The second messenger, bis-(3',5')-cyclic dimeric guanosine monophosphate (cyclic di-GMP), is involved in the control of multiple bacterial phenotypes, including those that impact host-pathogen interactions. Bioinformatics analyses predicted that Mycobacterium leprae, an obligate intracellular bacterium and the causative agent of leprosy, encodes three active diguanylate cyclases. In contrast, the related pathogen Mycobacterium tuberculosis encodes only a single diguanylate cyclase. One of the M. leprae unique diguanylate cyclases (ML1419c) was previously shown to be produced early during the course of leprosy. Thus, functional analysis of ML1419c was performed. The gene encoding ML1419c was cloned and expressed in Pseudomonas aeruginosa PAO1 to allow for assessment of cyclic di-GMP production and cyclic di-GMP-mediated phenotypes. Phenotypic studies revealed that ml1419c expression altered colony morphology, motility and biofilm formation of P. aeruginosa PAO1 in a manner consistent with increased cyclic di-GMP production. Direct measurement of cyclic di-GMP levels by liquid chromatography–mass spectrometry confirmed that ml1419c expression increased cyclic di-GMP production in P. aeruginosa PAO1 cultures in comparison to the vector control. The observed phenotypes and increased levels of cyclic di-GMP detected in P. aeruginosa expressing ml1419c could be abrogated by mutation of the active site in ML1419c. These studies demonstrated that ML1419c of M. leprae functions as diguanylate cyclase to synthesize cyclic di-GMP. Thus, this protein was renamed DgcA (Diguanylate cyclase A). These results also demonstrated the ability to use P. aeruginosa as a heterologous host for characterizing the function of proteins involved in the cyclic di-GMP pathway of a pathogen refractory to in vitro growth, M. leprae.

INTRODUCTION

The bacterial metabolite bis-(3',5')-cyclic dimeric guanosine monophosphate (cyclic di-GMP) serves as a signal-transducing second messenger (Römling et al., 2013). Cyclic di-GMP has been described in a broad range of bacterial species and is associated with regulation of a variety of bacterial activities including survival (Bharati et al., 2012), virulence (Kulasakara et al., 2006), cell differentiation (Abel et al., 2011) and biofilm formation (Hickman et al., 2005; Irie et al., 2012; Borlee et al., 2010). Pathogen-produced cyclic di-GMP has also been found to stimulate the host innate immune response (Karaolis et al., 2007). The formation of cyclic di-GMP is catalysed via diguanylate cyclase (DGC), an enzyme that utilizes GTP as the substrate and possesses a conserved GG(D/E)EF motif that is part of the active-site domain (Chan et al., 2004; Wassmann et al., 2007). Approximately 50% of proteins with characterized DGC activity (Römling et al., 2013) are regulated via allosteric control and the binding of cyclic di-GMP to a conserved inhibitory site (I-site) motif (RxXD) directly upstream of the GGDEF motif (Christen et al., 2006; Wassmann et al., 2007). Phosphodiesterase (PDE) proteins that possess EAL or HD-GYP domains are responsible for the depletion of cyclic di-GMP and conversion into 5'-phosphoguanylyl-(3',5')-guanosine (pGpG) or two molecules of guanosine monophosphate (GMP), respectively (Ryan et al., 2006; Schmidt et al., 2005). Recently, oligoribonuclease (Orn) has also been identified to participate in the conversion of pGpG into GMP in Pseudomonas aeruginosa (Cohen et al., 2015; Orr et al., 2015). The activities of GGDEF and EAL domains can be incorporated into a single bifunctional protein or separated on individual proteins. More importantly,
sensory input domains typically located in the N-terminus of proteins containing DGC or PDE domains receive environmental signals that activate the DGC or PDE activity (Römling et al., 2013; Sondermann et al., 2012). The potential combinatorial diversity of various sensory input domains coupled to individual DGC or PDE activity or bifunctional proteins allows organisms such as *Pseudomonas* spp. to utilize cyclic di-GMP in the regulation of multiple activities such as biofilm formation, motility and virulence (Hickman et al., 2005; Kulasakara et al., 2006; Merritt et al., 2007). The genome of *P. aeruginosa* PA01 encodes 41 proteins predicted to metabolize intracellular cyclic-di-GMP levels. This includes 17 GGDEF domain-containing proteins, 5 EAL-domain-containing proteins, 16 proteins that contain composite GGDEF-EAL domains and 3 proteins with HD-GYP domains (Kulasakara et al., 2006; Ryan et al., 2009; Römling et al., 2013; Mills et al., 2011).

In mycobacteria, the production and regulatory activity of cyclic di-GMP has been investigated in *Mycobacterium tuberculosis*, the causative agent of tuberculosis (Gengenbacher & Kaufmann, 2012). This bacterium encodes a single bifunctional DGC-PDE protein (Rv1354c) and both domains have been shown to be active (Gupta et al., 2010). Disruption of Rv1354c PDE activity decreases the pathogenicity and dormancy in *M. tuberculosis* (Hong et al., 2013). In addition, studies of cyclic di-GMP PDE activity of Rv1357c in the closely related *Mycobacterium bovis* BCG Pasteur 1173P2 demonstrated that cyclic di-GMP was associated with the regulation of lipid production and pellicle growth and promoted resistance to nitrosative stress (Flores-Valdez et al., 2015). Protein interaction studies have also suggested that cyclic di-GMP production in *M. tuberculosis* is involved in regulation of rhamnose biosynthesis, a key sugar in the formation of the mycobacterial cell wall (Deng et al., 2014). Moreover, studies of the *Mycobacterium smegmatis* homologue of rv1354c demonstrated that the cyclic di-GMP was involved in colony morphology and long-term survival during nutrient starvation (Sharma et al., 2014; Bharati et al., 2012; Gupta et al., 2015).

The production and regulatory role of cyclic di-GMP is unknown in *Mycobacterium leprae*, which is an acid-fast bacterium and the causative agent of leprosy or Hansen’s disease. Leprosy remains a public health concern in several low- and middle-income countries, with over 200 000 new cases reported each year (WHO, 2014). Leprosy is a chronic disease and the causative agent of leprosy or Hansen’s disease. Leprosy remains a public health concern in several low- and middle-income countries, with over 200 000 new cases reported each year (WHO, 2014). Leprosy is an infectious disease caused by the bacterium *M. leprae* and presents with a high bacillary load and skin and peripheral nerves (Walker & Lockwood, 2007). The chronic nature of leprosy and the ability of *M. leprae* to adapt to various host environments is contradictory to a bacterium with a degenerative genome, as well as a reduced number of transcription factors and regulatory proteins (Cole et al., 2001). Bioinformatics analyses of the annotated *M. leprae* genome revealed three putative DGC proteins, which all retain the conserved GGDEF domain. One of the putative DGC coding sequences of *M. leprae*, ML1419c, was previously shown to be expressed when *M. leprae* was experimentally infected in the mouse footpad (Williams et al., 2009). More importantly, immunological studies revealed a strong antigen-specific T cell response to peptides of ML1419c in paucibacillary patients and in the household contacts of multibacillary patients (Spencer et al., 2005). This demonstrated that *M. leprae* produces ML1419c in the early stages of leprosy; however, the physiological function of this protein was unknown.

The inability to culture *M. leprae* in vitro or subject it to genetic manipulation is a major impediment in characterizing the physiological function of *M. leprae* proteins. To overcome these limitations, we chose to express *ml1419c* in an alternative heterologous host, *P. aeruginosa* PA01, in which the effects of increasing and decreasing levels of cyclic di-GMP are well documented and control a range of phenotypes (Borlee et al., 2010; Hickman & Harwood, 2008; Hickman et al., 2005; Irie et al., 2012; Kulasakara et al., 2006; Merritt et al., 2007; Lee et al., 2007; Mills et al., 2011). Previously published studies have specifically shown that overexpression of the DGC encoded by *tpbB* (PA1120) increases cyclic di-GMP concentrations and corresponding biofilm formation in *P. aeruginosa* (Hickman & Harwood, 2008; Kulasakara et al., 2006). This study demonstrated the DGC activity of ML1419c from *M. leprae* in *P. aeruginosa* and provides strong evidence that *M. leprae* is capable of producing the second messenger cyclic di-GMP.

**METHODS**

**Bacterial strains, genomic DNA and growth conditions.** *P. aeruginosa* PA01 and recombinant strains were grown at 37°C in either Lennox LB medium or Vogel–Bonner minimal medium (VBM4) with L-arabinose (Gold Biotechnology) for inducible expression experiments. *Escherichia coli* strain BL21(DE3) (Invitrogen) was grown in Miller LB medium at 37°C. Gentamicin (100 µg ml⁻¹) (Gold Biotechnology) was used for selection of recombinant *P. aeruginosa* strains. Kanamycin (50 µg ml⁻¹) (Sigma Aldrich) and gentamicin (10 µg ml⁻¹) were used for selection of recombinant *E. coli* strains. Genomic DNA from *M. leprae* strains NHDP63 and Thai-53 was obtained from Biodefense and Emerging Infections Research Repository.

**Construction of plasmids and site-directed mutagenesis of ** *ml1419c*. Heterologous gene expression in *P. aeruginosa* PA01 was accomplished by using the arabinose-inducible vector pN105 containing Pbad promoter (Newman & Fuqua, 1999). *M. leprae* ml1419c was amplified by PCR from genomic DNA of *M. leprae* with Q5 high-fidelity DNA polymerase (New England Biolabs), the forward primer 5'-GGATTCCAGGGAGGATATTGGTTGAGACCGTGCGTAG-3' and the reverse primer 5'-GGACTAGTTCAACTAGTTGTTGTTGATCGT-3'. Underlined sequences represent EcoRI and SphI sites in

---

**Downloaded from www.microbiologyresearch.org by**

<table>
<thead>
<tr>
<th>IP: 54.70.40.11</th>
</tr>
</thead>
<tbody>
<tr>
<td>On: Thu, 13 Dec 2018 12:14:00</td>
</tr>
</tbody>
</table>
the forward and reverse primers, respectively. Boldface sequence represents inclusion of an optimized ribosome binding site (Wolgast et al., 2003). The 1692-bp \textit{ml1419c} fragment was cloned into p\textit{pN105} (Newman & Fuqua, 1999) using the EcoRI and SpeI sites to generate p\textit{MRB105}.

To express \textit{ml1419c} in \textit{E. coli} for protein purification, we performed PCR amplification using \textit{M. leprae} genomic DNA, as well as the forward primer 5'-AACCATGTTGGAACGGTGCTGATGCCG-3' and the reverse primer 5'-TTAACCTGGCAGTTGTGGTGAACGG-3'. Underlined sequences represent the NdeI and HindIII sites, respectively. The resulting PCR product was cloned into the expression vector pET-28a (+) at NdeI and HindIII sites to generate p\textit{MRB109}.

Site-directed mutagenesis of \textit{ml1419c} was accomplished with the QuikChange Lightning Site-Directed Mutagenesis (Agilent Technologies) following the manufacturer's recommendations. The forward primer 5'-GTTGGTGGGTAGGTTCGTCGCTCTGATCCTG-3' and the reverse primer 5'-CAGGATCAGAGCGACGAACCTACCCACCAC-3' were used to generate \textit{ML1419c} sequences encoding proteins with a deletion of the \textit{cG}GDE\textit{F}K\textit{EM} motif (\textit{ML1419c}_{\text{cGDEFKEM}}). The \textit{ml1419c}_{\text{cGDEFKEM}} construct was digested with EcoRI and SpeI endonucleases and cloned into the p\textit{pN105} to generate p\textit{MRB108}. The plasmid p\textit{pN1120} (Hickman & Harwood, 2008) expresses \textit{pBB} (PA1120), a well-characterized DGC from \textit{P. aeruginosa}, was used as a positive control (Kulasakara et al., 2006; Ueda & Wood, 2009) and p\textit{pN105} was used as a negative control.

All constructs were confirmed by nucleotide sequencing. Plasmids used in this study are shown in Table S1 (available in the online Supplementary Material).

**Protein and whole cell lysate isolation.** Recombinant \textit{ML1419c} was purified from \textit{E. coli} BL21(DE3). \textit{E. coli} transformed with p\textit{MRB109} was cultured to an \textit{OD}_{600} of ~0.4–0.6 and induced with 0.5 mM IPTG (EMD Millipore) for 3 h. Cells were lysed using an ultrasonic processor (Vibra-cell VC750) with an amplitude setting of 30 % and a 1 min linear gradient to 100 % buffer B and a 1 min linear gradient to 10 % solvent B (methanol).

** SDS-PAGE and Western blot analysis.** Purified \textit{E. coli} recombinant ML1419c (500 ng or 25 ng), \textit{P. aeruginosa} whole cell lysates (5 µg), soluble fractions (5 µg) and insoluble fractions of proteins (5 µg) were resolved under denaturing conditions on NuPAGE 4–12 % Bistris polyacrylamide gels (Invitrogen). Proteins were visualized by staining with Coomassie G-250 stain (Invitrogen). For Western blots, proteins were transferred to PVDF membranes and probed with rabbit anti-ML1419c polyclonal serum (1:200 000). The secondary antibody was anti-rabbit IgG conjugated to HRP (1:20 000) (Promega). Chemiluminescence was used to visualize reactive proteins by incubating membrane with luminol-based enhanced chemiluminescence HRP substrate (Thermo Scientific).

**Phenotypic assays.** Colony morphology and dye binding. \textit{P. aeruginosa} strains were grown on VBM as agar containing 40 µg ml\(^{-1}\) Congo red (Sigma Aldrich), 15 µg ml\(^{-1}\) Coomassie brilliant blue (Sigma Aldrich) and 100 µg ml\(^{-1}\) gentamicin and in the presence or absence of 1 % L-arginine at 30 °C for 48 h. The colony morphology was observed under a Leica MZ95 stereo microscope (Leica Microsystems).

**Motility assays.** Swimming and twitching motility were assessed as previously described (Dzielé et al., 2001; Darzins, 1993) with minor modifications. Swimming motility was assayed by stab inoculating 1 µl of a \textit{P. aeruginosa} overnight culture into low-viscosity Lennox LB agar (0.3 % Bacto agar) containing with 100 µg ml\(^{-1}\) gentamicin and 0.2 % L-arginine. The diameters of the swimming zone were measured after growth at 37 °C for 24 h. Twitching motility was assayed by stab inoculating a colony of \textit{P. aeruginosa} through Lennox LB agar containing 100 µg ml\(^{-1}\) gentamicin and 0.2 % L-arginine and cultured at 37 °C for 48 h. The migration of bacteria attached on the polystyrene plate surface was visualized by staining with 0.1 % crystal violet and the diameters were measured.

**Biofilm formation:** Biofilm formation was assessed as previously described (O’Toole, 2011) with modifications. The log-phase \textit{P. aeruginosa} cultures (\textit{OD}_{600} of 0.6–0.7) were adjusted to an \textit{OD}_{600} of ~0.1 in VBM containing 100 µg ml\(^{-1}\) gentamicin and 0.2 % L-arginine. Aliquots (150 µl) of diluted cultures were added to 96-well polystyrene plates (Nunc Microwell 96-well microplates #9243656, Thermo Scientific) in replicates of six and incubated at 37 °C in a sealed bag for 24 h. The 96-well plates were washed twice with water, stained with 0.1 % crystal violet for 10 min and washed twice with water. Bound crystal violet was solubilized with 30 % acetic acid and the absorbance was measured at 590 nm.

**Quantitative analysis of cyclic di-GMP by liquid chromatography–mass spectrometry (LC-MS).** Overnight cultures of \textit{P. aeruginosa} were diluted with VBM (1:100) containing 100 µg ml\(^{-1}\) gentamicin and 0.2 % L-arginine. An aliquot (2 ml) of \textit{P. aeruginosa} culture grown to \textit{OD}_{600} of ~0.6–0.7 was extracted with 100 µl of 0.6 % (final concentration) perchloric acid (Sigma Aldrich) (Hickman & Harwood, 2008) spiked with 100 nM [\(^{35}\)S]Adenosine (Omicron Biochemicals). The precipitate of the perchloric acid extraction was used for determination of protein concentration by the BCA assay and sample normalization (Irie & Parsek, 2014) and the extracts were neutralized with 20 µl of 2.5 M potassium bicarbonate. The neutralized supernatants were stored at ~80 °C.

Cyclic di-GMP extracts (10 µl) were applied to an Agilent 1200 HPLC system coupled to an Agilent 6520 quadrupole time-of-flight (Q-TOF) mass spectrometer. Samples were resolved on an Atlantis T3 C18 column (3 µm particle size, 2.1 × 150 mm, Waters) at a flow rate of 350 µl min\(^{-1}\) at 30 °C. The gradient consisted of 100 % solvent A (0.1 % acetic acid, 10 mM ammonium acetate in water) for 1 min followed by a 1 min linear gradient to 10 % solvent B (methanol), a 1.4 min linear gradient to 20 % buffer B and a 1 min linear gradient to 100 % buffer B and held for 1 min. The Q-TOF mass spectrometer was operated in positive ion mode at 2 GHz extended dynamic range, an m/z range from 100 to 1700 and at a scan rate of one spectrum per second. MS data were collected in profile and centroid mode. Electrospray ionization source parameters were 2.5 kV, 350 °C gas temperature, drying gas flow rate of 11 ml min\(^{-1}\) and a nebulizer flow rate of 45 psi. The cyclic di-GMP peak was confirmed by LC-tandem mass spectrometry (MS/MS) using collision energy of 19.8 eV. The cyclic di-GMP standard was purchased from BIOLOG Life Science Institute. The relative
RESULTS

Bioinformatics analyses and identification of cyclic di-GMP-related proteins in *M. leprae*

Identification of putative DGC proteins (GGDEF) or cyclic di-GMP PDE proteins (EAL or HD-GYP) encoded in the genome of *M. leprae* strain TN was achieved by interrogation of the NCBI CDD and the SMART databases. Bioinformatics analyses revealed that *M. leprae* harbours one putative hybrid DGC-PDE protein (ML1750c), two putative DGC proteins (ML1419c and ML0397c) and one putative PDE protein with an EAL domain (ML1752c) (Fig. 1). No proteins with a HD-GYP motif were identified. Results from these bioinformatics analyses of *M. leprae* strain TN agree with the large-scale census of cyclic di-GMP-related proteins (Römling et al., 2013).

The hybrid DGC-PDE protein ML1750c possesses an N-terminal GAF sensor domain, as well as GGDEF and EAL domains (Fig. 1a). This protein is homologous to *M. smegmatis* MSEG_2196 with 64.39% identity and *M. tuberculosis* Rv1354c with 62.30% identity, both of which were experimentally defined as possessing active GGDEF and EAL domains (Gupta et al., 2010; Bharati et al., 2012; Hong et al., 2013). Unexpectedly, *M. leprae* with its reduced genome encoded two additional DGC proteins, ML1419c and ML0397c (Fig. 1b, c), that are not encoded by *M. smegmatis* or *M. tuberculosis*. ML1419c possesses three sequential PAS-signalling domains N-terminal to the GGDEF domain. Two predicted haem-binding sites were also identified within the N-terminal region of the ML1419c PAS-signalling domains. ML0397c harbours a single N-terminal PAS sensor domain linked to a GGDEF domain and 10 C-terminal transmembrane domains as predicted by analysis with the TMHMM program. A single predicted PDE protein (ML1752c) is homologous to Rv1357c of *M. tuberculosis* (Gupta et al., 2010; Römling et al., 2013), but it has no homology to proteins encoded by *M. smegmatis* (Fig. 1d).

Multiple alignments of the *M. leprae* proteins containing GGDEF and EAL domains (Figs 1e and S1a) were performed with T-Coffee (Di Tommaso et al., 2011). This demonstrated conservation of key amino acids in putative active sites. The three predicted DGCs of *M. leprae* (ML1750c, ML1419c and ML0397c) all contain the A-site sequence, RxGGDEF (Ryjenkov et al., 2005). Thus, these proteins are expected to act as functional DGCs producing cyclic di-GMP. In addition, ML1419c and ML0397c possess an I-site, RxxD motif (Wassmann et al., 2007), that is located directly upstream of the A-site. Multiple alignment analyses for the proteins with predicted PDE activity and EAL domains revealed that the conserved residues of the EAL active site are present in ML1750c and ML1752c. This conservation included appropriately spaced residues of E, N, E, D, K and E, except for the last E residue of ML1752c that is replaced with a K residue (Römling et al., 2013) (Fig. S1b). These *in silico* data indicate that *M. leprae* has a greater capacity than *M. tuberculosis* or *M. smegmatis* for cyclic di-GMP production. Based on immunological data, one of the putative DGCs (ML1419c) is known to be produced *in vivo* by *M. leprae* (Spencer et al., 2005; Williams et al., 2009), and thus, it was selected for further assessment of DGC activity.

Conditional expression of *ml1419c* in *P. aeruginosa* PAO1

*M. leprae* cannot be cultured *in vitro*; thus, protein functions for this bacterium are typically studied in model organisms (Monot et al., 2009). The production and function of cyclic di-GMP has been extensively studied in *P. aeruginosa*, where phenotypes and mutants associated with this second messenger molecule are well described. In this study, *ml1419c* and a mutated construct of this gene, *ml1419c<sub>AGGDEF</sub>* were conditionally expressed in *P. aeruginosa* PAO1 under the control of the L-arabinose-responsive $P_{BAD}$ promoter (Newman & Fuqua, 1999). Recombinant protein production was assessed by SDS-PAGE and Western blot of whole cell lysates of *P. aeruginosa* strains (Fig. 2). Protein of the expected size (approximately 61 kDa) that reacted with anti-ML1419c polyclonal serum was observed in the whole cell lysates of *P. aeruginosa* expressing *ml1419c* or *ml1419c<sub>AGGDEF</sub> when grown in the presence of arabinose (Fig. 2b). Moreover, *ml1419c* and *ml1419c<sub>AGGDEF</sub> were expressed in *P. aeruginosa* and produced soluble protein (Fig. S2). In the absence of arabinose, no or low levels of recombinant protein production was observed. No products reactive to the anti-ML1419c polyclonal serum were observed for *P. aeruginosa* containing the pJN105 vector control regardless of the presence or absence of arabinose.
ml1419c alters P. aeruginosa colony morphology

To provide an initial assessment of whether expression of ml1419c resulted in cyclic di-GMP production in P. aeruginosa PAO1, we investigated the colony morphology of recombinant P. aeruginosa strains on VBMM Congo red and brilliant blue agar plates (Fig. 3). P. aeruginosa PAO1 typically forms round colonies with smooth surfaces and regular borders; however, increased intracellular cyclic di-GMP levels induced formation of small colonies with wrinkly or rugose colony morphology and increased Congo red and brilliant blue binding that is correlated to the increase of exopolysaccharide production (Starkey et al., 2009; Hickman et al., 2005). The recombinant P. aeruginosa PAO1 conditionally expressing ml1419c in the presence of arabinose resulted in small and wrinkled colonies exhibiting rugose morphology. These colonies were similar in appearance to the positive control of P. aeruginosa expressing tpbB, a well-studied DGC. Colony morphology of recombinant P. aeruginosa PAO1 harbouring the arabinose-inducible ml1419c expression plasmid is similar to wild-type colony morphology when arabinose is omitted from the growth medium (Fig. S3). The first two glycine residues of the GGDEF motif confirmed DGCs participate in binding of the GTP substrate and the glutamic acid binds Mg\(^{2+}\) that is necessary for DGC activity (Chan et al., 2004; Wassmann et al., 2007). Thus, it was expected that alteration of P. aeruginosa colony morphology by ml1419c expression would be abrogated by an in-frame deletion of 472GGDEF476 (ml1419c\(_{\Delta GGDEF}\)). P. aeruginosa expressing ml1419c\(_{\Delta GGDEF}\)
resulted in larger colonies that resembled the smooth colony morphology of wild-type *P. aeruginosa* PAO1 or vector control (Fig. 3). These data indicated that recombinant ML1419c functions as a DGC and the GGDEF domain of this protein is essential for this activity.

To provide better assessment of potential cyclic di-GMP production by ML1419c and the impact of this production on *P. aeruginosa*, we measured several quantifiable phenotypes (twitching motility, swimming motility and biofilm formation). Swimming was assessed by quantitatively measuring the swim zone diameter formed by bacteria from the point of inoculation in low-viscosity agar (O’Toole & Kolter, 1998). In contrast, twitching motility was quantified by measuring the migration of bacteria that were inoculated at a single point between the agar and the polystyrene petri dish (Darzins, 1993; Déziel *et al.*, 2001). Similar to *P. aeruginosa* expressing *tpbB*, the expression of *ml1419c* suppressed swimming and twitching motility (Figs 4a, b and S4a, b). This is consistent with increased production of cyclic di-GMP and corresponding inhibition of bacterial flagella and type IV pili function (Merritt *et al.*, 2007; Simm *et al.*, 2004). In concordance with decreased DGC activity, alteration of the GGDEF domain in ML1419c (ML1419c<GGDEF>) resulted in swimming and twitching motility that was similar to wild-type PAO1 and the vector control (Figs 4a, b and S4a, b). The suppression of the swimming and twitching phenotypes of *P. aeruginosa* expressing *ml1419c* was significant when compared to wild-type PAO1 and strains expressing the

---

**Fig. 2.** Expression of *ml1419c* from *M. leprae* in *P. aeruginosa* PAO1. Recombinant *P. aeruginosa* PAO1 containing *ml1419c*, *ml1419c<GGDEF>*, or the pJN105 vector (VC) were grown in the presence (+) or in the absence (−) of 0.2% L-arabinose. Whole cell lysates (5 µg) of the recombinant strains were analysed by SDS-PAGE with Coomassie blue staining (a) and Western blot (b). Purified recombinant ML1419c produced in *E. coli* was used as a positive control (lane P) for Coomassie blue staining (500 ng) and Western blot (25 ng).

**Fig. 3.** Colony morphology of *P. aeruginosa* PAO1 and recombinant strains. Strains were grown on VBMM agar containing Congo red, brilliant blue and 1% L-arabinose. Rugose colonies were observed in *P. aeruginosa* expressing *tpbB* and *ml1419c*. PAO1 wild-type, *P. aeruginosa* expressing *ml1419c<GGDEF>* and the vector control (VC) form round colonies with smooth surfaces. Scale bar corresponds to 1 mm.

**ml1419c** expression provides for quantifiable phenotypic differences associated with DGC activity

To provide better assessment of potential cyclic di-GMP production by ML1419c and the impact of this production on *P. aeruginosa*, we measured several quantifiable phenotypes (twitching motility, swimming motility and biofilm formation). Swimming was assessed by quantitatively measuring the swim zone diameter formed by bacteria from the point of inoculation in low-viscosity agar (O’Toole & Kolter, 1998). In contrast, twitching motility was quantified by measuring the migration of bacteria that were inoculated at a single point between the agar and the polystyrene petri dish (Darzins, 1993; Déziel *et al.*, 2001). Similar to *P. aeruginosa* expressing *tpbB*, the expression of *ml1419c* suppressed swimming and twitching motility (Figs 4a, b and S4a, b). This is consistent with increased production of cyclic di-GMP and corresponding inhibition of bacterial flagella and type IV pili function (Merritt *et al.*, 2007; Simm *et al.*, 2004). In concordance with decreased DGC activity, alteration of the GGDEF domain in ML1419c (ML1419c<GGDEF>) resulted in swimming and twitching motility that was similar to wild-type PAO1 and the vector control (Figs 4a, b and S4a, b). The suppression of the swimming and twitching phenotypes of *P. aeruginosa* expressing *ml1419c* was significant when compared to wild-type PAO1 and strains expressing the
ΔGGDEF mutation in ml1419c (Fig. 4a, b). It was also noted that suppression of motility was only observed when P. aeruginosa expressing ml1419c was grown in the presence of arabinose.

Another well-documented activity of cyclic di-GMP in P. aeruginosa is the induction of biofilm formation (Hickman et al., 2005; Lee et al., 2007). Elevated cellular levels of cyclic di-GMP increase the production of biofilm matrix components in P. aeruginosa including the Pel and Psl polysaccharides and a biofilm-associated adhesin (Hickman et al., 2005; Hengge, 2009; Borlee et al., 2010). Under conditions of arabinose induction, P. aeruginosa expressing ml1419c or tpbB produced significantly more biofilm as compared to wild-type PAO1 and the vector control (Fig. 4c). As observed with the other phenotypic assays, P. aeruginosa biofilm formation was significantly reduced when the GGDEF motif was deleted from ML1419c (ML1419c<sub>GAGGDEF</sub>) (Fig. 4c). The effect of ml1419c on multiple phenotypes of P. aeruginosa PAO1 associated with increased levels of cyclic di-GMP production provides strong evidence that the M. leprae ML1419c functions as a DGC. Additionally, the deletion of GGDEF motif of ML1419c reduced or eliminated ml1419c induction of phenotypes associated with elevated cellular levels of cyclic di-GMP.

**Detection of cyclic di-GMP in vivo by LC–MS**

To directly assess the DGC function of ML1419c in the P. aeruginosa genetic background, we performed LC–MS to detect and measure the relative abundance of cyclic di-GMP in P. aeruginosa extracts. Initial analyses of cyclic di-GMP standard and [13C]adenosine (applied as an internal standard) demonstrated that these two products eluted with retention times of 7.061 and 7.702 min and yielded m/z values of 691.1021 and 269.1065, respectively. MS/MS fragmentation of cyclic di-GMP resulted in transition ions of m/z 152.0577, 248.0786 and 540.0566. These transition ions were used to confirm that the parent ion m/z 691.1021 represented cyclic di-GMP (Fig. S5).

LC–MS analyses of wild-type PAO1 and recombinant P. aeruginosa strains revealed the presence of cyclic di-GMP (Fig. 5a–c). The quantitative analyses of the relative cyclic di-GMP levels were based on the normalized peak area of cyclic di-GMP to [13C]adenosine. P. aeruginosa expressing ml1419c significantly increased the abundance (approximately ninefold) of cyclic di-GMP detected as compared to the vector control (Fig. 5d). However, the cyclic di-GMP abundance of P. aeruginosa expressing ml1419c was lower than that of P. aeruginosa expressing tpbB (Fig. S6). In comparison, there were no significant differences in cyclic di-GMP abundance between the vector control and P. aeruginosa expressing ml1419c<sub>GAGGDEF</sub>. These data confirm that ML1419c of M. leprae functions to induce cyclic di-GMP synthesis in P. aeruginosa and that the GGDEF motif is part of the active-site domain as shown in other DGCs (Ryjenkov et al., 2005).

**DISCUSSION**

This current study demonstrated that ML1419c encoded on the genome of M. leprae possesses functional DGC activity resulting in the synthesis of cyclic di-GMP that can be measured by phenotypic or analytical assays. Thus, the M. leprae ML1419c protein was renamed as DqA (Diguanylate cyclase A). Given the inability to grow or manipulate M. leprae in
vitro, the use of *P. aeruginosa* as a heterologous expression host was critical for demonstrating the enzymatic activity of DgcA. In fact, the recombinant expression of the *M. leprae* dgcA was able to mimic the same phenotypes in *P. aeruginosa* that were induced by overexpression of a well-characterized *P. aeruginosa* encoded DGC, *tpbB* (Kulasakara et al., 2006; Ueda & Wood, 2009). DgcA from *M. leprae* possesses conserved DGC A-site (RFGGDEF) and I-site (RSRD) motifs. The consensus sequence of the A-site motif, RxGG(D/E)EF, has been well studied and is known to participate in protein dimerization, as well as substrate binding and catalytic activity (Wassmann et al., 2007; Chan et al., 2004). Thus, we hypothesized that deletion of the GGDEF domain would prevent enhanced production of cyclic di-GMP by recombinant ML1419c and the subsequent loss of *P. aeruginosa* phenotypes associated with increased DGC activity. Modification and expression of recombinant ML1419c, *ml1419c*<sub>GGDEF</sub>, resulted in decreased cyclic di-GMP levels as compared to that of *P. aeruginosa* expressing full-length *ml1419c*. Likewise, swimming and biofilm phenotypes associated with increased cyclic di-GMP production were altered.

Two of the DGC genes from *M. leprae* (*dgcA* and *ml0397c*) were previously found to be expressed during infection in animal models and leprosy patients (Williams et al., 2009; Spencer et al., 2005). These data indicate that, at various stages during infection, the *M. leprae* DGCs may play important roles in leprosy pathogenesis. The production of DgcA during the early stages of leprosy as described in previous studies (Williams et al., 2009; Spencer et al., 2005) and the ability of DgcA to induce a robust immune response (Geluk et al., 2011) indicate a potential role for this protein in the pathogen’s ability to sense and respond to environmental changes during the initial stages of infection. The putative sensing domain of DgcA from *M. leprae* possesses three PAS domains, and two of these PAS domains have conserved haem-binding sites. Thus, we hypothesize that DgcA likely responds to oxygen tension,
nitric oxide and/or carbon monoxide (Henry & Crosson, 2011). Importantly, the environmental cues that are perceived by the PAS sensor domains of DgcA would be expected to alter downstream gene expression and protein function via the activity of DgcA-derived cyclic di-GMP. Previous research groups have overexpressed GGDEF proteins similar to DgcA in order to study protein activity in the absence of activating signal (Kulasakara et al., 2006; Hickman & Harwood, 2008), and it is likely that elevated protein levels facilitate the dimerization and activation of these DGCs (Hallberg et al., 2016). Thus, it is possible that the DGC activity demonstrated by the recombiniton expression of *M. leprae* dgcA in *P. aeruginosa* was a result of protein abundance or interaction of the DgcA PAS domains with an environmental signal. Future structural studies of DgcA and targeted binding assays are required to define the ligands that bind to this protein and whether they induce or repress DGC activity.

Comparison of the genomes of the two primary mycobacterial pathogens, *M. leprae* and *M. tuberculosis*, reveals that *M. leprae* has a significantly smaller genome, a relatively large number of pseudogenes, fewer functional proteins and fewer transcription factors (Cole et al., 2001). Consequently, *M. leprae* is refractory to *in vitro* growth and has evolved into an obligate intracellular pathogen. Given the narrow biological niche of *M. leprae*, it is intriguing that this pathogen harbours three coding sequences for known or predicted DGCs (dgcA, ml10397c and ml1750c), whilst *M. tuberculosis* has only one protein, Rv1354c, a homologue of ML1750c. This same coding sequence is also the only DGC found in the genome of *M. smegmatis* (msmeg_2196) (Bharati et al., 2012), a nonpathogenic saprophyte commonly used as a model to define gene function of mycobacterial pathogens (Reyrat & Kahn, 2001; Singh & Reyrat, 2009). Rv1354c and MSMEG_2196 have been confirmed to have DGC and PDE activity, and phenotypes have been associated with rv1354c and msmev_2196.

During the course of our studies, we expressed *ml1419c* in *M. smegmatis*, which does not encode an *ml1419c* homologue. However, we found this to be an inadequate heterologous host system to elucidate the biochemical function of ML1419C. Expression of *ml1419c* in *M. smegmatis* under the control of an hsp60 promoter produced recombinant protein, which subsequently altered the colony morphology, pellicle formation and sliding motility of *M. smegmatis* (data not shown). However, we were unable to reproducibly detect cyclic di-GMP production with a reasonable amount of cells in this system. The requirement for 1–3 g of mycobacterial cells to conduct analytical measurement of cyclic di-GMP in *M. tuberculosis* or *M. smegmatis* (Bharati et al., 2012; Hong et al., 2013) does not provide a robust system to correlate cyclic di-GMP levels with phenotypes. Thus, given the ambiguity of cyclic di-GMP production in *M. smegmatis* and the reported cyclic di-GMP phenotype(s) of this bacterium (Bharati et al., 2012; Gupta et al., 2015), future efforts will focus on defining the function and significance of DGC activity in *M. leprae*, using a model system that allows for genetic analyses of individual DGCs. Towards this goal, all of the known and predicted DGCs of *M. leprae* (Fig. 1) are encoded in the genomes of *Mycobacterium leprae* (another bacterium restricted to *in vivo* growth) and *Mycobacterium haemophilum* (an opportunistic pathogen). Homologues of *ML1419c* and ML1750c are also encoded in several environmental actinobacteria such as *Mycobacterium rhodesiae*, *Mycobacterium chubuense*, *Mycobacterium chlorophenolicum*, *Mycobacterium rufum* and *Rhodococcus fascians*. Evolutionary reduction of the number cyclic di-GMP signalling pathways is generally believed to be inversely correlated with a bacterium’s need and ability to adapt to rapidly changing environmental conditions (Römling et al., 2013). Therefore, the conservation of DGC genes in *M. leprae* and the presence of these genes in opportunistic pathogens and several environmental *Mycobacterium* spp. is indicative of the potential role for DGCs in the signalling response required for *M. leprae* to survive as an obligate intracellular pathogen. Interestingly, like *M. leprae*, *M. haemophilum* is a pathogen of the skin and displays optimal growth at 30 °C (Sompolinsky et al., 1978). Recently, it has been reported that *M. haemophilum* can be genetically manipulated to express foreign genes (Tufariello et al., 2015), and efforts are now under way to use this *Mycobacterium* sp. as a model to further define the physiological functions and signalling events associated with DgcA and the additional predicted PDE and DGCs of *M. leprae*.

**ACKNOWLEDGEMENTS**

This work was supported by National Institutes of Health grants, National Institute of Allergy and Infectious Diseases grant R01-AI022553 and National Institute of Arthritis and Musculoskeletal and Skin Diseases grant R01-AR040312. We thank Grace Borlee and Ashli Simone for technical assistance and advice in the completion of this work.

**REFERENCES**


Edited by: M. Braunstein