Tuberculosis xxx (2015) 1–10

Contents lists available at ScienceDirect

Tuberculosis

journal homepage: http://intl.elsevierhealth.com/journals/tube

MOLECULAR ASPECTS

The mycobacterial PhoH2 proteins are type II toxin antitoxins coupled to RNA helicase domains

Emma S.V. Andrews, Vickery L. Arcus*

School of Science, University of Waikato, Hamilton 3240, New Zealand

ARTICLE INFO

Article history:
Received 8 January 2015
Accepted 29 March 2015

Keywords:
Toxin-antitoxin
RNA helicase
RNAse
PIN-domain
PhoH2

SUMMARY

PhoH2 proteins are found in a diverse range of organisms that span the bacterial tree and little is known about this large protein family. PhoH2 proteins have two domains: An N-terminal PIN domain fused to a C-terminal PhoH domain. The genome of *Mycobacterium tuberculosis* encodes 48 PIN domains and 47 of these constitute the VapC components of the 47 VapBC toxin-antitoxins. The 48th member of the *M. tuberculosis* PIN domain array is found in the single PhoH2 protein encoded in the genome. All characterized PIN domain proteins are RNases and the PhoH domains are predicted ATPases. This fusion of a PIN domain with an ATPase reflects a much wider association between PIN domains and PhoH domains across many prokaryote genomes. Here, we examine PhoH2 proteins from *M. tuberculosis*, *Mycobacterium smegmatis* and a thermophilic homologue from *Thermobispora bispora* and we show that PhoH2 is a sequence-specific RNA helicase and RNAse. In addition, *phoH2* from *M. tuberculosis* and *M. smegmatis* is part of a longer mRNA transcript which includes a small, unannotated open reading frame (ORF) upstream of the *phoH2* gene. This small gene overlaps with the beginning of the *phoH2* gene in a manner similar to the PIN domain toxin-antitoxin operons. We have annotated the upstream gene as *phoAT* and its putative promoter elements satisfy previously characterized consensus sequences at the −10 site. Conditional growth experiments carried out in *M. smegmatis* revealed a negative effect on growth by the expression of *M. tuberculosis* PhoH2 that was alleviated by co-expression of the PhoAT peptide. Thus in *M. tuberculosis*, PhoH2 represents a new variation on a type II PIN domain toxin-antitoxin systems such that the toxin-antitoxin is now coupled to an RNA helicase whose predicted biological function is to unwind and cleave RNA in a sequence specific manner.

1. Introduction

The causative agent of tuberculosis in humans is *Mycobacterium tuberculosis* and this organism must survive a number of different stressors (e.g. hypoxia and nutrient stress) in order to persist within its host. The genome of *M. tuberculosis* has 48 proteins belonging to the PIN domain family [4] and in 47 cases, these domains are part of a VapBC toxin-antitoxin operon. Many researchers have hypothesized that the VapBC operons are associated with the stress response and persistence [5,14,22,25]. The remaining PIN domain belongs to the PhoH2 superfamily.

The PhoH2 superfamily of proteins comprise two domains: An N-terminal PIN domain coupled to a C-terminal PhoH domain. PFam lists 1341 sequences with PhoH2 architecture (PIN-PhoH fusion) [20] across many diverse bacterial and archaeal phyla and the superfamily is annotated as COG 1875 [12,42] at NCBI. There are no homologues amongst the eukaryotes. Aside from a single PhoH domain structure deposited in the PDB from the Midwest Center for Structural Genomics (PDB code 3B85), almost nothing is known about this large family of prokaryotic proteins.

Thus far all characterized PIN domain proteins are RNases [2] and in 95% of cases exist as single domain proteins found in the genome with an upstream transcription factor that forms a toxin-antitoxin operon. In this context, the PIN domains are labeled VapC (toxin) and the upstream antitoxin is annotated as VapB, forming a VapBC toxin-antitoxin complex. One of the most extensively characterized VapC proteins from a VapBC system

Abbreviations: AAA/AAA+ ATPases, associated with diverse cellular activities; ATP, adenosine triphosphate; IPTG, isopropylthio-β-D-galactosidase; KH, domain K-homology domain; NTP, nucleotide triphosphate; ORF, open reading frame; PIN domain, PilT N-terminal domain; PVDF, polyvinylidene fluoride; RNAP, RNA polymerase; SF, superfamily; SPITC, 4-sulfophenyl isothiocynate; TRAM, tRNA modifying enzyme; Vap, virulence associated protein.

* Corresponding author. Tel.: +64 7 8384679.
E-mail address: varcus@waikato.ac.nz (V.L. Arcus).

http://dx.doi.org/10.1016/j.tube.2015.03.013
1472-9792/

(MSMEG_1283/4 from *Mycobacterium smegmatis*) is a post-transcriptional regulator of metabolism by virtue of its sequence-specific mRNase activity [34]. In the remaining ~5% of cases where PIN domain proteins are found they are fused with either TRAM domains, KH domains or AAA+ ATPase domains [4]. TRAM and KH domains are annotated as RNA-binding domains [3,36] and the AAA+ ATPase domains belong to the AAA+ superfamily of ring-shaped, P-loop NTPases [27]. AAA+ ATPase proteins generally oligomerize into hexameric rings and are involved in remodeling or translocating macromolecules in an energy dependent manner.

To examine the function of PhoH2 proteins, and the function of a PIN domain protein in the context of a protein fusion we investigated PhoH2 from *M. tuberculosis* and *M. smegmatis*. For practical purposes, we have also biochemically characterized a non-mycobacterial PhoH2 homologue from a thermophile (PhoH2 from *Thermobispora bispora*, 77% sequence identity to PhoH2 from *M. tuberculosis*).

2. Materials and methods

2.1. Cloning for protein expression

The gene sequences for *phoH2 (Rv1095)* and *phoAT-phoH2* were amplified from *Mycobacterium bovis* BCG strain Pasteur 1173P2 genomic DNA (100% identity to *phoH2* and *phoAT-phoH2* from *M. tuberculosis*). The products were digested with *Nco*I/*Bam*HI restriction enzymes and ligated into the pYUB1049 shuttle vector [23]. Ligations were transformed into *Escherichia coli* TOP10 cells and were selected by plating on low salt LB agar medium supplemented with 0.05% (v/v) Tween-80 and 50 μg/ml hygromycin B. Constructs were sequenced prior to transformation into *M. smegmatis* $mc^2$4517 cells using previously published methods [7,23]. The same approach was taken to clone *phoH2* (*MSMEG_5247*) and *phoAT-phoH2* from *M. smegmatis* into modified pYUB1049 plasmid pYUB28b [6] for expression in *M. smegmatis* $mc^2$4517. *MSMEG_5247* was also cloned into a modified in-house pET28b plasmid, pET28b–*Pst*I between *Pst*I/*Hind*III for expression in *E. coli*. Ligations were transformed into *E. coli* DH5α and plated out onto LB agar medium supplemented with 50 μg/ml kanamycin. Constructs were sequenced prior to transformation into *E. coli* protein expression strain BL21. For the cloning of the thermophilic PhoH2 homologue from *T. bispora* the amino acid sequence for *TBIS_3092* was synthesized by GeneArt® and cloned into pET28b–*Pst*I between the sites *Pst*I and *Hind*III as above described above.

2.2. Protein expression and purification for mass spectrometry and activity assays

For mycobacterial protein expression, a single transformed colony was used to inoculate a PA-0.5G seeder culture supplemented with 0.05% (v/v) Tween-80 and 50 μg/ml kanamycin and hygromycin B and was cultured at 37 °C for 48 h. These seeder cultures were used at a 1:100 dilution to inoculate a ZYP-5052 autoinduction expression media and the cultures were grown at 37 °C for 96 h. Cells were harvested by centrifuging at 4600 rpm at 4 °C for 20 min. For purification, cells were resuspended in lysis buffer (50 mM Na-phosphate buffer pH 7.4, 200 mM NaCl, 20 mM imidazole), sonicated on ice and harvested by centrifugation. The soluble fractions containing His-tagged PhoAT-PhoH2 from *M. tuberculosis*, PhoAT-PhoH2 or PhoH2 from *M. smegmatis* were purified by IMAC on a HisTrap HP column (GE Healthcare, UK). When required the protein fractions were purified further by size exclusion chromatography, using an S200 10/300 Superdex™ column (GE Healthcare, UK) in the same buffer. For *E. coli* protein expression, a single transformed colony was used to inoculate an LB seeder culture supplemented with 50 μg/ml kanamycin. This culture was grown for 24 h at 37 °C and was used at a 1:100 dilution to inoculate an LB expression culture supplemented with 50 μg/ml kanamycin. These cultures were incubated at 37 °C and were induced with a final concentration of 1 mM IPTG at an OD_{600} of 0.4–0.6 and further incubated with shaking at 37 °C overnight. Cells from large scale expression cultures were harvested and purified as described for mycobacterial protein expression in 50 mM TRIS buffer pH 8, 200 mM NaCl. For ATPase activity assays, PhoH2 proteins were purified in 50 mM Na-HEPES buffer pH 7.4, 200 mM NaCl.

2.3. RT-PCR

Total RNA from *M. bovis* BCG strain Pasteur 1173P2 and *M. smegmatis* $mc^2$155 was extracted using a phenol-GITC method for the isolation of RNA. To prepare cDNA, 1 μg of RNA was treated with 1 U of RQ1 RNase-Free DNase (Promega, USA) in the presence of 10 mM TRIS-HCl pH 7, 0.5 mM $MnCl_2$ according to manufacturer's instructions. Synthesis of cDNA was carried out using the Superscript III First-Strand Synthesis system (Invitrogen, USA) according to kit instructions using 100 ng random hexanucleotides and reactions with (+) and without (−) Superscript III to exclude DNA contamination. *M. bovis* BCG and *M. smegmatis* cDNA (±RT) were used as templates for PCR reactions with primer combinations that spanned the upstream sequence of *Rv1095* and *MSMEG_5247*, including observed possible alternative 5′ translational start sites upstream of *phoH2* using the primers listed in Supplementary Table 1. PCR products were analyzed by agarose gel electrophoresis and the experiment was repeated at least 3 times to confirm the presence or absence of +RT products.

2.4. Construction of ΔphoAT phoH2 knockout in M. smegmatis

An unmarked deletion of *phoAT-phoH2* was created by a two-step allelic exchange mutagenesis [37]. For this purpose a construct containing 735 bp and 736 bp regions flanking the *phoAT-phoH2* gene on the left and right respectively using primers listed in Supplementary Table 1, was cloned into pX33 to yield pX33 *phoAT-phoH2* LFRF. This plasmid was transformed into *M. smegmatis* $mc^2$155 and transformants were selected at 28 °C in the presence of 5 μg/ml gentamycin. For deletion of *phoAT-phoH2*, strains carrying pX33 *phoAT-phoH2* LFRF were grown in the presence of gentamycin at 42 °C to select for integration of the plasmid into the chromosome of *M. smegmatis* via a single crossover event. Colonies were screened for integration by exposure to 250 mM catechol. Selected colonies were grown in LBT medium at 37 °C and aliquots of these cultures were plated onto low salt (2 g/l NaCl) LBT plates containing 10% sucrose and incubated at 42 °C to select for a second crossover event leading to the loss of the plasmid and deletion of *phoAT-phoH2*. Colonies were screened for loss of the plasmid with 250 mM catechol and candidate mutants were screened by PCR using primers that flanked the deletion site.

2.5. Conditional growth experiments

To create tetracycline inducible expression constructs the genes for *phoH2* and *phoAT-phoH2* from *M. tuberculosis* and *M. smegmatis* were amplified from genomic DNA using the pMIND primers listed in Supplementary Table 1. The PCR products were digested and ligated between the *Bam*HI/*Spe*I sites of the tetracycline-inducible plasmid pMIND [13] and were transformed into *E. coli* TOP10 cells for selection on low salt LB medium supplemented with 50 μg/ml hygromycin B. The forward primers introduced a synthetic RBS sequence upstream of the *phoH2* gene to ensure translation.

Please cite this article in press as: Andrews ESV, Arcus VL, The mycobacterial PhoH2 proteins are type II toxin antitoxins coupled to RNA helicase domains, Tuberculosis (2015), http://dx.doi.org/10.1016/j.tube.2015.03.013