Journal of Structural Biology 197 (2017) 201–209

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Structural and functional characterization of a cold adapted TPM-domain with ATPase/ADPase activity

María L. Cerutti [a,b], Lisandro H. Otero [a,b], Clara Smal [a], Leonardo Pellizza [a], Fernando A. Goldbaum [a,b], Sebastián Klinke [a,b,*], Martín Aran [a,*]

[a] Fundación Instituto Leloir, IIBBA-CONICET, Av. Patricias Argentinas 435, C1405BWE Buenos Aires, Argentina
[b] Plataforma Argentina de Biología Estructural y Metabolómica PLABEM, Av. Patricias Argentinas 435, C1405BWE Buenos Aires, Argentina

ARTICLE INFO

Article history:
Received 16 September 2016
Received in revised form 27 October 2016
Accepted 29 October 2016
Available online 1 November 2016

Keywords:
TPM domain
X-ray crystallography
ATPase activity
NTPDases
Psychrophiles
Cold-adaptation

ABSTRACT

The Pfam PF04536 TPM_phosphatase family is a broadly conserved family of domains found across prokaryotes, plants and invertebrates. Despite having a similar protein fold, members of this family have been implicated in diverse cellular processes and found in varied subcellular localizations. Very recently, the biochemical characterization of two evolutionary divergent TPM domains has shown that they are able to hydrolyze phosphate groups from different substrates. However, there are still incorrect functional annotations and uncertain relationships between the structure and function of this family of domains. BA41 is an uncharacterized single-pass transmembrane protein from the Antarctic psychrotolerant bacterium *Bizionia argentinensis* with a predicted compact extracytoplasmic TPM domain and a C-terminal cytoplasmic low complexity region. To shed light on the structural properties that enable TPM domains to adopt divergent roles, we here accomplish a comprehensive structural and functional characterization of the central TPM domain of BA41 (BA41-TPM). Contrary to its predicted function as a beta-propeller methanol dehydrogenase, light scattering and crystallographic studies showed that BA41-TPM behaves as a globular monomeric protein and adopts a conserved Rossmann fold, typically observed in other TPM domain structures. Although the crystal structure reveals the conservation of residues involved in substrate binding, no putative catalytic or intramolecular metal ions were detected. Most important, however, extensive biochemical studies demonstrated that BA41-TPM has hydrolase activity against ADP, ATP, and other di- and triphosphate nucleotides and shares properties of cold-adapted enzymes. The role of BA41 in extracellular ATP-mediated signaling pathways and its occurrence in environmental and pathogenic microorganisms is discussed.

1. Introduction

The TPM_phosphatase Pfam family PF04536, formerly known as "DUF477" and "Repair_PSII", is a superfamily of poorly characterized protein domains, named after the TLP18.3, Psb32 and MOLO-1 proteins.

Although some progress has been made in understanding the cellular role of TPM domains, there are very few TPM-containing proteins functionally and/or structurally characterized to date. The TPM containing proteins AtTLP18.3 from *Arabidopsis thaliana* and Psb32 from the cyanobacterium *Synechocystis* sp. are homologous proteins that associate with the thylakoid Photosystem II (PSII) complex and provide partial protection to resist photoinhibition under fluctuating high-light conditions (Sirpio et al., 2007). A detailed biochemical and structural characterization of the TPM domain of AtTLP18.3 showed that it has acid phosphatase activity against different substrates, including synthetic phosphopeptides, leading to the proposal that this protein may participate in regulatory dephosphorylation processes in the thylakoid lumen (Wu et al., 2011). The crystal structure of the AtTLP18.3 TPM domain [PDB code 3PTJ] reveals a Rossmann fold architecture with eight α-helices and four β-strands forming a three-layer (αβα) sandwich, sharing high structural similarity with bacterial phosphotransferases and exopolyphosphatases (Wu et al., 2011). Moreover, the

Abbreviations: pNPP, p-Nitrophenylphosphate; pSer, O-phospho-L-serine; AMP-PCP, β,γ-methyleneadenosine 5′-triphosphate; SLS, static light scattering; DLS, dynamic light scattering; CD, circular dichroism; NTPDase, nucleoside triphosphate diphosphohydrolase; LCR, low-complexity region.

* Corresponding authors at: Fundación Instituto Leloir., Av. Patricias Argentinas 435, Buenos Aires C1405BWE, Argentina (S. Klinke and M. Aran).
E-mail addresses: sklinke@leloir.org.ar (S. Klinke), maran@leloir.org.ar (M. Aran).

http://dx.doi.org/10.1016/j.jsb.2016.10.010
1047-8477/

crystal structure of AtTLP18.3 in complex with serine [PDB code 3PW9] allowed for the identification of the residues involved in ligand binding. Later, MOLO-1 was identified as a positive regulator of levamisole-sensitive acetylcholine receptors at the *Caenorhabditis elegans* neuromuscular junction (Boulin et al., 2012). Three-dimensional homology modeling for the MOLO-1 TPM domain revealed the presence of diverging loops, which may have account to its recent re-classification as a founder member of the new Pfam family PF17175 (modulator of levamisole receptor-1).

NMR structures of the TPM domains of two functionally uncharacterized prokaryotic proteins were also described: CG2496 from *Corynebacterium glutamicum* [PDB code 2KPT] and PG0361 from *Porphyromonas gingivalis* [PDB code 2KW7] (Eletsky et al., 2012). The TPM domain of CG2496 was shown to interact with methiothepin, a potential lead compound for antibiotics (Stark et al., 2014). Most recently, a functional analysis of the TPM-containing protein Rv2345 from the pathogenic bacterium *Mycobacterium tuberculosis* revealed phosphatase and ATPase activities (Sinha et al., 2015).

Later on, our group obtained the solution and crystal structures of BA42 [PDB codes 2MPB and 4OA3, respectively] (Aran et al., 2014), a protein from *Bizionia argentinensis*, a recently discovered psychrotolerant species isolated from Antarctic surface seawater (Bercovich et al., 2008). Interestingly, this structure is the first one of the PF04536 Pfam family to be comprised of a stand-alone TPM domain, where both a new topological variant of the central β-sheet typically observed in the TPM domain structures, and a double metal binding site stabilizing a pair of crossing loops were revealed.

As most members of TPM_phosphatase Pfam family, all these characterized proteins (except for BA42) are predicted to be single pass transmembrane proteins, with their TPM domain facing either the extracellular or luminal milieu. Despite sharing very low sequence identities, these TPM structures present a similar αβα "sandwich" fold not previously found in other protein domains.

In order to gain insight into the relationship between structure and function of this family of domains, we present here the X-ray structure and the biophysical and biochemical characterization of the central TPM domain of BA41 from *B. argentinensis*, a new member of the TPM_phosphatase PF04536 family, hereafter renamed BA41-TPM. The reported structure reveals that the protein folds with a conserved α4β4 Rossmann fold previously observed in other TPM domains, but lacking divalent metals. Remarkably, activity assays demonstrate that BA41-TPM hydrolyze nucleotides di- and triphosphate with broad base specificity and share properties of cold-adapted enzymes. These findings will lead to further explorations of the extracellular nucleotide signaling pathway in environmental and pathogenic microorganisms.

2. Materials and methods

2.1. Bioinformatic analysis

Gene and protein sequences were obtained from the NCBI database (www.ncbi.nlm.nih.gov). The conserved domain search was performed using the sequence of BA41 as query against the Pfam database (Finn et al., 2016) and the Conserved Domain Database (CDD) (Marchler-Bauer et al., 2015). The signal peptides and transmembrane helices from BA41 and other TPM-containing proteins were predicted using the Signal P4.1 server (Petersen et al., 2011) and the TMHMM Server v. 2.0 (Krogh et al., 2001). The low complexity regions were predicted using the SEG server (Wootton, 1994). Sequence phylogenomic searches were run with by BLAST-P (Altschul et al., 1997), EggNOG 4.5 (Huerta-Cepas et al., 2016) and the Clusters of Orthologous Groups of proteins (COGs) database (Galperin et al., 2015). The structural homologs of BA41 were identified using the Protein Data Bank (http://www.rcsb.org/pdb/home/home.do). Amino acid sequences were aligned using the Clustal Omega web server v1.2.1 (Goujon et al., 2010; Sievers et al., 2011) and the alignment was visualized using GeneDoc v2.7. Alignment was performed considering the sequence comprised between the signal peptides and the transmembrane domain.

2.2. Cloning

The sequence corresponding to BA41-TPM (residues 29–178) was amplified by PCR using *B. argentinensis* strain JUB59 genomic DNA as template, 5′-GAGAACCTGTACTTTCAGGGTATGCAGCCTG TTTTAGGACAATTTAC-3′ as 5′-primer and 5′-GGGGACCACTTTGTA CAAGAAAGCTGGGTTAACCTTGATATT CACCTGTTAAAAC-3′ as 3′-primer. The PCR product was purified, amplified with the forward primer 5′-GGGGACAAGTTTGTACAAAAAAGCAGGCTCGGAGAA CCTGTACTTTCAG-3′ and the reverse primer 5′-GGGGACCACTTTG TACAAGAAAGCTGGGTTA-3′ and recombined using the Gateway® BP Clonase® II enzyme mix into the pDONR 201 vector (Invitrogen). The plasmid DNA from positive clones was purified, confirmed by DNA sequencing and recombined into the pDEST-527 expression vector using the Gateway® LR Clonase® II enzyme mix. The final construct contained a six-histidine tag (His-tag) followed by a Tobacco Etch Virus (TEV) cleavage site N-terminally from the BA41-TPM sequence.

2.3. Protein expression and purification

BA41-TPM was expressed in BL21(DE3) *E. coli* cells. The cells were grown at 37 °C in LB medium containing 100 μg mL^{-1} ampicillin to an OD600 nm of 0.8 and recombinant protein expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside for 16 h at 30 °C. Cells were centrifuged and the pellet was resuspended in lysis buffer (50 mM Tris–HCl, 0.5 M NaCl, 40 μg mL^{-1} phenylmethylsulfonyl fluoride, pH 8.0) and disrupted by sonication. The material was centrifuged and the soluble fraction containing the His-tagged protein was filtered through a 0.45 μm membrane and loaded onto a HisTrap HP column (all columns from GE Healthcare) equilibrated with binding buffer (50 mM Tris–HCl, 0.5 M NaCl, 20 mM imidazole, pH 8.0). The recombinant protein was eluted using a linear gradient of elution buffer (50 mM Tris–HCl, 0.5 M NaCl, 0.5 M imidazole, pH 8.0). The appropriate protein fractions were pooled and dialyzed against cleavage buffer (25 mM Tris–HCl, 0.1 M NaCl, pH 8.0) for 2 h at 4 °C to reduce the imidazole concentration for TEV cleavage. After this step, 1 mM DTT and 33 μg of His-tagged TEV protease (Sigma–Aldrich) per mg of protein were added to the dialysis bag and the reaction was allowed to proceed for 16 h at 4 °C. After cleavage, the TEV protease and the histidine tag were removed by passage through the HisTrap HP column. The eluted protein was concentrated to approximately 4.0 mg mL^{-1} in Amicon devices (Millipore), aliquoted and stored at −70 °C. The quality of the final preparation was checked by SDS–PAGE (15% gel) and UV spectrophotometry. For the crystallographic studies, BA41-TPM was further purified on a Superdex-75 column with isocratic elution in gel filtration buffer (50 mM Tris–HCl, 0.25 M NaCl, pH 8.0). The major peak was then concentrated to approx. 20 mg mL^{-1} and at the same time exchanged with crystallization buffer (10 mM Tris–HCl, 25 mM NaCl, pH 8.0). The protein concentration was estimated by measuring its absorbance at λ = 280 nm. The theoretical molar extinction coefficient of the purified protein (ε_{280} - = 21,430 M^{-1} cm^{-1}) was estimated from its sequence using the ProtParam tool from the ExPASy server (Artimo et al., 2012).

2.4. Light scattering

The average MW of BA41-TPM was determined by static light scattering (SLS) on a Precision Detectors PD2010 90° light scattering instrument tandemly connected to an HPLC apparatus, including a Waters 486 UV detector and an LKB 2142 differential refractometer. The chromatographic runs were performed in a Superdex 75 GL 10/300 column with a buffer containing 25 mM Tris–HCl, 250 mM NaCl, pH 8.0 at a flow rate of 0.4 mL min^{-1}. The elution was monitored by measuring its SLS signal at 90°, its UV absorption at $\lambda = 280$ nm, and its refractive index (RI). Data were analyzed with the Discovery32 software supplied by Precision Detectors. The MW of each sample was calculated relating its SLS and RI signals and comparing this value with the one obtained for Bovine serum albumin (MW 66.5 kDa). The effect of divalent metals on the secondary structure content of BA41-TPM was assessed by running the protein sample (120 μM) before and after EDTA treatment. The EDTA treatment consisted in incubating the protein with 1 mM EDTA for 15 min and its subsequent removal by extensive dialysis.

Dynamic light scattering (DLS) measurements were performed at 25 °C with a Zetasizer Nano-S apparatus (Malvern Instruments). For this purpose, 50 μL of the protein were diluted to 0.5 mg ml^{-1} in 10 mM Tris–HCl, 100 mM NaCl, pH 8.0 and placed in a low volume quartz cuvette. Hydrodynamic diameter (D_H) distributions and MWs were calculated using the DTS v.7.11 software provided by the supplier. Data represent an average of at least four measurements.

2.5. Circular dichroism experiments

Far-UV circular dichroism (CD) measurements were carried out in a Jasco J815 spectropolarimeter using a 0.1-cm path length cell. BA41-TPM samples were diluted to 5 μM in 10 mM Tris-HCl, 50 mM NaCl, pH 8.0. Far-UV CD spectra were recorded at 5 and 90 °C, and the assessment of thermal unfolding reversibility was performed by fast cooling back to 5 °C. Each CD spectrum is the average of ten replicate scans. Data were converted to molar ellipticity [θ] per dmol of protein (in units of ° cm^2 $dmol^{-1}_{prot}$) to avoid differences between samples. Thermal unfolding curves were followed by monitoring ellipticity at $\lambda = 216$ nm. The samples were slowly heated with a Peltier system (Jasco) and the temperature scanning was done on the 5–90 °C range at a 2 °C min^{-1} ramp rate and 0.5 °C sampling with 5 s waiting time. The effect of divalent metals on BA41-TPM thermostability was assessed by measuring the unfolding curve before and after EDTA treatment. The apparent T_m was calculated as the temperature midpoint of the thermal transition.

2.6. Crystallization, X-ray data collection and structure resolution

Initial crystallization conditions were determined on 96-well plates using a Honeybee 963 robot (Digilab) and commercial kits from Jena Bioscience and Hampton Research, in a sitting drop vapor diffusion configuration and exploring different protein concentrations (14–28 mg mL^{-1}). After one week of equilibration at 20 °C, two conditions out of the 480 tested showed tiny bars: (i) 24% (w/v) PEG 5000 MME, 0.1 M imidazole pH 7.0, 2% (w/v) PEG 400, and (ii) 18% (w/v) PEG 8000, 0.1 M sodium cacodylate pH 6.5, 0.2 M zinc acetate. Crystals were then optimized by the hanging drop method by mixing 1 μL of the protein stock at 14 mg mL^{-1} with an equal amount of crystallization solution. A slightly modified condition (i) (23% (w/v) PEG 5000 MME, 0.1 M imidazole pH 6.1, 2% (w/v) PEG 400) yielded the best crystals after several weeks. Samples were cryoprotected in mother liquor added with 10% (w/v) PEG 400 and then cryocooled in liquid nitrogen in Hampton Research loops.

Preliminary native X-ray diffraction data were collected on our Bruker D8 QUEST microfocus diffractometer. The final datasets were then obtained on single crystals at 100 K at the PROXIMA 1 protein crystallography beamline at the SOLEIL Synchrotron (France) with a PILATUS 6 M detector (Dectris). Diffraction data were processed to a maximum resolution of 1.40 Å with XDS (Kabsch, 2010) and scaled with Aimless (Evans and Murshudov, 2013). A total of 5% of the recorded reflections were flagged for cross validation. Detailed information on data collection parameters and processing statistics are presented in Table 1.

The BA41-TPM structure was solved by the molecular replacement method using the MrBUMP program (Keegan and Winn, 2007) from the CCP4 package (Winn et al., 2011). For this purpose, the PG0361 domain from *P. gingivalis* (PDB code 2KW7) (Eletsky et al., 2012) was selected as search model, and the procedure continued by automated modification of the probe with MolRep (Vagin and Teplyakov, 2010), molecular replacement search with Phaser (McCoy et al., 2007), and initial model building with Buccaneer (Cowtan, 2006). Two copies of the molecule were located in

Table 1
X-ray diffraction data collection and refinement statistics.

Statistics		
PDB code		5ANP
Data collection		
Synchrotron source		SOLEIL
Beamline		PROXIMA 1
Number of frames		1200
Oscillation step (deg)		0.2
Detector distance (mm)		196
Wavelength (Å)		0.8266
Exposure per frame (s)		0.2
Indexing and scaling		
Cell parameters	a (Å)	34.62
	b (Å)	58.14
	c (Å)	60.20
	α = γ (deg)	90.00
	β (deg)	96.04
Space group		$P2_1$
Resolution limit (Å)		1.40
Number of unique reflections		46,527
Average multiplicity[a]		4.6 (4.6)
<I/σ(I)>		11.6 (2.2)
R_{meas}		0.094 (0.824)
Completeness (%)		99.5 (99.5)
No. of chains per asymmetric unit		2
Solvent content (%)		33
Overall B-factor (Wilson plot, $Å^2$)		13
Refinement		
Resolution range (Å)		34.43–1.40
Number of protein atoms		2225
Number of ligand atoms		–
Number of water molecules		290
R		0.176
R_{free}		0.189
Rms deviations from ideal values (Engh and Huber, 1991)		
Bond lengths (Å)		0.010
Bond angles (deg)		1.0
Average B-factor ($Å^2$)		17
MolProbity validation (Chen et al., 2010)		
Clashscore		1.58
MolProbity score		0.90
Ramachandran plot		
Favored (%)		98.6
Allowed (%)		1.4
Disallowed (%)		–

[a] Values in parentheses correspond to the highest resolution shell (1.40–1.44 Å).

the asymmetric unit and their packing was checked with success. Several cycles of manual model building and refinement were then performed with Coot (Emsley et al., 2010) and Buster (Bricogne et al., 2014), respectively. The final model was validated with MolProbity (Chen et al., 2010). Detailed statistics on the refinement process are shown in Table 1.

2.7. Activity assays

The capacity of the BA41 TPM domain to hydrolyze organophosphate substrates was analyzed with the use of p-Nitrophenylphosphate (pNPP), O-phospho-L-serine (pSer), and different nucleotides (all of analytical quality from Sigma–Aldrich). Reaction mixtures contained 50 mM sodium acetate pH 4.0 or HEPES pH 7.0, 2.5 mM magnesium chloride, and the respective substrates in a total volume of 0.1 mL. All reactions were initiated by adding 33 μg of purified protein and then incubated at 37 °C for 2 h. For the pNPP substrate, the reactions were terminated by the addition of 0.25 mL of 0.5 M sodium hydroxide and the absorbance was measured at $\lambda = 405$ nm. For the pSer and nucleotide substrates, the release of inorganic phosphate (Pi) was measured colorimetrically by the modified malachite green method (Baykov et al., 1988). The malachite green dye reagent containing 0.045% (w/v) malachite green in 6 N sulfuric acid, 7.5% (w/v) ammonium molybdate, and 11% Tween 20 in 1:0.25:0.02 ratio was added to the reaction and allowed to develop for 30 min. Absorbance was measured at $\lambda = 620$ nm and the net amount of released Pi was calculated from a standard phosphate curve. All reactions were measured in a DTX 880 Multimode Detector microplate reader (Beckman Coulter). Control reactions contained all components except the protein. The effect of temperature was determined in the 4–65 °C range. The dependence on the divalent cation was evaluated by the addition of different metals to the reaction mixture in a concentration of 2.5 mM. The presence of residual metals in the reaction buffer was evaluated by addition of 5 mM EDTA. Kinetics measurements were analyzed by fitting the data to the Michaelis-Menten equation using the nonlinear regression equation $Vi = (V_{max}\ [S])/(km + [S])$, where Vi is the initial velocity and [S] is the concentration of the substrate. To compare the hydrolysis of different substrates, the results were normalized (when required) to nmol Pi min^{-1} mg $enzyme^{-1}$.

2.8. Protein Data Bank deposition

The model coordinates and structure factor amplitudes were deposited at the Protein Data Bank under the code 5ANP.

3. Results and discussion

3.1. Functional annotation of BA41

Through a structural genomics project we have previously identified suitable targets for structure determination among the proteins coded in the genome of *B. argentinensis*. In this context, we have recently solved the solution and crystal structure of BA42, a 145-residue protein coded by the ORF 42 contig 3, which is the first structure of a member of the PF04536 family comprised of a stand-alone TPM domain as above mentioned (Aran et al., 2014). With the aim to identify other TPM-containing proteins in the genome of the aforementioned bacterium, we performed a computational screen from the Conserved Domain Database at NCBI. The survey revealed an uncharacterized gene encoding a protein with a predicted TPM-domain (E-value = 4.8×10^{-36}), namely the hypothetical protein BZARG_1550, hereafter renamed BA41 (accession number EGV44687). BA41 is a 277-residue polypeptide composed of a predicted signal sequence (residues 1–33), an extracytoplasmic TPM domain (residues 47–172), a transmembrane domain (residues 187–206), and an intracellular glycine-rich low complexity region (residues 209–277) (Fig. 1A). Interestingly, BA41 has orthologs in bacterial models, such as *Escherichia coli*, and relevant pathogenic microorganisms like *Bordetella pertussis*, *Clostridium tetani*, *Brucella abortus* and *Pseudomonas aeruginosa* (Fig. S1).

Multiple amino acid sequence alignment of BA41 against other TPM-containing proteins such as AtTLP18.3, Rv2345, CG2496 and PG0361, illustrates that TPM domains invariably start at a conserved V-X-D motif from the strand β1 and end after the helix α4 (here arbitrarily set at the conserved Leu172 residue in the BA41 numbering) (Fig. 1B).

Previous studies showed that the TPM domains of *M. tuberculosis* Rv2345 and *A. thaliana* AtTLP18.3 are capable of hydrolyzing phosphate groups from different substrates (Sinha et al., 2015; Wu et al., 2011). The crystal structure of the TPM domain of AtTLP18.3 in complex with serine [PDB code 3PW9] identified three key residues as to be involved in substrate binding, namely Val101, Asp102 and Lys112. Although the structure of the TPM domain of Rv2345 is currently unknown, the corresponding conserved residues Thr36, Asp37 and Arg48 were suggested to be responsible for the observed hydrolase activity (Sinha et al., 2015). Likewise, the triad (V/T)D(R/K) of putative catalytic residues present in these two TPM domains is also conserved in *B. argentinensis* BA41, corresponding to the residues: Tyr48, Asp49 and Lys59 (Fig. 1B).

Despite these observations suggest that BA41 could exhibit phosphoric hydrolase activity, phylogenomic databases classify this protein as a member of the cluster of the orthologous group COG1512, whose functional annotation is as a methanol dehydrogenase or as beta-propeller domains of methanol dehydrogenase type. Methanol dehydrogenases are quino-enzymes found in the periplasm of many Gram-negative methylotrophic or autotrophic bacteria (Anthony, 1986). They belong to the broad class of eight-bladed β propeller quinoproteins comprising a range of alcohol and aldehyde dehydrogenases enzymes that are involved in distinctive catabolic pathways (Keltjens et al., 2014). Thus, given this functional annotation is inconsistent with the TPM_phosphatase Pfam classification, we decided to carry out a complete set of structural, biophysical and biochemical studies with a truncated form of the BA41 protein that were suitable for crystallographic experiments: BA41-TPM (Fig. 1A).

3.2. Biophysical characterization of BA41-TPM

With the aim of determining the oligomeric state of BA41-TPM in solution, the purified protein was analyzed by light scattering techniques. The SEC-SLS analyses demonstrated that BA41-TPM elutes as a single and symmetric peak with an average MW of 15.8 kDa (Fig. 2A), in good agreement with the MW calculated from its amino acid sequence (16.7 kDa). Complementary DLS measurements were also consistent with the expected value for a globular monomer, giving an average particle size diameter of 3.96 nm (Fig. 2A, inset).

To date, BA42 from *B. argentinensis* is the only protein comprising a TPM domain for which an extensive biophysical characterization has been performed. CD thermal unfolding experiments demonstrated that BA42 presents a reversible thermal unfolding-folding process, and that its thermal stability is modulated by Ca^{2+}, being incremented upon metal binding (Aran et al., 2014). Thus, we subsequently evaluated the secondary structure and thermal stability of BA41-TPM and compared it to the behavior of the stand-alone BA42 domain. The far-UV CD spectrum of BA41-TPM showed a negative plateau between $\lambda = 209$–220 nm, which is compatible with a mixed α-helical and β-sheet secondary structure

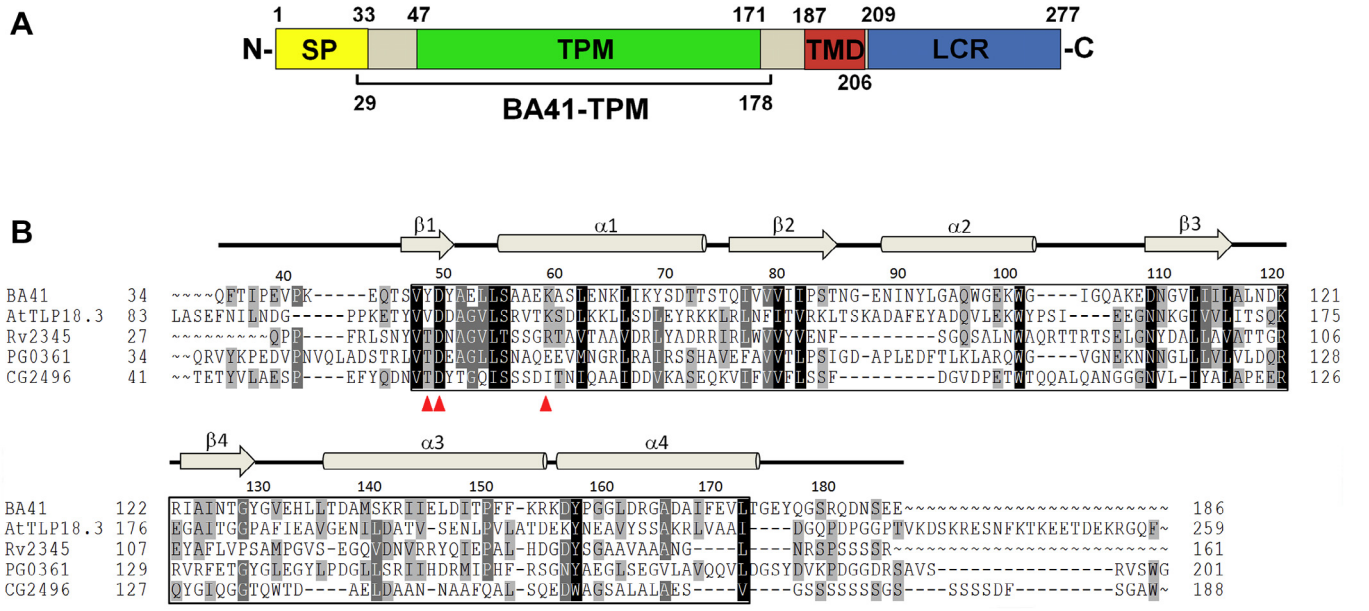

Fig. 1. Sequence alignment of BA41 with other characterized TPM-containing proteins. (A) Schematic domain organization of BA41. SP: Signal peptide, TPM: TPM domain, TMD: Transmembrane domain, LCR: Low complexity region. The region corresponding to the truncated form of BA41 used in this work is indicated as BA41-TPM. (B). Multiple sequence alignment of BA41 and other well characterized TPM-containing proteins: BA41, hypothetical protein BZARG_1550 from *Bizionia argentinensis* (accession number EGV44687); PG0361, conserved domain protein from *Porphyromonas gingivalis* (accession number AAQ65571); AtTLP18.3, thylakoid lumen protein 18.3 from *Arabidopsis thaliana* (accession number AEE33146); Rv2345, transmembrane protein from *Mycobacterium tuberculosis* (accession number CMA73573); CG2496, chromosome segregation ATPase from *Corynebacterium glutamicum* (accession number BAB99668). TPM domains (boxed) were defined as the region comprising the Val47 residue of the V-X-D motif and the conserved Leu172 residue in the BA41 numbering. Black, grey and light grey backgrounds represent 100, 80 and 60% conservation within similarity groups, respectively. The location of the secondary structure elements at the top of the alignment refers to the BA41 crystal structure obtained in this work. Red triangles indicate the residues involved in substrate binding as reported for AtTLP18.3 (Wu et al., 2011). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

content (Fig. 2B, inset). Contrary to the process observed in BA42, heating BA41-TPM to 90 °C produced an irreversible denaturation of the protein, as assessed by the absence of recovery of ellipticity upon cooling back to 5 °C (Fig. 2B, inset). The apparent melting point temperature (*Tm*) of BA41-TPM was calculated to be 59 °C (Fig. 2B).

The crystal structures of BA42 and the TPM domain of AtTLP18.3 revealed the presence of structural Ca^{2+} ions. Therefore, in order to evaluate the effect of divalent metals on the BA41-TPM structure, a series of SEC-SLS and CD experiments were carried out in the presence and absence of EDTA. Overall, no changes in the SEC pattern nor in the thermal unfolding curves after EDTA treatment were observed (Fig. 2A and B), thus indicating that either the protein does not contain any metals in its structure or these metals are not important for maintaining the structural stability. Thus, in order to discern between these two possibilities crystallographic studies on BA41-TPM were accomplished.

3.3. High-resolution crystal structure of BA41-TPM

The BA41-TPM crystal structure was determined at 1.40 Å resolution by means of the molecular replacement method in the monoclinic $P2_1$ spacegroup, as described in Materials and Methods. The final model was refined to $R = 0.176$ and $R_{free} = 0.189$ with excellent stereochemistry (Table 1). In this sense, Ramachandran statistics show that almost 99% of the residues lie in the favored region of the plot without outliers. The asymmetric unit contains two individual molecules (A and B) that are essentially identical, with a root mean square deviation (rmsd) of 0.67 Å for 139 aligned C^{α} atoms. An analysis of the existing interfaces in the crystal packing of BA41-TPM did not reveal specific interactions that may yield stable quaternary structures, as judged by the PDBePISA server (Krissinel and Henrick, 2007). This result confirms the expected behavior of the protein as a monomer as seen in the SLS-SEC and DLS experiments. In both chains, there is lack of electron density for the first 5–6 residues. Additionally, the residue range 105–108 from molecule B, which corresponds to an exposed loop, could not be traced in the crystal model. With the exception of these regions, the rest of the polypeptide chain presents continuous density that is consistent with the high resolution of the diffraction data. At the end of the refinement process, a total of 290 water molecules were located in the model. Neither cofactors nor metal ions were found in the electron density. Detailed statistics of the refinement process can be found in Table 1.

The polypeptide chain of BA41-TPM adopts a compact Rossmann fold with an approximate size of 45 × 40 × 35 Å (Fig. 3A), which is compatible with the expected TPM domain predicted by the Pfam database. The protein bears a central mixed four-stranded β-sheet with $\uparrow\beta_1\uparrow\beta_2\uparrow\beta_3\downarrow\beta_4$ topology that is surrounded by three α-helices on one side (α1, α3 and α4) and a single α-helix on the opposite side (α2). The near equivalent content of α-helical and β-sheet elements is in strong agreement with the recorded far-UV CD spectra in solution above described.

A search for homologue proteins with known three-dimensional structures using the PDBeFold server (Krissinel and Henrick, 2004) revealed only four hits (rmsd values ranging from 1.43 to 2.04 Å), which are all members of the PF04536 family: (i) PG0631 domain from *P. gingivalis*, (ii) AtTLP18.3 from *A. thaliana*, (iii) BA42 from *B. argentinensis*, and (iv) CG2496 from *C. glutamicum* (Table S1).

The high structural similarity between BA41 and the AtTLP18.3 TPM domain, together with the conservation of the residues involved in the binding of the serine ligand in AtTLP18.3 (Val101, Asp102 and Lys112, Fig. 3B), suggest that BA41 may present hydro-

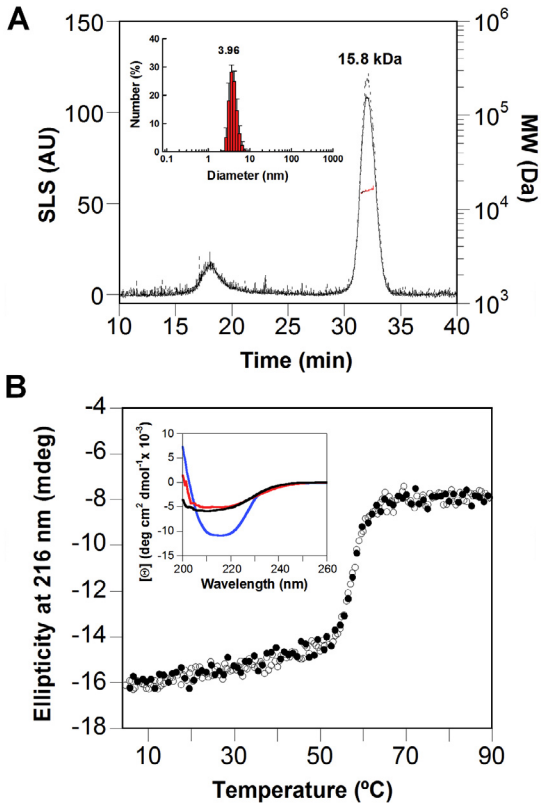

Fig. 2. Biophysical characterization of BA41-TPM. (A) SEC-SLS analysis of BA41-TPM. The trace of the calculated MW is shown in red. The number above the peak corresponds to the estimated MW. Inset. Hydrodynamic diameter distribution of BA41-TPM measured by DLS. The number above the peak corresponds to the estimated D_H. (B) Thermal unfolding of BA41-TPM. The secondary structure transitions were measured in the presence (closed circles) and absence (open circles) of EDTA. All measurements were carried out in 10 mM Tris–HCl, 50 mM NaCl, pH 8.0. Inset. Irreversible thermal-induced denaturation of BA41-TPM. Far UV CD spectra of the protein (5 μM) recorded at 5 °C (blue line), 90 °C (red line), and after cooling back form 90 °C to 5 °C (black line). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

lase activity like the other members of this Pfam family. As a divergence, the single Ca^{2+} ion located about 10 Å from the substrate binding site in AtTLP18.3 was not found in the electron density in the BA41 structure.

In addition, we were also able to obtain well-diffracting crystals (approx. 1.70 Å maximum resolution, data not shown) in the presence of 200 mM zinc acetate among other chemicals (see crystallization condition (ii), Section 2.6 in Materials and Methods). Once the structure was solved, the electron density maps for this particular crystallization condition showed the presence of several Zn^{2+} cations, however, all of them corresponded to intermolecular bridging ions without any putative catalytic or intramolecular role. Additionally, DLS experiments demonstrated that Zn^{2+} cations induce BA41-TPM oligomerization (Table S2), which confirms that the metal acts as a key reagent in nucleation for crystal growth rather as a specific ligand. In fact, metal ions, such as Cu^{2+} and Zn^{2+}, are well known to promote protein oligomerization, as in the case of the amyloid-β peptide (reviewed in (Hane and Leonenko, 2014)).

Taken together, the biophysical and structural results demonstrate that the TPM domain of BA41 behaves as a stable monomer, which does not contain a pyrroloquinoline quinone prosthetic group, nor it ensembles into the oligomeric structures typically found in β-propeller alcohol dehydrogenase proteins. Additionally, no alcohol dehydrogenase activity for BA41-TPM was detected on standard assays against a mix of alcohols and aldehydes (data not shown), providing further evidence of an inaccurate functional annotation for this protein.

3.4. Enzymatic activity of the BA41 TPM domain

As mentioned before, the TPM domains of Rv2345 and AtTLP18.3 are capable of removing phosphate groups from different substrates (Sinha et al., 2015; Wu et al., 2011). The TPM domain of AtTLP18.3 was shown to have phosphatase activity against pNPP, pSer and several synthetic phosphorylated oligopeptides. Its optimal activity was found to be at pH 4.0 and 37 °C and poorly dependent on the presence of divalent metals. In addition,

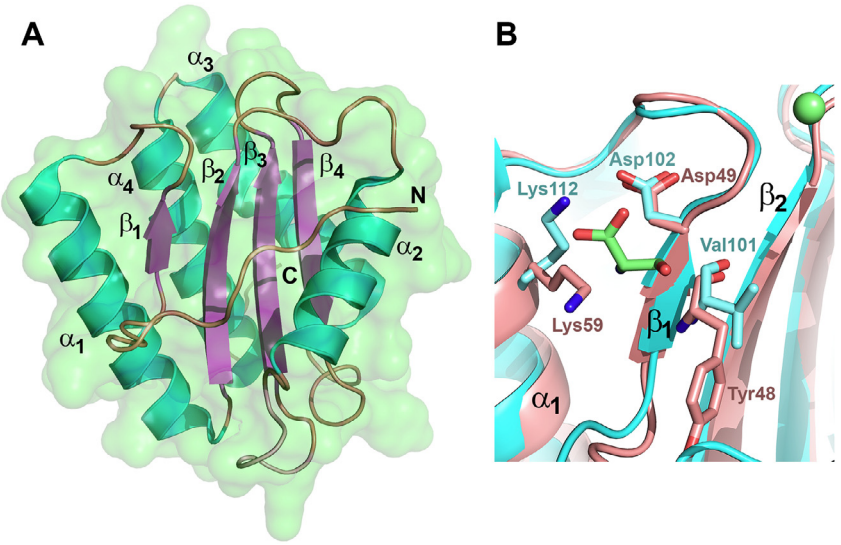

Fig. 3. Crystal structure of BA41-TPM. (A) Cartoon representation (α-helices, green; β-strands, pink; loops, brown). Secondary structure elements as well as the N- and C-termini are indicated. The molecular surface is depicted in the background in light green. (B) Superposition between BA41-TPM (salmon) and AtTLP18.3 (cyan) at its serine substrate binding site, in a similar orientation to (A). The most relevant side chains are represented in sticks. The calcium ion and the serine ligand molecule from AtTLP18.3 are depicted in green. The figure was prepared using PyMOL (Schrödinger, 2010). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the TPM domain of Rv2345 revealed hydrolase activity against pNPP but also against ATP (Sinha et al., 2015).

Based on these previous experimental data, we decided to evaluate the hydrolase properties of BA41-TPM against the substrates pNPP, pSer and ATP at neutral and acidic pH conditions. Interestingly, BA41-TPM displayed the highest activity with ATP at pH 7.0 and 4.0, respectively (Fig. 4A). However, the protein showed no activity against pNPP, and 40–65% lower activity with pSer against the observed with ATP at pH 4.0 and 7.0, respectively. Thus, indicating that BA41-TPM behaves more as an ATPase than a phosphatase enzyme.

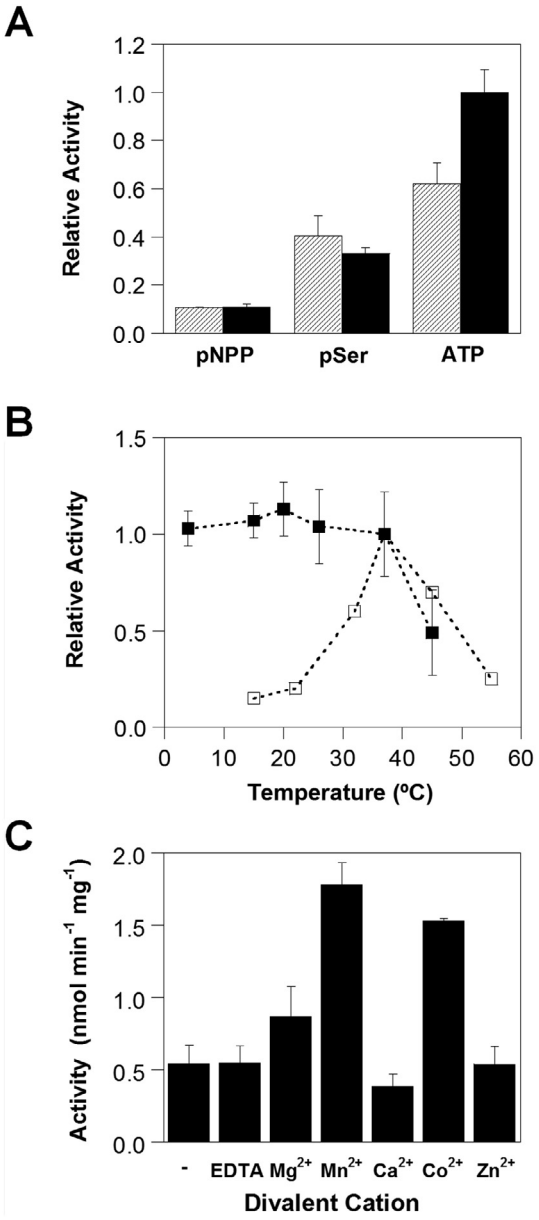

Fig. 4. Biochemical characterization of BA41-TPM. (A) Hydrolase activity with different substrates and pH conditions. pNPP, pSer and ATP substrates were assayed at pH 4.0 (dashed bars) and 7.0 (closed bars). Values are expressed as the means ± standard deviations and are relative to the highest activity (ATP at pH 7.0). (B) Effect of the temperature on the activity of the psychrophilic BA41 and the mesophilic AtTLP18.3 TPM domains. BA41-TPM activity (closed symbols) was assayed in the presence of 2 mM ATP, 2.5 mM $MgCl_2$, pH 7.0. The results are the means of at least three experiments performed in duplicate. The activity of the AtTLP18.3 TPM domain (open symbols) was estimated by extrapolating the data reported by Wu et al. (Wu et al., 2011). Values are expressed relative to the activity at 37 °C for each enzyme. (C) Effect of divalent metals on the BA41-TPM activity. ATPase enzymatic assays were done in the absence or presence of 2.5 mM divalent cations at pH 7.0. EDTA (5 mM) was added to the reaction buffer in absence of metals. The results are the means of at least three independent experiments performed in duplicate. Error bars represent SD.

In order to compensate for the slow reaction rates at low temperatures, psychrophiles enzymes exhibit up to tenfold higher specific activity in this temperature range than their mesophilic counterparts (Gerday et al., 1997). BA41-TPM displays this typical peculiarity of cold-adapted enzymes, with high activity at low to middle temperatures (4–37 °C). By contrast, the TPM domain from the mesophilic AtTLP18.3 shows virtually no activity below 20 °C and a marked maximum at 37 °C (Fig. 4B). Although the highly spontaneous hydrolysis of ATP at temperatures over 50 °C invalidated measurements beyond this value, activity assays with the more stable pSer substrate demonstrated that BA41-TPM has no activity at 70 °C (Fig. S2), which is in line with the observed irreversible denaturation of the protein in this range of temperature (see Fig. 2B). Given that no information on the thermal unfolding of the TPM domain of AtTLP18.3 is currently available, it can be inferred that either both proteins have similar global stability or, even if they had differential stability, their active sites are disassembled in a similar temperature range.

As most ATPases, BA41-TPM activity is modulated by divalent cations (Fig. 4C). Although it exhibited a basal activity in the absence of metals ions in the reaction buffer, the addition of Mn^{2+}, Co^{2+} or Mg^{2+} induced a significant increase in its ATPase activity (3.3, 2.8, and 1.6-fold, respectively). In contrast, the ATPase activity of BA41-TPM in the presence of Ca^{2+} or Zn^{2+}remained invariable.

In seawater, Mg^{2+} is one the fourth most abundant ion (1290 ppm), while Mn^{2+} and Co^{2+} are present at only 0.0004 ppm (Turekian, 1968). Considering BA41 is synthetized by a marine microorganism, we presumed that the physiological role of Mn^{2+} or Co^{2+} as its cofactors could be irrelevant. In this sense, as activation exerted by the non-physiological Sr^{2+}, Zn^{2+}, Cd^{2+} and Co^{2+} cations have been reported for recombinant cellular nucleoside triphosphate diphosphohydrolases (NTPDases) (Zebisch and Strater, 2007), we continued our biochemical characterization by using magnesium as cofactor in the enzymatic reaction.

We subsequently assayed the ability of BA41-TPM to hydrolyze different nucleotides (Table 2). The order of the relative hydrolase activity was found to be ribonucleosides diphosphate (NDPs) > ribonucleosides triphosphate (NTPs) > ribonucleosides monophosphate (NMPs) ≈ deoxynucleotides triphosphate (dNTPs), indicating a greater efficiency in the hydrolysis of beta-phosphate groups of nucleotides. In this regard, it is important to note that in enzymatic assays with NTPs we could not discriminate between the hydrolysis of gamma and beta phosphate groups and, therefore, the efficiency of BA41-TPM to hydrolyze gamma phosphate groups may be lower than the experimentally observed. Additionally, the enzyme hydrolyzed both purine and pyrimidine nucleotides with similar efficiency revealing no base specificity. Interestingly, low relative activity was observed against dNTPs, suggesting selectivity for ribose. In addition, BA41-TPM showed less efficiency to hydrolyze NMPs, pNPP and pSer suggesting that the sugar or aromatic rings adjacent to the oxygen atom of the phosphoester bond to be hydrolyzed produce a steric hindrance effect that inhibits or drastically reduces enzyme catalysis.

The BA41-TPM activity against varying concentrations of ADP and ATP substrates displayed a typical Michaelis-Menten kinetic (Fig. S3). The K_m values for ATP and ADP were similar (Table 3), suggesting no differences in substrate affinities. However, the V_{max} was ~100% higher for ADP than for ATP, which explains the preference for NDPs as above noted.

Most of the ecto-ATPases that occur on the cell surface and hydrolyze extracellular nucleotides belong to the ecto-nucleoside

Table 2
Substrate specificity of BA41-TPM.

Substrate	Base			
	A	C	G	U/T
Ribonucleotides				
NMP	48	47	47	39
NDP	163	ND	166	ND
NTP	100	ND	94	100
Deoxyribonucleotides				
dNTPs	34	60	47	46
Miscellaneous				
pNPP	2			
pSer	40			
AMP-PCP	21			

The Pi released from ATP was set at the 100% level.
ND: not determined.

Table 3
Kinetics parameters of BA41-TPM.

Substrate	K_m (mM)	V_{max} (nmol min^{-1} mg^{-1})	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} M^{-1})
ATP	0.17 ± 0.07	0.53 ± 0.07	0.44 ± 0.06	2.59×10^3
ADP	0.24 ± 0.06	0.99 ± 0.09	0.83 ± 0.08	3.45×10^3
AMP	LA	LA		

LA, low activity.
The *V*max and K_m (mean ± SD) were obtained from the nonlinear regression fit of data on Fig. S3. The k_{cat} was calculated assuming a MW of 16.7 kDa.

triphosphate diphosphohydrolases (NTPDases) family. NTPDases are ubiquitously found in higher eukaryotes, but they have also been identified in a wide range of human parasites. These proteins comprise a family of membrane-bound enzymes characterized by the presence of five apyrase conserved regions and by the ability to catalyze the hydrolysis of di- and triphosphate nucleosides to the monophosphate form, with varying specificity for the nucleobase (Zimmermann et al., 2012). In this sense, it is interesting to highlight that BA41 presents K_m values that fall among those observed in the highly efficient ecto-nucleotidases ATPases (Sansom et al., 2008b; Stroppolo et al., 2001). In contrast, the V_{max} and turnover rates of BA41 are several orders of magnitude lower than those reported for these proteins. In fact, the catalytic efficiency of BA41-TPM resembles that reported for the thylakoidal AtTLP18.3 TPM domain, whose V_{max} values ranged between 3 and 14 nmol min^{-1} mg^{-1} (Wu et al., 2011).

Given this background, TPM domains do not seem to be optimized for high hydrolase activities. They would rather play roles in sensing pathways or be modulated by yet undiscovered factors. In this regard, members of the ecto-NTPDase family were shown to form oligomeric complexes with increased catalytic activity in membrane fractions and to interact with other transmembrane proteins, positively affecting ATP transport and extracellular hydrolysis (reviewed in (Zimmermann et al., 2012).

The high-resolution structure of some NTPDases in complex with substrate showed that their dual ATP/ADP specificity is achieved at the expense of base promiscuity, engaging the same catalytic site for hydrolysis of di- and triphosphate nucleosides (Zebisch et al., 2013). Although BA41 shares characteristics with members of this family of ATPases, unfortunately all co-crystallization and soaking attempts to obtain the crystal structure of BA41-TPM in complex with different nucleotides (ATP, ADP, AMP-PNP) and divalent cations (Mg^{2+}, Mn^{2+}) were unsuccessful, thus preventing further analysis of the active site of BA41.

Another striking feature in BA41 is its predicted intracellular Gly-rich low-complexity region (see Fig. S1). Low-complexity regions (LCRs) are regions of disorder resulting from little diversity in the amino acid sequence composition that are predicted to correspond to disordered structures (Simon and Hancock, 2009; Tompa, 2003). Inquiringly, these Gly-rich stretches are not only present in BA41 orthologues, but in the C-terminal LCR regions of the PG0361 and CG2496 proteins. In general, terminal LCRs are expected to be involved in flexible and rapidly reversible binding, as those occurring in stress responses, translation and transport processes (Coletta et al., 2010).

4. Concluding remarks

Extracellular nucleotides are potent signal molecules in a wide variety of organisms. In mammals, ATP and ADP act as signal messengers of the named purinergic pathway, mediating neuronal processes, host inflammation and immune responses. Parasites, such as *Toxoplasma gondii*, *Schistosoma mansoni*, *Trypanosoma* spp., and the bacterium *Legionella pneumophila* are believed to utilize this same pathway to weaken the host response to infection (Sansom et al., 2008a). Additionally, extracellular ATP was also involved in bacterial adhesion and biofilm formation in some pathogenic strains (Xi and Wu, 2010).

Further *in vivo* experiments with BA41 mutants in *Bizionia argentinensis* and pathogenic bacterial microorganisms would help to elucidate the putative role of this protein in extracellular nucleotide signaling transduction pathways and its potential implications in their mechanisms of virulence.

Acknowledgments

The authors acknowledge access to the PROXIMA 1 beamline at the SOLEIL Synchrotron (France) and to the Argentinian Ministry of Science, Technology and Productive Innovation (MINCyT) for travel support. This work was supported by the Argentinian Research Council (CONICET) and the Argentinian Agency for Science and Technology Promotion (ANPCyT).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jsb.2016.10.010.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Anthony, C., 1986. Bacterial oxidation of methane and methanol. Adv. Microb. Physiol. 27, 113–210.

Aran, M., Smal, C., Pellizza, L., Gallo, M., Otero, L.H., Klinke, S., Goldbaum, F.A., Ithurralde, E.R., Bercovich, A., Mac Cormack, W.P., Turjanski, A.G., Cicero, D.O., 2014. Solution and crystal structure of BA42, a protein from the Antarctic bacterium Bizionia argentinensis comprised of a stand-alone TPM domain. Proteins 82, 3062–3078.

Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., de Castro, E., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E., Grosdidier, A., Hernandez, C., Ioannidis, V., Kuznetsov, D., Liechti, R., Moretti, S., Mostaguir, K., Redaschi, N., Rossier, G., Xenarios, I., Stockinger, H., 2012. ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res. 40, W597–W603.

Baykov, A.A., Evtushenko, O.A., Avaeva, S.M., 1988. A malachite green procedure for orthophosphate determination and its use in alkaline phosphatase-based enzyme immunoassay. Anal. Biochem. 171, 266–270.

Bercovich, A., Vazquez, S.C., Yankilevich, P., Coria, S.H., Foti, M., Hernandez, E., Vidal, A., Ruberto, L., Melo, C., Marenssi, S., Criscuolo, M., Memoli, M., Arguelles, M., Mac Cormack, W.P., 2008. *Bizionia argentinensis* sp. nov., isolated from surface marine water in Antarctica. Int. J. Syst. Evol. Microbiol. 58, 2363–2367.

Boulin, T., Rapti, G., Briseno-Roa, L., Stigloher, C., Richmond, J.E., Paoletti, P., Bessereau, J.L., 2012. Positive modulation of a Cys-loop acetylcholine receptor by an auxiliary transmembrane subunit. Nat. Neurosci. 15, 1374–1381.

Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O.S., Vonrhein, C., Womack, T.O., 2014. BUSTER Version 2.10.1. Global Phasing Ltd., Cambridge, United Kingdom.

Chen, V.B., Arendall 3rd, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 12–21.

Coletta, A., Pinney, J.W., Solis, D.Y., Marsh, J., Pettifer, S.R., Attwood, T.K., 2010. Low-complexity regions within protein sequences have position-dependent roles. BMC Syst. Biol. 4, 43.

Cowtan, K., 2006. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr. Sect. D Biol. Crystallogr. 62, 1002–1011.

Eletsky, A., Acton, T.B., Xiao, R., Everett, J.K., Montelione, G.T., Szyperski, T., 2012. Solution NMR structures reveal a distinct architecture and provide first structures for protein domain family PF04536. J. Struct. Funct. Genomics 13, 9–14.

Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K., 2010. Features and development of Coot. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 486–501.

Engh, R.A., Huber, R., 1991. Accurate bond and angle parameters for X-ray protein structure refinement. Acta Crystallogr. A 47, 392–400.

Evans, P.R., Murshudov, G.N., 2013. How good are my data and what is the resolution? Acta Crystallogr. Sect. D Biol. Crystallogr. 69, 1204–1214.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., Bateman, A., 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 44, D279–D285.

Galperin, M.Y., Makarova, K.S., Wolf, Y.I., Koonin, E.V., 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. Nucleic Acids Res. 43, D261–D269.

Gerday, C., Aittaleb, M., Arpigny, J.L., Baise, E., Chessa, J.P., Garsoux, G., Petrescu, I., Feller, G., 1997. Psychrophilic enzymes: a thermodynamic challenge. Biochim. Biophys. Acta 1342, 119–131.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., Lopez, R., 2010. A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res. 38, W695–W699.

Hane, F., Leonenko, Z., 2014. Effect of metals on kinetic pathways of amyloid-beta aggregation. Biomolecules 4, 101–116.

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., Rattei, T., Mende, D.R., Sunagawa, S., Kuhn, M., Jensen, L.J., von Mering, C., Bork, P., 2016. EggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res. 44, D286–D293.

Kabsch, W., 2010. Xds. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 125–132.

Keegan, R.M., Winn, M.D., 2007. Automated search-model discovery and preparation for structure solution by molecular replacement. Acta Crystallogr. Sect. D Biol. Crystallogr. 63, 447–457.

Keltjens, J.T., Pol, A., Reimann, J., Op den Camp, H.J., 2014. PQQ-dependent methanol dehydrogenases: rare-earth elements make a difference. Appl. Microbiol. Biotechnol. 98, 6163–6183.

Krissinel, E., Henrick, K., 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. Sect. D Biol. Crystallogr. 60, 2256–2268.

Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797.

Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. 305, 567–580.

Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M., Hurwitz, D.I., Lanczycki, C.J., Lu, F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Bryant, S.H., 2015. CDD: NCBI's conserved domain database. Nucleic Acids Res. 43, D222–D226.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674.

Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786.

Sansom, F.M., Robson, S.C., Hartland, E.L., 2008a. Possible effects of microbial ecto-nucleoside triphosphate diphosphohydrolases on host-pathogen interactions. Microbiol. Mol. Biol. Rev. 72, 765–781 (Table of Contents).

Sansom, F.M., Riedmaier, P., Newton, H.J., Dunstone, M.A., Muller, C.E., Stephan, H., Byres, E., Beddoe, T., Rossjohn, J., Cowan, P.J., d'Apice, A.J., Robson, S.C., Hartland, E.L., 2008b. Enzymatic properties of an ecto-nucleoside triphosphate diphosphohydrolase from *Legionella pneumophila*: substrate specificity and requirement for virulence. J. Biol. Chem. 283, 12909–12918.

Schrödinger, L., 2010. The PyMOL Molecular Graphics System, Version 1.3r1.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539.

Simon, M., Hancock, J.M., 2009. Tandem and cryptic amino acid repeats accumulate in disordered regions of proteins. Genome Biol. 10, R59.

Sinha, A., Eniyan, K., Sinha, S., Lynn, A.M., Bajpai, U., 2015. Functional analysis of TPM domain containing Rv2345 of *Mycobacterium tuberculosis* identifies its phosphatase activity. Protein Expr. Purif. 111, 23–27.

Sirpio, S., Allahverdiyeva, Y., Suorsa, M., Paakkarinen, V., Vainonen, J., Battchikova, N., Aro, E.M., 2007. TLP18.3, a novel thylakoid lumen protein regulating photosystem II repair cycle. Biochem. J. 406, 415–425.

Stark, J.L., Copeland, J.C., Eletsky, A., Somerville, G.A., Szyperski, T., Powers, R., 2014. Identification of low-molecular-weight compounds inhibiting growth of corynebacteria: potential lead compounds for antibiotics. ChemMedChem 9, 282–285.

Stroppolo, M.E., Falconi, M., Caccuri, A.M., Desideri, A., 2001. Superefficient enzymes. Cell. Mol. Life Sci. 58, 1451–1460.

Tompa, P., 2003. Intrinsically unstructured proteins evolve by repeat expansion. BioEssays: News Rev. Mol. Cell. Dev. Biol. 25, 847–855.

Turekian, K.K., 1968. Oceans. Michigan University.

Vagin, A., Teplyakov, A., 2010. Molecular replacement with MOLREP. Acta Crystallogr. Sect. D Biol. Crystallogr. 66, 22–25.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A., McNicholas, S.J., Murshudov, G.N., Pannu, N.S., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A., Wilson, K.S., 2011. Overview of the CCP4 suite and current developments. Acta Crystallogr. Sect. D Biol. Crystallogr. 67, 235–242.

Wootton, J.C., 1994. Non-globular domains in protein sequences: automated segmentation using complexity measures. Comput. Chem. 18, 269–285.

Wu, H.Y., Liu, M.S., Lin, T.P., Cheng, Y.S., 2011. Structural and functional assays of AtTLP18.3 identify its novel acid phosphatase activity in thylakoid lumen. Plant Physiol. 157, 1015–1025.

Xi, C., Wu, J., 2010. DATP/ATP, a multifunctional nucleotide, stimulates bacterial cell lysis, extracellular DNA release and biofilm development. PLoS ONE 5, e13355.

Zebisch, M., Strater, N., 2007. Characterization of rat NTPDase1, -2, and -3 ectodomains refolded from bacterial inclusion bodies. Biochemistry 46, 11945–11956.

Zebisch, M., Krauss, M., Schafer, P., Lauble, P., Strater, N., 2013. Crystallographic snapshots along the reaction pathway of nucleoside triphosphate diphosphohydrolases. Structure 21, 1460–1475.

Zimmermann, H., Zebisch, M., Strater, N., 2012. Cellular function and molecular structure of ecto-nucleotidases. Purinergic Signal. 8, 437–502.