Crystal Structures of Two Transcriptional Regulators from *Bacillus cereus* Define the Conserved Structural Features of a PadR Subfamily

Guntur Fibriansah 1,∗, Ákos T. Kovács 2,∗b, Trijntje J. Pool 1, Mirjam Boonstra 2, Oscar P. Kuipers 2, Andy-Mark W. H. Thunnissen 1∗

1 Laboratory of Biophysical Chemistry, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands, 2 Department of Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands

Abstract

PadR-like transcriptional regulators form a structurally-related family of proteins that control the expression of genes associated with detoxification, virulence and multi-drug resistance in bacteria. Only a few members of this family have been studied by genetic, biochemical and biophysical methods, and their structure/function relationships are still largely undefined. Here, we report the crystal structures of two PadR-like proteins from *Bacillus cereus*, which we named bcPadR1 and bcPadR2 (products of gene loci BC4206 and BCE3449 in strains ATCC 14579 and ATCC 10987, respectively). BC4206, together with its neighboring gene BC4207, was previously shown to become significantly upregulated in presence of the bacteriocin AS-48. DNA mobility shift assays reveal that bcPadR1 binds to a 250 bp intergenic region containing the putative BC4206–BC4207 promoter sequence, while *in situ* expression of bcPadR1 decreases bacteriocin tolerance, together suggesting a role for bcPadR1 as repressor of BC4206–BC4207 transcription. The function of bcPadR2 (48% identical in sequence to bcPadR1) is unknown, but the location of its gene just upstream from genes encoding a putative antibiotic ABC efflux pump, suggests a role in regulating antibiotic resistance. The bcPadR proteins are structurally similar to LmrR, a PadR-like transcription regulator in *Lactococcus lactis* that controls expression of a multidrug ABC transporter via a mechanism of multidrug binding and induction. Together these proteins define a subfamily of conserved, relatively small PadR proteins characterized by a single C-terminal helix for dimerization. Unlike LmrR, bcPadR1 and bcPadR2 lack a central pore for ligand binding, making it unclear whether the transcriptional regulatory roles of bcPadR1 and bcPadR2 involve direct ligand recognition and induction.

Introduction

Members of the PadR-like protein family (accession no. Pfam PF03551) form a large and widespread group of bacterial transcription factors that regulate diverse processes including multi-drug resistance, virulence and detoxification. They are named after the phenolic acid decarboxylation repressor, which is a transcription factor that represses genes associated with the phenolic acid stress response in gram-positive bacteria such as *Bacillus subtilis*, *Pediococcus pentosaceus*, and *Lactobacillus plantarum* [1,2]. Besides the founding member, only a few of the PadR-like proteins have been or are currently under investigation. These include AphA, which activates virulence gene expression in *Vibrio cholerae* [3], LmrR and LadR, which are repressors of genes encoding a multi-drug resistance pump in *Lactococcus lactis* and *Listeria monocytogenes*, respectively [4,5], and Pex, a regulator of circadian rhythms in *Synechococcus elongates* [6]. Crystal structures have been reported of AphA, LmrR and Pex, confirming that these proteins share a common fold consisting of two domains: a highly conserved N-terminal winged helix-turn-helix (wHTH) domain of approximately 80–90 amino acid residues, involved in DNA-binding, and a variable C-terminal domain of one or more α-helices, involved in dimerization [7–9]. PadR-like proteins are related to members of the multiple antibiotic resistance regulator (Mar) protein family, which have similar wHTH DNA-binding domains but significantly larger and distinct C-terminal dimerization domains [10]. The size of the C-terminal domain is also a major discriminator among different PadR-like proteins, which have been classified into two distinct subfamilies [4]. PadR-like proteins of subfamily 1 (hereafter abbreviated as PadR-s1, approximately 180 amino acids), which include AphA, PadR and LadR, have relatively large C-terminal domains of about 80–90 amino acids containing multiple α-helices, while subfamily 2 PadR-like proteins (PadR-s2, approximately 110 amino acids), which includes LmrR, have significantly smaller C-terminal domains of about 20–30 amino acids forming a single α-helix.
Currently, the Protein Data Bank contains structures of only a few PadR-s2 proteins. Among these proteins, LmrR is the only one that has been studied both structurally and functionally: the structures of the other proteins were solved by structural genomics centers without a published record of their functional properties. LmrR regulates the expression of LmrCD, a major multidrug ABC efflux pump of *L. lactis* [11], via a mechanism involving multidrug binding and induction [5,9]. Forming a negative feedback loop, LmrR also regulates transcription of its own gene, which is situated just upstream from the *lmrCD* genes [12]. Crystal structures of LmrR in complex with different ligands revealed that LmrR contains an unusual multidrug binding site within a large central hydrophobic pore at its dimer interface [9]. Multidrug recognition is dominated by aromatic stacking interactions of the flat heterocyclic cores of the ligands with opposing indole groups of two dimer-related tryptophan residues (W96/W96′), the prime indicating the dimer-related subunit). Interestingly, the drug-binding tryptophan residues in LmrR, which is located in the C-terminal helix, is conserved in other PadR-s2 proteins, including those with known structure. However, in these latter proteins the equivalent tryptophans are buried in a completely closed dimer interface, making it highly unlikely that they have a similar ligand-binding role [9]. Unfortunately, assessment of the significance of the structural differences between LmrR and the other PadR-s2 proteins is prohibited by the lack of complete functional data.

To investigate whether LmrR is an exceptional member of the PadR-s2 family or whether its structural ligand-binding features are preserved in other members, requires structure/function analysis of LmrR homologs that function in multidrug or antibiotic resistance. Such a homolog of LmrR was recently identified in *Bacillus cereus* as the product of gene locus BC4206. Together with its neighboring gene BC4207, which encodes a putative membrane protein, BC4206 was shown to become highly upregulated if *B. cereus* is treated with low concentrations of enterocin AS-48, a broad-spectrum antimicrobial cyclic peptide produced by *Enterococcus faecalis* S-48 [13,14]. Here we report the crystal structure of the BC4206 gene product from *B. cereus* strain ATCC 14579, a subfamily-2 PadR-like protein which we named bcPadR1. In addition, we report the crystal structure of a bcPadR1 homolog from *B. cereus* strain ATCC 10987, the product of gene locus BCE3449, which we named bcPadR2. This latter protein was selected for structural characterization because its gene is positioned in the vicinity of two genes encoding a putative ABC efflux pump (gene loci BCE3447 and BCE3446). The proximity of these genes, similar as in the *lmr*-*lmrCD* locus, could indicate that bcPadR2 regulates the expression of the ABC efflux pump, perhaps following a similar ligand binding and induction mechanism as LmrR. Preliminary functional analysis, combined with our structural analysis and comparison with related proteins, reveal the roles of various conserved residues in the two subfamily-2 PadR-like proteins from *B. cereus*. Based on our results, we discuss possibilities for ligand-induced regulation of these proteins.

**Materials and Methods**

**Cloning**

The open reading frames encoding *bcPadR1* and *bcPadR2* were amplified by PCR from the genomic DNA of *Bacillus cereus* strains ATCC 14579 and ATCC 10987, respectively, and cloned in the vector pNSC8048 [15]. Plasmid pNSC8048 carries the *nisI* promoter, chloramphenicol selectable marker and a C-terminal strep-tag (SRWSPHPQFEK). The forward and reverse primers for *bpadR1* (locus BC4206) were 5′-CGACCATGGGGGCGAGC-CAAATGTAAAAGGTGTA-3′ and 3′-GCTTCTAGAGCTCCCGCTTGAAATAAGTT-3′, respectively, while forward and reverse primers for *bpadR2* (locus BCE3449) were 5′-CGACCATGGAAAAATTAACTGAAATGCTGAAAG and 3′-GCTTCTAGAGCTTTGACTTCAGACGTTATATAT-3′, respectively. These primers were designed to create Neurospora crassa restriction sites of the resulting PCR products (showed as bold and underlined letters). The PCR fragments and pNSC8048 vector were digested with NcoI and XbaI restriction enzymes and the vector was dephosphorylated further with calf intestine alkaline phosphatase (CIP, Fermentas). Ligation was performed using T4 ligase (Roche) to yield pNSC-*padR1* and pNSC-*padR2*. DNA sequences of the plasmids were all verified by sequencing (Macrogen, The Netherlands).

**Expression and purification**

To express *bcPadR1* and *bcPadR2*, plasmids pNSC-*padR1* and pNSC-*padR2* were transformed into *L. lactis* NZ9000 cells, which is a derivative of strain MG1363 carrying the *nisI/nisK* genes crucial for the induction and overexpression of the target protein under the nisin inducible system. The transformed *L. lactis* NZ9000 cells were grown in 100 ml of rich medium (M17, 0.5% glucose and 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin) were diluted to 50 mM NaCl with buffer A (20 mM Tris-HCl, pH 8.0, 1 mM EDTA, supplemented with 10 mg/ml lysozyme and 1 tablet of a Complete Protease Inhibitor cocktail (Roche). After 1 hour incubation at 30 °C, MgSO4 (10 mM) and DNAseI (10 μg/ml) were added into the suspension and incubation was continued for another 5 min. The cells were disrupted by sonication and remaining cell debris was removed by centrifugation (12,000 rpm for 10 min at 4 °C; JLA 10.500 rotor, Beckman), washed with 50 mM Tris-HCl, pH 7.4, and re-centrifuged (8000 rpm for 5 min at 4 °C; 5810R rotor, Eppendorf). The pellets were stored at −20 °C until further use.

For purification, the cell pellets were resuspended in lysis buffer, containing 100 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, supplemented with 10 mg/ml lysozyme and a tablet of a Complete Protease Inhibitor cocktail (Roche). After 1 hour incubation at 30 °C, MgSO4 (10 mM) and DNAseI (10 μg/ml) were added into the suspension and incubation was continued for another 5 min. The cells were disrupted by sonication and remaining cell debris was removed by centrifugation (12,000 rpm for 10 min at 4 °C; 5810R rotor, Eppendorf). *bcPadR1* or *bcPadR2* were purified from the supernatant using Streptactin Sepharose column chromatography, according to the protocol provided by the manufacturer (IBA Biotechnology GmbH). The protein-containing elution fractions (in 100 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, 2.5 mM dithiobiotin) were diluted to 50 mM NaCl with buffer A (20 mM Tris-HCl, pH 8.0, 1 mM EDTA and 1 mM dithiobiotin) and then applied to a heparin column (3 ml Heparin FastFlow, GE Healthcare), connected to an AKTA Explorer FPLC system (GE Healthcare), to remove bound DNA from the protein. The retained protein was eluted by applying a linear gradient (2.5–100% in 15 column volumes) of 2 M NaCl in buffer A. The fractions containing *bcPadR1* or *bcPadR2* were pooled and concentrated to ~8 mg/ml in buffer containing 10 mM Tris-HCl, pH 8.0, 250 mM NaCl, 1 mM EDTA and 1 mM DTT using an Amicon Ultra 3 KDa cut-off concentrator (Millipore). The protein purity was confirmed by silver-stained SDS PAGE.

**Expression of *bcPadR1* in *B. cereus* and determination of antimicrobial activity of AS-48**

Plasmid pNSC-*padR1* was electroporated into *B. cereus* ATCC 14579 harboring the *nisI* genes on plasmid pNZ9530 [16] and expression was induced with 20 ng/mL of nisin. The antimicrobial activity of AS-48 was determined as the minimal inhibitory.
concentration (MIC) value against *B. cereus* following previous practice [17]. Bacterial growth was followed in the presence of various concentrations of AS-48 and monitored every 15 minutes using a TECAN GENios Absorbance Reader (TECAN). Without AS-48 addition, the final OD<sub>600 nm</sub> value (at the end of the exponential phase) for the *B. cereus* culture was about 1.1 ± 0.1. An inhibition curve was made by plotting OD<sub>600 nm</sub> at the end of incubation versus AS-48 concentration. The minimal inhibitory concentration value was determined from the inhibition curve by interpolation. The lowest concentration of AS-48 at which less than 1% of the total increase in the OD<sub>600</sub> measured in the absence of AS-48, had occurred, was taken as the MIC value.

Electrophoretic mobility shift assay

Electrophoretic mobility shift assays (EMSAs) were carried out essentially as described by Susanna et al [18]. The promoter regions of *B. cereus* BC4206 and BC4029 genes were obtained by PCR with oligos (for BC4206 5'-TTTGCACGTTCTTCAGC-3' and 5'-GATGCAACCTTCTAA-3'; for BC4029 5'-TTTGGCTGTTTTCGAGTGT-3' and 5'-CTTTTACCGCATTT-GACTGT-3') and the resulting fragments (250 bp and 347 bp for BC4206 and BC4029, respectively) were end-labeled with [γ-32P]ATP using T4 polynucleotide kinase (Roche Nederland B.V., The Netherlands). Purified bPadR1 protein and its DNA probe were premixed on ice in binding buffer (20 mM Tris HCl (pH 8.0), 5 mM MgCl<sub>2</sub>, 100 mM NaCl, 0.5 mM dithiothreitol, 0.05 mg/ml poly(dI-dC), 0.05 mg/ml bovine serum albumin, and 8.7% glycerol) [18]. Reaction mixtures contained poly(dI-dC) that is known to eliminate non-specific DNA binding. Samples were incubated at 30°C and were loaded on a 6% polyacrylamide gel after 20 min incubation. Gels were run in 1× TBE buffer (0.089 mM Tris, 0.089 mM boric acid, 0.022 mM EDTA) at 90 V for 60 minutes, dried in a vacuum dryer and autoradiographed using phosphoscreens and a Cyclone PhosphorImager (Packard Instruments, Meriden, CT).

**Gel filtration.** Analytical gel filtration was performed with a Superdex 200 PC 3.2/30 gel filtration column (GE Healthcare), mounted on a SMART-system (GE Healthcare). The running buffer contained 20 mM Tris-HCl, pH 8.0, 200 mM NaCl, 1 mM EDTA and 1 mM dithiouridyl. The injected protein samples volumes were 20 μl. As molecular weight markers a mixture of five proteins was used: aldolase ~158 kDa, bovine serum albumin ~67 kDa, ovalbumin ~43 kDa, chymotrypsin ~25 kDa, and ribonuclease ~13.7 kDa.

**Crystallization.** Prior to crystallization trials, bPadR1 was subjected to Thermofluor-based stability optimization experiments [19] that were performed on a thermal cycler (myIQ, Bio-Rad) using Sypro-Orange dye (Invitrogen). The results revealed that bPadR1 was most stable in a buffer containing 20 mM MES pH 6.0, 250 mM NaCl, 1 mM EDTA and 1 mM DTT. Accordingly, both bPadR1 and bPadR2 were buffer exchanged to the aforementioned buffer using a PD-10 desalting column (GE Healthcare). The proteins were concentrated further to ~10 mg/ml and either used immediately for crystallization or stored at –80°C.

Initial crystallization trials were set up as vapour-diffusion sitting drops at 298 K using various commercially available screens [i.e., JCSG+ (Molecular Dimensions), Wizard Screen (Emerald Biosystem), Structure Screen (Emerald Biosystem) and Cryo Screen (Hampton Research)] and a Mosquito crystallization robot (TTP LabTech) for drop dispensing. Droplets were composed of 0.2 μl protein stock solution and 0.2 μl crystallization solution and were equilibrated against 60 μl of the crystallization solution. Trigonal pyramid crystals of bPadR1 were obtained from JCSG+ Screen condition no. 45, containing 0.17 M ammonium sulfate, 25.5% (v/v) PEG 4000 and 15% (v/v) glycerol. These crystals were used immediately for data collection without further optimization. For bPadR2, similar crystals grew from JCSG+ condition no. 60 (0.1 M imidazole, pH 8.0, and 10% (v/v) PEG 8000) which was optimized manually using the hanging-drop vapour diffusion method. The final crystallization solution contained 0.1 M imidazole pH 8.0 and 9% (v/v) PEG 8000.

**Data collection and structure determination.** Prior to diffraction data collection, a single bPadR1 crystal with dimensions of ~160 μm × 160 μm × 180 μm was transferred to a cryo-protectant solution containing 0.17 M ammonium sulfate, 27% (v/v) PEG 4000 and 20% (v/v) glycerol, and subsequently dipped into liquid nitrogen. A similar procedure was applied to a bPadR2 crystal (~100 μm × 100 μm × 160 μm) with 0.1 M imidazole, pH 8.0, 10% (v/v) PEG 8000 and 20% (v/v) ethylene glycol as cryo-protectant. Diffraction data up to 2.5 Å and 2.2 Å resolution for bPadR1 and bPadR2, respectively, were collected at 100 K using beam line ID14-2 at the ESRF, Grenoble, equipped with an ADSC Quantum 4 detector (see Table 1 for data collection statistics). The two data sets were indexed and integrated in space group P4<sub>1</sub> with the program XDS [20], reindexed to space group P4<sub>2</sub>2 and scaled and merged with the program Pointless and Scala [21], respectively, from the CCP4 software package [22]. Data collection statistics are shown in Table 1.

**Table 1. Data collection and refinement.**

<table>
<thead>
<tr>
<th></th>
<th>bPadR1</th>
<th>bPadR2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9330</td>
<td>0.9330</td>
</tr>
<tr>
<td>Space group</td>
<td>P4&lt;sub&gt;2&lt;/sub&gt;2</td>
<td>P4&lt;sub&gt;2&lt;/sub&gt;2</td>
</tr>
<tr>
<td>Unit cell dimensions (Å)</td>
<td>a = b = 73.5, c = 84.1</td>
<td>a = b = 43.9, c = 120.1</td>
</tr>
<tr>
<td>Resolution (Å)*</td>
<td>44–2.5 (2.64–2.5)</td>
<td>44–2.2 (3.23–2.2)</td>
</tr>
<tr>
<td>Total observations</td>
<td>65429</td>
<td>49981</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>8474</td>
<td>6500</td>
</tr>
<tr>
<td>R&lt;sub&gt;merge&lt;/sub&gt;*</td>
<td>20.6 (2.9)</td>
<td>27.9 (3.3)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>100 (100)</td>
<td>99.7 (98.1)</td>
</tr>
<tr>
<td>R&lt;sub&gt;free&lt;/sub&gt;*</td>
<td>0.074 (0.611)</td>
<td>0.042 (0.511)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>37–2.5</td>
<td>41–2.2</td>
</tr>
<tr>
<td>R-factor/F三位 (%)*</td>
<td>20.2/22.6</td>
<td>22.3/27.7</td>
</tr>
<tr>
<td><strong>Number of atoms</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-H protein</td>
<td>826</td>
<td>825</td>
</tr>
<tr>
<td>Water</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Ligand</td>
<td>28 (glycerol, sulfate)</td>
<td>0</td>
</tr>
<tr>
<td>Average B (Å&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>58.3</td>
<td>67.8</td>
</tr>
<tr>
<td><strong>RMSD</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.003</td>
<td>0.010</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>0.702</td>
<td>1.010</td>
</tr>
<tr>
<td>Ramachandran plot</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% most favoured</td>
<td>98.0</td>
<td>97.9</td>
</tr>
<tr>
<td>% disallowed</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

*Values in parentheses refer to the highest resolution shell. The R-factor is calculated for all measured reflections in the specified resolution range, while the R<sub>free</sub> is calculated for reflections belonging to a random test set not used in refinement (10% of the data). doi:10.1371/journal.pone.0048015.t001
The structures of bcPadR1 and bcPadR2 were solved by molecular replacement using the program Balbes [23] running on the YSBL web server [http://www.ysbl.york.ac.uk/YSBLPrograms/]. The best solutions for the two proteins were found using the structure of a transcriptional regulator from Entenococcus faecalis V583 (PDB entry 3hli) as a search model and by changing the space groups to P41;2;2 and P41;2;2 for bcPadR1 and bcPadR2, respectively. The solutions were subjected to automated model building using ARP/wARP [24] via the ARP/wARP web service [http://cluster.embli-hamburg.de/ARPwARP/remote-http.html], and completed by applying several cycles of manual model building using COOT [25] and restrained-refinement using Phenix.refine [26]. TLS (Translation/Libra-
mation) was included using Phenix TLSlib [27]. The final models of bcPadR1 and bcPadR2 contain 103 residues (amino acids 3–105) and 99 residues (amino acids 3–66 and 72–106), respectively. The geometry of the models was validated by Molprobity [28].

**Structure and sequence analysis.** Structure and sequence analysis was carried out using the following software: SuperPose [29], for superimposition and calculation of root mean square deviations (RMSDs); ESPript [30] and Jalview [http://www.jalview.org] for preparation of sequence alignments; and Pymol (Schrodinger, LLC) for the depiction of structures and preparation of figures.

**Accession numbers**

Structure factors and the coordinates of the final models of bcPadR1 and bcPadR2 have been deposited in the Protein Data Bank [http://www.rcsb.org] with accession numbers 4ESB and 4ESF, respectively.

**Results**

**Gene targets of bcPadR1 and bcPadR2**

Genes associated with resistance mechanisms that are transcriptionally controlled by PadR-like proteins are generally encoded in the vicinity of the padR genes. Figure 1A shows the chromosomal regions around the two B. cereus padR genes, in comparison to the lmrR-lmrCD operon. The gene encoding bcPadR1 (locus BC4206 in B. cereus ATCC 14579) is located immediately downstream from BC4207, encoding a putative membrane protein of unknown function. These genes are predicted to share a common promoter, located immediately upstream from BC4206. Transcriptome analysis revealed that BC4206 and BC4207 are co-upregulated in response to enterocin AS-48 [14], indicating that these genes indeed form a bicistronic transcriptional unit. The same study also revealed that plasmid-based overexpression of BC4206 increased the resistance of B. cereus ATCC 14579 to AS-48, confirming the role of this gene in bacteriocin resistance, although its mechanism of operation is unknown. The BC4206-BC4207 transcriptional unit is conserved among B. cereus and B. thuringiensis species, but not present in the closely related B. anthracis [14]. To further examine the regulatory role of BC4206, we also overexpressed this gene in B. cereus ATCC 14579. Overexpression results in an increased sensitivity against AS-48 (MIC value of 1.0 μg/ml compared to a MIC value of 2.5 μg/ml for the wild-type strain, p-value<0.005, >6 cultures as determined with Student’s t-test) suggesting that bcPadR1 acts as a repressor of the BC4206-BC4207 transcriptional unit, when AS-48 is not present in the medium. Electrophoretic mobility shift assays further reveal that purified bcPadR1 specifically binds to the intergenic region between BC4205 and the BC4206-BC4207 operon (Figure 1B and 1C). Incubation of the DNA probe with 0.5 μM bcPadR1 results in shifted bands, while higher bcPadR1 concentration results in stronger, but more diffused binding (Figure 1C). This might indicate the binding of two or more bcPadR1 dimers to the probe. The DNA probe used in these experiments contains the upstream regulatory region of BC4205 that codes for a putative ssp photoproduc-like. The expression of BC4205 is not altered in response to AS-48 addition [14], so its expression is likely to be unrelated to the AS-48 induced expression of BC4206-BC4207 operon. In vitro DNA binding of bcPadR1 was not affected by the addition of AS-48 suggesting an indirect sensing of AS-48 in the medium, although we cannot exclude the possibility that slightly altered in vitro conditions (e.g. different binding buffer) are needed to observe any effect of AS-48 addition on DNA binding.

The gene encoding bcPadR2 (locus BCE3449 in B. cereus ATCC 10987) lies downstream from two genes, BCE3447 and BCE3446, which encode a putative ABC-efflux pump. The product of BCE3447 forms the ATPase subunit of the transporter, while the product of BCE3446 forms the membrane permease. Interestingly, the amino acid sequence of the BCE3447-BCE3446 encoded pump is similar to that of the daunorubicin/doxorubicin resistance ABC transporter DrrAB from Streptomyces peucetius (overall sequence identity of 32%) [31,32]. The sequence identity of the BCE3447-BCE3446 encoded pump with LmrCD is 27%, but LmrCD is about twice as large as BCE3447-BCE3446 and DrrAB. DNA analysis indicates that the BCE3449-BCE3448-BCE3447-BCE3446 gene organization is conserved in various Bacillus species, but without functional expression data we cannot confirm that these genes form an operon or are coregulated.

**Sequence analysis**

bcPadR1 and bcPadR2 contain 105 and 106 amino acid residues, respectively, and have all the characteristics of the PadR-s2 protein family. Figure 2 shows a multiple sequence alignment of the two proteins with other PadR-s2 proteins of which structures are currently deposited in the PDB, including LmrR. Pair-wise sequence identities range from 20% (bcPadR2 compared with a transcriptional regulator from Clostridium thermocellum, PDB entry 1XMA) to 61% (bcPadR2 compared with a transcriptional regulator from Entenococcus faecalis V583, PDB entry 3HHH). bcPadR1 and bcPadR2 share a sequence identity of 45%, and are 26% and 30% identical in sequence to LmrR, respectively. Using an alignment-derived profile hidden Markov model, all homologs of bcPadR1 and bcPadR2 currently deposited in the Uniprot database were identified (>2150 sequences) and used to derive conserved sequence motifs. The highly conserved residues of the PadR-s2 proteins are mainly located in the wHTH domain and cluster in three regions, (i) at the N-terminus of helix α2 [motif YGYE/L/I], (ii) in helix α3 [motif EGT(L/I)YPxLxRLE] and (iii) in the region encompassing strand β2 of the wing and the N-terminus of helix α2 [motif G(P/R)RKYYx(L/I)TxRxG]. The identified sequence motifs partly overlap with the consensus sequence previously derived for all PadR-like proteins [4]. In addition to the above mentioned conserved sequence motifs, other highly conserved residues are found in more isolated positions at the N-terminus, helix α1 and the C-terminal helix of the subfamily
Figure 1. Genomic neighborhood of the bcPadR1 and bcPadR2 encoding genes and analysis of promoter binding. (A) Organization of the BC4206(bcPadR1)-BC4207 operon of B. cereus ATCC 14579, the putative BCE3449(bcPadR2)-BCE3448-BCE3447-BCE3446 regulon of B. cereus ATCC 10987 and the lmrR-lmrCD regulon of L. lactis. The relative scale of the genes and intergenic regions is proportional to nucleotide length. Boxes indicate the operon/regulon boundaries. The PadR-encoding genes are colored in yellow, while their (putative) target genes, encoding resistance-associated membrane proteins, are in green. (B) The sequence of the intergenic region between BC4205 and BC4206 that is used for the EMSA experiments. Putative -35 and -10 promoter sequences of BC4206 are underlined, the start codons of BC4205 and BC4206 are highlighted in bold, and the amino acids of the coded proteins are indicated above the DNA sequence. The putative binding site of bcPadR1 (homologous to the canonical ATGT/ACAT inverted sequence motif) is highlighted in blue. (C) EMSA experiments with bcPadR1. DNA fragments encompassing the promoter regions of BC4206 (lanes 1–10) and BC4029 (lanes 11–13, negative control) were prepared by PCR and end labeled with 32P. DNA binding was assayed as described in the Materials and Methods. Lanes 1, 10 and 11 contain the free DNA probe. Samples run on lanes 2 and 12 contain 0.5 μM; lane 3 1 μM; lanes 4–9 and 11 2 μM of purified bcPadR1 protein. Lanes 5 to 9 contains samples including increasing concentrations of AS-48 from 0.5 μM to 0.69 μM.

doi:10.1371/journal.pone.0048015.g001
2 PadR-like proteins, including the central tryptophan residue in the C-terminal helix which has a central role in multidrug-binding in LmrR.

Structure determination

To allow easy production of pure protein for crystallization, bcPadR1 and bcPadR2 were expressed with a C-terminal streptavidin-tag (SRWSHPQFEK) using an L. lactis nisin-overexpression system. Protein purification was carried out in two steps using Strep-Tactin and DNA affinity chromatography. Gel filtration experiments revealed that the recombinant bcPadR1 and bcPadR2 proteins have an apparent molecular weight of 43 kDa, similar as observed previously for LmrR and consistent with the presence of elongated dimers in solution (the calculated molecular weight of monomeric strep-tagged bcPadR1 and bcPadR2 is 13.3 kDa and 13.7 kDa, respectively).

BcPadR1 was crystallized in space group P4_12_21, with one protein molecule per asymmetric unit and a solvent content of 71%. Similarly, bcPadR2 was crystallized in space group P4_12_21, with one protein molecule per asymmetric unit and a solvent content of 42%. The bcPadR structures were determined by molecular replacement using the structure of the PadR-like protein from E. faecalis V583 as a search model. Model building and refinement were somewhat hampered by relatively high atomic displacement parameters (B-factors), indicative of some global disorder or flexibility of the proteins, even though the overall fit to the electron density is well defined. The final refined structures of bcPadR1 and bcPadR2 have an R-factor/R free of 20.2/22.6% at 2.5 Å resolution and 22.3/27.7% at 2.2 Å resolution, respectively, with good geometry (see Table 1 for details of the refinement statistics). The two structures are highly similar with a Cα-backbone root mean square deviation (RMSD) of 1.45 Å (for 96 common residues). The N-termini (residues 1–2 in both bcPadR proteins) and the C-terminal strep-tags (residues 106–115 in bcPadR1 and residues 108–116 in bcPadR2) were not included in the models due to weak or absent electron density. In addition, residues 67–71 in the β-wing of bcPadR2 were found to be disordered and therefore left out from the model.

Overall structures of bcPadR1 and bcPadR2

In the crystals, bcPadR1 and bcPadR2 form similar dimers, with a dimeric 2-fold axis that coincides with a crystallographic 2-fold
axis. The approximate dimensions of the dimers are 80 Å × 34 Å × 30 Å (Figure 3A). The bcPadR1 and bcPadR2 subunits contain a typical wHTH DNA-binding domain, consisting of helices α1, α2, the DNA recognition helix α3 and strands β1 and β2 (together forming the wing), and a dimerization domain, including the C-terminal helix α4. Helix α4 forms a protruding arm which in the dimer interacts with helices α4’ and α1’ of the dimer-mate. Dimerization of bcPadR1 and bcPadR2 buries approximately 1740 Å² solvent-accessible surface area per subunit which is related to ~30% of the total surface area. The dimeric interface involves mainly van der Waals contacts between hydrophobic residues, most of which are conserved between the two bcPadR proteins (see Figure 2). In addition a few hydrogen bonds are observed at the dimer interface, but these are not conserved.

**BcPadR1 and bcPadR2 show a closed dimeric interface**

Although similar in overall fold, it is evident that the two bcPadR structures do not show the same dimeric interface as LmrR (Figures 3 and 4). Unlike LmrR, but like the other PadR-s2 proteins with known structure, the bcPadR1 and bcPadR2 dimers have a completely closed dimer interface. The differences at the dimer interface are associated with changes in the relative orientation of helices α1 and α4 with respect to the core of the wHTH domain (Figure 4B). In addition, helix α4 in bcPadR1 and bcPadR2 shows significant bending towards the dimeric centre, whereas the equivalent helix in LmrR is almost perfectly straight (Figure 4C). As a consequence, helices α4 and α4’ in the bcPadR dimers can approach each other closely at the dimer interface, allowing a coiled-coil interaction mode. In LmrR the C-terminal helices do not make any mutual contacts, but, instead, interact strongly with helices α1 and α2 of the dimer mate. Intriguingly, despite these significant differences in dimeric organization, the residues in bcPadR proteins and LmrR that form the inter-subunit interactions are located at similar positions in the structures (Figure 2). It is unclear how the differences in amino acid sequence at these regions affect the conformation of the polypeptides in the dimer. In particular, the highly conserved tryptophan residues in the C-terminal helices of the bcPadR dimers (Trp91/Trp91’ and Trp93/Trp93’ in bcPadR1 and bcPadR2, respectively), which are the equivalent of Trp96/Trp96’ in LmrR, are completely buried in the dimer interface. The closed dimer conformation and inaccessibility of the conserved tryptophan residues of the bcPadR proteins are incompatible with a drug-binding mechanism as observed in LmrR.

**Structural similarity with non-PadR-like transcription factors**

A structural similarity search against available structures in the Protein Data Bank (PDB) using the DALI server [33] reveals that bcPadR1 and bcPadR2 are not only similar to other PadR-like proteins, but also to various non-PadR-like transcription factors including members of the MarR family, SmtB/ArsR family and the MecI/BlaI family (Table 2). These proteins all contain
wHTH-domains for DNA binding and form functional dimers. The similarities of the non-PadR-like structural homologs of \(bc\) PadR1 and \(bc\) PadR2 are mainly limited to the wHTH domains, which can be superimposed to those of \(bc\)PadR1 and \(bc\)PadR2 with RMSDs (for Cα-backbone superpositions) of 1.1–1.9 Å (Figure 4A). In contrast, the dimerization topologies of the non-PadR-like structural homologs of \(bc\) PadR1 and \(bc\) PadR2 are highly diverse, often including additional helices at the N- and C-terminal regions. One notable exception is the transcription factor RTP (replication terminator protein) from \(Bacillus subtilis\), which plays a crucial role in the coordinated termination of DNA replication [34]. Despite the absence of significant sequence homology, \(bc\) PadR1 and \(bc\) PadR2 show a highly similar overall topology and dimeric organization as RTP, including the presence of a single C-terminal helix for dimerization (Figure 3). The striking similarities of the \(bc\)PadR and RTP structures strongly suggest that these proteins use a common mode of DNA binding.

**Putative DNA-binding interactions**

To understand how \(bc\)PadR1 and \(bc\)PadR2 may bind to DNA, the \(bc\)PadR structures were superimposed to structures of close homologs in complex with DNA (Figure 5). At present no structure is available of a PadR-like protein bound to its cognate DNA. However, DNA-bound structures are available for two of the non-PadR-like close homologs identified by the DALI search, i.e., the replication terminator protein RTP (PDB entry 1F4K) and the transcriptional repressor BlaI (PDB entry 1XSD) [34,35]. RTP and BlaI exploit common principles to fulfill their role as DNA-binding regulatory proteins, which most probably are also used by the \(bc\)PadR proteins. The operator sequences of RTP and BlaI

---

**Table 2. DALI-identified structural homologs of \(bc\)PadR1 and \(bc\)PadR2.**

<table>
<thead>
<tr>
<th>PDB-ID</th>
<th>Z-score</th>
<th>RMSD (Å)</th>
<th>Cα aligned</th>
<th>Nr. res</th>
<th>% id.</th>
<th>Protein name</th>
<th>Protein family</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3HHH</td>
<td>16.1/17.7</td>
<td>1.6/1.7</td>
<td>101/101</td>
<td>103/103</td>
<td>42/61</td>
<td>PadR-s2</td>
<td>u.</td>
<td></td>
</tr>
<tr>
<td>3L7W</td>
<td>14.2/13.6</td>
<td>2.0/2.7</td>
<td>99/100</td>
<td>106/106</td>
<td>27/27</td>
<td>PadR-s2</td>
<td>u.</td>
<td></td>
</tr>
<tr>
<td>3R2</td>
<td>14.0/14.0</td>
<td>2.5/1.6</td>
<td>99/100</td>
<td>112/112</td>
<td>30/26</td>
<td>PadR-s2</td>
<td>u.</td>
<td></td>
</tr>
<tr>
<td>1XMA</td>
<td>13.9/14.8</td>
<td>2.6/1.5</td>
<td>97/98</td>
<td>103/101</td>
<td>22/20</td>
<td>PadR-s2</td>
<td>u.</td>
<td></td>
</tr>
<tr>
<td>1LNW</td>
<td>12.8/12.0</td>
<td>2.7/3.0</td>
<td>93/88</td>
<td>141/141</td>
<td>13/16</td>
<td>MexR</td>
<td>MarR</td>
<td>[40]</td>
</tr>
<tr>
<td>3F8B</td>
<td>12.5/11.7</td>
<td>3.1/2.7</td>
<td>100/98</td>
<td>105/101</td>
<td>23/30</td>
<td>LmrR</td>
<td>PadR-s2</td>
<td>[9]</td>
</tr>
<tr>
<td>1YG2</td>
<td>11.0/10.6</td>
<td>3.2/2.8</td>
<td>83/82</td>
<td>169/169</td>
<td>22/24</td>
<td>AphA</td>
<td>PadR-s1</td>
<td>[8]</td>
</tr>
<tr>
<td>1F4K</td>
<td>10.6/11.0</td>
<td>3.1/2.5</td>
<td>98/98</td>
<td>115/115</td>
<td>12/15</td>
<td>RTP</td>
<td>RTP</td>
<td>[34]</td>
</tr>
<tr>
<td>1SMTP</td>
<td>10.1/9.5</td>
<td>2.6/2.9</td>
<td>78/77</td>
<td>98/101</td>
<td>14/19</td>
<td>SmTB</td>
<td>SmTB/ArsR</td>
<td>[41]</td>
</tr>
<tr>
<td>1SD6</td>
<td>9.7/9.6</td>
<td>3.3/3.1</td>
<td>84/84</td>
<td>119/119</td>
<td>15/19</td>
<td>Blal</td>
<td>MecI/Blal</td>
<td>[35]</td>
</tr>
</tbody>
</table>

Abbreviations: PadR-s1, PadR-like subfamily-1; PadR-s2, PadR-like subfamily-2; n.k., not known; u, unpublished.

*The DALI-search was performed using a single subunit of each bcPadR dimer.

doi:10.1371/journal.pone.0048015.t002
contain specific inverted repeat motifs and the proteins bind such that their dimeric dyads are aligned with the inverted repeat dyads of the DNA operators. In this way, the recognition helices ($a_3$) of the dimer-related wHTH domains can insert into two successive major grooves of the DNA, while the $b$-wings lie across the phosphate backbone, within the vicinity of adjacent minor groove clefts. DNA binding is stabilized by electrostatic interactions between the highly positively charged surface of the DNA binding site and the negatively charged DNA backbone. Additional DNA:protein interactions comprise various non-bonding contacts, including hydrophobic interactions, and a few hydrogen bond interactions. Residues in the $a_3$-helix are responsible for specific DNA binding and recognition of the operator DNA sequence, by penetrating into the major groove and forming hydrogen bonds with bases of the nucleotides.

The $bcPadR1$ dimer can be superimposed to the DNA-bound complexes of RTP and BlaI with RMSD values of 2.9 Å (for 179 common C$\alpha$ atoms) and 3.4 Å (for 155 common C$\alpha$ atoms). Overall, the dimers of $bcPadR1$ and $bcPadR2$ are highly similar to RTP and BlaI, but there are small differences in the relative dispositions of the dimer-related subunits which affect the relative orientations of the $a_3$ helices. Since DNA-binding by transcription factors is often associated with specific conformational changes in both protein and DNA, the superpositions cannot be used to accurately predict the details of the DNA binding mechanism of $bcPadR1$ and $bcPadR2$. However, analysis of the models in combination with a structure-based sequence alignment of $bcPadR1$, $bcPadR2$, RTP and BlaI (Figure 5) leads to a few interesting observations. First, like in RTP and BlaI, the complementary surfaces in $bcPadR1$ and $bcPadR2$ that form the putative DNA binding sites are highly common C$\alpha$ atoms divided equally between both subunits) and 4.1 Å (for 153 common C$\alpha$ atoms), respectively. Likewise, the $bcPadR2$ dimer can be fitted to the DNA-bound complexes of RTP and BlaI with RMSD values of 2.9 Å (for 179 common C$\alpha$ atoms) and 3.4 Å (for 155 common C$\alpha$ atoms). Overall, the dimers of $bcPadR1$ and $bcPadR2$ are highly similar to RTP and BlaI, but there are small differences in the relative dispositions of the dimer-related subunits which affect the relative orientations of the $a_3$ helices. Since DNA-binding by transcription factors is often associated with specific conformational changes in both protein and DNA, the superpositions cannot be used to accurately predict the details of the DNA binding mechanism of $bcPadR1$ and $bcPadR2$. However, analysis of the models in combination with a structure-based sequence alignment of $bcPadR1$, $bcPadR2$, RTP and BlaI (Figure 5) leads to a few interesting observations. First, like in RTP and BlaI, the complementary surfaces in $bcPadR1$ and $bcPadR2$ that form the putative DNA binding sites are highly

Figure 5. Modeling of the DNA-bound $bcPadR$ complexes. (A) Superposition of the $bcPadR1$ (light-green) and $bcPadR2$ (red) dimers onto DNA-bound RTP (orange, PDB code 1F4K). (B) Superposition of the $bcPadR1$ (light-green) and $bcPadR2$ (red) dimers onto DNA-bound BlaI (cyan, PDB code 1XSD). Below the superpositions are structure-based sequence alignments including secondary structures. Residues of RTP and BlaI which participate in DNA binding are indicated with small spheres using the following color scheme: green, interacting with the phosphate backbone; blue, interacting with the base moiety and yellow, interacting with the ribose backbone). doi:10.1371/journal.pone.0048015.g005
Figure 6. The bound sulfate anion in the bcPadR1 structure. (A) Interactions between bcPadR1 residues and the sulfate ion. Also shown is the DNA-bound RTP structure (PDB entry 1F4K) onto which the bcPadR1 dimer was superimposed. The bcPadR1 residues and the sulfate ion are shown in ball-and-stick representation and labeled, while the DNA strand of the RTP-DNA complex is drawn using lines. (B) Interactions between the RTP residues and the DNA phosphate backbone adjacent to the sulfate anion in the superimposed bcPadR1 structure. The RTP residues and DNA strands are shown in ball-and-stick representation and labeled, while the bcPadR1 residues and sulfate are drawn using lines.

doi:10.1371/journal.pone.0048015.g006

positively charged, due to the presence of various lysine and arginine residues. A number of these lysine and arginine residues from part of the previously described conserved sequence motifs in the PadR-s2 proteins, in particular those located in helix α3 and the β-wing, strongly implying a direct role in DNA binding. Furthermore, the conserved YGYE motif in bcPadR1 and bcPadR2, which is located at the N-terminal end of helix α2, is partly preserved in RTP (residues 33–36 with sequence YGLK). In RTP, the side chain of Tyr33, together with the backbone amides of Gly34 and Leu35, forms hydrogen bonds with the phosphate-backbone of the DNA. As these interactions are crucial for the strong DNA binding affinity of RTP [36], it is likely that the conserved residues in bcPadR1 and bcPadR2 have a similar DNA binding role. This is corroborated by the observation of a bound sulfate anion in the bcPadR1 structure, which forms interactions with the highly conserved residues Gly25, Tyr26, Glu42 Arg71, and Lys72, likely mimicking the interactions with a phosphate moiety of the DNA backbone (Figure 6). Finally, the α3 helices of the bcPadR proteins show a limited number of conserved residues relative to the equivalent helices in RTP and BlaI. Interestingly, two residues of the conserved EGTLYPxLxRLE motif are identical to sequence-specific DNA binding residues in RTP or BlaI, i.e., Tyr46 and Arg51, strongly suggesting that these residues in helix α3 of the bcPadR proteins have a similar DNA binding role. However, the limited overall sequence conservation of the α3 helices is also indicative of the differences that must exist between the specific cognate operator sequences of these DNA-binding proteins.

Discussion

In this study, we have determined the crystal structures of bcPadR1 and bcPadR2, which are the products of gene loci BC4206 and BCE3449 of B. cereus strains ATCC 14579 and ATCC 10987, respectively. Functional data point toward the hypothesis that bcPadR1 autoregulates the transcription of its own gene, as bcPadR1 binds to the promoter region of gene BC4206 and overexpression of BC4206 gene results in increased sensitivity against AS-48. The structures of bcPadR1 and bcPadR2 are similar to LmrR, containing a conserved N-terminal wHTH DNA binding domain and a C-terminal helix involved in dimerization. The conserved structures and sequence relationships confirm that all three proteins belong to the same subfamily of small PadR proteins (subfamily 2 according to Huillet et al. [4], which we abbreviated as PadR-s2.) The identification of conserved sequence motifs in the wHTH domains, in particular in the α3 DNA recognition helix, further suggests that the PadR-s2 proteins share a common cognate DNA operator sequence. Operator sites of LmrR in the lmrR and lmrCD promoter regions have been identified previously [5,12], and were shown to contain an ATGT/ACAT inverted sequence motif, separated by a short spacer of about 8–10 nucleotides. The inverted repeat is identical to the conserved DNA sequence motif revealed for members of the other subfamily of PadR proteins (PadR-s1 family) [1,37]. Indeed, similar sequence motifs, although somewhat degenerate, are present in the promoter of the BC4206–4207 operon [ATGT-(N)8-ATAT, data not shown].

Functional studies and footprinting analysis are required to fully establish the identity of the target DNA operator sites and to identify other possible target genes controlled by bcPadR1 and bcPadR2.

Compared to related transcription factors of the MarR, SmtB/ArsR and MecI/BlaI families, the bcPadR proteins show a “minimal” fold, lacking extra N- and C-terminal helices associated with dimerization. In the non-PadR-like proteins the extra helices often play a role in the mechanisms of transcriptional induction, allowing modulation of DNA binding activities through direct binding of a ligand. Also the structural features in LmrR that are related with its role as a multi-drug induced transcription regulator, i.e., a large central pore with two opposing tryptophan residues for multi-drug binding, are notably absent in bcPadR1 and bcPadR2. These observations suggest that the DNA binding activities of bcPadR1 and bcPadR2 are not directly modulated by ligands. In line with this hypothesis, in vitro DNA binding of bcPadR1 was not altered by the presence of AS-48. This is not surprising since AS-48 is known to target the cell membrane, inducing membrane pore formation [14], and is unlikely to enter the cytoplasm. Perhaps, the bcPadR proteins undergo a conformational change, which opens the dimer interface, thereby forming a ligand binding site. Alternatively, transcriptional regulating by the bcPadR proteins may involve an indirect induction mechanism with the participation of another protein. Such a mechanism has been proposed for the phenolic acid stress
response in *L. plantarum*, which is regulated by the founding member of the PadR-like transcription factors and involves a protein belonging to the universal stress protein family [30,31]. Examining such possibilities for the mechanism of action of the *bc*PadR proteins has to await experimental data characterizing their regulatory functions and identifying the target genes which they control.

**References**


**Acknowledgments**

We thank the staff of the ESRF (Grenoble) for providing facilities for diffraction measurements and for assistance.

**Author Contributions**

Conceived and designed the experiments: GF ATK OPK AMWHT. Performed the experiments: GF ATK TJP MB AMWHT. Analyzed the data: GF ATK TJP MB AMWHT. Contributed reagents/materials/analysis tools: OPK AMWHT. Wrote the paper: GF ATK OPK AMWHT.