Two exopolyphosphatases with distinct molecular architectures and substrate specificities from the thermophilic green-sulfur bacterium Chlorobium tepidum TLS

Tomás Albi and Aurelio Serrano

The genome of the thermophilic green-sulfur bacterium Chlorobium tepidum TLS possesses two genes encoding putative exopolyphosphatases (PPX; EC 3.6.1.11), namely CT0099 (ppx1, 993 bp) and CT1713 (ppx2, 1557 bp). The predicted polypeptides of 330 and 518 aa residues are Ppx-GppA phosphatases of different domain architectures – the largest one has an extra C-terminal HD domain – which may represent ancient paralogues. Both ppx genes were cloned and overexpressed in Escherichia coli BL21(DE3). While CtPPX1 was validated as a monomeric enzyme, CtPPX2 was found to be a homodimer. Both PPX homologues were functional, K⁺-stimulated phosphohydrolases, with an absolute requirement for divalent metal cations and a marked preference for Mg²⁺. Nevertheless, they exhibited remarkably different catalytic specificities with regard to substrate classes and chain lengths. Even though both enzymes were able to hydrolyse the medium-size polyphosphate (polyP) P₁₃–₁₈ (polyP mix with mean chain length of 13–18 phosphate residues), CtPPX1 clearly reached its highest catalytic efficiency with triplyphosphorylase and showed substantial nucleoside triphosphatase activity, while CtPPX2 preferred long-chain polyPs (>300 Pi residues) and did not show any detectable NTPase activity. These catalytic features, taken together with the distinct domain architectures and molecular phylogenies, indicate that the two PPX homologues of Chl. tepidum belong to different Ppx-GppA phosphatase subfamilies that should play specific biochemical roles in nucleotide and polyP metabolisms. In addition, these results provide an example of the remarkable functional plasticity of the Ppx-GppA phosphatases, a family of proteins with relatively simple structures that are widely distributed in the microbial world.

INTRODUCTION

Inorganic polyphosphates (polyPs) are naturally occurring linear polymers of tens to hundreds of orthophosphate residues linked by high-energy phosphoanhydride bonds. Despite being found in every living being in nature – from bacteria to mammals (Kulaev et al., 2005; Kornberg et al., 1999) – and likely conserved from prebiotic times, the major attention to polyP has been on its role in heterotrophic, pathogenic bacteria (mainly γ-proteobacteria and actinobacteria) and yeasts. The polyPs’ ubiquity suggests that they perform important roles in the cell that have been changing during evolution. In prokaryotes, polyP has usually been described just as a polyanion similar to ATP or other phosphate metabolites acting as a reservoir of energy (Kulaev et al., 2005) or Pi (Urech et al., 1978; Schuddemat et al., 1989). Beyond that, polyPs have been proved in a variety of ways to be essential for cell growth, responses to stresses and stringencies, for survival, and for the virulence of pathogens (Ogawa et al., 2000; Rashid & Kornberg, 2000; Kim et al., 2002; Shi et al., 2004; Zhang et al., 2005; Rao et al., 2009; Nikel et al., 2013).

PolyPs are synthesized in bacteria by polyP kinase (PPK; EC 2.7.4.1), which catalyses the readily reversible conversion of the terminal γ-phosphate of ATP to polyP (Rao et al., 2009). PolyP may be utilized as a substrate by transferases and hydrolases as well. They are degraded to Pi by

Abbreviations: GP₄, guanosine 5'-tetraphosphate; GPPase, guanosine pentaphosphate phosphohydrolase; Mₘ, molecular mass; NTPase, nucleoside triphosphatase; P₃, triphosphate; Prc, cyclic trimetaphosphate; P₁₃–₁₈, polyphosphate mix with mean chain length of 13–18 phosphate residues; polyP, polyphosphate; P₃₋₅, long-chain polyphosphate mix (>300 phosphate residues); pNPP, p-nitrophenylphosphatase; polyPase, Ppx-GppA phosphatase; PPK, polyphosphate kinase; PXX, exopolyphosphatase; TTM, triphosphate tunnel metalloenzyme.

The GenBank/EMBL/DDBJ accession numbers for the ppx1 and ppx2 construct sequences are HG764584.1 and HG764585.1, respectively.

Six supplementary figures and three supplementary tables are available with the online version of this paper.
either endopolyphosphatases (PPN; EC 3.6.1.10) (Sethuraman et al., 2001) or exopolyphosphatases (PPX; EC 3.6.1.11). These latter hydrolase and progressively release the terminal orthophosphate from polyP, which contains three or more phosphohydryde bonds. Based on the primary structure, two major non-homologous classes of PPX enzymes could be defined. Firstly, the prototypic cytoplasmic ScPPX1 from yeast (ScPPX1) and its orthologues in fungi and protists, which belong to the DHH-DHHA2 phosphoesterase family (Pfam PF02833), which also includes the well-characterized prokaryotic family II pyrophosphatases. ScPPX1 is an extremely active phosphohydrolase with approximately 40 times the specific activity of the Escherichia coli polyphosphate and it is able to efficiently hydrolise polyP of 3 up to 1000 Pi residues (Lichko et al., 2003). A second polyphosphatase class includes the Ppx-GppA polyphosphatases (polyPases) (Pfam PF02541) (Reizer et al., 2002). They are widely distributed among bacteria and archaea (Cardona et al., 2002; Kristensen et al., 2008), such as the polyPase PPX1 and guanosine pentaphosphatase GPPA of E. coli. The polyPase EcPPX1 of E. coli is encoded by the ppxI gene which together with the ppxA gene forms an operon (Akiyama et al., 1992). This polyPase progressively and nearly completely hydrolises the terminal residues of polyP to Pi with a strong preference for long-chain substrates. EcPPX1 is a 58 kDa enzyme that forms dimers in solution (Rangarajan et al., 2006) and requires Mg\(^{2+}\) for maximal activity. Alternatively, the second sequence-related E. coli exopolyphosphatase, designated GPPA (Keasling et al., 1993), shows both polyPase and guanosine pentaphosphatase phosphohydrolase (GPPase; EC 3.6.1.40) activities. GPPase enzymes liberate Pi by progressive hydrolisis of the terminal phosphoanhydride bonds of long-chain polyP (1000 residues) or by hydrolisis of pppGpp to generate pGpp, an intracellular alarmone or second messenger that controls the bacterial stringent response, an adaptative process induced in response to nutrient starvation (Rao et al., 2009; Mechold et al., 2013).

Hydrolisis of the shortest polyP, tripolyphosphate (P\(_3\)), has been reported in crude extracts of bacteria, yeasts, protists and animal tissues. These ubiquitous tripolyphosphatase activities usually have been associated with a range of proteins lacking sequence similarities with Ppx-GppA polyPases, and described as promiscuous activities, towards substrates other than their natural ones, of enzymes such as the inorganic pyrophosphatases (Jetten et al., 1992; Baykov et al., 1999; Kohn et al., 2012), adenosylmethionine synthase (Markham et al., 1980; Pérez Mato et al., 2001), DHH-DHHA2 exopolyphosphatases (Rodrigues et al., 2002), the human metastasis regulator protein H-Prune (Tammenkoski et al., 2008) and the CYTH superfamily of tunnel metalloenzymes, which was named after its two founding members, the bacterial CyaB adenylate cyclase and the mammalian thiamine triphosphatase (Bettendorff & Wins, 2013). A group of CYTH proteins named triphosphate tunnel metalloenzymes (TTMs) has been found recently in some bacteria (e.g. Clostridium thermocellum, Nitrosomonas europaæa) (Keppetipola et al., 2007; Delvaux et al., 2011) and the plant Arabidopsis thaliana (Moeder et al., 2013), and was reported to be composed of highly specific inorganic tripolyphosphatases. However, the specific metabolic roles of TTM proteins and their contribution, together with the more widespread Ppx-GppA phosphatases, to the ubiquitous triphosphatase activity have not been studied so far.

The presence of Ppx-GppA phosphatase paralogues has been reported so far only for the Gram-positive actinobacteria Corynebacterium glutamicum ATCC 13032 (Lindner et al., 2009) and Mycobacterium tuberculosis H37Rv (Choi et al., 2012). In both cases, two ppx genes encoding putative polyPases with a single domain architecture (Ppx-GppA, Pfam PF02541) and similar predicted molecular mass (M\(_m\)) values (approximately 35 kDa) were reported, but in neither case was a full kinetic characterization of the two paralogous proteins carried out. Interestingly, a peculiar scenario of two polyPase isoforms with some biochemical differences, probably generated by proteolytic processing of a single PXX protein precursor, was reported for the actinobacterium Microlunatus phosphovorus (Lichko et al., 2002), which has since been shown to have a single ppx gene (see below). Reported here will be the first, to the best of our knowledge, comparative study of two ppx paralogous genes of the anaerobic, phototrophic bacterium Chlorobium tepidum that encode functional polyPases of different domain architectures. Functional characterization showed dramatic differences in substrate specificity against short- and long-chain polyPs and nucleotides. The remarkable structural and catalytic differences found between these bacterial PPX homologues strongly support them as members of two distinct subfamilies of Ppx-GppA exopolyphosphatases with specific roles in nucleotides and phosphate metabolisms.

**METHODS**

Reagents, strains and plasmids. Linear sodium polyPs, PPI, P\(_3\), P\(_{13–18}\) (polyP mix with mean chain length of 13–18 phosphate residues) and water-insoluble Maddrell salt (crystalline long-chain polyP of very high M\(_m\)), cyclic trimetaphosphate (P\(_{6c}\)), NTPs and guanosine 5'-tetraphosphate (GP\(_4\)) were purchased from Sigma. When necessary, polyP was washed twice with 3.5 ml 70 % (v/v) ethanol, dried overnight in a vacuum desiccator and resuspended in 600 μl distilled water. Very-long-chain polyPs with a mean chain length of approximately 800 phosphate residues (P\(_{LC}\)) were obtained by fractionation of solubilized Maddrell salt, prepared as described by Becke-Goehringer (1961), on a 2 % (w/v) polyacrylamide/0.8 % (w/v) agarose gel. All other chemicals were of analytical grade.

The strain Chl. tepidum TLS-1 was kindly provided by Professor Dr Michael T. Madigan (Southern Illinois University, Carbondale, IL, USA). E. coli DH5α was used as a host for cloning and propagation, and E. coli BL21 (DE3) was used for overexpression of cloned genes. Plasmids pGEM-T Easy and pQE-80L, used as cloning and expression vectors, respectively, were purchased from Promega and Qiagen.

DNA manipulation. Genomic DNA of Chl. tepidum (strain TLS-1/ ATCC 49652) was extracted using the method described by Wahlund et al. (1991). The PCR-amplified products and plasmids were extracted with DNA gel extraction and plasmid miniprep kits from...
Cloning of two *Chl. tepidum* genes encoding putative Ppx-GppA phosphatases. In the published *Chl. tepidum* TLS genome sequence (TIGR 2002) (Eisen et al., 2002), the complete ORFs for two paralogous genes encoding putative polyPases, *ppx1* (gi_21645997) and *ppx2* (gi_21647723), were inferred. For expression in *E. coli*, these ORFs were amplified by high-fidelity PCR using two pairs of specific primers, for which directional cloning introduced upstream and downstream restriction sites BamHI and PsI, respectively, as is shown in Table S1 (available in the online Supplementary Material). The PCR-amplified DNA fragments corresponding to the *ppx1* and *ppx2* genes were recovered and cloned into pGEM-T Easy vector for sequencing.

Construction of recombinant plasmids and expression in *E. coli*. The *ppx* genes were digested with BamHI and PsI, and then ligated into pQE-80L. In this way, a His6 tag was added to the N-terminal end of the native proteins. The recombinant plasmids were transformed into *E. coli* BL21(DE3), and the cells were incubated at 37 °C in 1 L Luria–Bertani medium supplemented with 100 µg ampicillin ml⁻¹ with vigorous shaking. The cultures were induced with 1 mM IPTG when the OD₆₀₀ of the culture increased to approximately 0.7 and then incubated at 30 °C for 4 h with shaking at 200 r.p.m.

Purification of the recombinant polyPases CtPPX1 and CtPPX2. Cells were harvested and resuspended in buffer A (200 mM NaCl, 5 mM MgCl₂, 10 mM imidazole, 25 mM Tris/HCl, pH 8.0), then lysed by sonication at 4 °C. Cell debris was removed by centrifugation. The crude extract was loaded onto a pre-equilibrated His-Trap HP 1 ml Ni-NTA column (GE-Healthcare). Subsequently, non-target proteins were removed by washing the column with buffer B (200 mM NaCl, 5 mM MgCl₂, 50 mM imidazole, 25 mM Tris/HCl, pH 8.0) until no more protein elution was observed. Finally, recombinant CtPPX1 and CtPPX2 were eluted by applying a linear gradient with a target concentration of 100 % buffer C (200 mM NaCl, 5 mM MgCl₂, 500 mM imidazole, 25 mM Tris/HCl, pH 8.0) at a flow rate of 2 ml min⁻¹. Fractions containing the purified proteins were pooled and dialysed three times against with 50 mM Tris/HCl (pH 6.5) buffer plus 5 mM MgCl₂ to remove imidazole and phosphate salts, then concentrated by ultrafiltration (Amicon Ultra 3 kDa membranes), and finally checked for polyPase activity.

Analytical gel filtration chromatography. Native *Mₘ* of CtPPX1 and CtPPX2 were determined using an FPLC gel filtration chromatography column (Superose 12 HR 10/30, 10 × 300 mm; GE-Healthcare). Proteins were eluted with 200 mM NaCl, 5 mM MgCl₂, 50 mM Tris/HCl (pH 6.5) buffer at a flow rate of 2 ml min⁻¹. Native *Mₘ* values were calculated by column calibration with six standard proteins of known *Mₘ* values: thyroglobulin (Thy, 669 kDa), ferritin (Fer, 443 kDa), β-amylase (β-Amy, 200 kDa), alcohol dehydrogenase (ADH, 150 kDa), carbonic anhydrase (CA, 29 kDa) and cytochrome c (Cytc, 12.4 kDa).

Peptide mass fingerprinting and validation of CtPPX proteins by MALDI-TOF MS. Protein samples corresponding to high-purity CtPPXs were derived from SDS-PAGE. Proteins were digested with trypsin, and the resulting peptides were extracted and loaded onto a suitable MALDI matrix, and later processed by a MALDI-TOF mass spectrometer (AutoFlex; Bruker-Daltonics Proteomics Service of the Instituto de Bioquímica Vegetal y Fotosíntesis, CSIC-Universidad de Sevilla), which generated peptide mass spectra in the mass range 0.8–2.5 kDa. The Mascot-Matrix Science database was used to analyse the peaks lists for protein identification (Koening et al., 2008).

Sigma-Aldrich. Preparation and transformation of *E. coli* competent cells was performed according to Green & Sambrook (2012).

**RESULTS AND DISCUSSION**

Identification of *ppx* and *ppk* paralogous genes in the *Chl. tepidum* genome

The GenBank database was searched using the TBLASTN algorithm and the deduced amino acid sequences of *E. coli* *ppk1* and *ppx* genes as queries (Akiyama et al., 1992) to look for homologues in the genomes of phototrophic bacteria. Several possible *ppx* and *ppk1* genes encoding, respectively, polyPase and PPK-like proteins, most of them annotated as putative, were identified in the genomes of phylogenetically diverse phototrophic bacteria, including anoxygenic photobacteria and cyanobacteria.Remarkably, pairs of *ppx* and *ppk1* paralogous genes involved in polyP metabolism, likely generated by ancient gene duplications, were found in the genome of the therophilic green-sulfur bacterium *Chl. tepidum* TLS (Eisen et al., 2002). Subsequent analysis revealed that the two putative *ppx* genes – CT0099 (993 bp) and CT1713 (1557 bp), hereafter referred as *ppx1* and *ppx2*, respectively, were amplified by high-fidelity PCR into pGEM-T Easy vector for sequencing.

Exopolyphosphatase activity assays. Unless otherwise stated, enzymic activities were measured using a standard assay mixture containing 50 mM Tris/HCl (pH 6.5) buffer, 5 mM MgCl₂, 20 mM KCl, 1 mM P₁₃₋₁₈ (calculated for polyP with a mean chain length of 15 phosphate residues) and 10 µl purified CtPPX at adequate concentration, in a total volume of 1 ml. Other polyPs, PpI, NTPs and GP₄ were used in the assays instead of P₁₃₋₁₈ when the efficiencies of alternative substrates were tested. All reactions were performed at room temperature (25 °C). Nucleoside triphosphatase (NTPase), inorganic pyrophosphatase and polyPase activities were determined by colorimetric measuring of released Pi with the ascorbic acid–ammonium molybdate reagent (Ames, 1966; Gómez-García et al., 2007). One unit is defined as the amount of enzyme catalysing the release of 1 µmol P₁ min⁻¹ under the standard conditions given. Alkaline phosphatase activity was monitored spectrophotometrically at 405 nm by the cleavage of p-nitrophenylphosphate (pNPP) (1 mM) at pH 7.5. Each enzymic activity determination was carried out in triplicate and mean values ± s.e are provided.

Determination of kinetic parameters. The *Kₘ* of the purified enzymes were calculated using mixtures containing concentrations of P₃, Gp₂, or P₁₃₋₁₈ from 10 to 1400 µM, at pH 6.5, and 0.6–1.1 µg of the indicated purified PPX in an assay volume of 1.0 ml. Kinetic parameters were determined by non-linear curve fitting from the Michaelis–Menten plot using the spreadsheet Anemona.xlt (Hernández & Ruiz, 1998).

Effects of pH and metal cations on the activity of CtPPX proteins. For the studies on the effect of pH, CtPPX activities were measured in assay mixtures covering the pH range from 5.5 to 11.0 (in increments of 0.5 pH units). The buffers used for optimal pH determinations were MES (pH 5.5–7.0), MOPS (pH 7.0–8.0), Tris (pH 8.0–9.0), CHES (pH 9.0–10.0) and CHAPS (10.0–11.0) at 50 mM final concentration, adjusted to the indicated pH ranges with NaOH or HCl.

To investigate the effects of different divalent metal cations on the activity of CtPPX1 and CtPPX2, 5 mM of the corresponding chloride salts was added to the reaction mixture instead of the Mg²⁺ salt. For this study, 8 mM EDTA was also included in the reaction mixture to test whether free-metal cofactor availability is a fundamental requirement for CtPPX polypase activity.
respectively - which are located in different regions of the bacterial genome encode different Ppx-GppA phosphatase proteins. Homologues of both genes were identified in cyanobacteria, other phototrophic bacteria and a range of diverse heterotrophic prokaryotes (bacteria and archaea) (Gómez-García et al., 2003; our unpublished results). At the protein level, sequence analyses of CtPPX1 (330 aa; nominal mass 35 799 Da) and CtPPX2 (518 aa; nominal mass 58 436 Da) revealed a quite low level of amino acid identity to each other (approximately 27 % identity) in the overlapping N-region (approximately 320 aa). This region encloses the Ppx-GppA domain (Pfam PF02541) containing a number of conserved motifs and conserved catalytic and substrate-cofactor-binding residues involved in phosphatase activity, while the extra C-terminal region exclusive of CtPPX2 (approximately 190 aa) harbours an HD domain (Aravind & Koonin, 1998) (Fig. S1). The identities shared between CtPPX1 and CtPPX2 and other investigated Ppx-GppA proteins suggested distant evolutionary relationships between them: while CtPPX1 shared higher identities (35–40 %) with one of the homologous proteins of Cor. glutamicum and M. tuberculosis, CtPPX2 shared the highest identity (approximately 35 %) along its overall sequence length with the polyPase of the cyanobacterium Synechocystis, which also possesses a C-terminal HD domain (Table S2). In contrast, the two parologueous ppk1 genes of Chl. tepidum, CT0887 (ppk1-1; 2097 bp) and CT1049 (ppk1-2; 2,145 bp), encoded proteins that share a remarkably high level of identity with each other, approximately 67 %, suggesting a relatively recent gene duplication event in this case. Considering the high sequence homology between the two Chl. tepidum PPKs, as well as with other previously studied PPX1 proteins (Rao et al., 2009), we decided to focus on the biochemical characterization of the two distinct PPX homologues with the aim of providing insights to their specific biological roles.

Gene cloning and overproduction of recombinant CtPPX proteins

The putative ppx genes of Chl. tepidum were cloned from genomic DNA by PCR amplification. DNA fragments with the expected sizes of 993 and 1557 bp for ppx1 and ppx2 genes, respectively, were obtained (Fig. 1a). They were initially cloned into pGEM-T Easy vector and afterwards into the expression vector pQE-80L, so a His6 tag was eventually added to the N-terminus in the recombinant proteins. The generated plasmids pTAR1/Ctep and pTAR2/Ctep containing, respectively, the recombinant ppx1 and ppx2 genes were introduced into the protease-deficient E. coli strain BL21. By the addition of IPTG, overexpression of Chlorobium ppx genes induced in early exponential phase cultures increased polyPase specific activity levels by about 10-fold in the bacterial host. Cell extracts from induced E. coli cultures overexpressing CtPPX1 and CtPPX2 showed major protein bands of about 37 and 60 kDa in SDS-PAGE gels (Fig. 1b) and high exopolypophosphatase specific activity levels with P₁₃₋₁₈ as substrate, 0.4 and 0.5 μmol min⁻¹ mg⁻¹, respectively. In contrast, extracts from cells containing the pQE-80L plasmid with no insert did not show the aforementioned major protein bands in SDS-PAGE gels and, furthermore, exhibited clearly lower specific activity levels, about 0.05 μmol min⁻¹ mg⁻¹, probably due to the bacterial host PPX.
Milligram quantities of overproduced recombinant CtPPX1 and CtPPX2 were obtained from the cell extracts (soluble protein fractions) after purification by Ni-NTA affinity chromatography, following a standard procedure as described in Methods. The protein elution profile showed in both cases a main peak overlapped with the single peak of polyPase activity corresponding to the recombinant protein, which was eluted at an imidazole concentration of 180 mM (Fig. S2). The purified recombinant proteins were then dialysed to remove imidazole and phosphate salts, and concentrated by ultrafiltration. At this stage CtPPXs were purified to 95–98% homogeneity as checked by SDS-PAGE analysis (data not shown).

CtPPX1 and CtPPX2 have different native oligomeric states

Ni-NTA chromatography-purified CtPPX1 and CtPPX2 preparations were analysed by FPLC gel filtration chromatography on a Superose 12 HR column, which allowed a greater purification level, up to apparent electrophoretic homogeneity, to be achieved (Fig. 1b). In both cases, the elution profiles of protein and enzymic activity showed single, symmetrical overlapped peaks, whose corresponding fractions exhibited in SDS-PAGE gels a single protein band of approximately 37 or 59 kDa (Fig. 2). Native $M_n$ values of 38.8 kDa for CtPPX1 and 100.4 kDa for CtPPX2 were calculated. Therefore, CtPPX1 was validated as a catalytically active monomeric enzyme, which is a rather unusual scenario for Ppx-GppA phosphatases, while CtPPX2 is a homodimeric enzyme, with peak exopolyphosphatase activities in their FPLC elution profiles of about 35 and 60 homodimeric enzyme, with peak exopolyphosphatase activity corresponding to the recombinant protein, which was eluted at an imidazole concentration of 180 mM (Fig. S2). The purified recombinant proteins were then dialysed to remove imidazole and phosphate salts, and concentrated by ultrafiltration. At this stage CtPPXs were purified to 95–98% homogeneity as checked by SDS-PAGE analysis (data not shown).

At this final stage of the purification procedure, both CtPPXs were obtained as functional, highly purified enzymes with a single polypeptide of 37.2 (CtPPX1) or 59.9 (CtPPX2) kDa in SDS-PAGE gels (Fig. 1b). The observed $M_m$ values are slightly higher than those predicted from mRNAs, as expected for polyhistidine-tagged recombinant proteins. Additionally, the identities of the CtPPX1 and CtPPX2 polypeptides were confirmed by peptide mass fingerprinting covering about 50–60% of the natural sequences, and eventual identification by MALDI-TOF MS (Fig. S3). As both isolated proteins were obtained as active and highly pure preparations they were used for the kinetic characterization of the functional polyPases from Chl. tepidum.

Kinetic analyses reveal different catalytic features of CtPPX1 and CtPPX2

Preference for polyP of different chain lengths. The substrate specificities and kinetic parameters of recombinant CtPPX1 and CtPPX2 proteins were investigated, using polyPs of different chain lengths and other phosphorylated substrates. It is noteworthy that the CtPPXs hydrolysed linear polyP of very diverse chain lengths, from the simplest, $P_3$, to $P_{15}$ of several hundred (>300) Pi residues, but with clearly different catalytic preferences (Fig. 3a). The highest specific activity for CtPPX1 was reached with $P_3$ (approximately $590 \pm 40$ μmol min$^{-1}$ mg$^{-1}$) and progressively dropped with longer polyPs to approximately $170 \pm 4$ and $15 \pm 1$ μmol min$^{-1}$ mg$^{-1}$ with $P_{13-18}$ and $P_{15}$, respectively. The opposite pattern was found for CtPPX2, which has a residual activity with $P_3$ (approximately $6 \pm 0.5$ μmol min$^{-1}$ mg$^{-1}$) and high phosphatase activities with longer polyPs such as $P_{13-18}$ ($180 \pm 5$ μmol min$^{-1}$ mg$^{-1}$) or $P_{15}$ ($126 \pm 4$ μmol min$^{-1}$ mg$^{-1}$). No phosphatase activity was observed with either CtPPX when using pNPP, PPI or $P_A$ as substrate (see Fig. 3a). The $K_m$, $V_{max}$ and $k_{cat}$ values of the CtPPXs were calculated for each of the polyP substrates $P_3$, $P_{13-18}$ and $P_4$ (summarized in Table 1). The corresponding values could not be estimated for $P_{15}$ because it consists of a mixture of very long polyPs (mean value 800 Pi residues) with quite different chain lengths. The turnover number ($k_{cat}$) and catalytic efficiency ($k_{cat}/K_m$) values of CtPPX1 with $P_3$ as substrate were about 30- and 65-fold higher than those of CtPPX2. However, the same kinetic parameters of CtPPX1 for a medium-chain polyP such as $P_{13-18}$ were about threefold and sevenfold lower than those of CtPPX2 (Figs S4 and S5). Overall, these data indicated that CtPPX homologues specifically hydrolyse polyP of different chain lengths. CtPPX1 virtually operates as an inorganic tripolyphosphatase, while CtPPX2 clearly prefers very-long-chain polyPs. In this respect, it is interesting to note that bacterial and plant TTM proteins, which are structurally different from polyPases, have been found to be very active and specific tripolyphosphatases (Moeder et al., 2013). This raises a possible scenario of unrelated protein families playing apparently redundant biochemical functions in certain organisms.

CtPPX1 has NTPase activity. Once it has been established that purified CtPPX1 has a strong preference for short-chain polyPs such as $P_3$, it was tested whether this recombinant polyPase also possesses NTPase activity (EC 3.6.1.15). Other studies reported that E. coli PPX (Akiyama et al., 1993), Cor. glutamicum PPX2 (Lindner et al., 2009) and M. tuberculosis MTB-PPX1 (Choi et al., 2012) possess modest ATPase activities. It is noteworthy that CtPPX1 was found to hydrolyse ATP and UTP (70–95 μmol min$^{-1}$ mg$^{-1}$) at similar levels to those that the polyP $P_{13-18}$ usually used in the polyPase assays, and to a lesser degree GTP, CTP and TTP (20–30 μmol min$^{-1}$ mg$^{-1}$) (Fig. 3b), but not phosphorylated carbon metabolites (glucose 6-P, fructose, fructose 6-P, fructose 1,6-dP). Also of note, when the organic tetrapolyphosphate $P_4$ was used as a substrate for CtPPX1, higher levels of phosphatase activity (approximately $430 \pm 20$ μmol min$^{-1}$ mg$^{-1}$), similar to those determined for $P_3$, were achieved (Fig. 3b). This suggested that the nucleoside part of the NTPs cause hindrance of catalysis on
the terminal phosphate residue. In contrast, CtPPX2 showed no detectable phosphatase activity with any NTP, and only a residual activity was observed with GP$_3$ (approximately 5 µmol min$^{-1}$ mg$^{-1}$) (Fig. 3b). Kinetic parameters clearly showed that CtPPX1 was much more active and efficient than CtPPX2 with GP$_3$, by approximately 30-fold (Table 1, Figs S4 and S5). It remains to be seen whether the bacterial alarmones pppGpp and ppGpp, not commercially available so far, are substrates and/or inhibitors on the polyPase or NTPase activities of CtPPX proteins. It cannot be excluded, therefore, that ppGpp may produce an inhibitory effect on these polyPases, as was previously reported for the M. tuberculosis and E. coli PPXs (Choi et al., 2012; Kuroda et al., 1997). In this respect, it is interesting to note that GTP and to a lesser degree PPI were inhibitors of CtPPX1 tripolyphosphatase activity ($K_i$ values of 0.4 and 3.8 mM, respectively), while other NTPs were not. In contrast, none of the NTPs tested significantly inhibited CtPPX2 activity with P$_{13-18}$ as a substrate.

**Requirements for mono- and divalent metal cations.** CtPPX1 and CtPPX2 did not require K$^+$ for their enzymic activities, but like most previously characterized bacterial polyPases (Lindner et al., 2009; Choi et al., 2012; Lichko et al., 2002; Akiyama et al., 1993; Bonting et al., 1993) they were clearly stimulated (about threefold) by the addition of 20 mM KCl (data not shown). In contrast, the phosphohydrolase activity of both polyPases was absolutely dependent on the presence of divalent metal cations in the reaction mixture. Maximum activity was reached with 5 mM Mg$^{2+}$, and was dramatically reduced (down to 10%) by an excess of the chelating agent EDTA (Fig. 3c). This result agrees with the fact that most polyPases of micro-organisms are stimulated by divalent metal cations (Rao et al., 2009). The requirement for a divalent metal cofactor can be partially met to different extents by a number of divalent cations, Mn$^{2+}$, Co$^{2+}$ and Fe$^{2+}$ being the most effective among all those tested (Fig. 3c). For instance, the reaction rates with 5 mM Mn$^{2+}$ were approximately 37 and 65% of those obtained with Mg$^{2+}$ for CtPPX1 and CtPPX2, respectively (Fig. 3c). However, no additive effects were observed, since in the presence of 5 mM Mn$^{2+}$ an equal concentration of Mg$^{2+}$ did not activate CtPPXs further.

**Different pH activity profiles.** Although polyPase activities of CtPPX1 and CtPPX2 have similar slightly acidic pH optima (approximately 6.5) they exhibit remarkable differences in their pH-dependence profiles (Fig. 4). CtPPX1 activity with P$_3$ as a substrate showed a markedly steeper
activity levels were determined with 1 mM NTPs or GP₄. NTPase and guanosine tetraphosphatase activities. Phosphatase activity was determined using 1 mM polyPs of different chain lengths as substrates. No significant activity was detected with pNPP, PPi or dNTPs with either of the two enzymes. (b) Substrate specificities of NTPase and guanosine tetraphosphatase activities. Phosphatase activity levels were determined with 1 mM NTPs or GP₄. NTPase and polyPase activities were measured as described in Methods. (c) Metal cofactor specificity of CtPPX1 and CtPPX2. PolyPase activity towards P₃ (CtPPX1, black bars) or P₃c (CtPPX2, white bars) in the presence of 5 mM divalent cation cofactors. The 100 % values assigned to the optimum cofactor Mg²⁺ correspond to 591 ± 37 and 125 ± 12 µmol min⁻¹ mg⁻¹ for CtPPX1 and CtPPX2, respectively. A drastic reduction in enzyme activity was observed in the presence of an excess of the chelating agent EDTA. NA, No addition of divalent cation. No detectable activities were found in the presence of EDTA with no addition of divalent cation (not shown). All data are shown as the mean ±SEM obtained from three independent experiments. The limit of detection was approximately 0.004 µmol min⁻¹ mg⁻¹.

**Fig. 3.** Catalytic activities of recombinant CtPPX1 and CtPPX2. (a) Influence of polyP length on the phosphatase activity. The release of Pi by CtPPX1 (black bars) and CtPPX2 (white bars) was determined using 1 mM polyPs of different chain lengths as substrates. No significant activity was detected with pNPP, PPi or P₃c with either of the two enzymes. (b) Substrate specificities of NTPase and guanosine tetraphosphatase activities. Phosphatase activity levels were determined with 1 mM NTPs or GP₄. NTPase and polyPase activities were measured as described in Methods. (c) Metal cofactor specificity of CtPPX1 and CtPPX2. PolyPase activity towards P₃ (CtPPX1, black bars) or P₃c (CtPPX2, white bars) in the presence of 5 mM divalent cation cofactors. The 100 % values assigned to the optimum cofactor Mg²⁺ correspond to 591 ± 37 and 125 ± 12 µmol min⁻¹ mg⁻¹ for CtPPX1 and CtPPX2, respectively. A drastic reduction in enzyme activity was observed in the presence of an excess of the chelating agent EDTA. NA, No addition of divalent cation. No detectable activities were found in the presence of EDTA with no addition of divalent cation (not shown). All data are shown as the mean ±SEM obtained from three independent experiments. The limit of detection was approximately 0.004 µmol min⁻¹ mg⁻¹.

**CTPPX1 and CTPPX2 belong to different subfamilies of Ppx-GppA phosphatases**

The catalytic and structural differences found between the two polyPase homologues of *Chl. tepidum* prompted us to carry out a molecular phylogenetic study to clarify their evolutionary relationships with other members of the Ppx-GppA protein superfamily. Proteins containing the Ppx-GppA domain are members of the sugar kinase/actin/hsp-70 superfamily and are different in both sequence and structure from the functionally related RelA/SpoT enzymes, which modulate the stringent response via synthesis and degradation of (p)ppGpp (Cashel et al., 1996). Ppx-GppA proteins are ubiquitous among bacteria and archaea, and typically perform enzymic roles as polyPases and/or GPPases (Reizer et al., 1993). In contrast, the only group of Ppx-GppA proteins reported so far in eukaryotes – the so-called RTG2 proteins of fungi – are regulatory proteins with hitherto unknown polyPase/GPPase activities that may function as protein phosphatases (Jazwinski, 2005); they are involved in the retrograde response, an adaptive signalling pathway of altered mitochondria to the cell nucleus (Liao & Butow, 1993).

To analyse the molecular phylogenetic relationships of the two CtPPX homologues with other bacterial, archaean and eukaryotic Ppx-GppA proteins, a molecular phylogenetic tree was constructed using sequences from selected species representatives of the main bacterial/archaeal groups and the eukaryotic RTG2 proteins, with special emphasis on potential paralogy scenarios among Ppx-GppA proteins (Fig. 5, Table S3). A number of relevant issues came out from this analysis. Six major assemblies of Ppx-GppA proteins with diverse domain architectures and phylogenetic distributions were defined. CtpPX1 and CtpPX2 are arranged with all other chlorobian orthologues in separated compact clusters included in two major evolutionarily distant Ppx-GppA phosphatases subfamilies: the single-domain polyPases of low *Mₘ* (35–40 kDa), with dual triplyphosphatase–NTPase activity, and the larger two-domain Ppx-GppA-HD polyPases (approximately 60 kDa), which displayed a strong preference for long-chain polyP (Fig. 5). The first polyPase class presents a broad distribution among major bacterial clades (*Bacteroidetes*, *Chlorobia*, *Actinobacteria*, *α* and *δ*-Proteobacteria, *Clostridia*, *Sinergistetes* and *Nitrospirae*); however, the latter class is prevailing among diverse phototrophic prokaryotes (*Chlorobia*, *Chloroflexi*, *Cyanobacteria*, etc.).
Table 1. Some physico-chemical and catalytic properties of the recombinant polyphosphatases PPX1 and PPX2 from Chl. tepidum TLS

<table>
<thead>
<tr>
<th>Property</th>
<th>CtPPX1</th>
<th>CtPPX2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_n$ (kDa)</td>
<td>38.8</td>
<td>100.4</td>
</tr>
<tr>
<td>Native oligomer (FPLC)</td>
<td>37.2</td>
<td>59.9</td>
</tr>
<tr>
<td>Subunit (SDS-PAGE)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oligomeric state</td>
<td>Monomer</td>
<td>Homodimer</td>
</tr>
<tr>
<td>Optimum pH</td>
<td>6.5</td>
<td>6.5</td>
</tr>
<tr>
<td>Domain architecture</td>
<td>Ppx-GppA</td>
<td>Ppx-GppA-HD</td>
</tr>
<tr>
<td>Optimum metal cofactor</td>
<td>Mg$^{2+}$</td>
<td>Mg$^{2+}$</td>
</tr>
<tr>
<td>Preferred substrate</td>
<td>Short-chain polyP ($P_3$)</td>
<td>Long-chain polyP</td>
</tr>
</tbody>
</table>

| $P_3$ kinetic parameters*      |              |              |
| $K_m$ ($\mu$M)                | 97.7 ± 9.0   | 212.9 ± 19.4 |
| $V_{max}$ (µmol min$^{-1}$ mg$^{-1}$) | 643.1 ± 44.9 | 6.8 ± 0.4 |
| $k_{cat}$ (s$^{-1}$)           | 398.7 ± 15.6 | 13.5 ± 3.3  |
| Catalytic efficiency $k_{cat}/K_m$ (mM$^{-1}$ s$^{-1}$) | 4081 ± 78     | 63 ± 6      |

| GP$_4$ kinetic parameters*    |              |              |
| $K_m$ ($\mu$M)                | 242.2 ± 20.8 | 335.5 ± 23.5 |
| $V_{max}$ (µmol min$^{-1}$ mg$^{-1}$) | 497.4 ± 23.6 | 7.1 ± 0.2   |
| $k_{cat}$ (s$^{-1}$)           | 308.4 ± 38.9 | 14.2 ± 1.4  |
| Catalytic efficiency $k_{cat}/K_m$ (mM$^{-1}$ s$^{-1}$) | 1273 ± 66     | 42 ± 6      |

| $P_{13–18}$ kinetic parameters* |              |              |
| $K_m$ ($\mu$M)                  | 597.4 ± 68.6 | 264.4 ± 12.4 |
| $V_{max}$ (µmol min$^{-1}$ mg$^{-1}$) | 245.6 ± 9.1  | 227.3 ± 11.3 |
| $k_{cat}$ (s$^{-1}$)            | 157.0 ± 7.9  | 453.8 ± 37.4 |
| Catalytic efficiency $k_{cat}/K_m$ (µM$^{-1}$ s$^{-1}$) | 263 ± 10      | 1 716 ± 134 |

*Kinetic parameters were determined by non-linear curve fitting from the Michaelis–Menten plot using the spreadsheet Anemona.xlt (Hernández & Ruiz, 1998). When indicated, data are means ± SEM of three independent determinations.

Heliobacteria), methanogenic Euryarcheae (Methano-
microbiales), Bacilli, Spirochaetes and other bacterial clades
well adapted to oligotrophic and/or extreme environments
e.g. Thermus/Deinococcus). It should be noted at this
point that the two previously studied CtPPX1-like Ppx-
GppA paralogues from the actinobacteria Cor. glutamicum
and M. tuberculosis are highly active on $P_3$ and possess
ATPase activity (Lindner et al., 2009; Choi et al., 2012),
but in both cases a full kinetic characterization of their
polyPase and NTPase activities was not performed.
However, although the function of the HD domain still
remains unknown, a possible role for CtPPX2 in adaptive
environmental responses, as was proposed for long-chain
polyPs (Lindner et al., 2009), can also be envisaged, as it
was reported in a broad superfamily of HD-domain
hydrolases involved among other functions in the bacterial
stringent response (Kuroda et al., 1997). It should be noted
that the gene encoding the Ppx-GppA-HD polyPase ortho-
logue of the cyanobacterium Synechocystis sp. PCC6803 is
a component of the Pho regulon strongly induced by P
deprivation, showing conspicuous oscillations of transcript
levels driven by the daily cycle (Gómez-García et al., 2003,
2013).

Closely related to the HD-domain polyPase assembly and,
like them, having a strong preference for long-chain
polyPs, emerge the GPPase and GPPase-like polyPase
clades as two sister groups of functionally different Ppx-
GppA phosphatase paralogues, generated by ancient
duplication from a common ancestor (Fig. 5). They are

Fig. 4. pH profile curve and polyPase activity of recombinant
CtPPX1 and CtPPX2 proteins. Dependence on pH for the
polyPase activity in the presence of 5 mM MgCl$_2$ at 30 °C of
purified recombinant CtPPX1 with $P_3$ (●), CtPPX1 with $P_{13–18}$ (▲) and
CtPPX2 with $P_{LC}$ as a substrate (○), respectively. Note both
distinct activities show a well-defined activity peak around pH 6.5.
The 100% levels correspond to 587 ± 39, 166 ± 7 and 125 ± 10 µmol min$^{-1}$ mg$^{-1}$ for CtPPX1 with $P_3$ and $P_{13–18}$ as substrates and
distinct activity peaks around pH 6.5. The 100% levels correspond to 587 ± 39, 166 ± 7 and 125 ± 10 µmol min$^{-1}$ mg$^{-1}$ for CtPPX1 with $P_3$ and $P_{13–18}$ as substrates and
CtPPX2 with $P_{LC}$ as substrate, respectively. Values are the means
of three independent experiments and bars indicate SEM.
large single-domain Ppx-GppA proteins with a C-terminal extra region (55–60 kDa) highly active on long-chain polyPs and GPPase activities (Choi et al., 1993). The remaining three major subfamilies of Ppx-GppA proteins form a broad assembly including: (i) a cluster of large polyPases (approximately 60 kDa) with a C-terminal region without specific domain assignment found in α-proteobacteria only as paralogues of the CtPPX1-like small polyPases–NTPases; (ii) a second group of single-domain polyPases (35–45 kDa) highly active on long-chain polyPs but with very low or residual NTPase/GPPase activities (Choi et al., 2012), and found in Actinobacteria, α-Proteobacteria, Bacilli, Rickettsia, some primitive bacterial groups (Aquificae, Thermotogae) and Archaea; and (iii) a cluster of eukaryotic RTG2 signalling proteins of fungi and choanoflagellates (Liao & Butow, 1993) with no polyPase activity reported so far. Interestingly,
some peripheral basal sequences of bacterial endocellular parasites/symbionts of eukaryotes (e.g. *Protochlorymia amoebophila*) appear also to be included in the latter clade (Fig. 5, Table S3).

Pairs of polyPase paralogues seem to occur in evolutionarily diverse bacterial groups. In most cases, PPX paralogues belong to distinct Ppx-GppA subfamilies and exhibit different structural properties, as we report in this study, suggesting ancient paralogy events. However, in some cases closely related paralogues are found within the same Ppx-GppA subfamily, suggesting more recent gene duplications and possible functional diversification (see Fig. 5, Table S3). These findings support specific biochemical roles for these homologous proteins, mostly associated with signaling pathways and/or environmentally regulated metabolic processes. In any case, these recurrent evolutionary scenarios strongly suggest that Ppx-GppA proteins should play important roles in adaptive cellular metabolism. It is interesting to note in this respect that neither of the CtPPX paralogous genes seems to be organized in hypothetical polyP operons as is the case for *E. coli* (Akiyama et al., 1993), as was inferred from their genome localizations (Fig. S6).

The notable structural and evolutionary diversities of Ppx-GppA proteins should correlate with their remarkable functional plasticity, as this work has demonstrated. It should be noted that the structurally simplest CtPPX1-like polyPases represent the only one Ppx-GppA subfamily with paralogy relationships with several other distinct Ppx-GppA subfamilies, including polyPases highly active on long-chain polyPs (Fig. 5, Table S3). This, together with the extreme simplicity of their preferred substrate – P3 is the simplest polyP – strongly supports an ancient position within the Ppx-GppA superfamily. One can speculate a possible ancestral role of P3 in the origin of life as a precursor of NTPs, similar to that of polyPases highly active on long-chain polyPs (Lindner et al., 2009) strongly suggest. Nevertheless, P3 has never been reported in prokaryotes, in contrast to long-chain polyPs, although it is known as an intermediate in a number of biosynthetic pathways, e.g. as an intermediate of S-adenosylmethionine, and is generated in some enzyme processes (Bettendorff & Wins, 2013; Delvaux et al., 2011). In contrast to this, P3 has been shown as a major polyP in acidocalcisomes of several parasitic protists (Moreno et al., 2000), the vacuole of yeast (*Castro et al.*, 1995) and the halotolerant microalga *Dunalieia* (Pick & Weiss, 1991), and the acidocalcisome-like, mitochondrial and nuclear compartments of mammalian cells (Kumbe & Kornberg, 1995; Abramov et al., 2007; Müller et al., 2009; Seidlmayer et al., 2012). Moreover, most of the few eukaryotic DHK-DHHA2 polyphosphatases studied so far exhibit high tripolyphosphatase activity (Rodrigues et al., 2002; Fang et al., 2007; Tammenkoski et al., 2008), and some of them, like the H-Prune protein, are involved in gene regulation and cell proliferation (Tammenkoski et al., 2008). Remarkably, a soluble DHK-DHHA2 exopolyphosphatase involved in cellular osmoregulation of the protist *Trypanosoma cruzi* is, like CtPPX1, highly active with both P3 and GP4, and has very low activity with long-chain polyP (Fang et al., 2007). Taking into account the known roles of prokaryotic GPPases and eukaryotic RTG2 and Prune proteins in transcriptional gene activation, one can speculate on a possible cellular regulatory function for P3 and CtPPX1-like polyPases. In any case, it may be expected that with the development of novel more-sensitive methods it will be possible to determine P3 concentration and subcellular localization as an essential step towards the understanding of their possible biological roles.

**ACKNOWLEDGEMENTS**

The authors thank Professor M. T. Madigan (Dept of Microbiology, Southern Illinois University, USA) for generously providing a sample of *Chlorobium tepidum* TLS cells. This work was supported by research grants from the Spanish Government (BFU2004-00843, BFU2007-61887 and BFU2010-15622) and the Andalusian Regional Government (PAIDI group BIO-261), all of them partially funded by the EU FEDER programme. PAIDI group BIO-261 belongs to the CeiiA3 and AndalusiaTECH University Campuses of International Excellence. The authors thank Dr M. R. Gómez-García for helpful suggestions and discussions.

**REFERENCES**


Two exopolyphosphatase homologues from Chl. tepidum


Edited by: C. Dahl