Cyclic di-GMP: a second messenger required for long-term survival, but not for biofilm formation, in *Mycobacterium smegmatis*

Manish Kumar and Dipankar Chatterji

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India

Cyclic di-GMP (c-di-GMP) plays an important role in bacterial adaptation to enable survival in changing environments. It orchestrates various pathways involved in biofilm formation, changes in the cell surface, host colonization and virulence. In this article, we report the presence of c-di-GMP in *Mycobacterium smegmatis*, and its role in the long-term survival of the organism. *M. smegmatis* has a single bifunctional protein with both GGDEF and EAL domains, which show diguanylate cyclase (DGC) and phosphodiesterase (PDE)-A activity, respectively, in vitro. We named this protein MSDGC-1. Deletion of the gene encoding MSDGC-1 did not affect growth and biofilm formation in *M. smegmatis*, but long-term survival under conditions of nutritional starvation was affected. Most of the proteins that contain GGDEF and EAL domains have been demonstrated to have either DGC or PDE-A activity. To gain further insight into the regulation of the protein, we cloned the individual domains, and tested their respective activities. MSDGC-1, the full-length protein, is required for activity, as its GGDEF and EAL domains are inactive when separated.

**INTRODUCTION**

The social behaviour of bacteria is increasingly being understood, and it appears that cell–cell communication or quorum sensing among bacteria plays a very important role in maintenance and survival of organisms. However, the most interesting aspect of quorum sensing is the coordinated response of bacteria; this confers multicellularity to the bacterial population, and generates responses such as biofilm formation, antibiotic production, sporulation, virulence factor expression, etc. (Miller & Bassler, 2001). Second messengers play an important role in relaying extracellular information, and cAMP and ppGpp are two well-known examples of second messengers (Chatterji & Ojha, 2001; Camilli & Bassler, 2006; Shenoy & Viswasvariah, 2006). Another interesting nucleotide derivative, cyclic di-GMP (c-di-GMP), has been found to be involved in modulating cell surface and biofilm formation in several bacteria. This molecule was first reported about 20 years ago in *Gluconacetobacter xylinum* (formerly *Acetobacter xylinum*), as a positive allosteric regulator of cellulose synthase (Ross et al., 1987; Weinhouse et al., 1997; Tal et al., 1998). The derivative c-di-GMP is formed by the condensation of two GTP molecules, and it is hydrolysed to GMP via a linear intermediate, pGpG. The synthesis and breakdown of c-di-GMP is regulated by diguanylate cyclase (DGC) and phosphodiesterase (PDE) encoded by GGDEF and EAL domains, respectively. The GGDEF conserved motif forms the catalytically active site in DGC, and mutations in these conserved residues have resulted in reduced activity of the protein (Paul et al., 2004; Christen et al., 2005). On the other hand, EAL domain proteins degrade c-di-GMP to pGpG, the linear intermediate, which is further hydrolysed to two GMP molecules by another presumably ubiquitous enzyme, PDE-B (Chan et al., 2004; Schmidt et al., 2005; Tamayo et al., 2005). Interestingly, another family of proteins, the HD–GYP-domain proteins, has recently been shown to degrade c-di-GMP to GMP in *Xanthomonas* (Ryan et al., 2006).

In recent years, c-di-GMP has been implicated in different biological processes in bacteria (for reviews, see Jenal & Malone, 2006; Tamayo et al., 2007). An increased level of c-di-GMP has been found to stimulate matrix (exopolysaccharide/cellulose) production, and hence biofilm formation, and to regulate the transition between motile and sessile forms (Simm et al., 2004). Various studies have demonstrated a clear link between the level of c-di-GMP and alterations in cell-surface properties, motility and...
biofilm formation in bacteria (D’Argenio et al., 2002; Kirillina et al., 2004; Tischler & Camilli, 2004; Hickman et al., 2005; Behyan et al., 2006; Kazmierczak et al., 2006; Kuchma et al., 2007; Merritt et al., 2007). Similarly, the role of c-di-GMP in the surface attachment, cell cycle and development of Caulobacter crescentus has been demonstrated elegantly by the research group of Jenal (Aldridge et al., 2003; Paul et al., 2004) and others (Huitema et al., 2006). c-di-GMP has also been implicated in the regulation of virulence in bacterial pathogens. Tischler & Camilli (2005) were the first to demonstrate that c-di-GMP regulates the expression of virulence factors in Vibrio cholerae. The ability of V. cholerae to form a biofilm, and its colonization in the host, are governed by a c-di-GMP-mediated switch (Tamayo et al., 2005). An EAL domain in Salmonella controls bacterial survival in mice by resisting the host phagocyte oxidase (Hisert et al., 2005). Several other studies have indicated a role for c-di-GMP in the virulence of different bacteria, such as Pseudomonas aeruginosa (Kuchma et al., 2005; Kulasekara et al., 2006), Xanthomonas campestris (Behyan et al., 2006), Brucella melitensis (Lestrade et al., 2003) and Bordetella pertussis (Merkel et al., 1998). However, most of the studies demonstrating the role of c-di-GMP in various physiological aspects have been carried out in Gram-negative bacteria. In Gram-positive organisms, the only studies carried out so far have been with Staphylococcus aureus, and these have revealed that treatment with extracellular c-di-GMP inhibits cell–cell interaction and biofilm formation in a dose-dependent manner (Karaolis et al., 2005; Brouillette et al., 2005).

Like most other bacterial genera, the genus Mycobacterium includes both pathogens and saprophytes. The pathogenic species include Mycobacterium tuberculosis and Mycobacterium leprae, which cause tuberculosis and leprosy, respectively. These diseases are serious health problems: one-third of the population of the world is infected by tubercle bacilli. Unfortunately, M. leprae does not grow in vitro, and M. tuberculosis has a doubling time of 24 h; therefore, it is difficult to develop effective drugs against these organisms. Also, latent bacilli can persist in the host for years in an ill-defined physiological state, and become reactivated when the host is immunocompromised (Bloom & Murray, 1992). The development of effective drugs against such highly adaptive pathogens requires a better understanding of their physiology. Many mycobacterial species, including Mycobacterium avium, Mycobacterium fortuitum, Mycobacterium marinum and Mycobacterium smegmatis, form biofilms, although it is unclear whether M. tuberculosis is able to form a biofilm (Hall-Stoodley et al., 2004; Ojha & Hatfull, 2007). The biofilms have high resistance towards commonly used drugs (Teng & Dick, 2003). As described earlier, c-di-GMP plays an important role in biofilm formation and adaptation in several bacterial species; however, to the best of our knowledge, the presence of c-di-GMP has not been reported in mycobacteria.

When we searched for GGDEF and EAL domains, important signatures were obtained in all the mycobacterial species for which genome sequences are available. M. smegmatis, a non-pathogenic fast-growing species, has been 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). To understand the physiological role of c-di-GMP in mycobacteria, we cloned, overexpressed and knocked out the GGDEF–EAL domain protein of M. smegmatis. We report here that M. smegmatis encodes a protein containing GAF, GGDEF and EAL domains in tandem. We named this protein MSDGC-1, and demonstrate that it is able to synthesize and degrade c-di-GMP in vitro. Deletion of the gene encoding MSDGC-1 affects the long-term survival of bacteria under conditions of nutritional starvation. Further, full-length MSDGC-1 is required for both DGC and PDE-A activity, as isolated GGDEF and EAL domains have no detectable biological 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 mc²155 was grown in MB7H9 (Difco) broth with 2 % (w/v) glucose as the carbon source, and 0.05 % (v/v) Tween 80, or on MB7H9 medium containing 1.5 % (w/v) agar. Antibiotics were used at the following concentrations, as and when required (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 polymerase (Finzyme), following the manufacturer’s instructions. All clones generated were confirmed by sequencing (Microsynth). Restriction enzymes used for cloning were obtained from New England Biolabs.

**Bioinformatics analysis.** A search of the M. smegmatis genome for the annotation of proteins containing GGDEF and EAL domains was carried out using the TIGR database (http://cmr.jcvi.org/tigr-scripts/CMR/CmrHomePage.cgi). Two ORFs, MSMEG_2196 and MSMEG_2774, were annotated as putative GGDEF and EAL domain proteins. The protein sequence was subjected to Smart database (v4.0; http://smart.embl-heidelberg.de/) and NCBI Conserved Domains database (v2.13; 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) and MSMEG_2774 (1080 bp, 359 aa) were PCR-amplified from M. smegmatis genomic DNA using primers mAf and mAr, and MKDGCf and MKDGCR, respectively (Table S1). The amplicons MSMEG_2196 and MSMEG_2774 were cloned into the plasmid pET21b using NdeI–NotI and NdeI–HindIII restriction sites, respectively. Vector pET28a was used for N-terminal His₆-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 (Jain et al., 2006b). Briefly, cells were induced with IPTG, harvested, and lysed in

http://mic.sgmjournals.org

2943
Cloning, expression and purification of GAF, GGDEF and EAL domains of MSDGC-1. For the construction of GAF (aa 23–166), GGDEF (aa 175–334) and EAL (aa 348–590), PET DGC_2196 was used as a template, and primers mkGAF (F) and mkGAF (R), mkGGDEF (F) and mkGGDEF (R), and mkEAL (F) and 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 the pET21b vector at NheI and HindIII sites, thus resulting in pGAF, pGGDEF, pEAL and pGGDEF-EAL. N-terminal His6-tag cloning was done using pET28a vector digested with the same set of restriction enzymes. For expression and purification of the protein, the strategy was similar to that used for the full-length protein, as described above.

Enzymic assays. DGC and PDE-A assays were adapted from procedures described by others (Christen et al., 2005). The reaction mixtures with purified proteins MSDGC-1 or MSDGC-2, or PleD protein, contained 25 mM Tris/HCl, pH 7.8, 250 mM NaCl and 10 mM MgCl2 in a 50 μl volume, and the reactions were started by the addition of a mixture of 0.1 mM GTP and [γ-32P]GTP [0.01 μCi μl−1 (370 Bq μl−1)]; BRIT). Aliquots were withdrawn at regular time intervals, and the reaction was stopped by the addition of an equal volume of 50 mM EDTA. Reaction products (2.5 μl) were separated on polyethyleneimine-cellulose plates (Merck) in 1:1.5 (v/v) (NH4)2SO4 and 1.5 M KH2PO4 (pH 3.6), and plates were exposed to a phosphor-imager screen. [α-32P] labelled c-di-GMP, prepared as described below, was used as the substrate for determination of DGC-A activity.

Preparation of [α-32P]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-32P]-labelled and non-labelled c-di-GMP. For radiolabelled c-di-GMP, [α-32P]GTP [125 μCi (4.625 MBq); 3500 Ci mmol−1 (129.5 TBq mmol−1)] was incubated with 100 μg purified His6-tagged PleD protein. The soluble fraction of the E. coli cell lysate (100 μg) was used as a negative control. The protein was precipitated by placing it in boiling water for 5 min, followed by centrifugation, and filtration through a 0.2 μm filter. The supernatant was tested for the presence of c-di-GMP by TLC, and used further without purification. Non-labelled c-di-GMP was prepared by using 1 mM GTP as a substrate, and purified by HPLC, as described (Simms et al., 2004). The identity of c-di-GMP was confirmed by MALDI-TOF MS.

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 (see Table S1), resulting in pMVDCG_2196. This vector was electroporated into M. smegmatis mc155, and colonies were selected on MB7H9 medium containing kanamycin. Growth in liquid culture was followed by plating declumped bacteria onto MB7H9 agar plate containing 2 % (w/v) glucose as the carbon source. Bacterial clumping was minimized by vortexing the cells with 0.5 mm glass beads, as described (Primm et al., 2000).

Targeted replacement of MSMEG_2196 in M. smegmatis mc155. A recombination cassette was constructed to delete MSMEG_2196 from the chromosome of M. smegmatis mc155 (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, and a fragment spanning from base 1175 of MSMEG_2196 to base 321 downstream of MSMEG_2196, with an EcoRI fragment holding the aph gene from vector pUC4K between them. After the preparative cloning steps, the whole recombination cassette was transferred to the suicide vector pPR27 to obtain the final construct pDGCKO. M. smegmatis mc155 was transformed with pDGCKO. The sucrose-resistant, gentamicin-sensitive and kanamycin-resistant colonies were selected for further analysis. The disruption of the gene, and the recombination event, were verified by Southern hybridization and PCR in selected colonies. For PCR, the primers used were mkGGDEF (F) and mkEAL (R) and KO-F and KO-R.

Table 1. Strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Characteristics</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. smegmatis mc155</td>
<td>Parental strain</td>
<td>This study</td>
</tr>
<tr>
<td>ΔMSDG–1-1</td>
<td>M. smegmatis in which MSMEG_2196 is replaced with a Kan' cassette</td>
<td>This study</td>
</tr>
<tr>
<td>M smeg pMV</td>
<td>M. smegmatis transformed with pMV261 vector alone, Kan'</td>
<td>This study</td>
</tr>
<tr>
<td>M smeg pMV2196</td>
<td>M. smegmatis transformed with pMVDGC_2196, Kan'</td>
<td>This study</td>
</tr>
<tr>
<td>MSDGC–1 complement</td>
<td>MSDGC–1 complemented with MSMEG_2196, Kan' Hyg'</td>
<td>This study</td>
</tr>
<tr>
<td>pUC4K</td>
<td>Source of aph gene</td>
<td>Pharmacia Biotech</td>
</tr>
<tr>
<td>pET21b</td>
<td>Cloning vector (C-terminal His6 tag), Amp'</td>
<td>Novagen</td>
</tr>
<tr>
<td>pET28a</td>
<td>Cloning vector (N-terminal His6 tag), Amp'</td>
<td>Novagen</td>
</tr>
<tr>
<td>pPR27</td>
<td>Suicide vector, Gen'</td>
<td>Pelicic et al. (1997)</td>
</tr>
<tr>
<td>pMV261</td>
<td>Cloning vector with hsp60 promoter, Kan'</td>
<td>Stover et al. (1991)</td>
</tr>
<tr>
<td>pMV complement</td>
<td>MSDGC–1 cloned in pMV361, Hyg'</td>
<td>This study</td>
</tr>
<tr>
<td>pMV361</td>
<td>Cloning vector with hsp60 promoter, Hyg'</td>
<td>This study</td>
</tr>
<tr>
<td>pMVDCG_2196</td>
<td>MSDGC–1 cloned in pMV261, Kan'</td>
<td>This study</td>
</tr>
<tr>
<td>pET DGC_2196</td>
<td>MSDGC–1 cloned in pET21b, Amp'</td>
<td>This study</td>
</tr>
<tr>
<td>pET DGC_2774</td>
<td>MSDGC–2 cloned in pET21b, Amp'</td>
<td>This study</td>
</tr>
<tr>
<td>pET DGC_2774 Nter</td>
<td>MSDGC–2 cloned in pET28a, Kan'</td>
<td>This study</td>
</tr>
<tr>
<td>pGAF</td>
<td>GAF domain of MSDGC–1 cloned in pET21b, Amp'</td>
<td>This study</td>
</tr>
<tr>
<td>pGGDEF</td>
<td>GGDEF domain of MSDGC–1 cloned in pET21b, Amp'</td>
<td>This study</td>
</tr>
<tr>
<td>pEAL Nter</td>
<td>EAL domain of MSDGC–1 cloned in pET28a, Kan'</td>
<td>This study</td>
</tr>
<tr>
<td>pGGDEF-EAL</td>
<td>GGDEF–EAL domain of MSDGC–1 cloned in tandem</td>
<td>This study</td>
</tr>
</tbody>
</table>
KO-R (Table S1). For Southern hybridization, the genomic DNA of the putative mutant, as well as that of mc155, was digested with PvuII. The probe used was a fragment of approximately 1.8 kb DNA from MSMEG_2196, and it was labelled with [α-32P]ATP by using the Random Primer Labelling kit (Bangalore Genei), following the protocol of the manufacturer.

Western blot analysis. Western blot analysis for detection of MSDGC-1 was carried out with primary antibodies raised in rabbit against the purified His6-tagged protein. Proteins extracted from cells grown for different lengths of time were normalized, separated by SDS-PAGE, and blotted onto a PVDF membrane. The polyclonal sera were diluted by a factor of 2000 before use, and secondary antibodies were purchased from Sigma Aldrich. The blots were developed with 26 mg aminoethylcarbazole ml−1 and 0.01% (v/v) H2O2. The x subunit of RNA polymerase was used as an internal control.

Isolation and detection of c-di-GMP from M. smegmatis. The cells were grown in MB7H9 with 2.0% (w/v) glucose, and then harvested and washed. Nucleotides were extracted by acid, as described (Weinhouse et al., 1997). Briefly, cells were incubated on ice with 0.6 M HClO4, and extracts were neutralized with 5 M K2CO3. Samples were centrifuged to remove the precipitate of KClO4, and the supernatant was lyophilized. Lyophilized extracts were resuspended in HPLC buffer A (see below), and subjected to HPLC separation. Purified c-di-GMP, as described above, was used as a standard, and peak area was used for quantification.

HPLC and MS. Samples (20 μl) were injected onto a C-18 column (250 x 4.6 cm; Supelco), and separated by reverse-phase HPLC (Agilent 1100). Buffer A (100 mM KH2PO4, and 4 mM tetrabutyl ammonium hydrogen sulfate, pH 5.9) and buffer B [75% (v/v) buffer A and 25% (v/v) methanol] were used in the gradient programme, as described by Ryjenkov et al. (2005). Nucleotides were detected at the wavelength of 254 nm. For MS, samples were diluted 1:10 in a solution containing 10 mg dihydroxybenzoic acid ml−1, 0.1% (v/v) trifluoroacetic acid, and 50% (v/v) acetonitrile. A total of 1 μl was spotted onto a MALDI-TOF MS plate, and the samples were analysed in reflectron and positive-ion modes with an Ultraflex TOF-TOF mass spectrometer (Bruker Daltonics). The resulting spectra were calibrated using the standard calibration kit provided by the manufacturer.

Long-term starvation cultures. AMSDGCI, the wild-type strain of M. smegmatis mc155, and the supplemented strain, were grown in MB7H9 containing 0.02% (w/v) glucose and 0.05% (v/v) Tween 80 until saturation. The antibiotics were omitted from the culture to rule out any possible effects that they might have had on long-term survival. In the knockdown strain and the complemented strain, the antibiotic marker was integrated chromosomally, and therefore it was not lost. Bacterial cultures were declumped before plating on agar, as described earlier. The number of c.f.u. was determined at regular intervals of time up to 20 days.

Biofilm formation assay. The biofilm formation assay was performed as described (Mathew et al., 2006). Cultures grown to stationary phase were diluted 1:100 to inoculate 30 ml Sauton’s medium in 90 mm diameter polystyrene Petri dishes, and the medium in 90 mm diameter polystyrene Petri dishes, and the Sauton’s medium. This culture was then diluted to a final OD600 of 0.05 in Sauton’s medium, and distributed into 96-well polystyrene plates at time zero. 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 film 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 A550 was measured using a microtitre plate reader.

RESULTS

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

The domain organization of sequences MSMEG_2196 and MSMEG_2774 was obtained from the Smart database, as depicted in Fig. 1. For convenience, we named the sequences MSDGC-1 and MSDGC-2, respectively, and these names are referred to throughout the text. When searching for GGDEF and EAL domains in other mycobacterial species, we found approximately 70 sequences containing either or both GGDEF and EAL conserved domains. Representatives of different domain architectures in various mycobacterial proteins are depicted in Fig. 1. Table 2 lists the number of proteins and their domain organization in different mycobacteria. Intriguingly, most mycobacterial species appear to have at least one protein with GAF–GGDEF–EAL domain organization (Fig. 1), except M. bovis BCG strain Pasteur, whereas Mycobacterium gilvum PYR-GCK and Mycobacterium vanbaalenii PYR-1 have the highest numbers of these proteins among mycobacteria, i.e. 22 and 15, respectively (Table 2). Proteins from different mycobacterial species with GAF–GGDEF–EAL domain architecture share >60% amino acid identity among themselves. We also observed several sequences with a GGEEF motif, in which D had been changed to E, and a few EAL domains with mutated GGDEF domains in the N-terminal region. For example, M. vanbaalenii PYR-1 has a protein with GGDEF and EAL domains in which GGDEF is altered to EAL.

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

In order to study the DGC and PDE-A activity of MSDGC-1 and MSDGC-2, the proteins were expressed in E. coli BL21 (DE3) cells, and purified to homogeneity, as judged by SDS-PAGE profiles (Fig. 2) and MS analysis (data not shown). MSDGC-1 showed c-di-GMP synthesis and degradation activity in vitro (Fig. 3). On the TLC plate, the spot with the same Rf value as that of c-di-GMP was apparently followed by another spot with the Rf of linear di-GMP (pGpG). A reaction mixture containing PleD protein (Paul et al., 2004), which is a known DGC protein of C. crescentus, was used as a positive control, and E. coli cell lysate was used as a negative control. For the identification of the reaction products, the samples that were not radioactively labelled were either directly analysed by MALDI-TOF MS (Fig. S1, available with the online version of this paper), or subjected to HPLC (Fig. 4), followed by collection of peaks and analysis by MALDI-TOF MS. The peak at 16.1 min had the same retention time as that of GTP, and a prominent [M + H]+ at an m/z ratio of...
524 was detected. Similarly, the peak at 18.5 min had the same retention time as that of purified c-di-GMP, and had the identical \( m/z \) value at \( 691\ [M+H]^+ \). The HPLC fraction at 16.4 min gave a signal at \( m/z \) 709 \([M+H]^+\), and this corresponded to the addition of a water molecule to c-di-GMP (Fig. 4). This is in agreement with EAL having PDE activity, which hydrolyses c-di-GMP to pGpG. Purified c-di-GMP was prepared from PleD (Paul et al., 2004), and used as a control. However, longer incubation times did not yield GMP, thus indicating that only PDE-A activity was present.

N-terminal His\(_6\)-tagged MSDGC-2 did not show c-di-GMP synthesis in vitro (Fig. 3b).

**Overexpression of MSDGC-1 alters the growth rate and colony morphology**

MSDGC-1 was subcloned in shuttle vector pMV261 under the control of the heat shock promoter hsp60. We
compared the growth of the MSDGC-1-overexpressing strain with that of the wild-type in MB7H9 broth supplemented with 2% (w/v) glucose and 0.05% (v/v) Tween 80. It can be seen from Fig. 5(a) that the overexpressing strain grew at a slower rate than the wild-type. On solid agar medium, the overexpressing strain showed a different colony morphology compared with the control, which was the wild-type containing the vector only (Fig. 5b). The normal colony of M. smegmatis had a non-uniform margin, rugose surface, and comparatively very large size, whereas the overexpressing strain had a uniform margin, smooth surface and round shape, and was smaller in size.

**Deletion of MSDGC-1 affects the long-term survival of M. smegmatis under conditions of nutritional starvation**

To establish the role of MSDGC-1 in the physiology of M. smegmatis, the gene MSMEG_2196 was deleted by allelic exchange. In the suicidal vector pPR27, the conserved regions of GGDEF and EAL of MSDGC-1 were replaced with a kanamycin-resistance marker. The advantage of using a pPR27-based construct was the ability of the plasmid to replicate at low temperatures only. This allows the bacteria to undergo recombination during growth at 30°C (Pelicic et al., 1997). Subsequent selection for kanamycin-resistant, sucrose-resistant colonies at 39°C, a temperature at which the plasmid cannot replicate, produces the double cross-over recombinant only. Colonies were further tested by PCR and Southern hybridization for confirmation of the knockout event. We obtained an amplicon size of 1.2 kb in the wild-type, and 2.0 kb in the knockout strain ΔMSDGC-1 (Fig. S2, available with the online version of this paper). In Southern hybridization, the probe hybridized to fragments of approximately 3.9 kb and 0.5 kb in the parental strain, and fragments of approximately 4.7 kb and 0.5 kb in the knockout strain. The increase in the latter strain resulted from the 800 bp increase in size of the region produced by insertion of the kanamycin cassette (Fig. S2).

**Fig. 2.** Purification of His<sub>6</sub>-tagged proteins. Lanes: 1, marker; 2, MSDGC-1; 3, MSDGC-2; C-terminal His<sub>6</sub>-tagged MSDGC-1 is approximately 67 kDa, and N-terminal His<sub>6</sub>-tagged MSDGC-2 is approximately 50 kDa. The purity of the preparations was determined by SDS-PAGE, and the gels were stained with Coomassie brilliant blue. Molecular mass (kDa) is shown on the left.

**Fig. 3.** (a) M. smegmatis protein MSDGC-1 has both DGC and PDE-A activities. Purified His<sub>6</sub>-tagged protein was assayed for the ability to synthesize c-di-GMP. The protein (5 μM) was incubated with assay buffer containing [α-<sup>32</sup>P]GTP, as described in Methods. Samples were withdrawn at different time points, and analysed on polyethyleneimine-cellulose TLC. 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 control (g). (b) A similar protocol was followed for MSDGC-2, and c-di-GMP synthesis was not observed. (c) The soluble protein fraction of E. coli cell extract (100 μg) was used as a negative control.
M. smegmatis has only one protein that is predicted to contain a GGDEF-EAL domain; therefore, the resulting knockout strain should be without c-di-GMP (c-di-GMP0). To test this feature, we analysed the total nucleotide extract from the wild-type and ΔMSDGC-1 by using HPLC coupled with MALDI-TOF MS. At 18.5 min, a sharp peak corresponding to c-di-GMP was observed in wild-type extracts, whereas no signal was observed in ΔMSDGC-1 (Fig. 6). Further, we collected these fractions, and identified them as c-di-GMP by using MALDI-TOF MS (data not shown).

When we checked ΔMSDGC-1 for its ability to form a biofilm, we did not observe any difference as compared

![Fig. 4. Detection of c-di-GMP and pGpG by HPLC analysis. A DGC reaction mixture without protein (a), with MSDGC-1 (b), and with purified c-di-GMP (c), was separated by reverse-phase HPLC (see Methods). Right, MALDI-TOF MS analysis of relevant HPLC fractions (GTP, c-di-GMP and pGpG, as indicated by arrows in a, b and c, respectively) In the positive-ion detection mode, major ions were detected at m/z values of: 524 [M+H]+ and 562 [M+K]+ for GTP; 691 [M+H]+ and 729 [M+K]+ for c-di-GMP; and 709 [M+H]+ and 747 [M+K]+ for pGpG.](image-url)
with the wild-type strain (Fig. S3, available with the online 
version of this paper). Also, growth in planktonic culture, 
and sliding motility on the soft agar, remained unchanged 
in the knockout strain (data not shown). Further, we tested 
the ability of the mutant to survive under long-term 
starvation of the carbon source. It can be observed from 
Fig. 7 that the culture of the knockout strain had 100-fold 
less viable bacteria compared with that of the wild-type 
when grown in MB7H9 medium containing 0.02 % (w/v) 
glucose and 0.05 % (v/v) Tween 80. Bacterial cultures were 
declumped before plating to minimize errors in counting 
caused by clumping (Fig. S4, available with the online 
version of this paper). However, it could be argued that 
this phenotype was caused by the polar effect of the 
disruption of MSDGC-1, and not by the loss of function of 
the gene. To rule out this possibility, we complemented the 
ΔMSDGC-1 strain with the functional gene through the

**Fig. 5.** Overexpression of MSDGC-1 alters the growth and colony 
morphology of *M. smegmatis*. (a) The cultures were grown in 
MB7H9 with 2 % (w/v) glucose and 0.05 % (v/v) Tween 80. The 
arrows indicate the time point when cultures were shifted to 42 °C 
for induction. ▲, Wild-type containing pMV261 vector (control); 
○, strain overexpressing MSDGC-1. The values represent the 
mean (±sd) of two independent experiments done in triplicate. (b) 
When cultured on agar, the strain overexpressing MSDGC-1 
produced small, round, smooth colonies, whereas *M. smegmatis* 
mc²155 wild-type strain (WT) produced large colonies with 
irregular margins and a rough surface.

**Fig. 6.** The ΔMSDGC-1 knockout strain is a c-di-GMP⁰ strain. 
The cells were grown for 72 h in MB7H9 medium with 2 % (w/v) 
glucose, and then centrifuged. The total nucleotides extracted from 
*M. smegmatis* mc²155 (b), ΔMSDGC-1 (c), and the MSDGC-1 
complemented strain (d), were run on reverse-phase HPLC. 
Purified c-di-GMP was run as control (a). The relevant portions of 
the chromatograms are shown. The fractions corresponding to c-
di-GMP peak (indicated by arrows) were collected, and analysed 
by MALDI-TOF MS to confirm their identity.
pMV361 vector containing MSDGC-1, and observed the reversal of the mutant phenotype (Fig. 7).

The level of c-di-GMP increases in the stationary phase of growth

The level of c-di-GMP in bacteria varies depending on the environmental conditions, and the phase of growth. We determined the level of MSDGC-1 production and c-di-GMP concentration in three phases of growth, i.e. exponential, early stationary and late stationary (Fig. 8a). Whole-cell lysates of \textit{M. smegmatis} mc\textsuperscript{155} were separated on SDS-PAGE, and analysed by Western blot using polyclonal antibodies, which had been raised against MSDGC-1 in rabbit. As expected, an immunoreactive protein band, approximately 67 kDa in size (Fig. 8b), was detected, with increased band intensity in stationary-phase samples. The amount of MSDGC-1 was about three times higher in stationary-phase samples compared with that in exponential-phase samples. We also measured the levels of c-di-GMP in the samples. Larger amounts of cells (3 g wet weight) were used to extract the total nucleotide, and the samples were subsequently run on HPLC. The peak with the same retention time as standard c-di-GMP was collected, and identified by MALDI-TOF MS. The peak areas were plotted for comparison, and the results were in agreement with an increased level of the protein MSDGC-1 in stationary-phase samples. The early- and late-stationary-phase cultures had (mean \pm SD) relative peak areas of 0.85 \pm 0.25 and 1.90 \pm 0.18, respectively, which indicated that they contained levels of c-di-GMP that were 1.3 and 3.1 times greater, respectively, than that of the exponential-phase culture (relative peak area 0.60 \pm 0.16).

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

The protein MSDGC-1 is a three-domain protein, with GGDEF and EAL domains in tandem with the N-terminal GAF domain, and it possesses both DGC and PDE-A activity. The GAF (cGMP, adenylyl cyclase, FhlA) domain is a cyclic nucleotide (cGMP)-binding domain present in PDE (Hurley, 2003). To understand the regulation of MSDGC-1, we cloned the GAF, GGDEF and EAL individually, and functionally characterized each one. We were not able to overexpress and purify the GAF domain, despite several attempts in which we cloned the domain in both C- and N-terminal His\textsubscript{6}-tag orientations. The GGDEF domain protein was purified by the addition of Triton X-100 to the lysis buffer to improve its solubility. However, the protein was partially contaminated with other proteins (Fig. 9a), and we performed the DGC assay without further purification of the protein. We did not observe a spot for c-di-GMP on the TLC plate, and this indicated that the DGC

Fig. 7. \(\Delta\)MSDGC-1 is not able to survive under conditions of long-term starvation. Cultures were grown in MB\textsuperscript{7}H\textsuperscript{9} medium with 0.02\% (w/v) glucose and 0.05\% (v/v) Tween 80. Aliquots were withdrawn at different times after the culture had reached stationary phase (48 h growth), and plated on MB agar containing 2\% (w/v) glucose as the carbon source. The number of c.f.u. was determined. ▲, Wild-type; □, \(\Delta\)MSDGC-1; ○, complemented strain.

Fig. 8. The level of c-di-GMP increases in stationary phase. (a) Growth curve of \textit{M. smegmatis} grown in MB\textsuperscript{7}H\textsuperscript{9} with 2\% (w/v) glucose as the carbon source. Arrows indicate the time points at which cells were harvested, lysed, and used for Western blot analysis and determination of the level of c-di-GMP. (b) Western blot of whole-cell lysate probed with antibody raised against MSDGC-1. The \(\alpha\) subunit of RNA polymerase (RpoA) was used as an internal control because its level remains unchanged in different phases of growth. Purified MSDGC-1 and the cell lysate from \(\Delta\)MSDGC-1 were used as positive and negative controls, respectively.
activity of the GGDEF domain was greatly reduced (Fig. 9b, lane 2). Similarly, purification of the isolated EAL domain yielded a partially purified protein that did not show PDE-A activity (Fig. 9b, lane 5). Further, trans-complementation or addition of one domain to the other separately, and prolonged incubation, did not improve DGC or PDE-A activity (data not shown). This indicates that the full-length protein is required for both DGC and PDE-A activities. One may argue that the full-length protein has four cysteine residues distributed outside the GGDEF domain, and that these may help in correct folding and oligomerization of the protein, and hence the activity of the isolated GGDEF domain was affected. It was clear that correct folding of the full-length protein is important for GGDEF–EAL activities, and that fused domains are mandatory, from the observation that the presence of the two domains in solution did not restore the catalytic activity of either domain. It has also been suggested that the GAF domain may help in binding to GTP (the substrate) or c-di-GMP, and that it is required for the activity of GGDEF. The combined GGDEF–EAL domain showed poor DGC activity, and PDE-A activity was not detected (Fig. 9b); this implies that the GAF domain is important for the activity of the combined domain.

**DISCUSSION**

c-di-GMP has established itself in the ranks of secondary messengers, such as cAMP and ppGpp, and it plays various roles in microbial physiology. The number of pathways being found to be influenced by this molecule in bacteria is constantly increasing. Here, we present evidence that the protein that can synthesize and degrade c-di-GMP is present in *M. smegmatis*. Our study produced the following experimental findings: (i) c-di-GMP is present in a Gram-positive bacterium; (ii) the bacterium used in this study has only one protein with a GGDEF–EAL domain; (iii) the protein is bifunctional, having both DGC and PDE-A activities. MSDGC-1 is a multidomain protein with GAF, GGDEF and EAL domains in tandem. We would like to note here that all known examples of composite GGDEF–EAL proteins have either DGC or PDE activity, and from the primary sequence it is difficult to predict the dominant activity (Tal et al., 1998; Kazmierczak et al., 2006; Kulasakara et al., 2006). However, it has been demonstrated recently that proteins with GGDEF and EAL domains can have both DGC and PDE activities, depending on the environmental conditions. For example, a full-length bacteriochrome protein from *Rhodobacter* has PDE activity, and the activity lies in the C-terminal EAL domain; however, in the absence of an EAL domain, it shows light-dependent DGC activity (Tarutina et al., 2006). Similarly, ScrC, a GGDEF–EAL protein from *Vibrio parahaemolyticus*, has intrinsic DGC and PDE activities, and these activities are controlled by ScrAB. Primarily, ScrC acts as PDE-A in the presence of ScrA and ScrB, whereas it acts as DGC in the absence of ScrA and ScrB (Ferreira et al., 2008).

It has been proposed that inactive domains containing a deviation from the conserved motif have a regulatory role (Römling et al., 2005; Jenal & Malone, 2006), or that they could form a separate subclass of GGDEF domains (Pei & Grishin, 2001). An example of this is the CC3396 protein of *C. crescentus*, where the N-terminal domain has a degenerated GGEF motif that binds to GTP, which allosterically controls the PDE-A activity of the EAL domain (Christen et al., 2005).
MSDGC-1 requires the presence of all the three domains, and deletion of any single domain compromises its activity. We were interested to find out how the opposing activities in the same protein are regulated; therefore, it was appropriate to make a comparison with the biosynthesis of another second messenger, ppGpp, which is a stringent factor. In Gram-negative organisms, two proteins, RelA and SpoT, control the synthesis and degradation of ppGpp, respectively, and maintain the level of ppGpp, which interacts with RNA polymerase, and downregulates rRNA synthesis under conditions of nutritional starvation (Cashel et al., 1996). Apparently, the protein Rel, which has RelA and SpoT in tandem, and a regulatory domain at its C-terminal, manifests the intradomain interaction into an opposing catalytic activity as a function of environmental conditions (Jain et al., 2006b). However, in the case of MSDGC-1, in order to investigate the interdependence of the domains, we separated and functionally identified each of the individual domains. We did not detect the presence of c-di-GMP or pGpG as reaction products when individual GGDEF and EAL domain proteins were used. It was also noted that there was a substantial reduction in DGC activity when the GGDEF–EAL domains were in tandem.

The requirement of a neighbouring domain or the full-length protein for the full activity has been reported by others (Ryjenkov et al., 2005; Kazmierczak et al., 2006). MSDGC-1 has a GAF domain that has not yet been studied in the context of c-di-GMP, but it is believed to modulate the activity of proteins with GGDEF–EAL domains by binding to other molecules, such as cAMP, cGMP or c-di-GMP. In single GGDEF domains investigated to date, the presence of a GAF domain, or phosphorylation of the N-terminal receiver domain, such as PAS, increases the activity of the GGDEF domain (Paul et al., 2004; Ryjenkov et al., 2005). Removal of the GAF domain from MSDGC-1 makes the protein inactive, and this suggests that the protein has a role in regulation of c-di-GMP synthesis and degradation. Mutational studies are underway to further understand the mechanism of regulation.

Genes encoding proteins that contain GGDEF and EAL domains are abundant in bacterial genomes, and they are usually found along with genes encoding sensory or regulatory domains, such as GAF and PAS, REC and FlhA (Galperin et al., 2001). All mycobacterial species have a homologue of MSDGC-1 (approx. 60% identity at the amino acid level), and this suggests that the protein may have a similar role to play in these species. Some species have more than one protein containing a GGDEF–EAL domain, and the need for more than one protein for the modulation of c-di-GMP has been questioned for some time. Studies done in different bacterial systems with multiple numbers of these proteins have suggested that the level of c-di-GMP is maintained in localized compartments of the cell, and that it provides bacteria with the ability to survive in different environmental niches in a specific manner (Römling et al., 2005). This appears to be true for some species of mycobacteria. For example, species such as M. avium, which proliferate well in water and soil, and aerosols and tissues of humans and animals, have a greater number of proteins with the GGDEF–EAL domain compared to M. smegmatis and M. tuberculosis. Other mycobacterial species, such as M. gilvum and M. vanbaalenii PYR-1, which have been isolated from polyaromatic hydrocarbon bioremediation sites, have even greater numbers of these proteins. This strongly suggests that c-di-GMP could be involved in the survival and adaptation of mycobacteria. In contrast to the well-studied models, such as E. coli, Pseudomonas, Salmonella and Vibrio, M. smegmatis has only one functional GGDEF–EAL domain. Although we found two proteins with GGDEF–EAL domains in the M. smegmatis genome, only one (MSDGC-1) was enzymically active in our assay system. MSDGC-2 (GGDEF domain, E value $2 \times 10^{-12}$) has an SDSEF amino acid sequence instead of the conserved NVDEF motif, and this might result in a decrease in DGC activity. However, in Shewanella oneidensis, the GGDEF domain protein with the NVDEF motif has been demonstrated to have DGC activity (Thormann et al., 2006). Further, we did not detect c-di-GMP in the ΔMSDGC-1 strain, thus establishing that M. smegmatis has only one enzymically active GGDEF–EAL domain. Recently, microarray studies on the M. smegmatis transcriptome in the biofilm mode of growth, and in the stationary phase of planktonic culture, have shown that MSDGC-2 is significantly induced by fourfold in stationary phase and during maturation of biofilms, as compared with the exponential-growth phase (Ojha & Hatfull, 2007). At present, we do not know what role MSDGC-2 plays in vivo, and how it helps in maturation of biofilms.

An elevated intracellular level of c-di-GMP alters gene expression, and modulates cell-surface properties, and this leads to the formation of a biofilm and reduced motility (Beyhan et al., 2006; Méndez-Ortiz et al., 2006). Studies by those authors in E. coli and V. cholerae have revealed that, in the presence of high levels of c-di-GMP, the genes responsible for flagella synthesis, motility and chemotaxis are repressed, and the genes for responsible exopolysaccharide synthesis and membrane-associated genes are induced; it has been found that most bacteria respond in the same manner towards high levels of c-di-GMP, and exhibit a similar phenotype. However, so far no information is available in the literature regarding biofilm formation and the intracellular level of c-di-GMP in Gram-positive bacteria.

Our data provide evidence that c-di-GMP has a role to play in the long-term survival of bacteria under conditions of nutritional starvation. It was surprising to find that the growth of the knockout strain and wild-type strain was comparable in liquid culture, and also that their ability to form a biofilm was similar. We noticed that old cultures from MB7H9 agar plates were difficult to revive, although mycobacteria are well known for their long-term persistence in stressful conditions, such as nutritional deprivation, oxygen limitation and extreme pH; this persistence is
often referred to as dormancy (Parrish et al., 1998; Smeluders et al., 1999). Production of ppGpp, another second messenger, is the hallmark of the stringent response, and it plays a pivotal role in the persistence and survival of bacteria (for a recent review see Jain et al., 2006a). Studies have shown that ppGpp is required for the long-term survival of M. tuberculosis, which is a pathogen notorious for its persistence in the host under conditions of prolonged starvation and anaerobiosis (Primm et al., 2000).

Adaptation to the changing environment is key for the survival of bacteria, and it is important to have a system that can sense environmental fluctuations. The stationary phase of bacterial cultures is a very useful system that provides interesting information about bacterial adaptation. Exhaustion of nutrients, accumulation of toxic metabolites, and environmental stresses, such as low temperature, change in pH and high osmolarity, are factors that force bacterial entry into stationary phase. Intriguingly, this growth phase has some striking features, such as reduction of cell division, slowing down of metabolic activity, increased resistance to stress, and altered gene expression; these features enable adaptation and survival. It can be argued that stationary-phase conditions are closer to the conditions that the bacteria encounter in the natural environment. Under such stressful conditions, bacteria survive by producing spores, fruiting bodies, etc. However, in the case of mycobacteria, altered colony morphology and different staining properties have been observed in conditions of nutritional starvation (Nyka, 1974; Dahl et al., 2005). During stationary phase, M. smegmatis forms clumps that increase in size during prolonged starvation, and these may promote survival (Smeulders et al., 1999). Similarly, P. aeruginosa forms cell aggregates to survive in the presence of the toxic detergent SDS, and formation of these aggregates is directly linked to levels of c-di-GMP (Klebensberger et al., 2007). In our experiments, we observed increased levels of c-di-GMP in stationary phase, and this suggests that this second messenger is important in adaptation, survival and persistence.

A recent surge in research focusing on c-di-GMP has been observed upon the realization that this molecule is present upon the realization that this molecule is present. Although the protein MSDGC-1 is from a species of non-pathogenic Mycobacterium, it has homologues with >65% identity at the amino acid level in pathogenic species such as M. tuberculosis H37Rv (1354c), M. leprae (NT01 3427) and M. avium. It would be extremely useful to conduct experiments to investigate the role of c-di-GMP in the persistence and survival of mycobacterial pathogens.

**ACKNOWLEDGEMENTS**

We thank Professor Urs Jenal, Biozentrum, Switzerland, for providing the PleD clone for the c-di-GMP assay. This work was financially supported by DBT, Government of India. M.K. acknowledges DBT for the postdoctoral fellowship. The authors also would like to thank Binod Kumar Bharati for help in a few experiments, Y. S. Prakash for help in HPLC and mass spectroscopy, and R. Saraswati for suggestions. Thanks are also given to Professor Umesh Varshney and Krishna from the Department of Microbiology and Cell Biology, Indian Institute of Science, for valuable suggestions during the review of this manuscript, and help in microscopy.

**REFERENCES**


Edited by: M. Daffé