Journal of Microbiological Methods 90 (2012) 46–52

Contents lists available at SciVerse ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

DNA binding site analysis of *Burkholderia thailandensis* response regulators

Kristy L. Nowak-Lovato [a], Alexana J. Hickmott [a], Tuhin S. Maity [a], Martha L. Bulyk [b,c,d], John Dunbar [a], Elizabeth Hong-Geller [a,*]

[a] Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM 87545, United States
[b] Division of Genetics, Department of Medicine, Brigham & Women's Hospital and Harvard Medical School, Boston, MA 02115, United States
[c] Department of Pathology, Brigham & Women's Hospital and Harvard Medical School, Boston, MA 02115, United States
[d] Harvard-MIT Division of Health Sciences and Technology, Harvard Medical School, Boston, MA 02115, United States

ARTICLE INFO

Article history:
Received 19 March 2012
Received in revised form 27 March 2012
Accepted 27 March 2012
Available online 13 April 2012

Keywords:
Response regulator
Protein-binding microarray
Promoter binding site
Burkholderia
Two component systems
Gene regulation

ABSTRACT

Bacterial response regulators (RR) that function as transcription factors in two component signaling pathways are crucial for ensuring tight regulation and coordinated expression of the genome. Currently, consensus DNA binding sites in the promoter for very few bacterial RRs have been identified. A systematic method to characterize these DNA binding sites for RRs would enable prediction of specific gene expression patterns in response to extracellular stimuli. To identify RR DNA binding sites, we functionally activated RRs using beryllofluoride and applied them to a protein-binding microarray (PBM) to discover DNA binding motifs for RRs expressed in *Burkholderia*, a Gram-negative bacterial genus. We identified DNA binding motifs for conserved RRs in *Burkholderia thailandensis*, including KdpE, RisA, and NarL, as well as for a previously uncharacterized RR at locus BTH_II2335 and its ortholog in the human pathogen *Burkholderia pseudomallei* at locus BPSS2315. We further demonstrate RR binding of predicted genomic targets for the two orthologs using gel shift assays and reveal a pattern of RR regulation of expression of self and other two component systems. Our studies illustrate the use of PBMs to identify DNA binding specificities for bacterial RRs and enable prediction of gene regulatory networks in response to two component signaling.

Published by Elsevier B.V.

1. Introduction

Bacteria employ two-component signaling systems to couple the sensing of stress signals to adaptive changes in gene expression, thus ensuring tight regulation and coordinated expression of the genome in response to the environment (Beier and Gross, 2006; Cheung and Hendrickson, 2010; Laub and Goulian, 2007). Two-component systems represent the single largest paralogous family of signaling proteins in the bacterial kingdom and regulate diverse cellular processes, including chemotaxis, osmoregulation, metabolism, and transport. As the name implies, the prototypical two-component system is composed of two parts. First, a histidine kinase catalyzes autophosphorylation on a conserved histidine residue upon sensing changes in growth conditions and then transfers the phosphoryl group to the receiver domain of the second component, a response regulator (RR), which functions as a downstream effector protein, often as a transcription factor that regulates gene expression (Bourret et al., 1991; Pao and Saier, 1995). For example, under K^+ limitation, the histidine kinase KdpD activates the response regulator KdpE, which in its phosphorylated state, induces expression of the *kdpFABC* operon *via* increased affinity for a 23 base pair sequence in the *kdpFABC* promoter. The *kdpFABC* operon, which lies adjacent to the *kdpDE* operon, encodes an inducible high-affinity K^+ uptake system that scavenges K^+ to maintain ionic homeostasis in the cell (Gasell and Altendorf, 2001).

The rapid sequencing of bacterial genomes in the last several years has revealed a diversity of RRs with undefined regulatory functions. From 1123 distinct bacterial genomes, ~39,000 two-component proteins adjacent in the genome have been identified (Ulrich and Zhulin, 2010). The majority of RRs with DNA binding capability fall into three major families based on the structural similarity of their effector domains, (1) OmpR/PhoB family, winged helix-turn-helix motif (Kenney, 2002), (2) NarL family, helix-turn-helix motif (Baikalov et al., 1996), and (3) NtrC family, ATPase domain (Yang et al., 2004). Although the target genes of some RRs can be predicted based on genomic organization, such as KdpE control of *kdpFABC*, RRs can regulate multiple target genes scattered throughout a bacterial genome. The completion of sequenced bacterial genomes has enabled bioinformatics searches using consensus sequence motifs to predict DNA binding sites for specific RRs. Thus far, experimental confirmation of DNA binding sites for RRs has been limited. Aside from KdpE, DNA binding sites have been determined for the *Escherichia coli* RRs OmpR (Pratt and Silhavy, 1995), NarL (Baikalov et al., 1996; Maris et al., 2002), and PhoB (Makino et al., 1988), which regulate osmolarity, nitrate response, and phosphate availability, respectively. However, the great majority of bacterial RRs has been identified based only on sequence

* Corresponding author. Tel.: +1 505 665 2465; fax: +1 505 665 3024.
E-mail address: ehong@lanl.gov (E. Hong-Geller).

0167-7012/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.mimet.2012.03.019

homology, and their target DNA binding sites remain unknown or poorly characterized. Further experimental identification of target DNA binding sites and the cognate genes regulated by specific RRs can link extracellular inputs (*e.g.* nutrient deprivation, ion concentration, pH change) to a regulatory gene network and better define the molecular mechanisms activated in response to two-component signaling pathways.

We have chosen Gram-negative bacteria *Burkholderia* spp. as the model organism for discovery of RR DNA binding sites. The *Burkholderia* genus encompasses ~60 species, which exhibit a wide range of biological functions, including pathogenicity, bioremediation, and nitrogen fixation. The two best-characterized species, *Burkholderia pseudomallei* and *Burkholderia mallei*, the causative agents of human melioidosis and equine glanders, respectively, are categorized as Category B biothreat agents by the CDC. We have employed protein-binding microarray (PBM) technology to determine the DNA binding specificities of RRs expressed in *Burkholderia thailandensis*, a closely-related species to *B. pseudomallei* that is non-pathogenic in humans. The PBM is a rapid methodology to simultaneously screen all sequence variants of a defined length and obtain comprehensive binding site measurements of DNA–protein interactions *in vitro* (Berger and Bulyk, 2009; Berger et al., 2006; Mukherjee et al., 2004). PBMs have been successfully used to analyze transcription factor binding specificities in a wide variety of organisms, including *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, mice, and humans (Berger et al., 2008, 2006; Robasky and Bulyk, 2011). To date, use of PBMs in bacterial systems has been limited to a specific quorum sensing RR, LuxR, in the marine bacterium *Vibrio harveyi* (Pompeani et al., 2008), two nucleoid-associated proteins, H-NS and Lsr2, from *Salmonella enterica* and *Mycobacterium tuberculosis* (Gordon et al., 2011), and several TetR and MarR transcription factors in *Burkholderia xenovorans* (Maity et al., 2011). In this study, we demonstrate the successful application of PBMs to both known and previously-uncharacterized *Burkholderia* RRs, as a broadly applicable method to identify bacterial transcription factor binding sites for analysis of gene regulation in a wide range of bacterial species. We also perform comparative PBM analysis between a pair of RR orthologs in *B. thailandensis* and *B. pseudomallei* to investigate the overlap of DNA binding specificities in different *Burkholderia* species. We expect that identification of RR DNA binding sites in *Burkholderia* can provide molecular insights into how two-component systems monitor different environmental parameters and allow for prediction of cellular behavior across bacterial species.

2. Materials and methods

2.1. Cloning and expression of GST fusions to RRs

The RR genes were amplified from *Burkholderia* genomic DNA using sequence-specific primers by PCR in 50 μl reactions [1 μl 100 μM primer 1, 1 μl 100 μM primer 2, 50 ng genomic DNA isolated from *B. thailandensis* E264 or *B. pseudomallei* K96243, 5 μl 10× Pfu reaction buffer, 1 μl 100 mM dNTPs, 2.5 μl DMSO, 2.5 U of PfuUltra DNA polymerase (Agilent, Santa Clara, CA), and distilled H_2O for the remaining volume] using the following conditions, (1) 94 °C, 3 min, (2) 94 °C, 1 min; 50 °C, 1 min, 72 °C, 1 min for 30 cycles, and (3) 94 °C, 1 min; 50 °C, 1 min, 72 °C, 10 min. The primers introduced 5′ BamHI and 3′ HindIII restriction sites for cloning. The primer sets used were: (1) KdpE (BTH_I1025) (F) GATCGGATCCGAATGCCCATGAGTGAACCGACCGTCACC and (R) GATCAAGCTTTCAGCCCGCGCCGACGAGCCGGTAGCC, (2) PhoB (BTH_I1267) (F) GATCGGATCCATGCCCAGCAACATTCTCGTCATCGAA and (R) GATCAAGCTTTTACGCGTGTTTCGCGAGCCGGTA, (3) OmpR (BTH_I2094) (F) GATCGGATCCATGGAAACGAAAAACCCCTCCAAG, and (R) GATCAAGCTTTCAGGCCGCGCCGTCGGGGATGAA (4) NarL (BTH_I1849) (F) GATCGGATCCATGACCATACGCGTACTGTTGATCGAC, and (R) GATCAAGCTTTCAGGCCTCGGCCGGATGCGGCGC, (5) RisA (BTH_I2094) (F) GATCGGATCCATGGAAACGAAAAACCCCTCCAAG and (R) GATCAAGCTTTCAGGCCGCGCCGTCGGGGATGAA, (6) BTH_II2335 (F) GATCGGATCCATGACCACCGTTTCTTCCACGCCCCGC and (R) GATCAAGCTTCTACCGCCTGCGATGCTCCACCGCGAA, and (7) BPSS2315 (F) GATCGGATCCATGACTCCTGCCTCTTCCACGCCCCGC and (R) GATCAAGCTTCTACCGCCTGCGATGCTCGACCGCGAA.

The RR genes were cloned as N-terminal GST fusions into the pGEX-KG vector using T4 ligase (NEB, Ipswich, MA), transformed into BL21 *E. coli* competent cells, and induced for protein expression with 1 mM IPTG for 4 h. Cells were lysed with 1 mg ml^{-1} lysozyme on ice for 30 min, followed by treatment with 10 μg ml^{-1} DNase and 10 mM $MgCl_2$ for an additional 30 min, and centrifugation at 40,000 rpm for 1 h. GST fusion proteins were purified from the cleared supernatants by incubation with agarose beads cross-linked to glutathione for 1 h and eluted with 50 mM Tris–Cl (pH 8), 10 mM reduced glutathione. Protein samples were then dialyzed using a Slide-alyzer cassette (Thermo Scientific, Rockford, Il.) with a 10,000 MW cut-off to remove free glutathione, quantified using the BCA protein assay (Thermo Scientific), and stored at −80 °C in a final concentration of 30% glycerol.

2.2. Gel shift assays

To demonstrate BeFx-mediated enhancement of RR binding, the *pstS* and *nar* promoter regions were PCR-amplified for use as target DNA in gel shift assays. The following primers were used for PCR: (1) *pstS* promoter (F) ATCGGCCGGACAGGCCGG and (R) GAGACCTCCAGTGTGTGA and (2) *nar* promoter (F) GATCGGATCCCGACATCGTGAGACGAAGCCG and (R) GATCAAGCTTGACGATTCTCTCGAGACGAGG. For the *cstA* ($BTH\_II2252^{-156,-130}$) promoter and internal histidine kinase ($BTH\_II2334^{+447,+468}$) gel shift assays, each set of complementary oligonucleotides, (1) *cstA* (F) TGCTACGTAGCGGCCATACGTAGTTCC and (R) GGAACTACGTATGGCCGCTACGTAGCA, (2) BTH_II2334 (F) GGCTACGTGCGCTACGTCTGG and (R) CCAGACGTAGCGCACGTAGCC, and (3) non-specific oligos, (F) CGAGGGAGAATGATCGTTCTACCCTT and (R) AAGGGTAGAACGATCATTCTCCCTCG, was placed in a heat block at 95 °C for 5 min followed by removal of the heat block to the benchtop. The temperature of the heat block was allowed to decrease to room temperature to allow for oligonucleotide annealing.

Binding reactions (20 μl) containing indicated concentrations of GST–RR fusion proteins, 1 μM of target DNA sequences, 100 μM $BeCl_2$, 10 mM NaF, and 2 μl of 10× binding buffer (20 mM Tris–Cl pH 7.5, 0.5 mM EDTA, 5% glycerol, 1 mM DTT, 0.005% Triton X-100, 50 mM NaCl, 5 mM $MgCl_2$, and 2.5 mM $CaCl_2$), were incubated for 30 min at room temperature. Binding reactions were separated on 8% non-denaturing polyacrylamide gels run in 0.5× TBE buffer and visualized with Sybr Green DNA stain (Life Technologies, Grand Island, NY) using a ChemiDoc gel documentation system (Bio-Rad, Hercules, CA).

2.3. Protein-binding microarrays

A minimum of two PBMs were performed for each RR as previously described (Berger and Bulyk, 2009) with modifications. Briefly, microarrays were obtained from Agilent Technologies in a 4×44 K format, AMADID #015681 and #016060 (cat # G2514F). We performed primer extension from a universal 24-mer region to generate a double-stranded microarray platform. GST fusion proteins were diluted to a final concentration of 125 nM in a volume of 175 μl (PBS, 2% milk, 200 μg ml^{-1} BSA, 0.3 μg ml^{-1} salmon testes DNA) in individual chambers of a four chamber gasket coverslip. In addition, we included 2 μM $BeCl_2$, 200 μM NaF and 1× binding buffer in all incubation and buffer washing steps to maintain activation of the RRs and an optimized ionic environment during protein binding to the microarray. Microarrays were scanned (GenePix Pro 4200A, Axon Instruments, Sunnyvale, CA) to detect specific DNA–RR interactions at multiple