Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary X-ray analysis of mannosyl-3-phosphoglycerate phosphatase from *Thermus thermophilus* HB27

Susana Gonçalves, Ana M. Esteves, Nuno Borges, Helena Santos and Pedro M. Matias*

Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Avenida da República, 2780-157 Oeiras, Portugal

Correspondence e-mail: matias@itqb.unl.pt

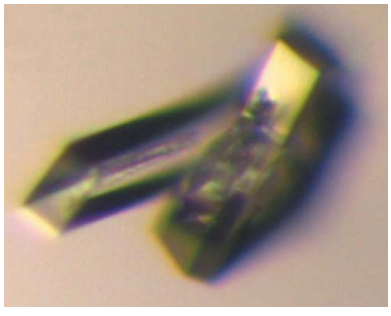
Received 23 November 2010
Accepted 20 January 2011

Mannosylglycerate (MG) is primarily known as an osmolyte and is widely distributed among (hyper)thermophilic marine microorganisms. The synthesis of MG *via* mannosyl-3-phosphoglycerate synthase (MpgS) and mannosyl-3-phosphoglycerate phosphatase (MpgP), the so-called two-step pathway, is the most prevalent route among these organisms. The phosphorylated intermediate mannosyl-3-phosphoglycerate is synthesized by the first enzyme and is subsequently dephosphorylated by the second. The structure of MpgS from the thermophilic bacterium *Thermus thermophilus* HB27 has recently been solved and characterized. Here, the cloning, expression, purification, crystallization and preliminary crystallographic analysis of MpgP from *T. thermophilus* HB27 are reported. Size-exclusion chromatography assays suggested a dimeric assembly in solution for MpgP at pH 6.3 and together with differential scanning fluorimetry data showed that high ionic strength and charge compensation were required to produce a highly pure and soluble protein sample for crystallographic studies. The crystals obtained belonged to the monoclinic space group $P2_1$, with unit-cell parameters a = 39.52, b = 70.68, c = 95.42 Å, β = 92.95°. Diffraction data were measured to 1.9 Å resolution. Matthews coefficient calculations suggested the presence of two MpgP monomers in the asymmetric unit and the calculation of a self-rotation Patterson map indicated that the two monomers could be related by a noncrystallographic twofold rotation axis, forming a dimer.

1. Introduction

Studies of the adaptation strategies of (hyper)thermophilic prokaryotes to hot marine environments highlight the role of unusual organic compounds that accumulate in the cytoplasm of these organisms mainly *via de novo* synthesis (Santos & da Costa, 2002). In addition to the fundamental interest in the physiological mechanisms of response to environmental constraints, a practical relevance arises from the utilization of these compounds as stabilizers of biomaterials (Cruz *et al.*, 2006; da Costa *et al.*, 1998; Faria *et al.*, 2008; Santos *et al.*, 2008). Within the class of sugar-derivative compounds, mannosylglycerate [α-D-mannopyranosyl-(1,2)-*O*-D-glycerate; MG] is primarily known as an osmolyte. In particular, MG plays a central role in the osmoadaptation of some halotolerant strains of *Thermus thermophilus* (Alarico *et al.*, 2005, 2007). Moreover, the superior ability of MG to protect proteins against heat denaturation has been extensively documented (Borges *et al.*, 2002; Faria *et al.*, 2008; Ramos *et al.*, 1997).

The pathways of MG synthesis have been characterized in detail (Borges *et al.*, 2004; Martins *et al.*, 1999; Empadinhas *et al.*, 2001). The most common route follows the two-step pathway generally used for the synthesis of other sugar-derivative compatible solutes, such as trehalose and glucosylglycerol (Elbein *et al.*, 2003; Hagemann *et al.*, 2001). This comprises the synthesis of a phosphorylated intermediary compound, mannosyl-3-phosphoglycerate (M-3-PG), by the retaining mannosyl-3-phosphoglycerate synthase (MpgS; EC 2.4.1.217), followed by its subsequent dephosphorylation by mannosyl-3-phosphoglycerate phosphatase (MpgP; EC 3.1.3.70) (Fig. 1). Alternatively, a single-step pathway has been identified in which mannosylglycerate



synthase (MgS; EC 2.4.1.–) catalyzes direct glycosyl transfer from GDP-mannose to D-glycerate, yielding MG (Martins *et al.*, 1999). Despite the increasing number of established DNA sequences coding for the two-step pathway of MG synthesis, knowledge of its regulatory mechanisms is still limited (Borges *et al.*, 2004; Müller *et al.*, 2005).

The structure of MpgS from *T. thermophilus* HB27 has recently been characterized by our team (Gonçalves *et al.*, 2010). MpgP, the enzyme that catalyzes the second reaction in the two-step route, is usually clustered in the *mpgS*/*mpgP* operon (Empadinhas *et al.*, 2001) and shares sequence-signature features with the haloalkanoid acid dehalogenase (HAD) superfamily of aspartate-nucleophile hydrolases. The HAD superfamily is topologically characterized by an α/β Rossmann-like fold with four conserved functional motifs located at the binding cleft of the Rossmann-like core and covered by a highly mobile catalytic unit composed of a helical loop designated squiggle and a juxtaposed β-hairpin named flap (Burroughs *et al.*, 2006). Cap insertions into the HAD core unit are usually grouped into three major classes: class I, with a four-helical bundle C1 cap, class II, with an $\alpha+\beta$-fold C2 cap, and class III, with a C0 cap corresponding to the single HAD core domain. As phosphate-monoester hydrolysis is the basis of the chemistry behind HAD activity, the addition of the cap module-insertion system to the HAD core domain provides a means of exploration of substrate and reaction space (Allen & Dunaway-Mariano, 2004; Lu *et al.*, 2005; Burroughs *et al.*, 2006). The high functional versatility of the HAD superfamily results in a vast repertoire of phosphoryl-group transfer reactions that are ubiquitously present in the three domains of life and are grouped into five major types according to the type of reaction: (i) phosphatase, (ii) ATPase, (iii) dehalogenase, (iv) phosphosugar mutase and (v) phosphonatase (Burroughs *et al.*, 2006; Allen & Dunaway-Mariano, 2009). MpgP falls within the mannosyl-3-phosphoglycerate phosphatase family of the Cof hydrolase clade and is characterized by a C2 type B cap module inserted within the cross-linker between strands S3 and S4 of the basal HAD core domain (HAD-IIB-MPGP family; http://cmr.jcvi.org; TIGR01486; SCOP c.108.1.10).

To date, two structural representatives of the HAD-IIB-MPGP family are known: the orthologous MpgP from *Pyrococcus horikoshii* (PDB entry 1wzc; Kawamura *et al.*, 2008) and the MpgP-related YedP from *Escherichia coli* (EC 3.1.3.–; PDB entry 1xvi; Y. Kim, A. Joachimiak, M. Cymborowski, T. Skarina, A. Savchenko & A. Edwards, unpublished work). The latter has an unknown function but exhibits protein–protein interaction features with the inner membrane protein YedI (UniProtKB P46125; Arifuzzaman *et al.*, 2006). Interestingly, in *E. coli* MG is taken up as mannosyl-6-phosphoglycerate, which is then used as a carbon source (Sampaio *et al.*, 2004). Insights into the catalytic mechanism of aspartyl-phosphatase activity have been obtained based on the structural information derived from the 'open' and 'closed' conformations of the respective apo and holo forms of *P. horikoshii* MpgP (Kawamura *et al.*, 2008). The mechanism proposed involves the formation of a phosphorane intermediate with trigonal bipyramidal geometry characteristic of the 'two-step' associative/dissociative nucleophilic substitution mechanism (Allen & Dunaway-Mariano, 2004; Burroughs *et al.*, 2006; Lahiri *et al.*, 2002).

Here, we describe the purification, crystallization and preliminary crystallographic characterization of MpgP from *T. thermophilus* HB27. Additionally, differential scanning fluorimetry (DSF; Niesen *et al.*, 2007) and size-exclusion chromatography (SEC) provided insights into how to obtain a stable and soluble form of MpgP that could be crystallized.

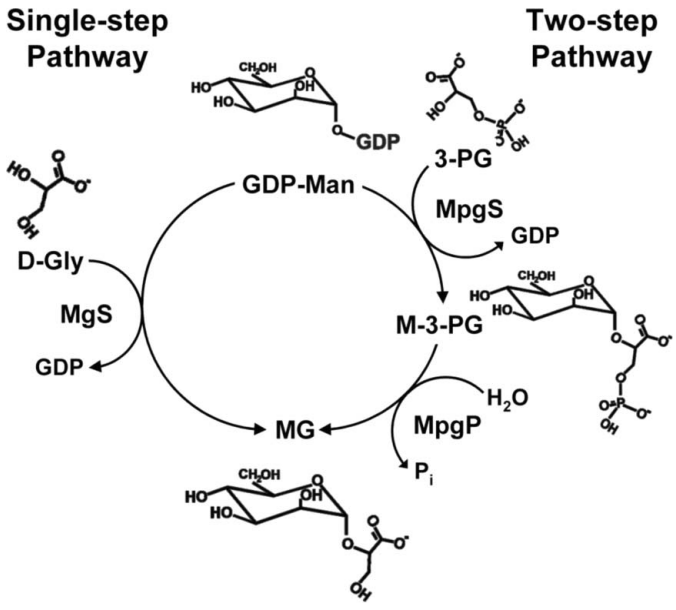


Figure 1
The two pathways for the synthesis of mannosylglycerate. Legend: GDP-Man, GDP-α-D-mannose; 3-PG, 3-D-phosphoglycerate; D-Gly, D-glycerate; MG, α-D-mannosylglycerate; P_i, inorganic phosphate; MpgS, mannosyl-3-phosphoglycerate synthase (EC 2.4.1.217); MpgP, mannosyl-3-phosphoglycerate phosphatase (EC 3.1.3.70); MgS, mannosylglycerate synthase (EC 2.4.1.–). Originally published in Gonçalves *et al.* (2010). Reproduced with permission from the American Society for Biochemistry and Molecular Biology.

2. Materials and methods

2.1. Cloning of the *T. thermophilus* HB27 *mpgP* gene

T. thermophilus HB27 genomic DNA was used as the template to amplify the *mpgP* gene (GenBank NC_005835.1) by polymerase chain reaction (PCR). The gene was then cloned into the pKK223-3 vector (PL-Pharmacia; GenBank M77749.1) between the *Eco*RI and *Pst*I restriction sites. Two primers (Thermo Scientific) were constructed and used for PCR amplification with *Pfu* polymerase (Fermentas). The sense primer mpgP-F (5′-GGC**GAATTC**ATGATCGTCTTCACCGACCTGGA-3′) carried the *Eco*RI restriction site (bold), whereas the antisense primer mpgP-R (5′-CCG**CTGCAG**TCAGGGCCCGCTCCCTCCTCGTCGG-3′) harboured the *Pst*I restriction site (bold).

The PCR amplification was carried out in a iCycler thermocycler (Bio-Rad) with a 25 µl reaction mixture consisting of 200 ng genomic DNA, 0.5 µ*M* primers, 0.25 m*M* dNTPs, 1.5 m*M* Mg_2SO_4 cofactor, *Pfu* buffer diluted to 1× from a 10× stock solution (Fermentas) and 1.25 U *Pfu* enzyme (Fermentas). This mixture was pre-incubated at 368 K for 5 min, followed by 30 cycles of sequential steps of denaturation (368 K for 1 min), annealing (338 K for 1 min) and primer extension (345 K for 1.5 min). In the final cycle the extension step was prolonged for 7 min. The PCR product was further digested with *Eco*RI and *Pst*I restriction enzymes and purified using the PCR DNA and Gel Band Purification Kit (GE Healthcare). The pKK223-3_mpgP vector construct was built with T4 DNA ligase (Promega) by a ligation reaction (10 µl total volume) between the DNA insert and the linearized pKK223-3 vector at the corresponding *Eco*RI and *Pst*I restriction sites with a 1:4 ratio of vector to insert. Positive recombinant colonies of *E. coli* DH5α cells carrying the vector construct were selected from LB–agar plates containing ampicillin (100 µg ml^{-1}) and the plasmid was extracted using the

PlasmidPrep Mini Spin Kit (GE Healthcare). Confirmation of the inserted gene was obtained by double-strand sequencing (STAB VIDA EUROMICS sequencing service, Caparica, Portugal; http://www.stabvida.com).

2.2. Expression and purification of *T. thermophilus* HB27 MgpP

The transformation, expression and purification of MgpP were performed as described previously (Empadinhas *et al.*, 2003) with some modifications. *E. coli* BL21 (DE3) strain (Novagen) bearing the plasmid construction pKK223-3_mpgP was grown in an orbital shaker at 200 rev min^{-1} and 310 K in YT medium supplemented with ampicillin (100 µg ml^{-1}) to an OD_{600} of 0.6–0.7, followed by induction with 1 m*M* IPTG (isopropyl β-D-1-thiogalactopyranoside) for 6 h at 310 K. The cells were harvested by centrifugation (7000*g*, 15 min, 277 K) and resuspended in buffer *A* (20 m*M* Tris–HCl pH 7.6, 5 m*M* DTT, 1 m*M* EDTA). The cells were disrupted in a French press followed by centrifugation (18 000*g*, 40 min, 277 K) to remove cell debris. The resulting cell extract was diluted (with 20 m*M* Tris–HCl pH 7.6, 1 m*M* EDTA) to about 10 mg ml^{-1} and heated for 20 min at 348 K to precipitate thermolabile contaminant proteins from *E. coli*. After centrifugation (25 000*g*, 45 min, 277 K), the supernatant was loaded onto a 70 ml Q-Sepharose column (GE Healthcare) equilibrated with buffer *A*. Elution was carried out with a linear gradient from buffer *A* to buffer *B* (20 m*M* Tris–HCl pH 7.6, 5 m*M* DTT, 1 m*M* EDTA, 1 *M* NaCl). Fractions containing MgpP eluted between 700 and 800 m*M* NaCl. These were pooled and dialyzed against buffer *A* before loading them onto a 6 ml Resource Q column (GE Healthcare) equilibrated with buffer *A*. Again, MgpP eluted between 700 and 800 m*M* NaCl from the linear gradient from buffer *A* to buffer *B*. The purest fractions were pooled and dialyzed against buffer *C* (20 m*M* MES–NaOH pH 6.3, 5 m*M* DTT, 1 m*M* EDTA) prior to loading them onto a 1 ml Mono S column (GE Healthcare) equilibrated with buffer *C*. A linear gradient was applied to buffer *D* (20 m*M* MES–NaOH pH 6.35, 5 m*M* DTT, 1 m*M* EDTA, 1 *M* NaCl), from which MgpP eluted as a single peak at about 760 m*M* NaCl. The eluted protein was concentrated to 10 mg ml^{-1} in a modified buffer *D* solution (20 m*M* MES–NaOH pH 6.35, 5 m*M* DTT, 1 m*M* EDTA, 760 m*M* NaCl) and used for subsequent crystallization and gel-filtration assays.

Protein purity was assessed by SDS–PAGE analysis (Fig. 2*a*), which showed a single band with an apparent molecular mass of 28.2 kDa and only a small percentage of contaminant. The MpgP activity was detected by visualizing the formation of MG derived from the hydrolysis of M-3-PG by thin-layer chromatography as described previously (Empadinhas *et al.*, 2003). The protein concentration was determined by the Bradford assay (Bradford, 1976). The molecular mass of MpgP was estimated by size-exclusion chromatography (SEC) using a 2.4 ml Superdex 200 3.2/30 precision column (GE Healthcare) equilibrated with buffer *E* (20 m*M* MES–NaOH pH 6.3, 760 m*M* NaCl, 1 m*M* EDTA and 5 m*M* DTT). The standards (GE Healthcare) used and their correspondent elution volumes (V_e) were ribonuclease (13.7 kDa; V_e = 1.89 ml), ovalbumin (43 kDa; V_e = 1.65 ml), albumin (66 kDa; V_e = 1.56 ml), aldolase (158 kDa; V_e = 1.43 ml) and ferritin (440 kDa; V_e = 1.23 ml). Blue dextran 2000 (GE Healthcare) was used to determine the void volume of the column (V_e = 1.01 ml). All of the assays were performed at a constant flow rate of 0.1 ml min^{-1}. The MpgP protein eluted at V_e = 1.59 ml, between the elution profiles for albumin and ovalbumin, suggesting that a homodimeric structure is the prevalent oligomeric state under the running conditions (Fig. 2*b*).

2.3. Differential scanning fluorimetry (DSF) and thermal shift assay

To address the solubility and stability of MpgP as a function of salt, pH and buffer composition, DSF analysis (Niesen *et al.*, 2007) was performed in a protein sample eluted from the MonoS column (1.2 mg ml^{-1}) containing the probe dye SYPRO Orange (Invitrogen) diluted to 50× from the original stock solution (5000×) and distributed in a Low 96-well White Multiplate PCR Plate (Bio-Rad). In each well, the protein-sample mixture (2 µl) was diluted 5× (to a final volume of 20 µl) by adding 18 µl of the corresponding buffer solution from a home-made 96-condition sparse-matrix screening kit. The buffer screening kit was prepared from the commercial JBS solubility kit buffers (Jena Bioscience). For each condition, NaCl was included at concentrations of 0, 150 and 500 m*M*. The last condition

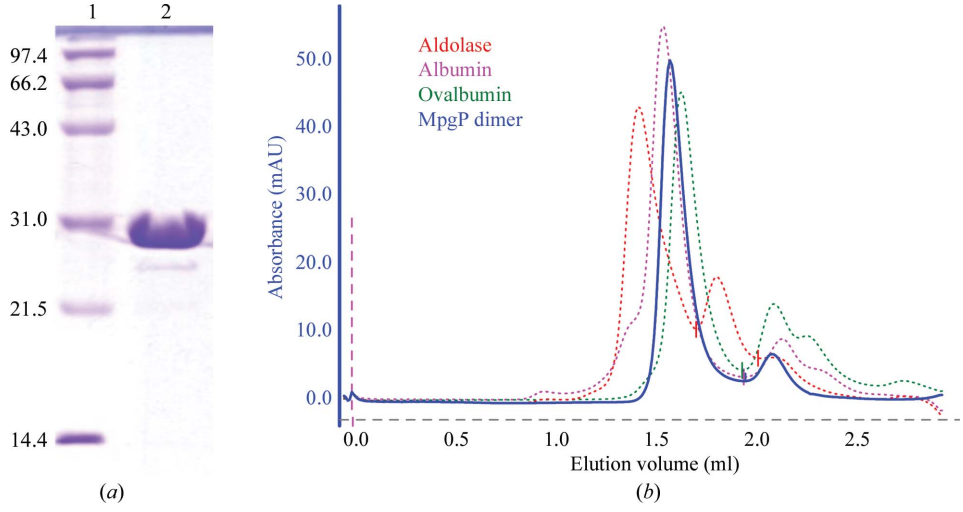


Figure 2
(*a*) 15% SDS–PAGE of pure recombinant *T. thermophilus* HB27 MpgP. Lane 1, Low Range Molecular Mass Markers (Bio-Rad) consisting of the protein markers phosphorylase b (97.4 kDa), bovine serum albumin (66.2 kDa), ovalbumin (43 kDa), carbonic anhydrase (31 kDa), soybean trypsin inhibitor (21.5 kDa) and lysozyme (14.4 kDa). Lane 2, MpgP monomer migrating according to its molecular mass (28.18 kDa). (*b*) Elution profile of MpgP loaded onto an analytical 2.4 ml Superdex 200 3.2/30 precision column. The MpgP elution profile is represented as a full curve in blue. Elution profiles of protein standards are shown as dashed curves for comparison. In addition to their standard elution peaks, degradation products are also eluted. mAU, milliunits of absorption.

contained 150 m*M* KCl instead of NaCl. The plate was sealed with Microseal type B film seal (Bio-Rad) and submitted to a temperature-gradient scan (293–363 K) in a iQ5 Real Time PCR machine (Bio-Rad). The fluorescence emission from the reporter dye was measured every 2 K of increasing temperature using the Cy3 fluorophore. The scanning protocol consisted of temperature stabilization at 293 K for 1 min in the first cycle, followed by a second cycle of 71 steps up to 363 K; in each step, the temperature gradient was held for 10 s before data acquisition. The third and last cycle consisted of lowering the temperature back down to 293 K to cool down the system before sample removal. The melting curves (Fig. 3) were analyzed with the *iQ5 Optical System Software* v.2.0 (Bio-Rad) and melting temperature values (T_m) were calculated with the Differential Scanning Fluorimetry Analysis tool (DSF v.2.5 for Bio-Rad IQ5; ftp://ftp.sgc.ox.ac.uk/pub/biophysics; Niesen *et al.*, 2007).

2.4. Effect of NaCl, EDTA and DTT on MpgP oligomerization and stability

MpgP protein eluted from the MonoS column was concentrated to 8.6 mg ml^{-1} in modified buffer *D* solution (see above) and was used for analytical SEC studies performed with a Superdex S75 3.2/30 precision column (GE Healthcare). Prior to column injection, the protein was diluted 5× (10 µl total volume) in the corresponding gel-filtration buffer solution used during each assay. The standard molecular-mass curve was determined using the globular protein standards (GE Healthcare) ribonuclease (13.7 kDa; V_e = 1.52 ml), chymotrypsinogen (25 kDa; V_e = 1.44 ml), ovalbumin (43 kDa; V_e = 1.26 ml) and albumin (66 kDa; V_e = 1.18 ml) in 20 m*M* MES–NaOH pH 6.3, 760 m*M* NaCl and 1 m*M* EDTA (Fig. 4*a*). Blue dextran 2000 (GE Healthcare) was used to determine the void volume of the column (V_e = 1.03 ml). The effect of NaCl, DTT and EDTA on the V_e of the protein standards was first assessed and was found to be negligible in comparison with the values obtained when using the manufacturer's recommended buffer (20 m*M* Na_2HPO_4/KH_2PO_4 pH 7.2, 150 m*M* NaCl). All of the assays were performed at a constant flow rate of 0.1 ml min^{-1} in the MES–NaOH pH 6.3 buffer system. The variation of the elution profiles with buffer composition is shown in Fig. 4.

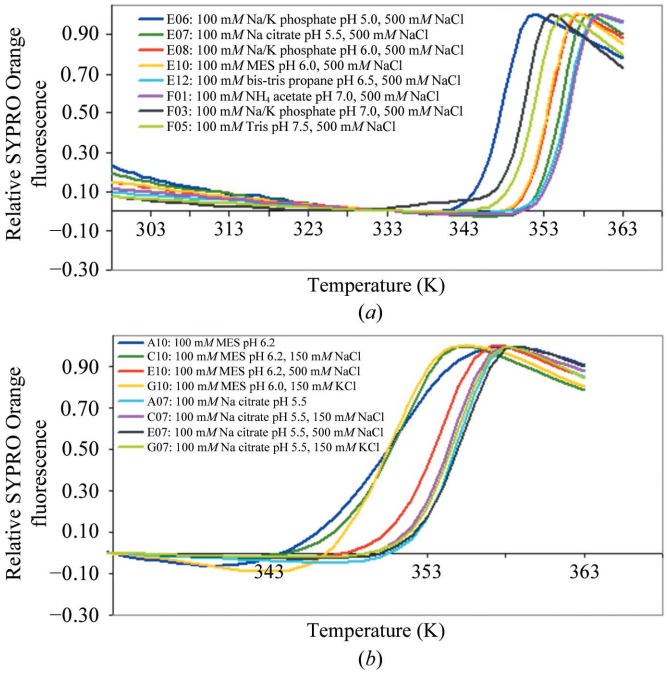


Figure 3
DSF thermal denaturation curves for MpgP. The buffer composition does not take into account the 10% dilution effect during protein-sample dilution (see §2). (*a*) MpgP stability as a function of pH at high ionic strength. The thermal unfolding profile is shown only within the pH interval 5.0–7.0 as the protein was less stable at other pH values. The chemical nature of the buffer also had an effect on the stability of MpgP as observed for the profiles between curves F01 and F03. (*b*) MpgP stability as function of salt. In the MES buffer system the stability of MpgP increases significantly with an increase in ionic strength (curves A10, C10 and E10), while the substitution of equivalent amounts of NaCl by KCl have little effect (curves C10 and G10). Sodim citrate pH 5.5 buffer (curves A07, C07, E07 and G07) was an exception since MpgP exhibited similar unfolding profiles regardless of the presence of salt.

2.5. MpgP crystallization and cryoconditions

Preliminary crystallization trials were performed on the nanoscale with the commercially available Morpheus crystallization kits (Molecular Dimensions; Gorrec, 2009) using a Cartesian Crystal-

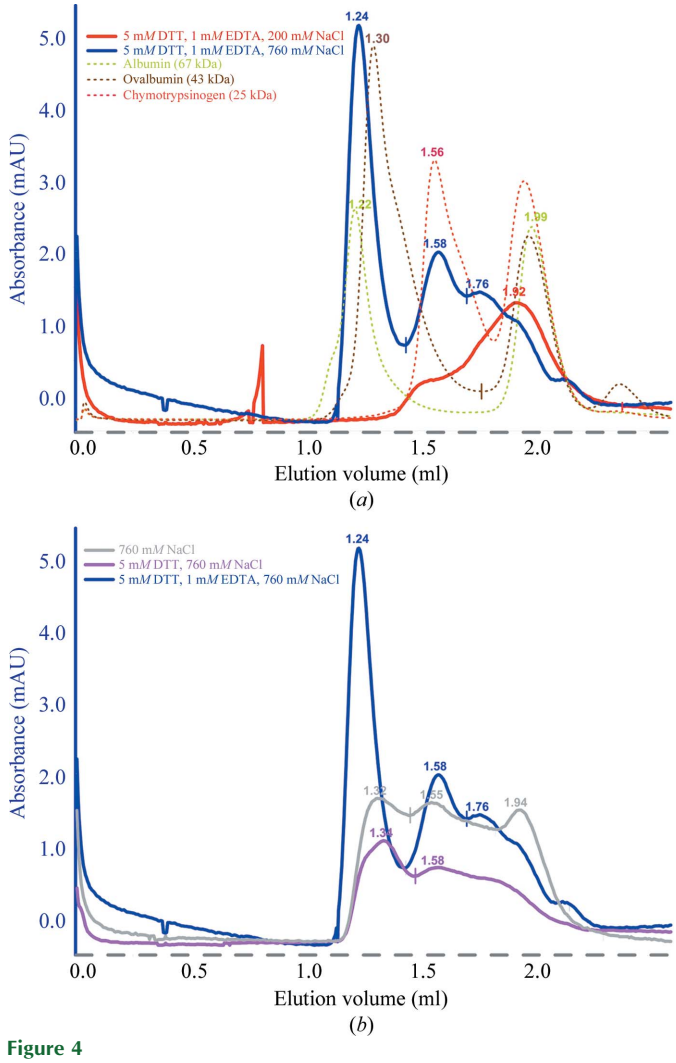


Figure 4
MpgP elution profiles from the SEC analysis with a Superdex 75 3.2/30 precision column. MpgP elution profiles are represented as full curves. The upper left text in each panel indicates the gel-filtration buffer composition in which MpgP was eluted and is colour-coded according to the colours of the respective MpgP elution curves. MES–NaOH pH 6.3 buffer was used in all gel-filtration runs. (*a*) Effect of NaCl in the elution of MpgP. The elution profiles of protein standards are shown as dashed curves for comparison. In addition to their standard elution peaks, lower molecular-mass degradation products are also eluted. (*b*) Effect of EDTA and DTT on the MpgP elution profile. mAU, milliunits of absorption.

Table 1
Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

Beamline	ESRF ID14-4
Detector	ADSC Q315r
Wavelength (Å)	0.95350
Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 39.52, b = 70.68, c = 95.42, β = 92.95
Resolution (Å)	50.00–1.90 (2.01–1.90)
No. of observations	148324 (18433)
Unique reflections	40742 (5910)
Multiplicity	3.6 (3.1)
Completeness (%)	98.0 (88.6)
$R_{\rm merge}$† (%)	4.4 (48.1)
$R_{\rm meas}$‡ (%)	5.2 (57.9)
$\langle I/\sigma(I)\rangle$	18.79 (2.37)

† $R_{\rm merge}$ is the merging R factor: $\sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl) \times 100\%$. ‡ $R_{\rm meas}$ is the redundancy-independent R factor (Diederichs & Karplus, 1997): $\sum_{hkl}[N/(N-1)]^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl) \times 100\%$. For each unique Bragg reflection with indices (hkl), $I_i(hkl)$ is the ith observation of its intensity and N is its multiplicity.

lization Robot Dispensing System (Genomics Solutions) and round-bottom Greiner 96-well CrystalQuick plates (Greiner Bio-One). Three crystallization drops per condition screened were prepared by mixing 100 nl reservoir solution with 100 nl protein solution. MpgP at 10 mg ml^{-1} in its concentration buffer was tested, as well as with added 5 m*M* $MgCl_2$, 5 m*M* sodium/potassium phosphate pH 6.8 and with added 5 m*M* $MgCl_2$, 5 m*M* sodium/potassium phosphate pH 6.8, 1 m*M* tris-(2-carboxyethyl)phosphine (TCEP). The drops were equilibrated against 100 µl reservoir solution. Promising plate-shaped and prism-like crystal clusters were found for all three of the MpgP protein samples in distinct conditions from the 96-well sparse matrix. In the scale-up trials a crystallization solution with the following composition was used (Gorrec, 2009): 0.02 *M* of each of the carboxylic acids sodium formate, ammonium acetate, trisodium citrate, sodium potassium L-tartrate and sodium oxamate as additives, 0.1 *M* MES–imidazole pH 6.5 (in a 1:1 molar ratio) as a buffering system and 30% of a precipitant mixture composed of 20% ethylene glycol (EDO) and 10% polyethylene glycol 8000 (PEG 8K). Native protein at 10 mg ml^{-1} with added 5 m*M* $MgCl_2$, 5 m*M* sodium/potassium phosphate and 2 m*M* DTT was mixed with the crystallization solution in a 1:1 ratio of protein to reservoir solution and 2 µl drops were set up at 293 K in 24-well crystallization plates using the sitting-drop vapour-diffusion technique. The drops were equilibrated against 500 µl reservoir solution. Prism-like crystals (Fig. 5) developed within 5 d, reaching dimensions of 30 × 50 × 400 µm. Since the crystallization solution is already cryoprotecting (Gorrec, 2009), native MpgP crystals were harvested and immediately flash-cooled in liquid nitrogen prior to data collection. The crystallized protein was confirmed as MpgP by N-terminal sequencing (Analytical Services Unit, Instituto de Tecnologia Química e Biológica, Oeiras, Portugal) of a dissolved crystal from the same drop as that used for X-ray diffraction data measurements.

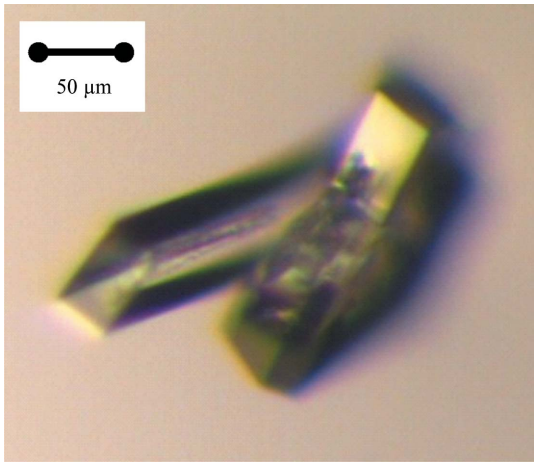


Figure 5
Native MpgP crystals used for data collection.

2.6. Data collection and processing

An X-ray diffraction data set was collected to 1.9 Å resolution at 100 K from a flash-cooled MpgP crystal using an ADSC Q315r CCD detector on beamline ID14-4 at ESRF, Grenoble, France. The diffraction images were integrated and scaled with the *XDS* program package (Kabsch, 2010). Further processing was carried out with the *CCP*4 program package (Collaborative Computational Project, Number 4, 1994). A summary of the data-processing statistics is presented in Table 1. The crystal belonged to the monoclinic space group $P2_1$, with unit-cell parameters a = 39.52, b = 70.68, c = 95.42 Å, β = 92.95°. Matthews coefficient calculations (Matthews, 1968) indicated the presence of two molecules in the asymmetric unit, with a corresponding $V_{\rm M}$ value of 2.38 Å^3 Da^{-1} and a predicted solvent content of 48%. A self-rotation Patterson function calculation revealed the presence of strong peaks nearly perpendicular to the crystallographic b axis, suggesting the presence of two monomers related by twofold noncrystallographic symmetry (NCS; Fig. 6).

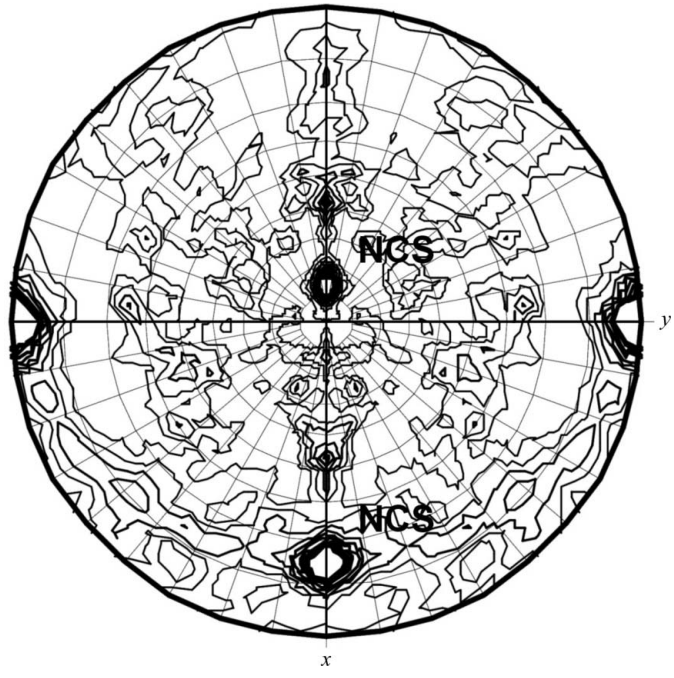


Figure 6
Self-rotation Patterson map in the κ = 180° section, calculated with *MOLREP* from *CCP*4 (Collaborative Computational Project, Number 4, 1994) using default parameters, with a radius of integration of 19.8 Å and data in the resolution range 39.5 > d > 3.0 Å. The unique b axis in space group $P2_1$ is oriented along the y axis in the figure. Two equivalent noncrystallographic symmetry peaks (NCS) are observed perpendicular to the crystallographic b axis and with an intensity of 68% relative to the origin, suggesting the presence of an MpgP dimer in the asymmetric unit, with the two monomers related by twofold NCS.

3. Results

3.1. MpgP thermal unfolding as a function of salt, pH and buffer composition

Application of the differential scanning fluorimetry (DSF) methodology (Niesen *et al.*, 2007) to the screening of 96 sparse-matrix buffer conditions allowed a quick analysis of the protein stability as a function of salt, pH and buffer. The results provided useful information regarding the choice of the storage buffer for the protein. In most of the conditions tested for MpgP the observed thermal unfolding profile was represented by a single sigmoidal curve apparently corresponding to a two-state transition. The best results were obtained for conditions containing high salt (500 m*M* NaCl) and within the pH interval 5.5–7.0 (Fig. 3), corresponding to $T_{\rm m}$ values ranging between 354 and 357 K. Interestingly, when sodium citrate pH 5.5 was present in the buffer system the protein exhibited a similar unfolding profile ($T_{\rm m}$ = 355 K) regardless of the presence of salt. This suggests that the carboxylic groups in sodium citrate have a stabilizing effect on MpgP, probably competing with salt in a type of interaction analogous to MonoS elution. Interestingly, the MpgP that eluted from the MonoS column was already in a buffer composition resembling those giving highest stability as found from the DSF assay and was also able to crystallize readily.

3.2. The role of NaCl, EDTA and DTT in MpgP dimerization and solubility

At the beginning of the optimization process of MpgP purification, the protein showed a tendency to aggregate and could not be concentrated above 1.2 mg ml^{-1}. This issue was overcome by direct elution from the MonoS column at about 760 m*M* NaCl, 5 m*M* DTT and 1 m*M* EDTA, thus allowing the protein to be concentrated to at least 10 mg ml^{-1}. As indicated by its elution profile from the analytical Superdex 200 3.2/30 precision column, this soluble form is probably a dimer in this buffer system, with an estimated molecular mass $MM_{\rm cal}$ of 65 $\pm$ 9 kDa (Fig. 2*b*). The effects of NaCl, DTT and EDTA were then assessed by size-exclusion chromatography (SEC) using an analytical Superdex 75 3.2/30 precision column (GE Healthcare) in the MES–NaOH pH 6.3 buffer system (Fig. 4). The elution profile of MpgP varied greatly with the ionic strength and with the presence of EDTA; the MpgP dimer ($V_{\rm e}$ = 1.24 ml) only eluted at 760 m*M* NaCl, 1 m*M* EDTA, while in the absence of either one of these two components low-molecular-mass products eluted. This suggests that strong electrostatic interactions were established between the protein (either folded or unfolded) and the column matrix. This is not surprising, since the theoretical isoelectric point of the MpgP monomer is 7.8 and, as confirmed by its strong binding affinity to the MonoS column, in MES–NaOH pH 6.3 buffer the dimer is very likely to have a total net positive charge in solution which is sufficiently high to establish this type of interaction. The addition of DTT also improved the solubility of the protein by preventing the formation of nonspecific disulfide bridges. However, after heating a protein sample to the optimal temperature for catalysis (348 K; Empadinhas *et al.*, 2003), the resulting elution profile suggested dimer destabilization together with protein degradation (data not shown). The nonconserved cysteine residue (Cys144) found in a variable region of the *T. thermophilus* MpgP sequence may therefore play a role in the dimerization of MpgP.

Our results suggest that *T. thermophilus* MpgP requires a high ionic strength to become properly folded and is most stable in solution as dimers, the formation of which possibly involves a disulfide bridge.

4. Concluding remarks

The biochemical data obtained from the SEC and DSF assays provided a more detailed view into the requirements for the highest stability of MpgP in solution. High ionic strength, together with compounds rich in carboxyl groups such as EDTA and sodium citrate, raised the solubility of MpgP to well above 1.2 mg ml^{-1}. The crystal structure of *P. horikoshii* MpgP (PDB entry 1wzc) suggests a monomeric assembly in solution, which is in agreement with its gel-filtration profile (Kawamura *et al.*, 2008). The crystal structure of the MpgP-related YedP from *E. coli* (PDB entry 1xvi) suggests a dimeric assembly based on *PISA* calculations (*Protein Interfaces, Surfaces and Assemblies* service, EBI; Krissinel & Henrick, 2005, 2007). Our biochemical results suggest that MpgP from *T. thermophilus* HB27 is most likely to be biologically active as a dimer. Attempts to solve the three-dimensional structure of *T. thermophilus* HB27 MpgP by molecular replacement using the two known MpgP crystal structures (PDB entries 1wzc and 1xvi) as search models were not successful. Since this protein does not contain any methionines, we are considering the preparation of a synthetic SeMet derivative by mutating a suitable number of leucine residues to methionines using the three-dimensional structure of *P. horikoshii* MpgP as a guide. Alternatively, heavy-atom derivatization will be explored. We expect to be able to correlate the MpgP crystal structure with the biochemical features described here. In addition, we aim to obtain additional information regarding the mechanism of the phosphoryl-transfer reaction, as well as its substrate-specificity features, which are usually modulated by the cap system and may be representative of the HADSF-IIB-MPGP family.

We would like to thank the ESRF for support during data collection, in particular the ID14-4 beamline staff. This work was funded by FCT grant PTDC/QUI/71142/2006. SG acknowledges FCT for grant SFRH/BD/23222/2005. The technical assistance of Ana Isabel Mingote (ITQB) is gratefully acknowledged.

References

Alarico, S., Empadinhas, N., Mingote, A., Simões, C., Santos, M. S. & Da Costa, M. S. (2007). *Extremophiles*, **11**, 833–840.
Alarico, S., Empadinhas, N., Simões, C., Silva, Z., Henne, A., Mingote, A., Santos, H. & Da Costa, M. S. (2005). *Appl. Environ. Microbiol.* **71**, 2460–2466.
Allen, K. N. & Dunaway-Mariano, D. (2004). *Trends Biochem. Sci.* **29**, 495–503.
Allen, K. N. & Dunaway-Mariano, D. (2009). *Curr. Opin. Struct. Biol.* **19**, 658–665.
Arifuzzaman, M. *et al.* (2006). *Genome Res.* **16**, 686–691.
Borges, N., Marugg, J. D., Empadinhas, N., Da Costa, M. S. & Santos, H. (2004). *J. Biol. Chem.* **279**, 9892–9898.
Borges, N., Ramos, A., Raven, N. D., Sharp, R. J. & Santos, H. (2002). *Extremophiles*, **6**, 209–216.
Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Burroughs, A. M., Allen, K. N., Dunaway-Mariano, D. & Aravind, L. (2006). *J. Mol. Biol.* **361**, 1003–1034.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cruz, P. E., Silva, A. C., Roldão, A., Carmo, M., Carrondo, M. J. & Alves, P. M. (2006). *Biotechnol. Prog.* **22**, 568–576.
Da Costa, M. S., Santos, H. & Galinski, E. A. (1998). *Adv. Biochem. Eng. Biotechnol.* **61**, 117–153.
Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.
Elbein, A. D., Pan, Y. T., Pastuszak, I. & Carroll, D. (2003). *Glycobiology*, **13**, 17R–27R.
Empadinhas, N., Albuquerque, L., Henne, A., Santos, H. & Da Costa, M. S. (2003). *Appl. Environ. Microbiol.* **69**, 3272–3279.
Empadinhas, N., Marugg, J. D., Borges, N., Santos, H. & Da Costa, M. S. (2001). *J. Biol. Chem.* **276**, 43580–43588.

Faria, T. Q., Mingote, A., Siopa, F., Ventura, R., Maycock, C. & Santos, H. (2008). *Carbohydr. Res.* **343**, 3025–3033.
Gonçalves, S., Borges, N., Esteves, A. M., Victor, B. L., Soares, C. M., Santos, H. & Matias, P. M. (2010). *J. Biol. Chem.* **285**, 17857–17868.
Gorrec, F. (2009). *J. Appl. Cryst.* **42**, 1035–1042.
Hagemann, M., Effmert, U., Kerstan, T., Schoor, A. & Erdmann, N. (2001). *Curr. Microbiol.* **43**, 278–283.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kawamura, T., Watanabe, N. & Tanaka, I. (2008). *Acta Cryst.* D**64**, 1267–1276.
Krissinel, E. & Henrick, K. (2005). *CompLife 2005*, edited by M. R. Berthold, R. Glen, K. Diederichs, O. Kohlbacher & I. Fischer, pp. 163–174. Berlin, Heidelberg: Springer-Verlag.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Lahiri, S. D., Zhang, G., Dunaway-Mariano, D. & Allen, K. N. (2002). *Biochemistry*, **41**, 8351–8359.
Lu, Z., Dunaway-Mariano, D. & Allen, K. N. (2005). *Biochemistry*, **44**, 8684–8696.
Martins, L. O., Empadinhas, N., Marugg, J. D., Miguel, C., Ferreira, C., Da Costa, M. S. & Santos, H. (1999). *J. Biol. Chem.* **274**, 35407–35414.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Müller, V., Spanheimer, R. & Santos, H. (2005). *Curr. Opin. Microbiol.* **8**, 729–736.
Niesen, F. H., Berglund, H. & Vedadi, M. (2007). *Nature Protoc.* **2**, 2212–32221.
Ramos, A., Raven, N., Sharp, R. J., Bartolucci, S., Rossi, M., Cannio, R., Lebbink, J., Van Der Oost, J., De Vos, W. M. & Santos, H. (1997). *Appl. Environ. Microbiol.* **63**, 4020–4025.
Sampaio, M. M., Chevance, F., Dippel, R., Eppler, T., Schlegel, A., Boos, W., Lu, Y.-J. & Rock, C. O. (2004). *J. Biol. Chem.* **279**, 5537–5548.
Santos, H. & da Costa, M. S. (2002). *Environ. Microbiol.* **4**, 501–509.
Santos, H., Lamosa, P., Faria, T. Q., Pais, T. M., de La Paz, M. L. & Serrano, L. (2008). *Thermophiles: Biology and Technology at High Temperatures*, edited by F. Robb, G. Antranikian, D. Grogan & A. Driessen, pp. 9–24. London: CRC Press.