki of 1.6 ± 0.3 mM (Figure
319 4B; IC_{50} = 3.8 ± 0.6 mM), and the most effective ligands were c-di-GMP and ppGpp, for

which Ki was 0.5 ± 0.04 mM and 0.5 ± 0.1 mM, respectively (Figure 4C-F; IC_{50} = 1.2 ± 0.2 mM and 1.3 ± 0.1 mM, respectively). Evidently, the more highly phosphorylated guanosine derivatives were the preferred ligands for XanR.

The cleft between the two Rossman-like domains (RD1/RD2) is commonly seen to harbor the ligand-binding site in LTTRs. We used the final XanR model for blind docking of select ligands to predict the energetically most favorable interaction. ppGpp has too many conformational degrees of freedom and is unsuitable for this purpose, so we used c-di-GMP as a predicted high-affinity ligand and guanosine as an example of a less favored ligand. Both c-di-GMP and guanosine were predicted to bind preferentially in the interface between RD1 and RD1 using the Monte Carlo-based method EDock, which is optimized for docking to predicted protein structures (Figure 5). This is consistent with the expected ligand-binding site in LTTRs. That the blind docking predicts binding of c-di-GMP to the conserved ligand-binding site within LTTRs is consistent with direct binding of ligands to XanR and with the ability of ligands to modulate DNA binding.

Gene regulation by XanR

The ability of XanR to control activity of the promoter driving expression of the *xan* operon was examined in *E. coli*. A destabilized version of enhanced green fluorescent protein (EGFP), which is characterized by shorter *in vivo* half-life, was expressed in *E. coli* from the low copy plasmid pACYC184. Expression of *egfp* using the *xan* promoter did not result in measurable fluorescence, perhaps because the *xan* promoter does not function in *E. coli* (data not shown). We therefore installed a consensus *E. coli* promoter upstream of the *xanR-xan* intergenic DNA and *egfp*, generating plasmid pACYC_xEGFP. *E. coli* cells

carrying this construct produced EGFP, although the extended 5'-UTR likely resulted in relatively weak fluorescence (Figure 6A). For *E. coli* cells transformed with both pACYC_xEGFP and pET_xanR, expressing XanR, no green fluorescence was visible, regardless of whether or not IPTG was added to induce *xanR* expression; however, quantitation revealed that addition of IPTG resulted in the lowest fluorescence (Figure 6B-D). This indicates that XanR represses the *xan* promoter and that leaky expression of *xanR* was sufficient to repress *egfp* expression. The specificity of XanR was assessed by comparing mRNA abundance in cells expressing XanR or the unrelated transcription factor PecS from *Pectobacterim atrosepticum* (Deochand, et al. 2016). Reduced *egfp* mRNA abundance was observed in cells harboring pET100_XanR (Figure 6E). Notably, *egfp* mRNA levels increased significantly on addition of serine hydroxamate, which mimics starvation by inhibiting charging of seryl-tRNA synthetase, a consequence of which is accumulation of (p)ppGpp (Ferullo and Lovett 2008, Sivapragasam and Grove 2016). That conditions under which the XanR ligand (p)ppGpp accumulates results in increased *egfp* expression is consistent with repression by XanR.

Discussion

Members of the LTTR protein family most often function as activators, and DNA binding may be modulated by association with specific ligands (Maddocks and Oyston 2008). Ligands generally bind in a cleft between the two Rossman-fold type subdomains of each monomer, and the resulting conformational changes are communicated to the DNA-binding domains *via* the connecting helix (Monferrer, et al. 2010, Ruangprasert, et al. 2010). Typically, the homotetrameric LTTR can bind three possible DNA sites in gene

promoters (Porrua, et al. 2007). Simultaneous binding to a high-affinity site and one of two low-affinity sites leads to DNA bending in absence of inducer, while binding of inducer causes the protein to bind instead to the other low-affinity site in concert with the high-affinity site; this change in binding mode is driven by inducer-mediated conformational changes that change the distance between dimeric DNA-binding domains (Ezezika, et al. 2007, Lerche, et al. 2016, Monferrer, et al. 2010). This results in relaxation of the DNA bend and an increase in transcriptional activity.

Repression of target genes by LTTRs has also been reported; for example, *Bacillus subtilis* CcpC negatively regulates expression of genes encoding enzymes in the citric acid cycle (Kim, et al. 2003). In this case, tetrameric apo-CcpC was proposed to bind two sites to generate a DNA bend that disfavors RNA polymerase binding. Binding of the inducing ligand citrate causes a relaxation of the DNA bend angle by releasing the second CcpC dimer from its low-affinity site, retaining CcpC binding only to the high-affinity site. This binding mode allows simultaneous RNA polymerase binding. By comparison, XanR represses the *xan* promoter and forms two complexes with very low mobility in absence of DNA competitor, possibly representing formation of complexes in which the DNA is bent. Ligand binding results in gradual disappearance of the XanR-DNA complex, perhaps reflecting complete disruption of DNA binding. Considering the conventional response of LTTRs to ligand, however, the observed disappearance of the XanR-DNA complex could also be due to ligand-mediated release of one XanR dimer from a low-affinity site. This might result in the other XanR dimer binding to a single site and generating a complex that is not stable to electrophoresis.

Exposure of *R. solanacearum* to H_2O_2 leads to increased *xdhA* expression (Flores-Cruz and Allen 2009). Such increased expression might derive from an altered mode of DNA binding by oxidized XanR, which is compatible with binding of RNA polymerase (but indistinguishable by EMSA). Alternatively, H_2O_2 may induce a stringent response and the attendant production of (p)ppGpp (Chang, et al. 2002), which in turn serves as a ligand for XanR. The precise mechanism for H_2O_2-mediated upregulation of the *xan* operon notwithstanding, the marked resistance of *R. solanacearum* to H_2O_2 may be linked to facilitated production of (p)ppGpp and c-di-GMP, which would promote stringent response and biofilm formation, respectively, thus contributing to survival.

Purine salvage has been linked to (p)ppGpp synthesis and/or virulence in a number of bacterial species. The requirement for precursors to (p)ppGpp synthesis under conditions of stress or nutrient limitation rationalizes this connection. The development of a positive feed-back loop in which genes encoding enzymes in the purine salvage pathway is promoted by phosphorylated guanosine derivatives is one mechanism by which cells can sustain synthesis of such guanosine-derived signaling molecules under unfavorable conditions. In *E. coli*, cellular levels of (p)ppGpp reach mM levels during stringent response (Cashel 1975). A Ki of ~0.5 mM for disruption of DNA binding by XanR is therefore in the physiologically relevant range. Our data indicate that the positive feed-back loop in which expression of genes encoding purine salvage enzymes is promoted by phosphorylated guanosine derivatives has evolved independently in *R. solanacearum*, and that it constitutes an example of similar evolutionary solutions to a common challenge.

Funding

This work was supported by the National Science Foundation (MCB-1714219 to A. Grove), the LSU Ronald E. McNair Research Scholars Program, and a grant from the NIH-IMSD Program to LSU.

Acknowledgments

We thank A. Pande for helpful discussions.

References

Brown DG, Allen C. *Ralstonia solanacearum* genes induced during growth in tomato: an inside view of bacterial wilt. Mol Microbiol 2004;**53**: 1641-60.

Cashel M. Regulation of bacterial ppGpp and pppGpp. Annu Rev Microbiol 1975;**29**: 301-18.

Cer RZ, Mudunuri U, Stephens R *et al.* IC50-to-Ki: a web-based tool for converting IC50 to Ki values for inhibitors of enzyme activity and ligand binding. Nucleic Acids Res 2009;**37**: W441-5.

Chang DE, Smalley DJ, Conway T. Gene expression profiling of *Escherichia coli* growth transitions: an expanded stringent response model. Mol Microbiol 2002;**45**: 289-306.

Deochand DK, Meariman JK, Grove A. pH-Dependent DNA Distortion and Repression of Gene Expression by *Pectobacterium atrosepticum* PecS. ACS Chem Biol 2016;**11**: 2049-56.

Eastgate JA, Thompson L, Milner J *et al.* Identification of a nonpathogenic *Erwinia amylovora guaB* mutant. Plant Pathol 1997;**46**: 594-9.

Ezezika OC, Haddad S, Clark TJ *et al.* Distinct effector-binding sites enable synergistic transcriptional activation by BenM, a LysR-type regulator. J Mol Biol 2007;**367**: 616-29.

Ferullo DJ, Lovett ST. The stringent response and cell cycle arrest in *Escherichia coli*. PLoS Genet 2008;**4**: e1000300.

Flores-Cruz Z, Allen C. *Ralstonia solanacearum* encounters an oxidative environment during tomato infection. Mol Plant Microbe Interact 2009;**22**: 773-82.

Gaca AO, Colomer-Winter C, Lemos JA. Many means to a common end: the intricacies of (p)ppGpp metabolism and its control of bacterial homeostasis. J Bacteriol 2015;**197**: 1146-56.

Genin S, Denny TP. Pathogenomics of the *Ralstonia solanacearum* species complex. Annu Rev Phytopathol 2012;**50**: 67-89.

Grove A, Kushwaha AK, Nguyen KH. Determining the role of metal binding in protein cage assembly. Methods Mol Biol 2015;**1252**: 91-100.

Hanko EKR, Paiva AC, Jonczyk M *et al.* A genome-wide approach for identification and characterisation of metabolite-inducible systems. Nat Commun 2020;**11**: 1213.

Jacobs JM, Babujee L, Meng F *et al.* The in planta transcriptome of *Ralstonia solanacearum*: conserved physiological and virulence strategies during bacterial wilt of tomato. MBio 2012;**3**.

Jenal U, Reinders A, Lori C. Cyclic di-GMP: second messenger extraordinaire. Nat Rev Microbiol 2017;**15**: 271-84.

Jewett MW, Lawrence KA, Bestor A *et al.* GuaA and GuaB are essential for *Borrelia burgdorferi* survival in the tick-mouse infection cycle. J Bacteriol 2009;**191**: 6231-41.

Jiang YL, Wang XP, Sun H *et al.* Coordinating carbon and nitrogen metabolic signaling through the cyanobacterial global repressor NdhR. Proc Natl Acad Sci U S A 2018;**115**: 403-8.

Kelley LA, Mezulis S, Yates CM *et al.* The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 2015;**10**: 845-58.

Kim JK, Kwon JY, Kim SK *et al.* Purine biosynthesis, biofilm formation, and persistence of an insect-microbe gut symbiosis. Appl Environ Microbiol 2014;**80**: 4374-82.

Kim SI, Jourlin-Castelli C, Wellington SR *et al.* Mechanism of repression by *Bacillus subtilis* CcpC, a LysR family regulator. J Mol Biol 2003;**334**: 609-24.

463 Lerche M, Dian C, Round A *et al.* The solution configurations of inactive and activated DntR
464 have implications for the sliding dimer mechanism of LysR transcription factors. Sci
465 Rep 2016;**6**: 19988.
466 Liechti G, Goldberg JB. *Helicobacter pylori* relies primarily on the purine salvage pathway for
467 purine nucleotide biosynthesis. J Bacteriol 2012;**194**: 839-54.
468 Liu K, Bittner AN, Wang JD. Diversity in (p)ppGpp metabolism and effectors. Curr Opin
469 Microbiol 2015;**24**: 72-9.
470 Lundgren BR, Connolly MP, Choudhary P *et al.* Defining the Metabolic Functions and Roles in
471 Virulence of the rpoN1 and rpoN2 Genes in *Ralstonia solanacearum* GMI1000. PLoS
472 One 2015;**10**: e0144852.
473 Maddocks SE, Oyston PC. Structure and function of the LysR-type transcriptional regulator
474 (LTTR) family proteins. Microbiology 2008;**154**: 3609-23.
475 Mandal S, Das RK, Mishra S. Differential occurrence of oxidative burst and antioxidative
476 mechanism in compatible and incompatible interactions of *Solanum lycopersicum* and
477 *Ralstonia solanacearum*. Plant Physiol Biochem 2011;**49**: 117-23.
478 Mansfield J, Genin S, Magori S *et al.* Top 10 plant pathogenic bacteria in molecular plant
479 pathology. Mol Plant Pathol 2012;**13**: 614-29.
480 Martinez-Gil M, Ramos C. Role of Cyclic di-GMP in the Bacterial Virulence and Evasion of the
481 Plant Immunity. Curr Issues Mol Biol 2018;**25**: 199-222.
482 Monferrer D, Tralau T, Kertesz MA *et al.* Structural studies on the full-length LysR-type
483 regulator TsaR from *Comamonas testosteroni* T-2 reveal a novel open conformation
484 of the tetrameric LTTR fold. Mol Microbiol 2010;**75**: 1199-214.

Muraoka S, Okumura R, Ogawa N *et al.* Crystal structure of a full-length LysR-type transcriptional regulator, CbnR: unusual combination of two subunit forms and molecular bases for causing and changing DNA bend. J Mol Biol 2003;**328**: 555-66.

O'Neill MC. Escherichia coli promoters. I. Consensus as it relates to spacing class, specificity, repeat substructure, and three-dimensional organization. J Biol Chem 1989;**264**: 5522-30.

Porrua O, Garcia-Jaramillo M, Santero E *et al.* The LysR-type regulator AtzR binding site: DNA sequences involved in activation, repression and cyanuric acid-dependent repositioning. Mol Microbiol 2007;**66**: 410-27.

Ross W, Vrentas CE, Sanchez-Vazquez P *et al.* The magic spot: a ppGpp binding site on *E. coli* RNA polymerase responsible for regulation of transcription initiation. Mol Cell 2013;**50**: 420-9.

Ruangprasert A, Craven SH, Neidle EL *et al.* Full-length structures of BenM and two variants reveal different oligomerization schemes for LysR-type transcriptional regulators. J Mol Biol 2010;**404**: 568-86.

Sivapragasam S, Deochand DK, Meariman JK *et al.* The Stringent Response Induced by Phosphate Limitation Promotes Purine Salvage in *Agrobacterium fabrum*. Biochemistry 2017;**56**: 5831-43.

Sivapragasam S, Grove A. *Streptomyces coelicolor* XdhR is a direct target of (p)ppGpp that controls expression of genes encoding xanthine dehydrogenase to promote purine salvage. Mol Microbiol 2016;**100**: 701-18.

Sivapragasam S, Grove A. The Link between Purine Metabolism and Production of Antibiotics in *Streptomyces*. Antibiotics (Basel) 2019;**8**: 76.

Sivapragasam S, Pande A, Grove A. A recommended workflow for DNase I footprinting using a capillary electrophoresis genetic analyzer. Anal Biochem 2015;**481**: 1-3.

Traxler MF, Zacharia VM, Marquardt S *et al.* Discretely calibrated regulatory loops controlled by ppGpp partition gene induction across the 'feast to famine' gradient in *Escherichia coli*. Mol Microbiol 2011;**79**: 830-45.

Vandesompele J, De Preter K, Pattyn F *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 2002;**3**: RESEARCH0034.

Xi H, Schneider BL, Reitzer L. Purine catabolism in *Escherichia coli* and function of xanthine dehydrogenase in purine salvage. J Bacteriol 2000;**182**: 5332-41.

Zhang W, Bell EW, Yin M *et al.* EDock: blind protein–ligand docking by replica-exchange monte carlo simulation. J Cheminform 2020;**12**.

Figures

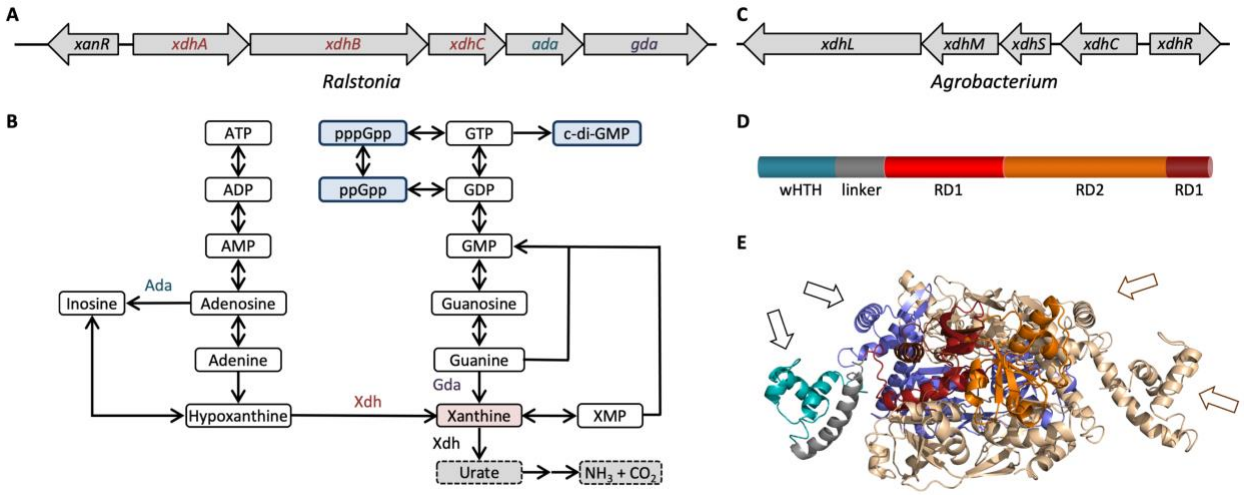


Figure 1. XanR is predicted to regulate production of enzymes involved in purine salvage. A. Genomic locus consisting of *xanR* (*Rsc2094;* new locus tag *RS_RS10505*) oriented divergently from an operon predicted to encode Xdh subunits (*xdhA, xdhB*), the Xdh assembly chaperone (*xdhC*), adenosine deaminase (*ada*), and guanine deaminase (*gda*). B. Abbreviated purine salvage pathway identifying reactions catalyzed by enzymes encoded in the *xan* operon. C. Genomic locus from *Agrobacterium fabrum* illustrating the TetR family transcription factor XdhR encoded divergently from an operon encoding the assembly chaperone XdhC and the three Xdh subunits XdhS, XdhM, and XdhL. D. Domain organization of XanR with the DNA-binding wHTH motif in cyan, the linker in gray, and the inducer-binding region in red and orange. Rossman-like domain 1 (RD1) comprises sequence downstream of the linker as well as sequence near the C-terminus (red and brown) whereas RD2 comprises contiguous sequence (orange). E. Model of XanR generated using Phyre2 in intensive mode; 97% of residues were modeled based on three identified templates (all of which with 100% confidence) to generate the model of a subunit in the extended conformation. Eight residues

538 at the C-terminus were modeled ab initio and should be considered an approximation only.
539 One extended subunit is colored according to the domain organization in A and the paired
540 compact subunit (based on template 2ESN) is in blue; these subunits generate a set of
541 recognition helices for binding consecutive DNA major grooves (black open arrows). The
542 other pair of subunits and arrows identifying recognition helices are colored light brown.
543
544
545

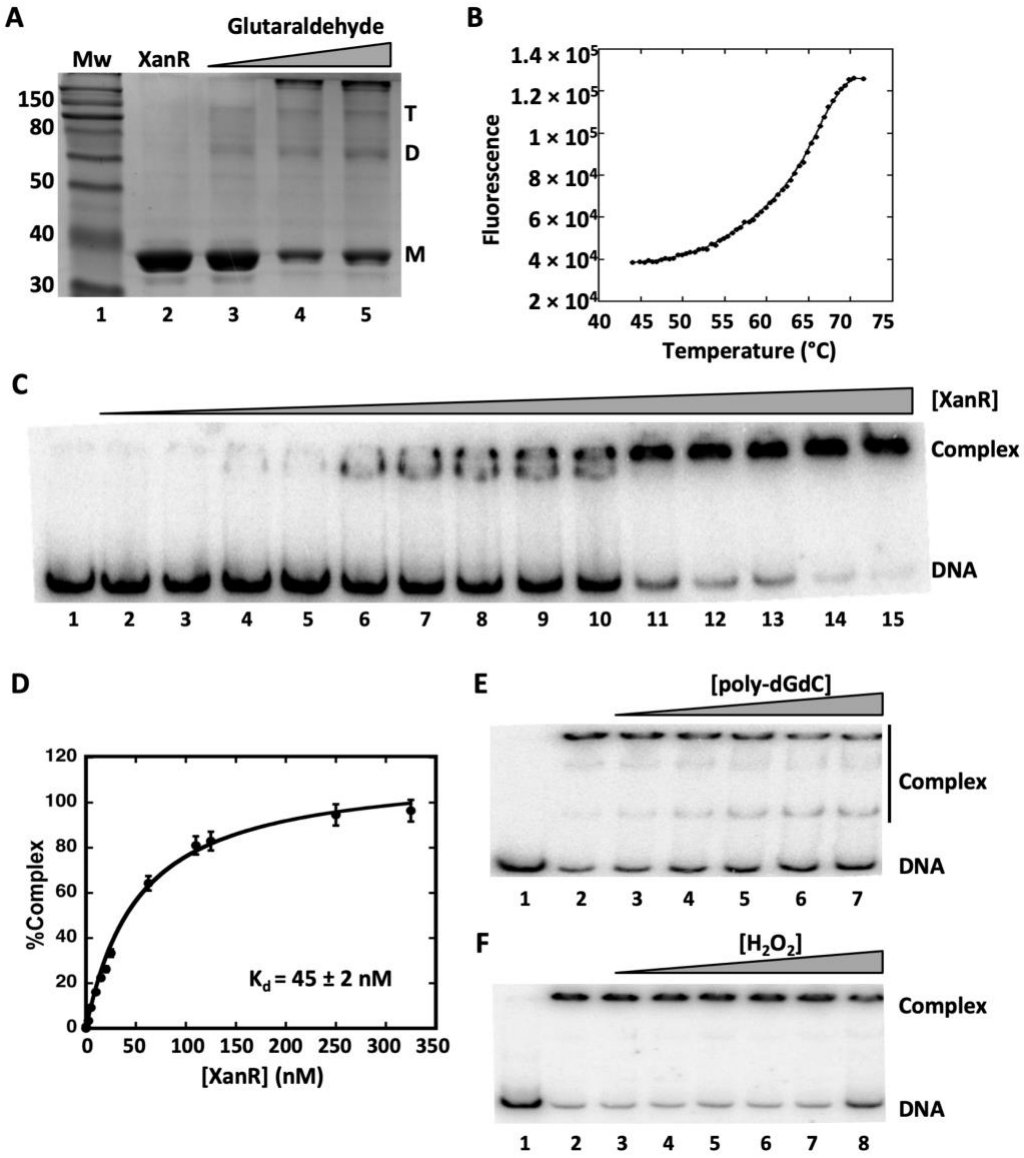


Figure 2. XanR binds *xanR-xan* intergenic DNA. A. SDS-PAGE gel showing purified XanR (lane 2) and reactions containing 0.005%, 0.01%, and 0.05% glutaraldehyde (lanes 3-5). Monomer (M), dimer (D), and tetramer (T) identified at the right. B. Thermal unfolding of XanR as measured by the fluorescence (arbitrary units) of SYPRO Orange binding to hydrophobic regions of unfolded protein as a function of temperature. The data represent the average from three measurements. C. EMSA with 0.05 nM *xanR-xan* intergenic DNA and increasing concentrations of XanR; lane 1, DNA only, lanes 2-15 with 0.0025 nM, 0.025 nM, 0.25 nM, 2.5

554 nM, 5 nM, 10 nM, 15 nM, 20 nM, 25 nM, 62.5 nM, 110 nM, 125 nM, 250 nM and 350 nM
555 tetrameric XanR, respectively. A 6% PAGE gel was used. D. Percent complex as a function of
556 [XanR]. Error bars represent standard deviation. E. A 4% PAGE gel in which all reactions
557 contained 0.05 nM DNA. Lane 1, DNA only. Reactions in lanes 2-7 contained 62.5 nM XanR to
558 which 0, 10 pg, 50 pg, 100 pg, 500 pg and 1 ng of poly-dGdC was added. F. Effect of oxidant
559 (H_2O_2). Reactions contained 0.05 nM DNA. Lane 1, DNA only. Reactions in lanes 2-8 contained
560 62.5 nM XanR and 0, 300 nM, 3 μM, 30 μM, 300 μM, 3 mM and 30 mM H_2O_2.
561
562
563
564
565
566
567

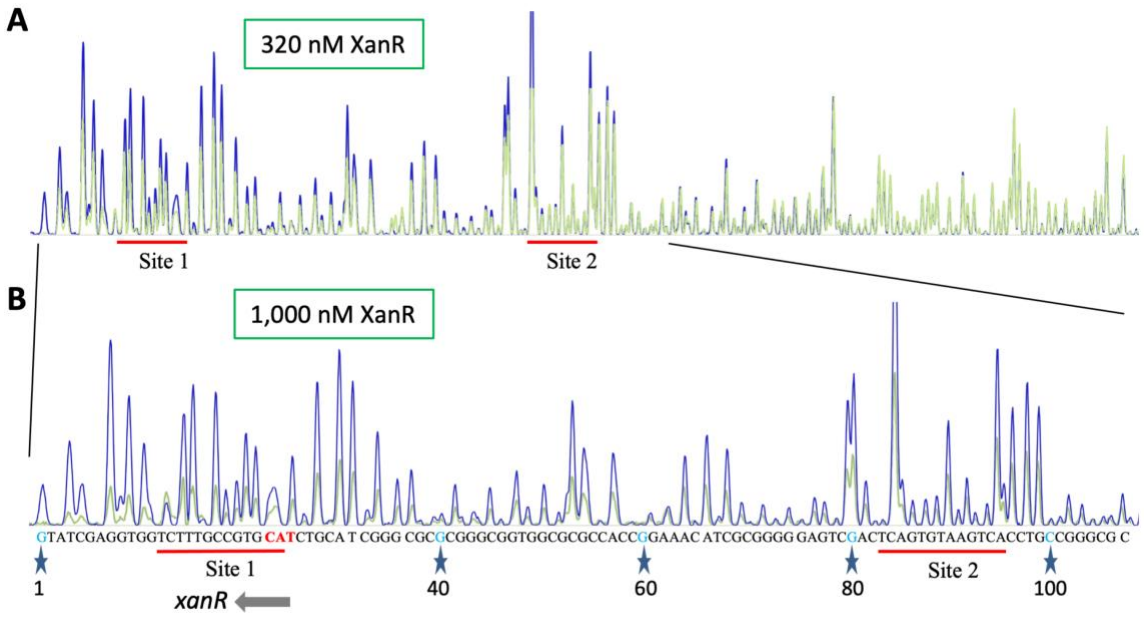
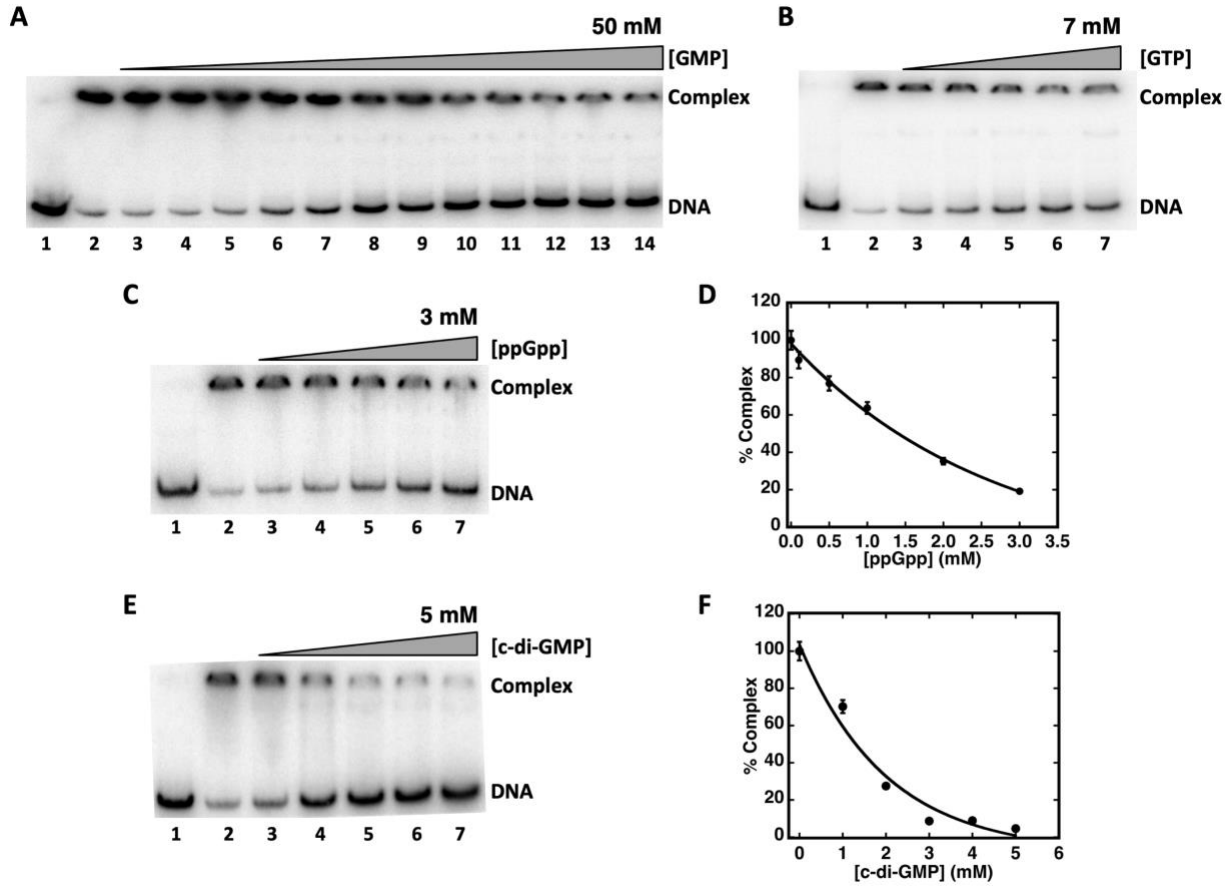


Figure 3. XanR partially protects a DNA region surrounding the *xanR* start codon. A. DNaseI digestion of *xanR-xan* DNA without (blue) or with XanR (320 nM tetramer; green). Two T-N_{11}-A motifs are identified by red lines, with Site 1 overlapping the *xanR* start codon. Representative of at least three independent experiments. B. Expanded view of the protected region (identified by black lines) with green line representing DNA with 1,000 nM XanR (tetramer). Note that protection is more pronounced at the higher protein concentration used in panel B compared to panel A. Grey arrow marks the direction of the *xanR* gene. The sequence of the protected region is provided below the electropherogram traces, with the *xanR* start codon in red.



582

583 Figure 4. DNA binding is attenuated by phosphorylated guanosine derivatives. A-C. EMSA
584 with 0.1 nM DNA and 62.5 nM tetrameric XanR. Lanes 1, DNA only. Lanes 2, DNA and XanR.
585 A. Lanes 3-14 with increasing concentrations of GMP (up to 50 mM). B. Lanes 3-7 with
586 increasing concentrations of GTP (up to 7 mM). C. Lanes 3-7 with increasing concentrations
587 of ppGpp (up to 3 mM). D. Percent complex as a function of [ppGpp]. E. EMSA with DNA (0.05
588 nM) and XanR. Lane 1, DNA only. Lane 2, DNA and XanR. Lanes 3-7 with increasing
589 concentration of c-di-GMP (up to 5 mM). E. Percent complex as a function of [c-di-GMP].
590 Error bars represent standard deviation; where error bars appear missing, they are smaller
591 than the symbol sizes.

592

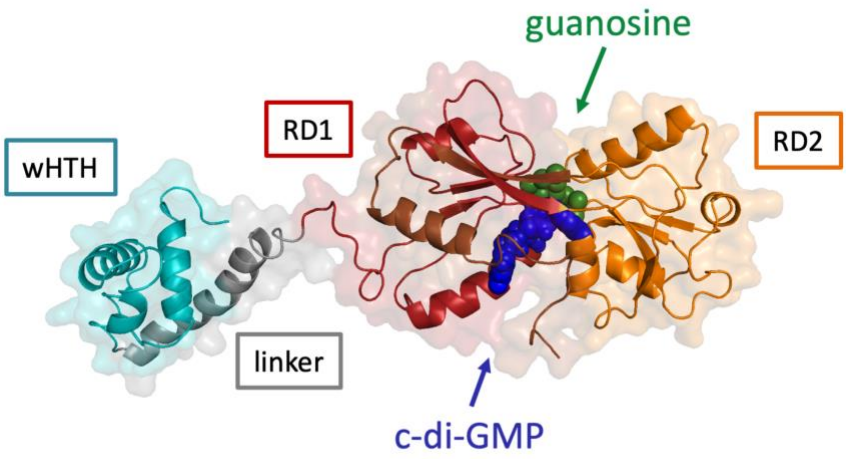


Figure 5. Prediction of ligand binding to XanR. A. Docking of c-di-GMP (blue space-filling representation) or guanosine (green) to the final model of XanR in the extended conformation; XanR is colored as in Figure 2 with functional regions identified. Residues implicated in binding to both ligands include P105, Y107, L108, F111, F112, I153, F169 and F276 from RD1 and D171, G249, and D251 from RD1. Figure was generated with PyMol.

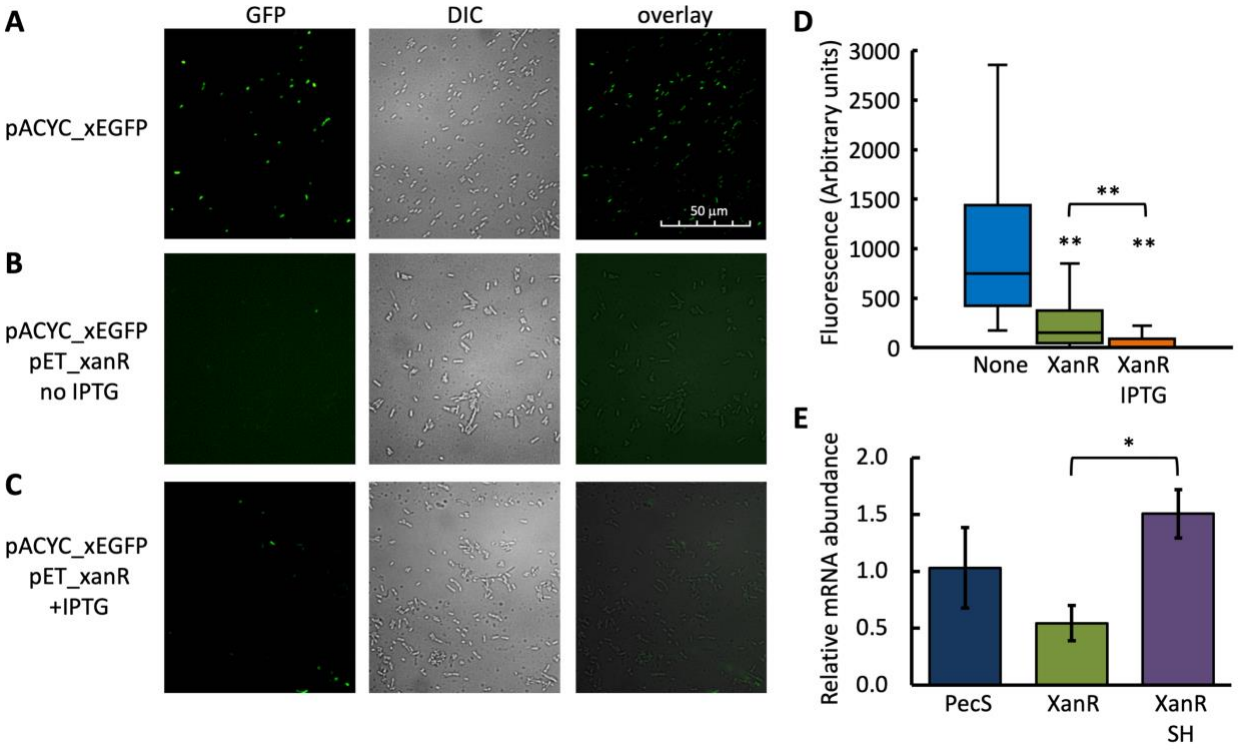


Figure 6. XanR functions as a repressor. *E. coli* cells transformed with pACYC_xEGFP only (A) or with both pACYC_xEGFP and pET_xanR (B-C). Leaky expression of *xanR* (B) or induction of *xanR* expression with 0.15 mM IPTG (C) is shown. Images represent GFP fluorescence (left), differential interference contrast (DIC; middle), and overlay of GFP and DIC (right). Scale bar (50 µm, shown in top panel) is same for all images. Images are representative of triplicate experiments. D. Quantitation of fluorescence. Blue, cells harboring pACYC_xEGFP only; green and orange, cells harboring pACYC_xEGFP and pET_xanR without and with IPTG, respectively. Asterisks represent statistically significant differences compared to cells harboring pACYC_xEGFP only based on a Student's t-test, except where bracket indicates otherwise ($p<0.001$). E. Relative mRNA abundance for cells harboring pACYC_xEGFP and pET_pecS (dark blue), pACYC_xEGFP and pET_xanR (green), and pACYC_xEGFP and

611 pET_xanR with serine hydroxamate (SH; purple). Data represent three biological replicates.

612 Asterisk represents statistically significant differences ($p<0.01$).

613