A full-length bifunctional protein involved in c-di-GMP turnover is required for long-term survival under nutrient starvation in *Mycobacterium smegmatis*

Binod K. Bharati,[1] Indra Mani Sharma,[1] Sanjay Kasetty,[1] Manish Kumar,[2] Raju Mukherjee[1] and Dipankar Chatterji[1]

[1]Molecular Biophysics Unit, Indian Institute of Science, Bangalore-560012, India

[2]Department of Microbiology and Cell Science, University of Florida, Gainesville, FL 32611-0700, USA

Correspondence
Binod K. Bharati
bk@mbu.iisc.ernet.in

The bacterial second messenger cyclic diguanosine monophosphate (c-di-GMP) plays an important role in a variety of cellular functions, including biofilm formation, alterations in the cell surface, host colonization and regulation of bacterial flagellar motility, which enable bacteria to survive changing environmental conditions. The cellular level of c-di-GMP is regulated by a balance between opposing activities of diguanylate cyclases (DGCs) and cognate phosphodiesterases (PDE-As). Here, we report the presence and importance of a protein, MSDGC-1 (an orthologue of Rv1354c in *Mycobacterium tuberculosis*), involved in c-di-GMP turnover in *Mycobacterium smegmatis*. MSDGC-1 is a multidomain protein, having GAF, GGDEF and EAL domains arranged in tandem, and exhibits both c-di-GMP synthesis and degradation activities. Most other proteins containing GGDEF and EAL domains have been demonstrated to have either DGC or PDE-A activity. Unlike other bacteria, which harbour several copies of the protein involved in c-di-GMP turnover, *M. smegmatis* has a single genomic copy, deletion of which severely affects long-term survival under conditions of nutrient starvation. Overexpression of MSDGC-1 alters the colony morphology and growth profile of *M. smegmatis*. In order to gain insights into the regulation of the c-di-GMP level, we cloned individual domains and tested their activities. We observed a loss of activity in the separated domains, indicating the importance of full-length MSDGC-1 for controlling bifunctionality.

INTRODUCTION

Recent advances in the study of the social behaviour of bacteria motivated us to investigate quorum sensing and its role in bacterial survival. Quorum sensing facilitates the coordinated response of a bacterial population in response to environmental stress. In both Gram-negative and Gram-positive bacteria, small signalling molecules or autoinducers are the key components, whose presence can be sensed by neighbouring cells and which form the basis of cell–cell communication (Taga & Bassler, 2003; Reading & Sperandio, 2006). Bacterial biofilm formation is one of the best-known quorum-sensing phenomena (Bassler, 1999; Miller & Bassler, 2001). Although free-living bacteria can form biofilms in their natural habitats, biofilms can also be destabilized, and small

molecules act as primary and secondary messengers, which are produced by bacteria to counter destabilization.

Cyclic AMP (cAMP) and ppGpp are two well-known examples of second messengers which play important roles in relaying extracellular information (Chatterji & Ojha, 2001; Camilli & Bassler, 2006; Shenoy & Visweswariah, 2006). A novel second messenger, c-di-GMP [bis-(3′-5′)-cyclic diguanosine monophosphate], was reported in *Gluconacetobacter xylinus*, and acts as a positive allosteric regulator of the cellulose synthase pathway (Ross *et al.*, 1987; Weinhouse *et al.*, 1997; Tal *et al.*, 1998). An increased level of c-di-GMP was found to stimulate the matrix (exopolysaccharide/cellulose) production required for biofilm formation, and was also found to regulate the transition between motile and sessile forms in Gram-negative bacteria (Kirillina *et al.*, 2004; Hickman *et al.*, 2005; Jenal & Malone, 2006; Kuchma *et al.*, 2007). Various studies have established a clear link between the level of c-di-GMP and alterations in cell-surface properties, cell

differentiation, motility and biofilm formation in bacteria (D'Argenio *et al.*, 2002; Tischler & Camilli, 2004; Kazmierczak *et al.*, 2006; Beyhan *et al.*, 2006; Tamayo *et al.*, 2007; Merritt *et al.*, 2007).

c-di-GMP is a condensation product of two GTP molecules, and it is hydrolysed to GMP via a linear intermediate product, pGpG. The synthesis and hydrolysis of c-di-GMP are carried out by diguanylate cyclase (DGC) and its cognate phosphodiesterase (PDE-A), respectively. The intracellular concentration of c-di-GMP is maintained by a balance between DGC and PDE-A activities. The GGDEF (Gly-Gly-Asp-Glu-Phe) and EAL (Glu-Ala-Leu) conserved motifs form major components of the two active sites in DGC and PDE-A, respectively. Mutations in these conserved residues result in reduced enzymic activities of the respective proteins (Paul *et al.*, 2004; Christen *et al.*, 2005). Studies have also shown the involvement of another domain, HD-GYP, with metal-dependent phosphodiesterase activity, which hydrolyses c-di-GMP to GMP (Galperin *et al.*, 1999, 2001; Dow *et al.*, 2006).

Studies demonstrating the role of c-di-GMP in various aspects of physiology have been mostly carried out in Gram-negative bacteria. *Staphylococcus aureus* is the only Gram-positive bacteria studied for c-di-GMP signalling in which c-di-GMP was found to inhibit cell–cell interactions and biofilm formation in a dose-dependent manner (Karaolis *et al.*, 2005; Brouillette *et al.*, 2005). Little is known, however, of c-di-GMP signalling in mycobacteria (Gupta *et al.*, 2010). To characterize the physiological importance of c-di-GMP in mycobacteria, we undertook studies on *MSMEG_2196* of *Mycobacterium smegmatis*, which encodes protein MSDGC-1. *M. smegmatis* is a non-pathogenic fast-growing species that is widely accepted as an excellent model system for studying various aspects of *M. tuberculosis* biology (Jacobs *et al.*, 1987; Smeulders *et al.*, 1999; Ojha *et al.*, 2000; Dahl *et al.*, 2005). MSDGC-1 has a tandem arrangement of GAF, GGDEF and EAL domains. We have systematically cloned, overexpressed and deleted *msdgc-1* to understand its role in *M. smegmatis*. We have also shown that MSDGC-1 is a bifunctional protein with the ability to synthesize and hydrolyse c-di-GMP. Deletion of the *msdgc-1* gene severely affects the long-term survival of mycobacteria under conditions of nutrient starvation (0.02 % glucose), and the full-length protein is required for DGC and PDE-A activities.

METHODS

Bacterial strains, plasmids and growth conditions. Plasmids and oligonucleotides used in this study are listed in Table 1 and Table S1 available with the online version of this paper. *Escherichia coli* strains DH5α and BL21(DE3) were grown in Luria–Bertani (LB) broth at 37 °C with agitation or on LB medium containing 1.5 % (w/v) agar. *M. smegmatis* mc^{2}155 (Snapper *et al.*, 1990) was grown in Middlebrook 7H9 (Difco) broth with 2 % (w/v) glucose as the carbon source and 0.05 % (v/v) Tween 80, or on 7H9 medium containing 1.5 % (w/v) agar. Antibiotics were used at the following concentrations (μg ml^{-1}): ampicillin (100), kanamycin (35) or gentamicin (20) for *E. coli*, and kanamycin (20), gentamicin (10) or hygromycin (30) for *M. smegmatis*. PCRs were carried out using DyNAzyme EXT DNA Polymerase (Finnzyme) following the manufacturer's instructions. All clones generated were confirmed by sequencing (Microsynth, Switzerland, or MWG, India). Restriction enzymes used for cloning were obtained from New England Biolabs.

Bioinformatics analysis. The *M. smegmatis* genome was searched using the TIGR database (http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage. cgi) for the annotation of proteins containing GGDEF and EAL domains.

Table 1. Strains and plasmids used in this study

Strain or plasmid	Characteristics	Reference or source
Strains		
M. smegmatis mc^{2}155	Wild-type strain	Snapper *et al.* (1990)
DGCKO	*M. smegmatis* in which *MSMEG_2196* has been replaced with the kan cassette, *kan^r*	This study
M. smegmatis pMV	*M. smegmatis* transformed with pMV261 vector alone, *kan^r*	This study
DGCOE	*M. smegmatis* transformed with pMVDGC_2196, *kan^r*	This study
DGCCO	DGCKO complemented with *MSMEG_2196*, *kan^r hyg^r*	This study
Plasmids		
pUC4K	Source of *aph* gene	Pharmacia Biotech
pET21b	Cloning vector (C-terminal His$_6$-tag), *amp^r*	Novagen
pET28a	Cloning vector (N-terminal His$_6$-tag), *kan^r*	Novagen
pPR27	Suicide vector, *gen^r*	Pelicic *et al.* (1997)
pMV261	Cloning vector with hsp60 promoter, *kan^r*	Stover *et al.* (1991)
pMV361	Cloning vector with hsp60 promoter, *hyg^r*	Stover *et al.* (1991)
pMVDGC_2196	MSDGC-1 cloned in pMV261, *kan^r*	This study
pET DGC_2196	MSDGC-1 cloned in pET21b, *amp^r*	This study
pET DGC_2774 Nter	MSDGC-2 cloned in pET28a, *kan^r*	This study
pGAF	GAF domain of MSDGC-1 cloned in pET21b, *amp^r*	This study
pGGDEF	GGDEF domain of MSDGC-1 cloned in pET21b, *amp^r*	This study
pEAL Nter	EAL domain of MSDGC-1 cloned in pET28a, *kan^r*	This study
pGGDEF-EAL	GGDEF–EAL domain of MSDGC-1 cloned in tandem	This study

Two ORFs, *MSMEG_2196* and *MSMEG_2774*, were annotated as putative GGDEF and EAL domain proteins. The protein sequence was subjected to Smart database (v5.1; http://smart.embl-heidelberg.de/) and NCBI Conserved Domains database (v2.30; www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) searches to identify the various domains present in the protein. Furthermore, we did a BLAST analysis of the GGDEF and EAL domains in a non-redundant NCBI database to find protein homologues in other mycobacteria.

Cloning, expression and purification of MSDGC-1 and MSDGC-2. Genes *MSMEG_2196* (1848 bp, 615 aa, MSDGC-1) and *MSMEG_2774* (1080 bp, 359 aa, MSDGC-2) were PCR-amplified from *M. smegmatis* genomic DNA using primers MKF/MKR and MKDGC F/MKDGC R, respectively (Table S1). The amplicons *MSMEG_2196* and *MSMEG_2774* were cloned into plasmid pET21b using *Nhe*I/*Not*I and *Nde*I/*Hin*dIII restriction sites, respectively. The pET28a vector was digested with the same set of restriction enzymes and used for N-terminal His$_6$-tag cloning of *MSMEG_2774*. The resulting plasmids pET DGC_2196, pET DGC_2774 and pET DGC_2774 Nter were transformed into *E. coli* BL21(DE3) for protein expression. Protein purification was done as described previously (Jain *et al.*, 2006). Briefly, cells were induced with 1 mM IPTG for 3 h, harvested, and lysed in lysis buffer containing 50 mM Tris/HCl (pH 7.9), 500 mM NaCl, 1 mM PMSF and 1 mg lysozyme ml^{-1}. After centrifugation, the lysate was loaded onto a Ni-NTA column and washed with 100 column volumes of wash buffer containing 20 mM imidazole. The protein was eluted with elution buffer containing 500 mM imidazole, dialysed against a buffer containing 50 mM Tris/HCl (pH 7.9), 500 mM NaCl and 5 mM β-mercaptoethanol, and stored at 4 °C for future use.

Enzymic assays. DGC and PDE-A assays were adapted from procedures described by others (Christen *et al.*, 2005). Briefly, the purified proteins MSDGC-1, MSDGC-2 and PleD (5 μM) were incubated with 50 mM Tris/HCl (pH 7.9), 500 mM NaCl, 10 mM MgCl$_2$, 5 mM β-mercaptoethanol, 0.1 mM GTP (cold) and [α-^{32}P]GTP [0.01 mCi μl^{-1} (370 Bq μl^{-1}), BRIT, Hyderabad, India] at 37 °C for 30 min in a 10 μl reaction volume. Reactions were stopped by boiling for 5 min followed by centrifugation at 12 000 r.p.m. (15 294 *g*) for 20 min at 4 °C. Supernatants (5 μl) were spotted on polyethyleneimine (PEI)-cellulose plates (Merck) and developed against 1 : 1.5 (v/v) (NH$_4$)$_2$SO$_4$ and 1.5 M KH$_2$PO$_4$ (pH 3.6), and plates were exposed to a phosphorimager screen.

Preparation of [α-^{32}P]c-di-GMP and cold c-di-GMP. The protocol used was adapted from that described by Christen *et al.* (2005) for the preparation of α-^{32}P-labelled and non-labelled c-di-GMP. For radiolabelled c-di-GMP, [α-^{32}P]GTP [125 μCi (4.625 MBq); 3500 Ci mmol^{-1} (129.5 TBq mmol^{-1})] was incubated with 100 μg purified His$_6$-tagged PleD protein containing 1 mM cold GTP. The protein was precipitated by boiling for 5 min followed by centrifugation and filtration through a 0.2 μm pore-size filter. The supernatant was tested for the presence of c-di-GMP by TLC. Non-labelled c-di-GMP was prepared by using 1 mM GTP as a substrate and purified by HPLC as described earlier (Simm *et al.*, 2004). The identity of c-di-GMP was confirmed by MALDI-TOF MS.

Cloning, expression and purification of GAF, GGDEF and EAL domains of MSDGC-1. For the construction of GAF (amino acids 23–166), GGDEF (amino acids 175–334) and EAL (amino acids 348–590), pET DGC_2196 was used as a template and primers mkGAF (F)/mkGAF (R), mkGGDEF (F)/mkGGDEF (R) and mkEAL (F)/mkEAL (R) (Table S1) were used, respectively. For the amplification of GGDEF–EAL, primers mkGGDEF (F) and mkEAL (R) were used. The required region was PCR-amplified and cloned into vector pET21b or pET28a at *Nhe*I and *Hin*dIII sites, resulting in pGAF, pGGDEF, pEAL and pGGDEF-EAL. The expression and purification of the various domains of MSDGC-1 were achieved using

a strategy similar to that described above for the full-length protein, with some modifications, and activities were checked *in vitro*. For the *in vivo* activity assay, *E. coli* BL21 cells were transformed with full-length MSDGC-1 as well as its various domains and allowed to grow to OD$_{600}$ 0.2. Thereafter, cells were fed with [^{32}P]H$_3$PO$_4$ and further allowed to grow till the OD$_{600}$ reached 0.6, followed by induction with 1 mM IPTG. After further incubation for 3 h, cells were harvested and lysed with 11 M formic acid. Total nucleotide extracts were separated on a PEI-cellulose TLC plate and visualized.

Substrate binding and steady-state kinetics. To calculate the steady-state kinetic parameters V_{max} and K_m of the purified protein, MSDGC-1 was assayed for DGC activity, as the separate domains did not show any activity. The radiolabelled GTP was used as a standard. For the determination of the initial velocity, different concentrations of substrate (1–200 μM GTP) were used with same amount of protein (5 μM). The reaction was stopped after 30 min by boiling for 5 min followed by centrifugation at 12 000 r.p.m. at 4 °C for 20 min. Resulting supernatants (1 μl) were spotted onto PEI-cellulose plates, developed in 1 : 1.5 (v/v) (NH$_4$)$_2$SO$_4$ and 1.5 M KH$_2$PO$_4$ (pH 3.6), and exposed as described above. The intensities of the spots were determined using the Multi Gauge v2.3 (Fujifilm) software from Science Lab and compared with the standard plot of radiolabelled GTP. K_m and V_{max} were determined from a linear plot of initial velocity ($1/v_o$) as a function of substrate concentration ($1/[S]$) and from a nonlinear Michaelis–Menten plot using GraphPad Prism (version 5.02).

Overexpression of MSDGC-1. The gene *MSMEG_2196* was subcloned into the *E. coli*–*Mycobacterium* shuttle vector pMV261 (Stover *et al.*, 1991) using primers Relco F and Relco R (Table S1), resulting in pMVDGC_2196. This vector was electroporated into *M. smegmatis* mc^{2}155 and colonies were selected on 7H9 agar medium containing kanamycin. Growth of the overexpression strain (DGCOE) in liquid culture was followed by plating on an 7H9 agar plate containing 2 % (w/v) glucose as the carbon source.

Targeted replacement of *MSMEG_2196* in *M. smegmatis* mc^{2}155. In order to investigate the role of c-di-GMP in the physiology of *M. smegmatis*, gene *MSMEG_2196* was deleted by allelic exchange. Conserved regions of GGDEF and EAL of MSDGC-1 were replaced with a kanamycin-resistance marker present in the suicide vector pPR27. A recombination cassette was constructed to delete *MSMEG_2196* from the chromosome of *M. smegmatis* mc^{2}155 (Table 1). The cassette consisted of a 3.2 kb DNA fragment spanning from base 246 upstream of *MSMEG_2196* to base 750 of *MSMEG_2196* containing *Xba*I and *Eco*RI restriction sites, and a fragment spanning from base 1175 of *MSMEG_2196* to base 321 downstream of *MSMEG_2196* containing *Eco*RI and *Not*I restriction sites. The DNA fragment containing the *kan*^{r} gene was obtained by digesting the pUC4K vector with *Eco*RI and cloned between upstream and downstream regions of the gene using the same restriction site. After the preparative cloning steps, the whole recombination cassette was transferred to the suicide delivery vector pPR27 to obtain the final construct pDGCKO. Competent *M. smegmatis* mc^{2}155 cells were transformed with pDGCKO. The sucrose-resistant, gentamicin-sensitive and kanamycin-resistant colonies were selected at 39 °C for further analysis. Disruption of the gene and the recombination event were verified by Southern blot analysis and PCR of selected colonies. For PCR, primers MKF and MKR were used for confirmation using genomic DNA as a template. For Southern blot analysis, the genomic DNA of the putative mutant (DGCKO) as well as that of mc^{2}155 was digested with *Pvu*II. The whole *MSMEG_2196* gene (~1.8 kb) from mc^{2}155 was used as a probe and labelled with [α-^{32}P]dATP by PCR. The complemented strain was generated by cloning a functional *msdgc-1* gene into vector pMV361. Hereafter, all strains used will be described as follows: wild-type or *M. smegmatis* mc^{2}155, the knockout strain as DGCKO, the overexpression strain as DGCOE and the complemented strain as DGCCO.

Extraction and detection of c-di-GMP from *M. smegmatis*. Wild-type, DGCKO and DGCCO cells were grown in 7H9 medium with 2.0 or 0.02 % (w/v) glucose at 37 °C with shaking. Cells were harvested at 24, 48 and 72 h and washed with sterile MilliQ water. Approximately 3 g of wet cell pellets was used to extract total nucleotides using a published method (Weinhouse *et al.*, 1997). Briefly, cells were incubated on ice with 0.6 M perchloric acid ($HClO_4$) and extracts were neutralized with 5 M potassium carbonate (K_2CO_3). Samples were centrifuged to remove the precipitate of $KClO_4$, and the supernatant was lyophilized. Lyophilized extracts were resuspended in MilliQ water and subjected to HPLC separation. Purified c-di-GMP (obtained from Nanyang Technological University, Singapore) was used as a standard, and its retention time was used as the reference.

HPLC analysis and MS. The extracted samples (100 μl) were injected onto a C-18 column (4.6 × 150 mm, Agilent Eclipse XDB-C18) and separated by reverse-phase HPLC (Agilent 1200). Buffer A (100 mM KH_2PO_4, 4 mM tetrabutyl ammonium hydrogen sulfate, pH 5.9) and buffer B [75 % (v/v) buffer A, 25 % (v/v) methanol] were used in the gradient program, as described by Ryjenkov *et al.* (2005). Nucleotides were detected at 254 nm. The HPLC fractions were desalted using a Sephadex G-10 column, lyophilized and spotted on a MALDI-TOF MS plate using 4-chloro-α-cyanocinnamic acid (CCA) as matrix. MALDI-TOF analysis was performed in an Ultraflex TOF mass spectrometer (Bruker Daltonics) using the negative-ion detection mode.

Long-term starvation cultures. The wild-type, DGCKO and DGCCO strains were grown in 7H9 containing 0.02 % (w/v) glucose and 0.05 % (v/v) Tween 80. The antibiotics were omitted from the culture to rule out any effect that they might have on long-term survival. Bacterial cultures were declumped by vortexing the cells with 0.5 mm glass beads before plating on a 7H9 agar plate, as described elsewhere (Primm *et al.*, 2000). The number of c.f.u. was determined at regular time intervals up to 20 days.

Biofilm formation and quantification assay. The biofilm formation assay was performed as described previously (Mathew *et al.*, 2006; Naresh *et al.*, 2010). Experiments were started from a culture of bacteria grown in 7H9 medium supplemented with 2 % glucose and 0.05 % Tween 80. The well-grown primary cultures (~OD_{600} 1.5) of the wild-type, DGCOE, DGCKO and DGCCO were washed with Sauton's medium (Himedia) to remove Tween and glucose. The cultures were diluted 1 : 100 to inoculate 30 ml Sauton's medium in 90 mm diameter polystyrene Petri dishes supplemented with 2 % glucose. Petri dishes were incubated for 5–7 days at 37 °C in a humidified incubator, and the formation of surface pellicles was monitored. Quantification of biofilm formation was performed as described earlier (O'Toole *et al.*, 1999). Briefly, well-grown primary cultures (~OD_{600} 1.5) were washed with Sauton's medium and diluted to a final OD_{600} of 0.05 in Sauton's medium, and were then distributed into 96-well polystyrene plates at time zero supplemented with 2 % glucose and incubated at 37 °C in a humidified incubator for 5–7 days. Each well received an inoculum of 200 μl and 10 wells were assayed for each strain at each time point. The samples were removed from the wells and the wells were washed twice with water. The adherent biofilm was stained with 1 % crystal violet and incubated for 45 min at room temperature. The residual dye was washed thoroughly with water and the plate was allowed to dry. The bound dye was solubilized in 300 μl 80 % (v/v) ethanol and the A_{550} was measured using a microtitre plate reader (SpectraMax 340PC[384], Molecular Devices).

Sliding motility assay. The spreading ability of mycobacteria on a moist surface was determined using a method described previously (Etienne *et al.*, 2002; Mathew *et al.*, 2006; Naresh *et al.*, 2010). Briefly, 7H9 medium (Difco) was solidified with 0.3 % high-grade agarose (Sigma) supplemented with 2 % glucose. Well-grown cultures (~OD_{600} 1.5) of the wild-type, DGCKO and DGCCO were adjusted to OD_{600} 0.5, and 10 μl volumes were inoculated into the centre of a plate. The spreading ability of mycobacteria was evaluated after incubation for 5–7 days at 37 °C in a humidified incubator.

RESULTS

Identification of GGDEF and EAL domain-containing proteins in *M. smegmatis*

Bioinformatics analysis revealed the existence of two genes, *MSMEG_2196* (MSDGC-1) and *MSMEG_2774* (MSDGC-2), encoding GGDEF–EAL and GGDEF domain proteins, respectively, when searches were performed against the *M. smegmatis* genome in the TIGR database. The domain organization of the sequences illustrated in Fig. 1(a) was obtained from the SMART database (http://smart.embl-heidelberg.de/smart/set_mode.cgi). The MSDGC-1 protein consists of GAF–GGDEF–EAL domains in tandem, while GGDEF is the only domain present in MSDGC-2. However, the N terminus of MSDGC-2 showed the presence of a transmembrane domain in the SMART database. Multiple
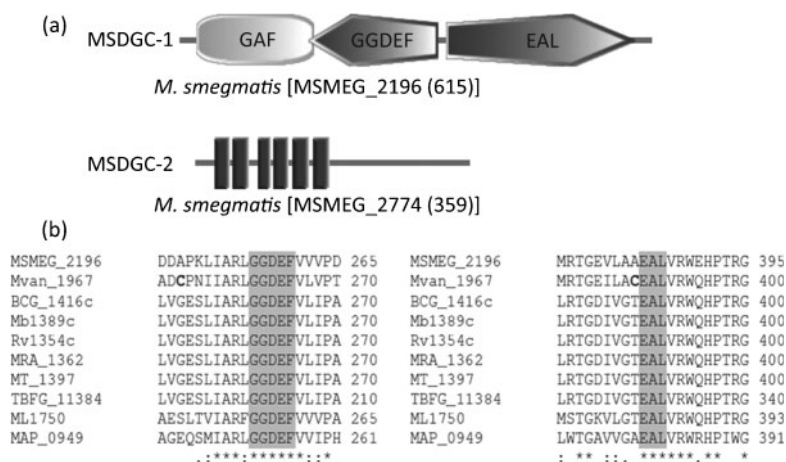

Fig. 1. (a) Representative domain architectures of proteins containing GGDEF, EAL and GAF motifs in *M. smegmatis*. The protein with GAF–GGDEF–EAL (MSDGC-1) is present in almost all mycobacterial species. The domain symbols were taken from the SMART database. The ORF and number of amino acids in the sequence are given after the organism name. Predicted transmembrane regions are shown as rectangles. (b) Multiple sequence alignment showing the presence of the conserved GGDEF and EAL motifs in different species of mycobacteria.

Table 2. Proteins with GGDEF and EAL domains in different mycobacterial species.

−, No proteins in this category.

Organism (total no. of proteins with GGDEF and/or EAL domains)	No. of proteins with the domain organization:							
	GAF–GGDEF–EAL	GGDEF–EAL	GGDEF*	GGEEF†	EAL	EAL with degenerate GGDEF	HD-GYP	Others‡
M. smegmatis mc²155 (2)	1	−	−	1	−	−	−	−
M. bovis BCG Pasteur (1)	−	−	−	−	1	−	−	−
M. tuberculosis H37Rv (2)	1	−	−	−	1	−	−	−
M. tuberculosis CDC1551 (1)	1	−	−	−	−	−	−	−
M. leprae TN (4)	1	−	1	−	1	−	−	PAS–GGDEF–AraH
Mycobacterium sp. JLS (5)	1	1	−	2	−	−	1	−
Mycobacterium sp. MCS (9)	2	2	−	4	−	−	1	−
M. avium 104 (7)	2	−	4	1	−	−	−	−
M. avium subsp. *paratuberculosis* K-10 (6)	1	−	2	−	1	1	1	−
M. vanbaalenii PYR-1 (15)	1	1	2	8	1	1	1	GAF–EAL–PAS–GGDEF
M. gilvum PYR-GCK (22)	1	3	3	10	1	2	−	GAF–EAL–PAS–GGDE

*GGDEF domain with one or more input sensory domain such as PAS, Rec, FlhA or GAF.
†GGDEF domain with conserved GGEEF or degenerate SDSEF motif.
‡Unusual domain organization arrangement, with EAL at the N terminus of GGDEF or an AraH output domain in the case of *M. leprae*.

sequence alignment of GGDEF and EAL domains showed the presence of conserved signature motifs in various mycobacterial proteins, as depicted in Fig. 1(b). Table 2 shows the number and domain organization of different GGDEF and EAL domain-containing proteins in mycobacteria. Interestingly, most mycobacterial species appear to have at least one protein with the GAF–GGDEF–EAL domain organization, except for the *Mycobacterium bovis* BCG Pasteur strain. Few mycobacterial species appear to have a HD-GYP domain for c-di-GMP degradation (Table 2). Proteins from different mycobacterial species with GAF–GGDEF–EAL domain architecture share >60 % amino acid identity among themselves.

M. smegmatis MSDGC-1 is a bifunctional protein *in vitro*

In order to investigate the regulation of c-di-GMP *in vitro*, MSDGC-1 and MSDGC-2 proteins were expressed in *E. coli* BL21(DE3) cells, and purified through an affinity column (Fig. 2) and checked for DGC and PDE-A activities. The MSDGC-1 protein showed both c-di-GMP synthesis and hydrolysis activities *in vitro* (Fig. 3a). The PEI-cellulose TLC was developed in 1 : 1.5 (v/v) saturated $(NH_4)_2SO_4$ and 1.5 M KH_2PO_4, pH 3.6, and a spot with the same R_f value (0.2) as that of c-di-GMP was observed that was apparently followed by another spot with the R_f value (0.4) of linear di-GMP (pGpG), as reported by Christen *et al.* (2006). A reaction mixture containing PleD protein, which is a known DGC protein of *Caulobacter crescentus*, was used as a positive control. PleD protein only shows c-di-GMP synthesis (Paul *et al.*, 2004); however, c-di-GMP was followed by an additional spot (Fig. 3a) that was not detected among the reaction products with MSDGC-1.

For further confirmation of the dual activity, MSDGC-1 protein was incubated with cold GTP, and the reaction product was analysed for the presence of c-di-GMP and pGpG using negative-ion mode MALDI-TOF MS (Fig. S1) or subjected to HPLC (Fig. 3b), followed by collection of peaks and subsequent MALDI-TOF MS analysis. The peak at 16.8 min (Fig. 3b, 1) was that of standard GTP, with an m/z of 522 $[M-H]^-$. Similarly, the peak at 19.2 min (Fig. 3b, 2) indicated the retention time of standard c-di-GMP and had an m/z value of 689.7 $[M-H]^-$. The MSDGC-1 reaction product (30 min) was subjected to HPLC, and a peak at 19.07 min (Fig. 3b, 3) indicated the retention time of c-di-GMP and an m/z value of 689.6 $[M-H]^-$. For the detection of pGpG (the hydrolysis product of c-di-GMP), the reaction mixture of MSDGC-1 was incubated for 45 min instead of 30 min, and we observed a peak at 16.8 min with an m/z value of 707.6 $[M-H]^-$, corresponding to pGpG (Fig. 3b, 4). Thus, it is evident that MSDGC-1, like other orthodox DGCs and cognate PDE-As, can use GTP to synthesize c-di-GMP, which can subsequently be hydrolysed into pGpG. However, longer incubation of c-di-GMP with MSDGC-1 did not yield
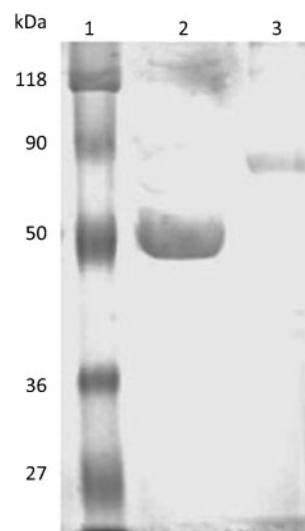

Fig. 2. Purification of His$_6$-tagged proteins. Lanes: 1, protein marker; 2, PleD; 3, MSDGC-1. C-terminal His$_6$-tagged MSDGC-1 is a 67 kDa protein and PleD is a 45 kDa protein.

GMP. Partially purified fractions of N-terminal His$_6$-tagged MSDGC-2 did not show c-di-GMP synthesis *in vitro* (data not shown).

The full-length protein is required for DGC and PDE-A activity

The MSDGC-1 protein is a three-domain protein with GGDEF and EAL domains in tandem with an N-terminal GAF domain, and it possesses both DGC and PDE-A activities. The GAF (cGMP, adenylyl cyclase, FhlA) domain is a cyclic nucleotide-binding domain present in PDE-As (Hurley, 2003). To understand the regulation of MSDGC-1, we attempted to express GAF, GGDEF and EAL individually, and functionally characterized each one. However, we were not able to overexpress and purify the GAF domain, despite several attempts. The combined GGDEF–EAL, EAL and GGDEF domains were purified by the addition of Triton X-100 to the lysis buffer to improve solubility (Fig. 4a). We performed the DGC assay of the GGDEF protein with a normalized protein concentration, but did not observe a spot corresponding to c-di-GMP on the TLC plate, which indicates that the DGC activity of the GGDEF domain was absent (Fig. 4b, lane 3). Similarly, purified EAL domain protein did not show PDE-A activity, as there was no spot corresponding to pGpG (Fig. 4b, lane 4). This indicates that the full-length protein is required for both DGC and PDE-A activities. In addition, the combined GGDEF–EAL domain without GAF showed poor DGC activity and PDE-A activity was absent *in vitro* (Fig. 4b, lane 5). This suggests that the GAF domain is important for the activities of the combined domain. In order to avoid any *in vitro* experimental error, the activity assay was also performed in a heterologous system. We noticed that only
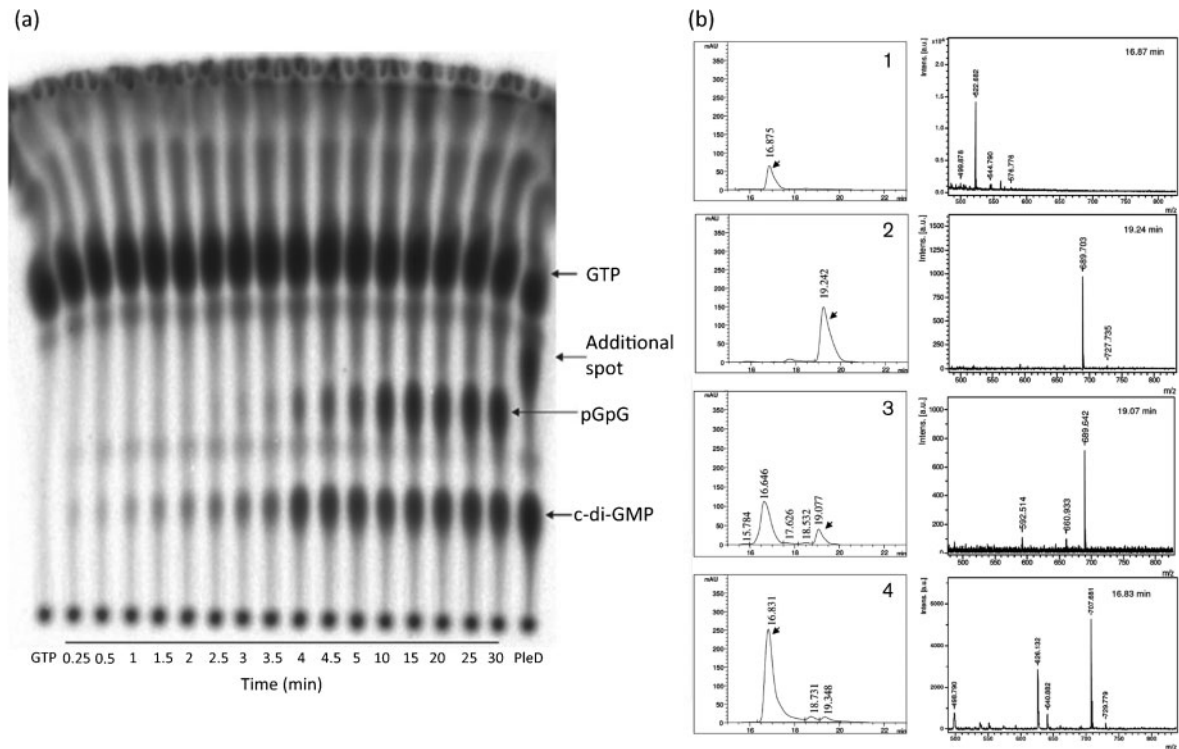
Fig. 3. (a) *M. smegmatis* protein MSDGC-1 has both DGC and PDE-A activities. Purified His$_6$-tagged protein was assayed for its ability to synthesize and hydrolyse c-di-GMP. Protein (5 µM) was incubated with assay buffer containing [α-^{32}P]GTP, as described in Methods. The cold GTP was kept constant in all the reactions (100 µM). Samples were withdrawn at different time points and analysed on a PEI-cellulose TLC plate. PleD, which is a known DGC protein from *C. crescentus*, was used as a positive control. GTP in assay buffer without protein was run as a negative control, as shown in the GTP lane. (b) Detection of c-di-GMP and pGpG from reaction mixture of MSDGC-1 by HPLC (left panels) and corresponding mass analysis (right panels). (1) Reaction mix without the addition of MSDGC-1, a peak corresponding to GTP at 16.88 min and *m/z* of 522.1 [M−H]$^-$; (2) standard c-di-GMP at 19.24 min and *m/z* of 689.71 [M−H]$^-$; (3) MSDGC-1 reaction product corresponding to c-di-GMP (19.07 min) and *m/z* [M−H]$^-$ of 689.6; (4) MSDGC-1 reaction product corresponding to pGpG (16.83 min) and *m/z* [M−H]$^-$ of 707.7.

full-length MSDGC-1 protein was active *in vivo* (Fig. S2, lane 5). At this stage it remains unclear why the GGDEF–EAL domain protein did not show activity in this experiment *in vivo*. Nevertheless, from these experiments it appears that full-length MSDGC-1 is required for dual activity.

Velocity (V_{max}) and Michaelis constants (K_m) of protein MSDGC-1

MSDGC-1 is a bifunctional protein with the ability to synthesize and hydrolyse c-di-GMP. Since separate GGDEF and EAL domains were found to be inactive, quantification of their individual activities remained difficult. We quantified the K_m and V_{max} values for the DGC activity of protein MSDGC-1 in steady-state kinetic experiments using densitometric analysis. Fig. 5(a) shows the synthesis of c-di-GMP and its degradation as a function of increasing GTP concentration. The steady-state kinetic parameters were derived from Michaelis–Menten and Lineweaver–Burk plots, as shown in Fig. 5(b). The K_m and V_{max} values

obtained from the Michaelis–Menten method were 22.3 µM and 0.99 ± 0.2 µM min^{-1}, respectively. Similarly, the K_m and V_{max} values obtained from the Lineweaver–Burk method were 25 µM and 0.95 ± 0.2 µM min^{-1}, respectively. The substrate (GTP) concentrations used in the Michaelis–Menten and Lineweaver–Burk methods were 1–200 and 5–200 µM, respectively. In this experiment, c-di-GMP was synthesized as a function of GTP concentration, and pGpG was produced as a function of increasing c-di-GMP concentration. K_m and V_{max} values obtained for c-di-GMP synthesis in our study followed Michaelis–Menten enzyme kinetics, and the cumulative rate due to c-di-GMP synthesis and degradation was used to calculate the initial velocity of c-di-GMP synthesis by MSDGC-1. Radiolabelled GTP was used to plot the standard curve. K_m and V_{max} values obtained in our study appear to be much higher than those reported by others (Christen *et al.*, 2005). One of the reasons for the higher values obtained in our case could be the lack of modulators such as small molecules or activation through phosphorylation, which perhaps are necessary for the optimal activity of MSDGC-1.
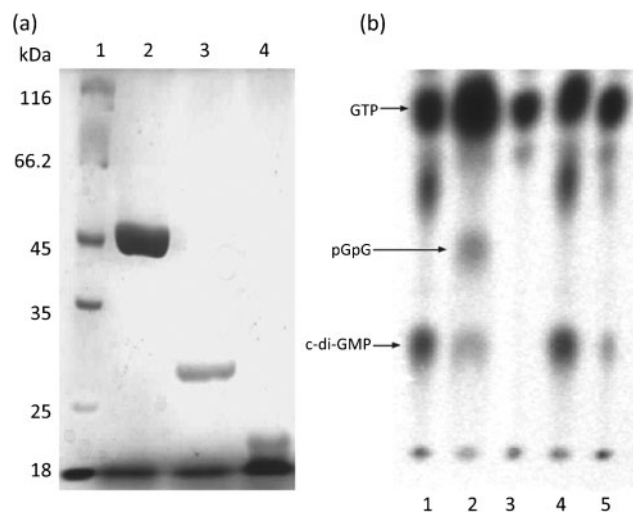
Fig. 4. (a) Purification profile of various domains of MSDGC-1. (a) SDS-PAGE (12 %) showing purified protein from an Ni-NTA column. Lanes: 1, protein marker; 2, GGDEF–EAL domain; 3, EAL domain; 4, GGDEF domain. (b) *In vitro* radiometric activity assays showing DGC and PDE-A activity of various domains of MSDGC-1. Lanes: 1, PleD; 2, full-length MSDGC-1; 3, GGDEF domain; 4, EAL domain; 5, GGDEF–EAL domain protein without GAF domain. The assay buffer used for determining DGC and PDE-A activity contained GTP ([α-^{32}P]GTP) and c-di-GMP ([α-^{32}P]c-di-GMP) as substrates, respectively, as described in Methods. Labelled c-di-GMP was prepared by incubating [α-^{32}P]GTP with PleD, and was used without further purification.

Generation and characterization of a Δ*MSMEG_2196* strain

In order to investigate the role of c-di-GMP in the physiology of *M. smegmatis*, the *MSMEG_2196* gene was deleted by allelic exchange. Conserved GGDEF and EAL regions of MSDGC-1 were replaced with a kanamycin-resistance marker present in the suicide vector pPR27. Colonies were further tested by PCR and Southern blot analysis for confirmation of the crossover event. For PCR, primers MKF and MKR were used, and a shift of ~800 bp was observed in the case of the knockout strain (DGCKO) as compared with the wild-type (data not shown). During Southern blot analysis, the probe hybridized to a fragment of approximately 3.9 kb in the wild-type strain and one of approximately 4.7 kb in the knockout strain (DGCKO). The increase of ~800 bp in DGCKO resulted from the insertion of the kanamycin cassette (Fig. S3).

M. smegmatis has single protein that is known to contain a GAF–GGDEF–EAL domain; therefore, the resulting knock-out strain should not contain any c-di-GMP (c-di-GMP^0). To test this feature, we analysed total nucleotide extracts by HPLC and MALDI-TOF from wild-type *M. smegmatis*, DGCKO and DGCCO cells grown to 72 h in enriched medium, as shown in Fig. 6. Standard c-di-GMP was run as a positive control, and gave a peak at 19.2 min (left panel,

Fig. 6a), and MALDI-TOF spectra confirmed the validity of the peak (right panel). A peak around 19 min was also detected in the chromatogram for the nucleotide extracts of the wild-type strain (Fig. 6b); however, it was absent in the knockout strain (Fig. 6c). This confirmed the presence of c-di-GMP in the wild-type and its absence in strain DGCKO. The validity of our conclusion was further supported by detection of the c-di-GMP peak in the complemented strain (Fig. 6d). Further phenotypic characteristics of the knockout strain were evaluated and compared with those of the wild-type. Growth in planktonic culture (Fig. 7b, 0.02 % glucose) and sliding motility on soft agarose remained unchanged in the knockout strain (Fig. S4a). Unlike in the case of Gram-negative bacteria, we did not observe any difference in the variants of *M. smegmatis* with respect to the ability to form biofilm as compared with the wild-type strain (Fig. S4b, c).

Deletion of *MSMEG_2196* affects long-term survival under nutrient starvation

The mutant (DGCKO) was assayed for its ability to survive for a prolonged period with a limited carbon source (0.02 % glucose). Bacterial cultures were declumped before plating to minimize errors in counting c.f.u. caused by clumping, and OD$_{600}$ was also measured to monitor the growth of different variants in 0.02 % glucose. The knockout strain (DGCKO) had a reduced viability of more than 50 % after 5 days, with a persistent decrease in the number of viable cells as compared with the wild-type (Fig. 7a). To test whether inactivating *MSDGC_2196* causes a polar effect, the DGCKO strain was complemented with the functional gene through the insertional plasmid pMV361 containing *MSMEG_2196*. The complemented strain (DGCCO) was able to survive carbon starvation similarly to the wild-type strain (Fig. 7a).

MSMEG_2196 was subcloned into a shuttle vector, pMV261, under the control of the heat-shock promoter (Hsp60) for overexpression studies. The growth profile of the MSDGC-1-overexpressing strain (DGCOE) was compared with that of the wild-type in 7H9 broth supplemented with 0.02 % (w/v) glucose and 0.05 % (v/v) Tween 80. The overexpressing strain grew more slowly than the wild-type (Fig. 7b). On solid agar medium, colonies produced by overexpression and knockout strains showed different colony morphologies compared with the control (wild-type strain with pMV261 vector). The wild-type *M. smegmatis* showed colonies with a non-uniform margin and rugose surface (wrinkled). The overexpression strain produced a raised colony with smooth margins. The knockout strain had irregular-shaped colonies that were flat and had undulating margins. The complemented strain showed a colony morphology similar to that of the wild-type (Fig. 7c), although not identical. However, it should be mentioned here that phenotypic differences may also arise from the fact that the hsp60 promoter was used in this study, instead of the gene's own promoter.
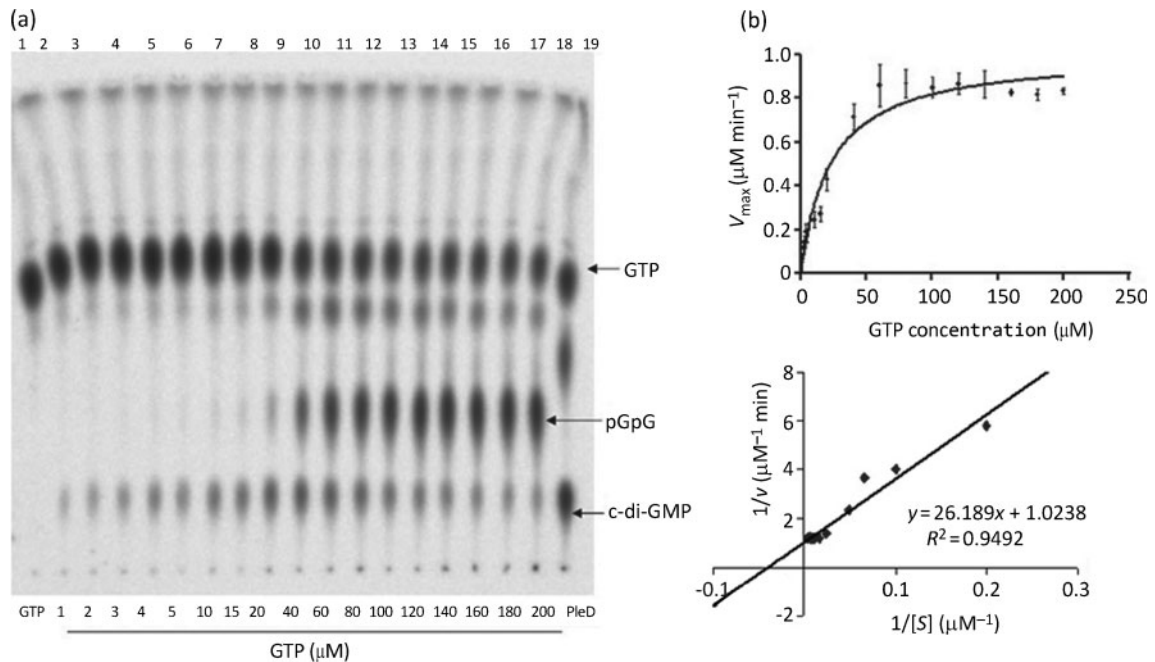
Fig. 5. (a) Velocity (V_{max}) and Michaelis constants (K_m) of the MSDGC-1 protein. Synthesis of c-di-GMP as a function of increasing concentrations of GTP. Lane 1, control (assay buffer); lanes 2–18, increasing concentrations of GTP from 1 to 200 µM. The concentration of the protein (MSDGC-1) was kept at 5 µM. The PleD reaction product was used as a positive control (lane 19). (b) Michaelis–Menten (upper panel) and Lineweaver–Burk plots (lower panel) of c-di-GMP synthesis with increasing concentrations of GTP.

DISCUSSION

In this study, we detected the presence of an additional second messenger, c-di-GMP, and the protein responsible for its synthesis and hydrolysis in a Gram-positive bacterium, *M. smegmatis*. Most of the known proteins with composite GGDEF and EAL domains show either DGC or PDE-A activity, whereby the inactive domain regulates the active partner. In the case of MSDGC-1, we separated and functionally characterized each of the individual domains *in vitro*. We did not detect the formation of c-di-GMP or pGpG as reaction products when individual GGDEF and EAL domain proteins were used. We noticed a substantial reduction in DGC activity of the GGDEF–EAL domain in the absence of the GAF domain. Activities of separated domains were also checked in a heterologous system, which mimics *in vivo* conditions. We were able to detect minimal c-di-GMP synthesis activity by full-length MSDGC-1 only. For the detection of c-di-GMP formation, an overexpression system was used, and any background c-di-GMP coming from the host with an empty vector remained undetectable (Fig. S2, lane 1) and acted as a control.

The requirement of a neighbouring domain or the full-length protein for complete activity has also been observed by others (Ryjenkov *et al.*, 2005; Kazmierczak *et al.*, 2006). MSDGC-1 consists of a GAF domain at its N terminus and is believed to modulate the activity of the protein with a

GGDEF–EAL domain by binding to other molecules such as cAMP, cGMP and c-di-GMP. In proteins with a single GGDEF domain, the presence of a GAF domain or phosphorylation of the N-terminal receiver domain has been shown to increase the activity of the GGDEF domain (Paul *et al.*, 2004; Ryjenkov *et al.*, 2005). Removal of the GAF domain from MSDGC-1 renders the protein inactive, and this suggests that the domain has a role in the regulation of c-di-GMP synthesis and degradation. Recently, a more comprehensive understanding of the regulation of a bifunctional protein, Lpl0329, in the Gram-negative pathogenic bacterium *Legionella pneumophila* has been obtained. In this case, phosphorylation-mediated control of dual activity has been demonstrated, in which upon phosphorylation, the DGC activity of the protein decreases but PDE-A activity remains (Levet-Paulo *et al.*, 2011). While studying the bifunctional behaviour of MSDGC-1 *in vitro*, we used the PleD reaction product to identify c-di-GMP as a positive control, since PleD is a bona fide DGC that synthesizes c-di-GMP. We consistently observed the presence of an additional spot in the PleD reaction product (indicated by an arrow in Fig. 3a) that was absent in the MSDGC-1 reaction product. This not only prompted us to identify the additional spot but also suggested that PleD and MSDGC-1 might use a different reaction mechanism to catalyse the synthesis of c-di-GMP from GTP as a substrate. In order to identify this spot, several experiments were carried out, and we concluded that the additional spot was a linear molecule
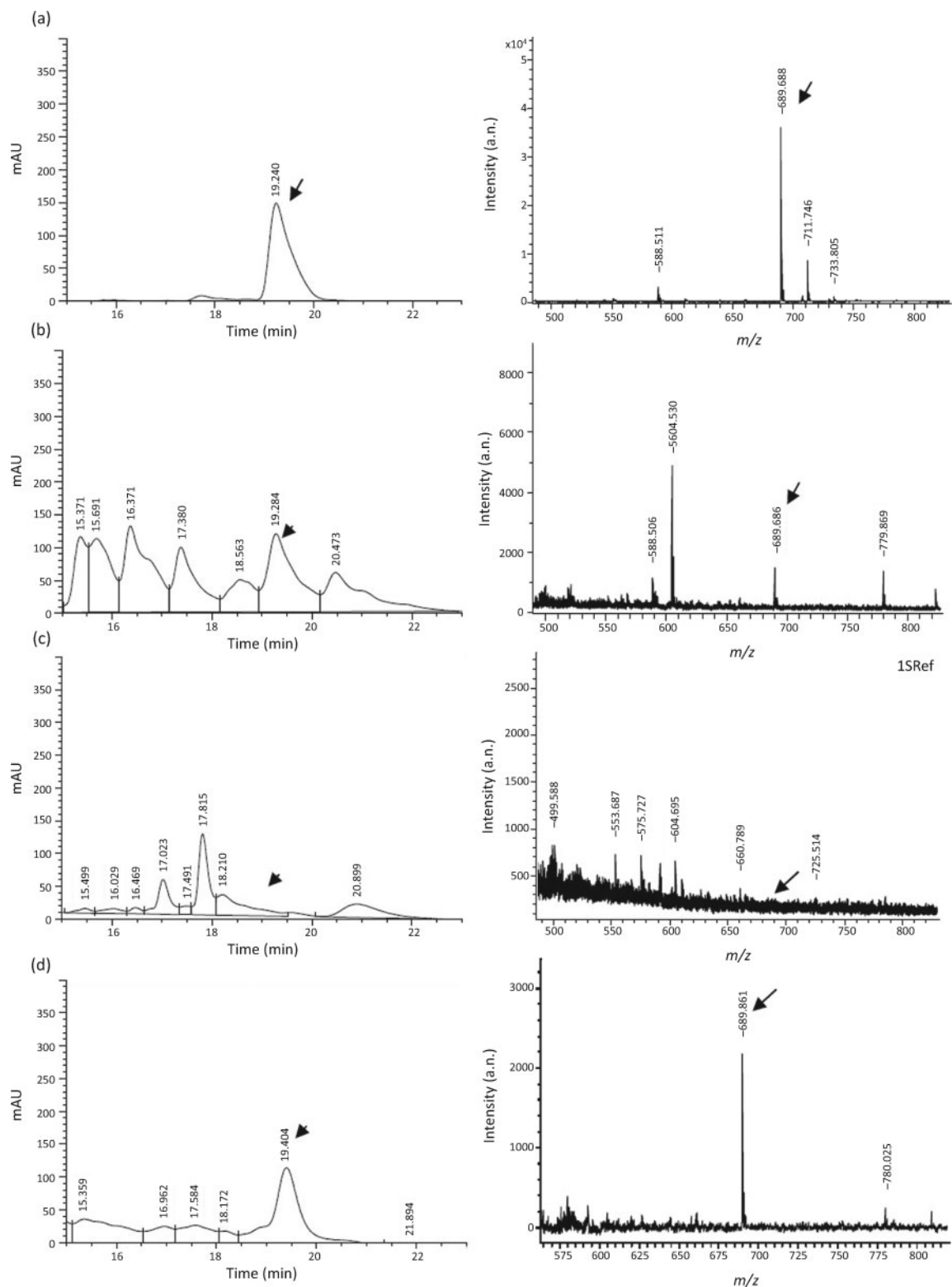
Fig. 6. Detection of c-di-GMP in *M. smegmatis* using reverse-phase HPLC (left panels) and MALDI-TOF (right panels). (a) Standard c-di-GMP showing the peak at 19.2 and *m/z* of 689.6; (b) total nucleotide extract from wild-type *M. smegmatis* grown for 72 h with 2 % glucose (the presence of c-di-GMP is marked with an arrow); (c) absence of c-di-GMP in both reverse-phase HPLC and MALDI-TOF when the total nucleotide extract was analysed as in (b) but devoid of MSDGC-1; (d) reappearance of c-di-GMP in the complemented strain.
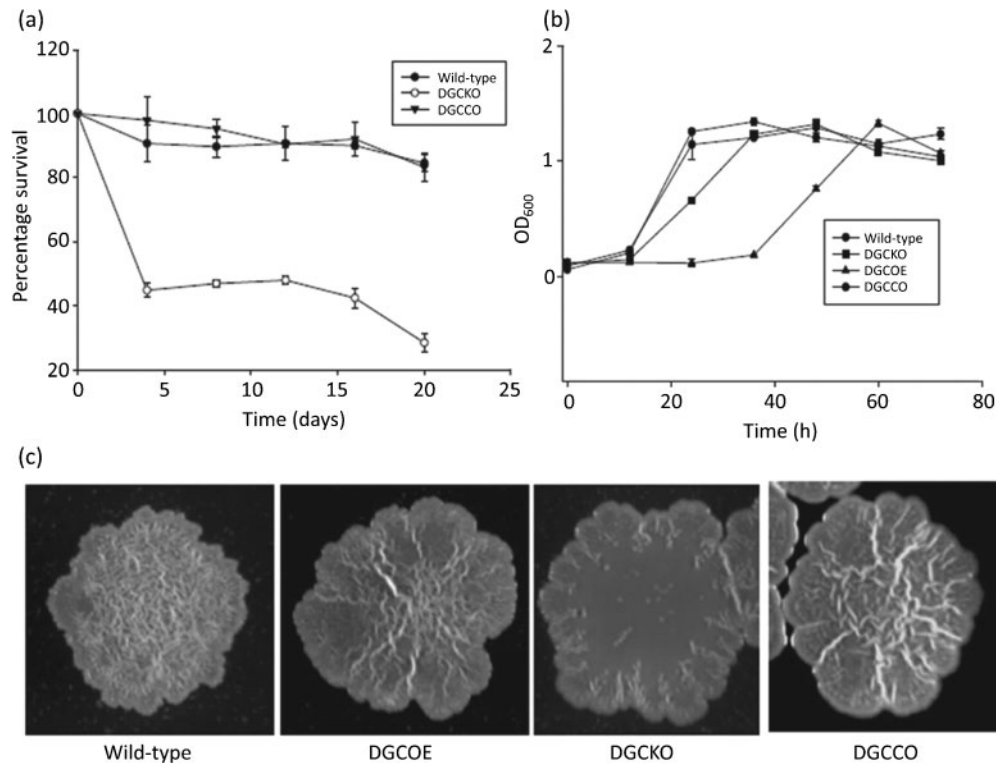
Fig. 7. Role of *msdgc-1* in long-term survival and phenotypic characterization. (a) Long-term survival analysis by c.f.u. determination; (b) growth profile; (c) colony morphology. Strains used in this study were: wild-type, *msdgc-1* deleted strain (DGCKO), *msdgc-1* complemented strain (DGCCO), and *msdgc-1* overexpression strain (DGCOE), where required. (a) and (b) were performed under carbon-limited conditions (0.02 % glucose), and (c) was performed under carbon-replete conditions (2 %) glucose. All experiments were done in triplicate.

which is an intermediate product of c-di-GMP processing by the PleD protein (Fig. S5).

All mycobacterial species have an orthologue of MSDGC-1, which suggests that the protein may play similar roles in these species. Some species have more than one protein with a GGDEF or EAL domain, and the need for more than one protein for the modulation of c-di-GMP has been questioned by some. Studies done in different bacterial systems with multiple copies have suggested that this may help to maintain the level of c-di-GMP in localized compartments of the cell, providing bacteria with the ability to survive in different environmental niches (Römling *et al.*, 2005). This appears to be true in case of the opportunistic pathogen *Mycobacterium avium*, which infects immuno-compromised patients. The *M. avium* genome encodes a greater number of proteins with a GGDEF–EAL domain than those of *M. smegmatis* and *M. tuberculosis*. Other mycobacterial species, such as *Mycobacterium gilvum* and *Mycobacterium vanbaalenii* PYR-1, which have been isolated from sites of bioremediation of polyaromatic hydrocarbons, have even greater numbers of these proteins. This strongly suggests that c-di-GMP is involved in the survival and adaptation of mycobacteria. It has also been shown that during exponential growth, PDE-A activity is favoured to

keep the c-di-GMP level low, and that during starvation, the level of c-di-GMP increases and as a result, c-di-GMP binds and activates the protein YcgR, which in turn binds to the flagellar motor directly and slows it down (Armitage & Berry, 2010; Merritt *et al.*, 2010; Boehm *et al.*, 2010; Paul *et al.*, 2010).

In vivo detection of c-di-GMP in *M. smegmatis* was difficult using radiometric TLC techniques due to the low concentrations encountered. However, we successfully detected the presence of c-di-GMP in stationary phase cultures using HPLC coupled with MALDI-TOF MS. Interestingly, we observed a higher level of c-di-GMP (about sevenfold) in the case of the wild-type strain at stationary phase under carbon-starved conditions (0.02 % glucose) compared with a culture with 2 % glucose (Fig. S6). The level of c-di-GMP was normalized using an equal weight of wet cells, and purified c-di-GMP was used to generate a standard curve. Our data suggest that *msdgc-1* is not involved in biofilm formation and that an increased level of c-di-GMP is required in the *M. smegmatis* stationary phase for survival under nutrient-depleted conditions. Perhaps it is not out of context to mention that the level of the stringent factor ppGpp is maintained by two opposing activities of a single polypeptide, Rel, in Gram-positive bacteria (Chatterji &

Ojha, 2001), which has been found to be overexpressed during late stationary phase. It will be interesting to see whether c-di-GMP and the stringent response have any links between them.

ACKNOWLEDGEMENTS

REFERENCES

Armitage, J. P. & Berry, R. M. (2010). Time for bacteria to slow down. *Cell* 141, 24–26.

Bassler, B. L. (1999). How bacteria talk to each other: regulation of gene expression by quorum sensing. *Curr Opin Microbiol* 2, 582–587.

Beyhan, S., Tischler, A. D., Camilli, A. & Yildiz, F. H. (2006). Transcriptome and phenotypic responses of *Vibrio cholerae* to increased cyclic di-GMP level. *J Bacteriol* 188, 3600–3613.

Boehm, A., Kaiser, M., Li, H., Spangler, C., Kasper, C. A., Ackermann, M., Kaever, V., Sourjik, V., Roth, V. & Jenal, U. (2010). Second messenger-mediated adjustment of bacterial swimming velocity. *Cell* 141, 107–116.

Brouillette, E., Hyodo, M., Hayakawa, Y., Karaolis, D. K. & Malouin, F. (2005). 3′,5′-cyclic diguanylic acid reduces the virulence of biofilm-forming *Staphylococcus aureus* strains in a mouse model of mastitis infection. *Antimicrob Agents Chemother* 49, 3109–3113.

Camilli, A. & Bassler, B. L. (2006). Bacterial small-molecule signaling pathways. *Science* 311, 1113–1116.

Chatterji, D. & Ojha, A. K. (2001). Revisiting the stringent response, ppGpp and starvation signaling. *Curr Opin Microbiol* 4, 160–165.

Christen, M., Christen, B., Folcher, M., Schauerte, A. & Jenal, U. (2005). Identification and characterization of a cyclic di-GMP-specific phosphodiesterase and its allosteric control by GTP. *J Biol Chem* 280, 30829–30837.

Christen, B., Christen, M., Paul, R., Schmid, F., Folcher, M., Jenoe, P., Meuwly, M. & Jenal, U. (2006). Allosteric control of cyclic di-GMP signaling. *J Biol Chem* 281, 32015–32024.

D'Argenio, D. A., Calfee, M. W., Rainey, P. B. & Pesci, E. C. (2002). Autolysis and autoaggregation in *Pseudomonas aeruginosa* colony morphology mutants. *J Bacteriol* 184, 6481–6489.

Dahl, J. L., Arora, K., Boshoff, H. I., Whiteford, D. C., Pacheco, S. A., Walsh, O. J., Lau-Bonilla, D., Davis, W. B. & Garza, A. G. (2005). The *relA* homolog of *Mycobacterium smegmatis* affects cell appearance, viability, and gene expression. *J Bacteriol* 187, 2439–2447.

Dow, J. M., Fouhy, Y., Lucey, J. F. & Ryan, R. P. (2006). The HD-GYP domain, cyclic di-GMP signaling, and bacterial virulence to plants. *Mol Plant Microbe Interact* 19, 1378–1384.

Etienne, G., Villeneuve, C., Billman-Jacobe, H., Astarie-Dequeker, C., Dupont, M. A. & Daffé, M. (2002). The impact of the absence of glycopeptidolipids on the ultrastructure, cell surface and cell wall properties, and phagocytosis of *Mycobacterium smegmatis*. *Microbiology* 148, 3089–3100.

Galperin, M. Y., Natale, D. A., Aravind, L. & Koonin, E. V. (1999). A specialized version of the HD hydrolase domain implicated in signal transduction. *J Mol Microbiol Biotechnol* 1, 303–305.

Galperin, M. Y., Nikolskaya, A. N. & Koonin, E. V. (2001). Novel domains of the prokaryotic two-component signal transduction systems. *FEMS Microbiol Lett* 203, 11–21.

Gupta, K., Kumar, P. & Chatterji, D. (2010). Identification, activity and disulfide connectivity of c-di-GMP regulating proteins in *Mycobacterium tuberculosis*. *PLoS ONE* 5, e15072.

Hickman, J. W., Tifrea, D. F. & Harwood, C. S. (2005). A chemosensory system that regulates biofilm formation through modulation of cyclic diguanylate levels. *Proc Natl Acad Sci U S A* 102, 14422–14427.

Hurley, J. H. (2003). GAF domains: cyclic nucleotides come full circle. *Sci STKE* 2003, PE1.

Jacobs, W. R., Jr, Tuckman, M. & Bloom, B. R. (1987). Introduction of foreign DNA into mycobacteria using a shuttle phasmid. *Nature* 327, 532–535.

Jain, V., Saleem-Batcha, R., China, A. & Chatterji, D. (2006). Molecular dissection of the mycobacterial stringent response protein Rel. *Protein Sci* 15, 1449–1464.

Jenal, U. & Malone, J. (2006). Mechanisms of cyclic-di-GMP signaling in bacteria. *Annu Rev Genet* 40, 385–407.

Karaolis, D. K., Rashid, M. H., Chythanya, R., Luo, W., Hyodo, M. & Hayakawa, Y. (2005). c-di-GMP (3′-5′-cyclic diguanylic acid) inhibits *Staphylococcus aureus* cell-cell interactions and biofilm formation. *Antimicrob Agents Chemother* 49, 1029–1038.

Kazmierczak, B. I., Lebron, M. B. & Murray, T. S. (2006). Analysis of FimX, a phosphodiesterase that governs twitching motility in *Pseudomonas aeruginosa*. *Mol Microbiol* 60, 1026–1043.

Kirillina, O., Fetherston, J. D., Bobrov, A. G., Abney, J. & Perry, R. D. (2004). HmsP, a putative phosphodiesterase, and HmsT, a putative diguanylate cyclase, control Hms-dependent biofilm formation in *Yersinia pestis*. *Mol Microbiol* 54, 75–88.

Kuchma, S. L., Brothers, K. M., Merritt, J. H., Liberati, N. T., Ausubel, F. M. & O'Toole, G. A. (2007). BifA, a cyclic-di-GMP phosphodiesterase, inversely regulates biofilm formation and swarming motility by *Pseudomonas aeruginosa* PA14. *J Bacteriol* 189, 8165–8178.

Levet-Paulo, M., Lazzaroni, J. C., Gilbert, C., Atlan, D., Doublet, P. & Vianney, A. (2011). The atypical two-component sensor kinase Lpl0330 from *Legionella pneumophila* controls the bifunctional diguanylate cyclase-phosphodiesterase Lpl0329 to modulate bis-(3′-5′)-cyclic dimeric GMP synthesis. *J Biol Chem* 286, 31136–31144.

Mathew, R., Mukherjee, R., Balachandar, R. & Chatterji, D. (2006). Deletion of the *rpoZ* gene, encoding the omega subunit of RNA polymerase, results in pleiotropic surface-related phenotypes in *Mycobacterium smegmatis*. *Microbiology* 152, 1741–1750.

Merritt, J. H., Brothers, K. M., Kuchma, S. L. & O'Toole, G. A. (2007). SadC reciprocally influences biofilm formation and swarming motility via modulation of exopolysaccharide production and flagellar function. *J Bacteriol* 189, 8154–8164.

Merritt, J. H., Ha, D. G., Cowles, K. N., Lu, W., Morales, D. K., Rabinowitz, J., Gitai, Z. & O'Toole, G. A. (2010). Specific control of *Pseudomonas aeruginosa* surface-associated behaviors by two c-di-GMP diguanylate cyclases. *MBio* 1, e00183–10.

Miller, M. B. & Bassler, B. L. (2001). Quorum sensing in bacteria. *Annu Rev Microbiol* 55, 165–199.

Naresh, K., Bharati, B. K., Avaji, P. G., Jayaraman, N. & Chatterji, D. (2010). Synthetic arabinomannan glycolipids and their effects on growth and motility of the *Mycobacterium smegmatis*. *Org Biomol Chem* 8, 592–599.

O'Toole, G. A., Pratt, L. A., Watnick, P. I., Newman, D. K., Weaver, V. B. & Kolter, R. (1999). Genetic approaches to study of biofilms. *Methods Enzymol* 310, 91–109.

Ojha, A. K., Mukherjee, T. K. & Chatterji, D. (2000). High intracellular level of guanosine tetraphosphate in *Mycobacterium smegmatis* changes the morphology of the bacterium. *Infect Immun* **68**, 4084–4091.

Paul, R., Weiser, S., Amiot, N. C., Chan, C., Schirmer, T., Giese, B. & Jenal, U. (2004). Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. *Genes Dev* **18**, 715–727.

Paul, K., Nieto, V., Carlquist, W. C., Blair, D. F. & Harshey, R. M. (2010). The c-di-GMP binding protein YcgR controls flagellar motor direction and speed to affect chemotaxis by a ''backstop brake'' mechanism. *Mol Cell* **38**, 128–139.

Pelicic, V., Jackson, M., Reyrat, J. M., Jacobs, W. R., Jr, Gicquel, B. & Guilhot, C. (1997). Efficient allelic exchange and transposon mutagenesis in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* **94**, 10955–10960.

Primm, T. P., Andersen, S. J., Mizrahi, V., Avarbock, D., Rubin, H. & Barry, C. E., III (2000). The stringent response of *Mycobacterium tuberculosis* is required for long-term survival. *J Bacteriol* **182**, 4889–4898.

Reading, N. C. & Sperandio, V. (2006). Quorum sensing: the many languages of bacteria. *FEMS Microbiol Lett* **254**, 1–11.

Römling, U., Gomelsky, M. & Galperin, M. Y. (2005). c-di-GMP: the dawning of a novel bacterial signalling system. *Mol Microbiol* **57**, 629–639.

Ross, P., Weinhouse, H., Aloni, Y., Michaeli, D., Weinberger-Ohana, P., Mayer, R., Braun, S., de Vroom, E., van der Marel, G. A. & other authors (1987). Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. *Nature* **325**, 279–281.

Ryjenkov, D. A., Tarutina, M., Moskvin, O. V. & Gomelsky, M. (2005). Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. *J Bacteriol* **187**, 1792–1798.

Shenoy, A. R. & Visweswariah, S. S. (2006). New messages from old messengers: cAMP and mycobacteria. *Trends Microbiol* **14**, 543–550.

Simm, R., Morr, M., Kader, A., Nimtz, M. & Römling, U. (2004). GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol Microbiol* **53**, 1123–1134.

Smeulders, M. J., Keer, J., Speight, R. A. & Williams, H. D. (1999). Adaptation of *Mycobacterium smegmatis* to stationary phase. *J Bacteriol* **181**, 270–283.

Snapper, S. B., Melton, R. E., Mustafa, S., Kieser, T. & Jacobs, W. R., Jr (1990). Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol Microbiol* **4**, 1911–1919.

Stover, C. K., de la Cruz, V. F., Fuerst, T. R., Burlein, J. E., Benson, L. A., Bennett, L. T., Bansal, G. P., Young, J. F., Lee, M. H. & other authors (1991). New use of BCG for recombinant vaccines. *Nature* **351**, 456–460.

Taga, M. E. & Bassler, B. L. (2003). Chemical communication among bacteria. *Proc Natl Acad Sci U S A* **100** (Suppl. 2), 14549–14554.

Tal, R., Wong, H. C., Calhoon, R., Gelfand, D., Fear, A. L., Volman, G., Mayer, R., Ross, P., Amikam, D. & other authors (1998). Three *cdg* operons control cellular turnover of cyclic di-GMP in *Acetobacter xylinum*: genetic organization and occurrence of conserved domains in isoenzymes. *J Bacteriol* **180**, 4416–4425.

Tamayo, R., Pratt, J. T. & Camilli, A. (2007). Roles of cyclic diguanylate in the regulation of bacterial pathogenesis. *Annu Rev Microbiol* **61**, 131–148.

Tischler, A. D. & Camilli, A. (2004). Cyclic diguanylate (c-di-GMP) regulates *Vibrio cholerae* biofilm formation. *Mol Microbiol* **53**, 857–869.

Weinhouse, H., Sapir, S., Amikam, D., Shilo, Y., Volman, G., Ohana, P. & Benziman, M. (1997). c-di-GMP-binding protein, a new factor regulating cellulose synthesis in *Acetobacter xylinum*. *FEBS Lett* **416**, 207–211.

Edited by: T. Parish