Publications
Atomic resolution crystal structure of VcLMWPTP-1 from *Vibrio cholerae* O395: Insights into a novel mode of dimerization in the low molecular weight protein tyrosine phosphatase family

Seema Nath, Ramanuj Banerjee, Udayaditya Sen

Crystallography and Molecular Biology Division, Saha Institute of Nuclear Physics, 1/AF, Bidhannagar, Kolkata 700064, India

**A R T I C L E   I N F O**

Article history:
Received 19 May 2014
Available online 6 June 2014

Keywords:
Bacterial protein phosphatases
Low molecular weight protein tyrosine phosphatases
X-ray crystallography
Active dimer

**A B S T R A C T**

Low molecular weight protein tyrosine phosphatase (LMWPTP) is a group of phosphotyrosine phosphatase ubiquitously found in a wide range of organisms ranging from bacteria to mammals. Dimerization in the LMWPTP family has been reported earlier which follows a common mechanism involving active site residues leading to an enzymatically inactive species. Here we report a novel form of dimerization in a LMWPTP from *Vibrio cholerae* O395 (VcLMWPTP-1). Studies in solution reveal the existence of the dimer in solution while kinetic study depicts the active form of the enzyme. This indicates that the mode of dimerization in VcLMWPTP-1 is different from others where active site residues are not involved in the process. A high resolution (1.45 Å) crystal structure of VcLMWPTP-1 confirms a different mode of dimerization where the active site is catalytically accessible as evident by a tightly bound substrate mimicking ligand, MOPS at the active site pocket. Although being a member of a prokaryotic protein family, VcLMWPTP-1 structure resembles very closely to LMWPTP from a eukaryote, *Entamoeba histolytica*. It also delineates the diverse surface properties around the active site of the enzyme.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Reversible protein phosphorylation is an important event involved in intracellular signal transduction pathways in response to triggering factors which regulate crucial metabolic activities in cell. The net result of incorporation of phosphate groups in proteins is diverse and this probably acts as a key step in cellular regulation [1]. The overall content of phosphotyrosine in cells is reciprocally controlled by protein tyrosine kinases (PTKs) which specifically phosphorylate tyrosines in proteins [2]. An abundant member of this class of proteins is the low molecular weight protein tyrosine phosphatases (LMWPTPs) ubiquitously found in all organisms ranging from prokaryotes to higher eukaryotes. LMWPTPs belong to class II PTPs which act on tyrosine phosphorylated proteins, low molecular weight aryl phosphates and natural and synthetic acyl phosphates [3].

Several structures of LMWPTP from eukaryotic organism such as human – HCPTPA (PDB: 5PNT) [4], bovine – BPTPA (PDB: 1ZI2) [5], mouse – RPTPA (PDB: 2P4U) [6], yeast – LTP1 (PDB: 1D1P) [7] and protozoan parasite *Entamoeba histolytica* – EhPtp (PDB: 3IDO) [8] are available. From the prokaryotic counterpart, structures of LMWPTP from gram-positive euubacteria *Staphylococcus aureus* – PtpA (PDB: 3ROF) [9], gram-negative euubacteria *Thermus thermophilus* – TT1001 (PDB: 2CWD), gram–positive proteobacteria *Bacillus subtilis* – YwIE (PDB: 4ETI) [10], gram–negative proteobacteria *Escherichia coli* – Wzb (PDB: 2WJA) and pathogenic *Mycobacterium tuberculosis* MPtpA (PDB: 1U2Q) [11] are available. It has been reported that the self association of mammalian LMWPTP (viz. *Bos taurus* LMWPTP, BPTP) produces inactive oligomers that are in equilibrium with its active monomers [12]. Among the prokaryotic LMWPTPs weak oligomerization has been found to exist in YwIE from *B. subtilis*, PtpB from the Gram-negative bacterium *Salmonella aureus* and *E. coli* Wzb. However the mode of dimerization is similar in both the cases and takes place through the direct involvements of the active site residues and the tyrosines of the DPY-loop, leading to a catalytically inactive species [13].

Here we report a novel mode of dimerization of VcLMWPTP-1, a 17.9 kDa (155 amino acids) protein from *Vibrio cholerae* O395. Fast protein liquid chromatography and glutaraldehyde crosslinking
reveal the existence of the dimeric species of the protein in solution. Kinetic studies of VcLMWPTP-1 using para-nitrophenyl phosphate (pNPP) as the substrate depicts that the protein is active, suggesting that the mode of dimerization in VcLMWPTP-1 is different from other LMWPTPs. In an attempt to investigate the molecular mechanism of this form of dimerization, we solved a high resolution crystal structure (1.45 Å) of VcLMWPTP-1 which depicts a mode of dimerization markedly different from those reported for other dimeric LMWPTPs. The extensive dimerization interface area of VcLMWPTP-1 is about double than that of other mode of dimerization and active site residues are not seen to be involved in the oligomerization. The active site of each monomer is totally accessible to the substrate which is evident from the crystal structure where the active site of each monomer is occupied by a tightly bound MOPS molecule in a substrate like manner. Comparison of the VcLMWPTP-1 structure and surface properties with similar structures in the PDB illuminates its structural convergence with LMWPTP from a eukaryote, *E. histolytica*.

2. Materials and methods

2.1. Cloning, expression and purification

VcLMWPTP-1 protein was cloned, over expressed and purified as mentioned earlier [14]. Active site mutant of VcLMWPTP-1, C8S was cloned using two-step PCR method and inserted after the start codon of pET24(+). The mutant was overexpressed as C-terminal His-tag and was purified in similar way of the wild type.

2.2. Determination of oligomers using FPLC

VcLMWPTP-1 (0.85 mg/ml, 1.75 mg/ml, 2.58 mg/ml and 4.25 mg/ml) in three different 50 mM Tris buffer, each at pH 7.6 containing 150 mM, 300 mM and 500 mM NaCl was fractionated by a Sephacryl S-100 (Amersham Biosciences) column (46 × 1.6 cm) at 0.9 MPa, pre-equilibrated with respective buffers and precalibrated with a protein mixture containing Lysozyme (MW14.3 kDa) Ovalbumin (MW 36.0 kDa) and Bovine serum albumin (MW 66.45 kDa) at room temperature. Fractions were collected at a flow rate of 0.4 ml per minute using an ÄKTAPrime chromatographic system. The elution profile was determined by monitoring the absorbance at 280 nm.

2.3. Crystalization, data collection and structure solution

Crystallization and data collection of VcLMWPTP-1 was reported earlier [14]. For phasing the 1.45 Å data of the crystal, the coordinates of *E. histolytica* (PDB: 3IDO) [8] was used for refinement with Phaser [15] in CCP4 [16] with *V. cholerae* (solvent content 43.63%), final TFZ = 13.3 and LLG = 157. Model building was done with Coot [17] and refinement was carried out using Phenix refine [18] with twin law -k,-h,-l. TLS refinement was performed during the final stages of refinement [19].

2.4. Structural analysis

Average B-factors for each residue were calculated using B average in CCP4 [20]. PISA webserver [21] and PIC server [22] were used for analysis of the structure and oligomeric state. The oligomeric state and other structural of the protein was analyzed using the. Sequence alignment of VcLMWPTP-1 with other LMWPTPs was done using ClustalO [23,24]. Figures were prepared using Pymol (http://www.pymol.org). The surface electrostatic potential was mapped using Chimera [25][−10 kT/e (red) to +10 kT/e (blue)].

2.5. Enzyme kinetics

Kinetic parameters were calculated for VcLMWPTP-1 using *p*-nitrophenyl phosphate (pNPP) as the substrate as described [26]. Briefly, pNPP at a concentration range of 1–40 mM was treated with 100 μM VcLMWPTP-1 and C8S mutant and quenched with 1 N NaOH after 10 min. The absorbance of the product, *p*-nitrophenol, thus formed is measured at 405 nm. The amount of *p*-nitrophenol was calculated from the standard curve of *p*-nitrophenol. For standard curve, stock solution of *p*-nitrophenol was diluted with 0.05 N NaOH and the absorbance of the samples was measured at 405 nm. To check the effect of temperature on the enzymatic activity samples were incubated at 5 °C interval in the water-bath prior to check the absorbance.

2.6. Glutaraldehyde crosslinking to capture the dimeric state in solution

For crosslinking assays, 2.3% freshly prepared solution of glutaraldehyde was added to a reaction mixture of 100 μl containing about 50 μg of the protein in 50 mM MOPS, pH 7.6, 300 mM NaCl and the reaction carried out at room temperature (25 °C). Samples were collected at 15 s, 30 s, 1 min and then up to 5 min at an interval of 1 min and quenched by addition of 10 μl of 1 M Tris–HCl, pH 8.0. Cross-linked proteins were analyzed through 15% SDS–PAGE.

3. Results and discussion

3.1. VcLMWPTP-1 forms dimer in solution

VcLMWPTP-1 elutes as a single peak in fast protein liquid chromatography (FPLC) (Fig. 1A) and calculations based upon standard curve shows that this species is the dimeric form of VcLMWPTP-1 (Fig. 1B and C). Moreover, presence of the dimeric form is observed irrespective of protein or salt (NaCl) concentration. Further confirmation of the dimeric form was performed through glutaraldehyde crosslinking where a gradual increase in the dimeric band with time could be seen in 15% SDS–PAGE (Fig. 1F).

3.2. Phosphatase activity of VcLMWPTP-1

Enzyme kinetic assays for VcLMWPTP-1 were performed at pH 4.8 and pH 7.6 at 25 °C using *p*-nitrophenyl phosphate as a substrate (Materials and Methods). The determined *Km* value under these conditions was 2.07 ± 0.2 mM and 2.03 ± 0.4 mM respectively (Fig. 1D), which is consistent with the *Km* values reported for other LMWPTPs like *P. aeruginosa* (1.0 mM) and *P. putida* (1.5 mM) [27]. This leads us to propose that the respective active site of the dimer is accessible to the substrate. So the dimer formed in case of VcLMWPTP-1 does not involve active site residues as commonly found in previously reported inactive dimeric species of LMWPTPs and pH variation has no effect on its catalytic activity. The enzyme shows optimum activity at 25 °C and the activity decreases with increasing temperature leading to complete loss of activity at 55 °C (Fig. 1E). Mutating the active site Cys8 to Ser (CBS) results in complete loss of enzymatic activity (Fig. 1E inset) as reported for other LMWPTPs [28].

3.3. Overall structure of VcLMWPTP-1 monomer

VcLMWPTP-1 crystallized in space group P31 using ammonium sulfate (AMS) as precipitant and the structure has been solved at 1.45 Å (Supplementary Table S1). The asymmetric unit contains four molecules of VcLMWPTP-1 and their arrangement suggests that molecules A and C and molecules B and D assemble together
to form two dimers. The high resolution electron density map allowed us to locate 149 residues out of 155 residues of each VcLMWPTP-1 molecule. Each VcLMWPTP-1 binds a MOPS molecule in its active site and one sulfate at the C-terminal loop. The fold and overall secondary structure is comprised of three layer βββ sandwiched architecture with a topology of Rossmann fold and the four stranded parallel β-sheet is sandwiched between three long (α1, α4 and α6) and three short (α2, α3, α5) helices (Fig. 2A).

3.4. Ligand bound ‘closed structure’ conformation

The active site is located at the crevice formed by loop-1 between β1 and α1, also called P-loop. It is surrounded by long, flexible loop-3 and short loop-6 and the DPY-loop, which is a portion of loop-8 preceding the last helix (α6) that contributes residue (D122) necessary for the catalytic action (Fig. 2A). From the well defined electron density it was evident that each VcLMWPTP-1 molecule tightly binds a ligand molecule, MOPS at the active site. Main-chain amide nitrogens of the residues constituting the P-loop and the side-chain amide nitrogens of Arg14 are engaged in the binding and stabilization of the sulfonate group of the ligand (Fig. 2B). A stereo representation of the electron density map around the active site of VcLMWPTP-1 molecule is shown in Supplementary Fig. S1. The piperazine ring of MOPS molecule is surrounded with several polar and aromatic residues (Thr40, Tyr43, His44, Asn47, Asp87, Glu89, Asn90, Asp122, Tyr124, Tyr125 and the disposition of catalytic Cys8 and Asp122 near the leaving group together resembles a substrate bound ‘closed structure’. This closure at the P-loop pocket is also evident from the structural alignment with an ‘open structure’ of apo-MPtpA (PDB: 2LUO) [30] to the ligand-bound structures (PDB: 1U2P, 1U2Q) [11] where the position of DPY-loop is shown to be closer to P-loop in case of the ligand-bound structure than the apo-form [30].

3.5. Multiple sequence alignment showing conserved motifs

Multiple sequence alignment and 3D superposition of LMWPTP structures determined from other organisms show that their active site is highly conserved (Supplementary Fig. S2). Among these structures VcLMWPTP-1 shares highest sequence identity (43%) with that of E. histolytica (PDB: 3JVI) [8] with an average RMSD of superposition 0.48 Å for 122 residues. Structural superposition indicates that the end of the distorted α4 and region preceding the catalytically important DPY-loop exhibits highest variability. Comparison of the structure of VcLMWPTP-1 to PtpA, LMWPTP from gram-positive eubacteria Staphylococcus aureus (PDB: 4ETI) and Wzb, LMWPTP from gram-negative proteobacteria E. coli (PDB: 2WJA) [10] also indicates significant structural difference around loop-3 which is closer to the P-loop. Besides the active site P-loop and DPY-loop, two more conserved sequences (motif) are evident from the alignment result. Details of the four motifs and possible role of the conserved residues are listed in Supplementary Table S2. The active site structure of PTPs follow the active site motif of CX4R, the only common feature between the large PTP family and LMWPTP [31]. Sequence alignment helps to dissect the
X5 portion of the P-loop sequence revealing ‘(T/L/M)GN(I/L)C’ as a consensus motif in this family, contrary to the notion of being tagged as ‘any amino acid’. It is also to be noted that the flanking residue(s) beside the P-loop is hydrophobic in nature. Motif-1 and Motif-4 are parts of P-loop and DPY-loop respectively. Motif-2 containing loop-2 is in close proximity to the active site whereas Motif-3 containing β3 is remotely distant from the active site.

Motif-2 region superposes well in 3D structures between VcLMWPTP-1 with eukaryotic LMWPTPs but a marked distortion with the prokaryotic counterpart is evident, especially with Wzb, YwlE and PtpA.

3.6. Dimerization of VcLMWPTP-1: role of a set of unique residues

Crystal structure of VcLMWPTP-1 shows that each dimer (A:C and B:D) faces an extended surface area that is further stitched by two sulfate ions at the interface region (Fig. 3A). Calculation of buried surface area (BSA) by PISA [21] shows 2910 Å² BSA for the B:D dimer and 2820 Å² BSA for the A:C dimer. The BSA of both the dimers is ~20% of their respective accessible surface area (ASA) and ~10% of their total ASA. PISA server also indicates that the average ΔG_biss value of both the dimers is ~14 kcal/mol -1 indicating they are stable in solution. At the interface region, twelve residues (Met9, Thr40, Ile41, Gly42, Tyr43, Lys71, Glu89, Ala92, Glu93, Arg97, Tyr124, and Tyr125) play pivotal role to stabilize the dimers through numerous hydrophobic and H-bonding interactions (Fig. 3B). Tyr43 and Lys71 from each chain are involved in cation-TT interaction with its dimeric counterpart. Sulfate ions which stitch the two monomers interacts with the main chain NH of Gly42 and the side chain NZ of Lys71 of one monomer and the main chain NH of Tyr43 of the other monomer. In contrast to the buried surface area (BSA) of VcLMWPTP-1 (average BSA 2865 Å²) the dimerization region of BPTP encompasses a much smaller BSA (1589 Å²) that consist only the residues from the active site and DPY-loop (Fig. 3C and D).

Analysis of the position of Lys residues at the dimeric interface of the VcLMWPTP-1 (Fig. 3E) shows only a pair of amine (Lys68 from each monomer at a distance of ~10 Å) which could crosslink with each other in presence of glutaraldehyde (Fig. 3E). Distance between all other inter-monomeric Lys residues and the N terminal amine are beyond the condensable distance. As this mode of dimerization is stabilized by hydrophobic and polar interactions, it would show a monomeric band in SDS–PAGE. However, if the dimer is crosslinked by glutaraldehyde a band near its dimer appears with time. Since the distance between Lys68 in the dimer is little larger than the condensable distance (6–8 Å), more dimers become crosslinked with gradual increase in time (Fig. 1F).

3.7. Insight into substrate specificity – diverged surface property around the active site

CASTp server [32] shows 256 Å³ as the volume of the active site cavity. Comparison of the structure and electrostatic surface around the active site of VcLMWPTP-1 with others reveals...
close resemblance with EhPtp and other mammalian LMWPTPs than with LMWPTPs of bacterial origin (Supplementary Fig. S3). VcLMWPTP-1 differs with prokaryotic LMWPTPs at two regions which are implicated in substrate recognition and binding. The W-loop, harboring a Trp-residue almost in every LMWPTP, is proposed to play an important role in substrate recognition [10]. However, corresponding residue is Tyr43 in VcLMWPTP-1 and Tyr49 in HCPTPA, important for substrate recognition for the latter. But in bacteria, Tyr43 is substituted by Leu44 in Wzb and Ser42 in YwlE. Asn53 in HCPTPA plays a crucial role in determining the substrate specificity [3], corresponding residue Asn47 in VcLMWPTP-1 may play a similar role. Again this is in sharp contrast with bacterial LMWPTPs like Wzb, Ptp where the corresponding residue K43 and Y44 respectively play indispensable role in substrate recognition [10]. Overall, these differences from prokaryotic LMWPTPs make VcLMWPTP-1 active site more like eukaryotic than of prokaryotic origin which may dictate its substrate specificity.

In conclusion, we report an exclusive dimeric species of VcLMWPTP-1 in solution through FPLC while atomic resolution X-ray structure shows that the dimeric species is 'novel' where the active site is not occluded. This dimer is catalytically active as demonstrated by kinetic studies while a MOPS molecule bound in a substrate like manner at the active site proves the active pocket is accessible for the substrate. VcLMWPTP-1 also distinguishes itself
from other bacterial LMWPTPs in terms of electrostatic charge distribution around the active site thereby suggesting close similarity in substrate(s) preference to eukaryotic LMWPTPs, than its bacterial homologue. Since the dimeric surface is closer to the active site it may further modulate the substrate recognition and specificity.

Acknowledgments

The laboratory of US is partly supported by the MSACR project, DAE, Govt. of India, SINP. RB is a Senior Research Fellow of the Council of Scientific and Industrial Research, India. SN is a Senior Research Fellow of the Department of Atomic Energy, India. US thank Dr. J. Dasgupta for her help during FPLC experiments.

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.bbrc.2014.05.129.

References

A Novel 8-nm Protein Cage Formed by Vibrio cholerae Acylphosphatase

Seema Nath, Ramanuj Banerjee and Udayaditya Sen

Crystallography and Molecular Biology Division, Saha Institute of Nuclear Physics, 1/AF, Bidhannagar, Kolkata 700064, India

Correspondence to:
http://dx.doi.org/10.1016/j.jmb.2013.09.014
Edited by T. Yeates

Abstract

Here we show the formation of an ~8-nm cage formed by the self-assembly of acylphosphatase from Vibrio cholerae O395 (Vc-AcP). The 12-subunit cage structure forms spontaneously and is stabilized through binding of sulfate ions at its exterior face and interfacial regions. Crystal structure and studies in solutions illuminate the basis for the formation of the cage, while a single (Cys20 → Arg) mutation (Vc-AcP-C20R) transforms Vc-AcP to a potent enzyme but disrupts the assembly into a trimer.

© 2013 Elsevier Ltd. All rights reserved.

Proteins can self-assemble in a highly specific and constrained way to form diverse architectures, from tubes to cages. Such architectures provide a suitable template for manipulation to create materials with a wide range of applications from biomedicine to material science [1–4]. Acylphosphatases (AcPs) are members of the hydrolase family, possessing ferredoxin-like folds (Fig. 1a), and are among the smallest enzymes known (molecular mass, ~10 kDa). Oligomerization of AcPs to form large aggregates is a common phenomenon, which has been reported for various sources [5,6]. Vc-AcP, a natural variant of Vibrio cholerae O395, shows nominal enzymatic activity due to the presence of a cysteine in place of catalytic arginine 20. A kinetic assay of Vc-AcP using acetylphosphate as the substrate has established $K_m$ and $K_{cat}$ values of $8.53 \pm 0.7$ mM and $9.29 \pm 0.3 \times 10^{-5}$ min$^{-1}$, respectively (Fig. 1b). A point mutant, Vc-AcP-C20R, however, turns out to be a potent enzyme with $K_m$ and $K_{cat}$ values of $2.36 \pm 0.6$ mM and $2.55 \pm 0.5$ min$^{-1}$, respectively (Fig. 1b), indicating that the mutant is nearly $10^5$ times more efficient in catalysis than the wild-type protein.

In our solution studies, Vc-AcP shows three different oligomeric states in solution, irrespective of the pH. The hydrodynamic radius (RH) and expected molecular mass (MM) of the oligomers vary from 2.5 ± 0.3 nm and 32 ± 8 kDa to 3.7 ± 0.3 nm and 68 ± 12 kDa, to 4.6 ± 0.3 nm and 119 ± 13 kDa (Fig. 1c), whereas Vc-AcP-C20R shows only one population with the RH and MM equivalent to that of the smallest oligomer of Vc-AcP (Fig. 1d). The crystal structure of the Vc-AcP-C20R mutant (1.96 Å) shows a triangular trimeric assembly (Fig. 1e, Supplementary Table 1), the size and mass of which matches well with that of the oligomeric state obtained in DLS (dynamic light scattering) experiments (Fig. 1d). The helix-mediated trimeric assembly (HT) is stabilized by intricate hydrogen bonding network through a set of residues (Y21, H22, H25, L28, and K29 from helix α1 of each monomer) that are specific for Vc-AcP compared to other AcPs.

Being evidently heterogeneous with respect to assembly state, wild-type Vc-AcP was not suitable for crystallization. Extensive pH and additive screening identified that tetrahedral anions such as sulfate, vanadate, phosphate, tungstate, and arsenate (~0.8 M) are capable of shifting the multimodal population to a monomodal state suitable for crystallization (Fig. 1f). This oligomeric state corresponds to the highest MM and RH. We determined the crystal structure of Vc-AcP (3.1 Å) revealing 12 monomers in the asymmetric unit, which are arranged to form a hollow cage-like structure obeying tetrahedral ($T_3^2$) cubic symmetry and having outer and inner diameters of ~8 nm and ~4.5 nm, respectively (Fig. 1h, Supplementary Table 1). This cage-like self-assembly is also supported by the program PISA [7], which reports a buried surface area of 26,040 Å$^2$ (40% of the total surface area) and an estimated standard state free energy of dissociation $\Delta G_{diss}$ of 105.5 kcal/mol. A total of 43 sulfate ions could be located at the junction points between monomers (Fig. 1h). Formation of cage-like
Fig. 1. (a) Cartoon representation of the AcP monomer showing its ferredoxin-like fold. The monomer is colored based on B-factor (from blue to red) with blue being the least flexible and red being the most flexible. (b) Hydrolysis of acetylphosphate by wild-type Vc-AcP and the single site mutant Vc-AcP-C20R. (c) DLS profile of Vc-AcP showing the population of three different oligomeric states with increasing molecular mass (MM, kDa) and hydrodynamic radius (R, nm). The data were taken without any additives. (d) DLS profile of the Vc-AcP-C20R mutant showing a monomodal population with hydrodynamic radius RH (2.5 ± 0.3 nm). (e) Crystal structure of the Vc-AcP-C20R trimeric assembly. Vc-AcP-C20R chains (yellow, green, and cadbury) are shown in cartoon, the central trimerization region is shown in sticks, and the central sulfate and active-site molybdate ions are shown in a ball-and-stick representation. Zoomed view of the triangular helix-mediated trimeric assembly depicts that the corners of the triangle are occupied by three molybdate ions (molybdenum in green and oxygen in red color), and the corners are clamped by the hydrogen bonding between the main-chain carbonyls of K29 and the side chains of T68. (f) Percent abundance of three different oligomeric states (trimer, hexamer, and dodecamers) of Vc-AcP in the presence of increasing concentration of ammonium sulfate. At 0.8 M ammonium sulfate, the solution turns into a monomodal population containing only dodecameric Vc-AcP. (g) Transmission electron microscopy images showing the formation of the assembly in solution. (h) Tetrahedral (shown in red) arrangement of Vc-AcP monomers with four 3-fold rotation axes (blue lines); three 2-fold axes (not shown in the figure) pass through the middle of two opposite edges (red line). (i–k) Three different junction points “HT,” “ST,” and “LT” of in the Vc-AcP protein cage, where “HT” is the Vc-AcP-C20R-like helix-mediated trimeric assembly, “ST” is formed by three monomers that associate through their beta-sheets, and “LT” is the junction point of the monomers through their loop 6 and loop 1. A gray sphere is drawn inside to emphasize the hollow interior. (l) Vc-AcP-C20R (yellow) is superposed on one of the monomers (magenta) of the Vc-AcP cage to show how R20 (shown in stick) clashes with H83 (shown in line) of other monomer and hinders the cage formation (the sulfate ion is shown in sticks).
assembly was also confirmed through transmission electron microscopy, where circular dotted structures with an average diameter of 7–9 nm could be observed (Fig. 1g). In the cage-like structure, 12 monomers are assembled in the form of four trimers, where each trimer resembles the helix-mediated Vc-AcP-C20R assembly (HT). The trimers occupy the corners of a tetrahedron and are so arranged that, at each face of a tetrahedron, three monomers, from three different trimers, interact with each other to form a beta-sheet mediated trimer (ST), whereas along each axis of the tetrahedron, they form a loop-mediated tetrameric junction (LT) (Fig. 1i–k).

It is an intriguing observation that a single point mutation converts the marginally active enzyme to a potent form, while simultaneously disrupting the cage assembly. The crystal structure provides some insight. Loop 6 is the most flexible portion of Vc-AcP. It interacts with loop 1 at the LT interface. Superposition of the Vc-AcP-C20R trimer on the Vc-AcP cage shows that, upon sulfate binding, the arginine 20 side chain of the mutant comes in close contact with H83 of the neighboring monomer, thereby destroying the favorable interaction at the LT interface and hindering the cage formation (Fig. 1l). Moreover, Q15 in loop 1 of Vc-AcP-C20R clashes with E82 of loop 6 of another monomer, but loop 6 of Vc-AcP takes an alternative route to escape this clash (Fig. 1l). These interactions suggest why cage formation is incompatible with the catalytically active conformation.

It remains to be proven whether this cage represents a true biological form of the protein, and if so, what cellular function or evolutionary advantage it might provide. The very low enzymatic activity adds an interesting piece to the puzzle. Beyond its putative biological function, the Vc-AcP particle presents a novel architecture whose hollow interior might be exploited as a carrier of drugs or other small molecules. Its structure could also be modified to produce designer cages with altered properties.

Protein Data Bank accession codes

Coordinates and structure factor files have been deposited with the accession codes 4HI1 and 4HI2.

Acknowledgements

The laboratory of U.S. is partly supported by the SPGHGD project, Department of Atomic Energy, Government of India, Saha Institute of Nuclear Physics. R.B. is a Senior Research Fellow of the Council of Scientific and Industrial Research, India. S.N. is a Senior Research Fellow of the Department of Atomic Energy, India. U.S. thanks Pulak Ray of Saha Institute of Nuclear Physics for help in electron microscopy.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jmb.2013.09.014.

Received 21 August 2013; Received in revised form 3 September 2013; Accepted 7 September 2013

Available online 17 September 2013

**Keywords:**
bacterial protein phosphatases;
dynamic light scattering;
protein crystallization;
protein self-assembly;
X-ray crystallography

**Abbreviation**
LT, loop mediated tetramer.

**References**


Cloning, purification, crystallization and preliminary X-ray analysis of two low-molecular-weight protein tyrosine phosphatases from *Vibrio cholerae*

Seema Nath, Ramanuj Banerjee, Susmita Khamrui and Udayaditya Sen

Cloning, purification, crystallization and preliminary X-ray analysis of two low-molecular-weight protein tyrosine phosphatases from *Vibrio cholerae*

Low-molecular-weight protein tyrosine phosphatases (LMWPTPs) are small cytoplasmic enzymes of molecular weight ~18 kDa that belong to the large family of protein tyrosine phosphatases (PTPs). Despite their wide distribution in both prokaryotes and eukaryotes, their exact biological role in bacterial systems is not yet clear. Two low-molecular-weight protein tyrosine phosphatases (*VcLMWPTP-1* and *VcLMWPTP-2*) from the Gram-negative bacterium *Vibrio cholerae* have been cloned, overexpressed, purified by Ni$^{2+}$–NTA affinity chromatography followed by gel filtration and used for crystallization. Crystals of *VcLMWPTP-1* were grown in the presence of ammonium sulfate and glycerol and diffracted to a resolution of 1.6 Å. *VcLMWPTP-2* crystals were grown in PEG 4000 and diffracted to a resolution of 2.7 Å. Analysis of the diffraction data showed that the *VcLMWPTP-1* crystals had symmetry consistent with space group *P*$_3$1 and that the *VcLMWPTP-2* crystals had the symmetry of space group *C*$_2$. Assuming the presence of four molecules in the asymmetric unit, the Matthews coefficient for the *VcLMWPTP-1* crystals was estimated to be 1.97 Å$^3$ Da$^{-1}$, corresponding to a solvent content of 37.4%. The corresponding values for the *VcLMWPTP-2* crystals, assuming the presence of two molecules in the asymmetric unit, were 2.77 Å$^3$ Da$^{-1}$ and 55.62%, respectively.

1. Introduction

Protein phosphorylation and dephosphorylation are involved in the regulation of many cellular processes such as cell growth, differentiation and metabolism (DeVinney *et al.*, 2000; Hunter, 1995; Mustelin *et al.*, 2005). The formation and hydrolysis of phosphate esters in proteins play key roles in signal transduction, through which external environmental stimuli are converted into internal cellular responses (Mustelin *et al.*, 2005; Neel & Tonks, 1997). Protein kinases and protein phosphatases are among many enzymes that catalyze such reversible reactions in a highly precise manner to control cellular activities. Defective or incorrect regulation of such systems result in loss of cell viability; thus, these proteins are ideal targets for drug design (Zhang, 2001; Hunter, 2000).

Protein tyrosine phosphatases (PTPs) belong to the protein phosphatase superfamily and catalyze the hydrolysis of phosphate esters on tyrosine residues in proteins. Based on their activity towards different phosphorylated amino acids, PTPs can be divided into two families: one class is exclusively active towards phosphorylated Tyr residues and the other acts on both phosphorylated Ser/Thr and Tyr residues in proteins. Low-molecular-weight PTPs fall into the first category (Ramponi & Stefani, 1997). The members of this family of proteins share very low sequence identities amongst themselves apart from the signature motif in the relatively flexible loop at the active site, $C_X_4CR$, where $X$ can be any amino acid. This loop is responsible for binding and hydrolyzing phosphorylated tyrosine residues and is thus known as the protein tyrosine phosphate-binding loop or P-loop (Zhang *et al.*, 1995).
Low-molecular-weight protein tyrosine phosphatases (LMWPTPs) are widely distributed in prokaryotes and eukaryotes (Kennelly & Potts, 1999; Cozzone et al., 2004) and play important roles in many biological processes. The reaction mechanism of eukaryotic LMWPTPs has been structurally, thermodynamically and kinetically characterized (Ramponi & Stefani, 1997). However, LMWPTPs from bacterial sources have been less explored in terms of their structure and function. Vibrio cholerae contains two LMWPTPs (accession Nos. A5F2Q3 and A5F307, hereafter termed VcLMWPTP-1 and VcLMWPTP-2, respectively). They have weak sequence identity (~30%), although their active-site signature motif CXGNXCR(S)P and the DPY loop, which play key roles in the hydrolysis of phosphorylated tyrosine, are conserved. However, the amino-acid residues that are located around the P-loop and the DPY loop differ significantly. It is believed that the residues around the P-loop are responsible for modulating substrate recognition, while the residues around the DPY loop are important in catalyzing the dephosphorylation reaction. Therefore, it seems that the physiological targets of these two LMWPTPs are quite different and that these two enzymes should have different dephosphorylation mechanisms. The primary sequences of these two LMWPTPs, especially VcLMWPTP-2, do not produce significant matches with other LMWPTP structures reported in the PDB, suggesting that the three-dimensional structures of these two PTPs may show distinct features that have not been observed in other LMWPTP structures reported to date. Furthermore, no literature is available on the substrate recognition, catalytic mechanism and kinetic parameters of these two PTPs. Importantly, the physiological substrates of these two LMWPTPs are not known and identification of these will definitely shed light on the functional roles of these PTPs in V. cholerae. The three-dimensional structures of these two proteins will also be useful to obtain insights into their catalytic functions at the atomic level. Moreover, knowledge of the structure and function of these phosphatases might be of use in drug design against this bacterial pathogen. Here, we report the cloning, overexpression, purification, crystallization and preliminary structural analysis of VcLMWPTP-1 and VcLMWPTP-2 at resolutions of 1.6 and 2.67 Å, respectively.

2. Materials and methods

2.1. Cloning and expression

The genes encoding VcLMWPTP-1 (155 amino acids) and VcLMWPTP-2 (166 amino acids) (accession Nos. A5F2Q3 and A5F307, respectively) were amplified from V. cholerae 0395 genomic DNA using the polymerase chain reaction with the following primers: for VcLMWPTP-1, forward primer 5’-CCAGCATATGCAGAGGTCATCTGGTGTTGTCG-3’ and reverse primer 5’-CAGGGATCCCTTAATCTGCTGGCCCTTGTTTTAG-3’; for VcLMWPTP-2, forward primer 5’-GGGCGCATATGGAAGTTAAGGGTTATCAG-3’ and reverse primer 5’-GGGGATCCCTTAATCTGAGATATAAATTTCGTTGCACGC-3’. The primers were synthesized (NeuProCell) with adaptor sites (shown in italics) and restriction-enzyme markers (labelled in kDa); lane 3, VcLMWPTP-2.

Figure 1

The homogeneity of the purified VcLMWPTP-1 and VcLMWPTP-2 proteins was checked by 12% SDS–PAGE. Lane 1, VcLMWPTP-1; lane 2, molecular-mass markers (labelled in kDa); lane 3, VcLMWPTP-2.

Figure 2

(a) Crystals of VcLMWPTP-1 grown in the presence of ammonium sulfate pH 5 at 293 K. The maximum dimensions of the crystals were 0.4 × 0.4 × 0.3 mm. (b) Crystals of VcLMWPTP-2 grown at 293 K (0.4 × 0.3 × 0.2 mm) appeared when 5% PEG 6000, 8% MPD pH 5.0 was used as a precipitant.
2.2. Purification

The cells were harvested by centrifugation at 4000 g for 20 min at 277 K. The cell pellet was resuspended in lysis buffer (buffer A; 50 mM HEPES pH 7.0, 300 mM NaCl, 1 mM phenylmethylsulfonyl fluoride and lysozyme) and the cells were disrupted by sonication. The crude lysate was centrifuged at 12,000 g for 40 min at 277 K. The supernatant was loaded onto an Ni<sup>2+</sup>–NTA column previously equilibrated with buffer A and subsequently washed with buffer A containing 5 and 10 mM imidazole. The protein was eluted using a gradient to 150 mM imidazole in buffer A. The 6×His tag was cleaved using restriction-grade thrombin (Novagen) and final purification of the protein from contaminating proteins, thrombin and cleaved 6×His tag was achieved by gel filtration using an S-100 (GE Healthcare) column pre-equilibrated with buffer B (50 mM HEPES pH 7.0, 300 mM NaCl, 0.5 mM DTT). The proteins thus purified were used for crystallization. The homogeneity of the purified protein was determined by SDS–PAGE using 15% (v/v) polyacrylamide gel (Fig. 1). The concentrations of both proteins were determined using the Bradford assay.

2.3. Crystallization of VcLMWPTP-1 and VcLMWPTP-2

For crystallization, thrombin-cleaved VcLMWPTP-1 (in a buffer consisting of 50 mM MOPS pH 7.6, 300 mM NaCl) and VcLMWPTP-2 (in a buffer consisting of 50 mM HEPES pH 7.0, 300 mM NaCl, 0.5 mM DTT) were concentrated to 6 mg ml<sup>−1</sup> using an Amicon ultra centrifugation unit (molecular-weight cutoff 10,000). Crystallization was performed by the hanging-drop vapour-diffusion method in 24-well crystallization trays (Hampton Research, Laguna Niguel, California, USA). Grid Screen Ammonium Sulfate, Grid Screen PEG 6000, Crystal Screen and Crystal Screen 2 from Hampton Research (Jancarik & Kim, 1991) were used to explore the initial crystallization conditions. 2 μl protein solution was mixed with 2 μl precipitant solution, inverted over a reservoir containing 600 μl precipitant solution and maintained at both 277 and 293 K. VcLMWPTP-1 crystallized in 2.4 M ammonium sulfate, 0.1 M citric acid pH 5.0, 2% glycerol at 277 K (Fig. 2a) and VcLMWPTP-2 crystallized in 0.2 M ammonium sulfate, 30% (w/v) PEG 8K at 293 K (Fig. 2b).

2.4. Data collection and processing

Crystals of VcLMWPTP-1 and VcLMWPTP-2 were looped out from the crystallization drops using a 20 μm nylon loop and flash-cooled in a stream of nitrogen (Oxford Cryosystems) at 100 K. A diffraction data set was collected on an in-house MAR Research image-plate detector of diameter 345 mm using Cu Kα radiation generated by a Bruker–Nonius FR591 rotating-anode generator equipped with Osmic MaxFlux confocal optics and operated at 50 kV.

![Figure 3](a) X-ray diffraction image of a VcLMWPTP-1 crystal; the edge of the detector corresponds to a resolution of 1.6 Å. (b) X-ray diffraction image of a VcLMWPTP-2 crystal, which diffracted to a resolution of 2.67 Å.

Table 1

<table>
<thead>
<tr>
<th>Space group</th>
<th>VcLMWPTP-1</th>
<th>VcLMWPTP-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit-cell parameters (Å&lt;sup&gt;3&lt;/sup&gt;)</td>
<td>P&lt;sub&gt;3&lt;/sub&gt;&lt;sup&gt;1&lt;/sup&gt;</td>
<td>C2</td>
</tr>
<tr>
<td>a = b = 87.47, c = 73.85</td>
<td>a = 121.38, b = 45.25, c = 88.56</td>
<td>α = γ = 90.0, β = 121.08</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>1.6</td>
<td>2.67</td>
</tr>
<tr>
<td>Molecules per asymmetric unit</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Matthews coefficient V&lt;sub&gt;M&lt;/sub&gt; (Å&lt;sup&gt;3&lt;/sup&gt;Da&lt;sup&gt;−1&lt;/sup&gt;)</td>
<td>2.26</td>
<td>2.77</td>
</tr>
<tr>
<td>Solvent content (%)</td>
<td>46</td>
<td>56</td>
</tr>
<tr>
<td>Total No. of reflections</td>
<td>158204 (22198)</td>
<td>36313 (4216)</td>
</tr>
<tr>
<td>No. of unique reflections</td>
<td>82757 (12178)</td>
<td>10945 (1286)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>95.2 (91.3)</td>
<td>93.1 (74.9)</td>
</tr>
<tr>
<td>R&lt;sub&gt;merge†&lt;/sub&gt; (%)</td>
<td>0.047 (0.268)</td>
<td>0.043 (0.183)</td>
</tr>
<tr>
<td></td>
<td>12.1 (3.0)</td>
<td>19.6 (6.1)</td>
</tr>
</tbody>
</table>

† R<sub>merge</sub> = Σ<sub>h,k,l</sub> [Σ<sub>i</sub> I(hkl) – (I(hkl))/Σ<sub>i</sub> I(hkl)], where I(hkl) is the observed intensity of the hkl measurement of reflection hkl and (I(hkl)) is the mean intensity of reflection hkl calculated after scaling.
and 65 mA. X-ray diffraction data were collected to a resolution of 1.6 Å from VcLMWPTP-1 crystals (Fig. 3a) and to a resolution of 2.67 Å from VcLMWPTP-2 crystals (Fig. 3b). Data were processed and scaled using iMOSFLM (Battye et al., 2011). Data-collection and processing statistics are given in Table 1.

3. Results and discussion

Both VcLMWPTPs were successfully purified for crystallization and biochemical assays. They share very low amino-acid sequence identity with other bacterial LMWPTPs of known structure. The sequence identity between VcLMWPTP-1 and VcLMWPTP-2 is also low (30%) and their sequences are distinctly different around the P-loop and the DPY loop, implying that these two PTPs are not just two redundant versions of the same protein and that they probably target different physiological substrates for catalysis. Therefore, we crystallized both of these proteins and collected diffraction data. The VcLMWPTP-1 crystals diffracted to 1.6 Å resolution and produced excellent-quality diffraction data (Table 1) with symmetry consistent with space group P31. Packing considerations based on the molecular mass of 17 kDa indicated the presence of four molecules in the asymmetric unit, corresponding to a Matthews coefficient \( V_M \) (Matthews, 1968) of 2.26 Å³ Da⁻¹ and a solvent content of 46%. The VcLMWPTP-2 crystals only diffracted to a resolution of 2.67 Å with moderate data quality (Table 1) and with C2 space-group symmetry. Packing considerations indicated a Matthews coefficient of 2.77 Å³ Da⁻¹, which corresponds to a solvent content of 56% considering two molecules of VcLMWPTP-2 (molecular mass of ~18 kDa) in the asymmetric unit.

A BLAST (Altschul et al., 1990) search for a homologous structure showed that the amino-acid sequence of VcLMWPTP-1 possesses the highest identity (43%) to that of protein tyrosine phosphatase from Entamoeba histolytica (PDB entry 3ido; Seattle Structural Genomics Center for Infectious Disease, unpublished work) followed by human low-molecular-weight phosphotyrosyl phosphatase (40% identity; PDB entry 5pnt; Zhang et al., 1998). VcLMWPTP-2 has the highest identity (29%) to protein tyrosine phosphatase from E. histolytica (PDB entry 3ido). Although the structure of E. histolytica protein tyrosine phosphatase (PDB entry 3ido) showed marginally better identity than human LMWPTP (PDB entry 5pnt), human LMWPTP gave slightly better results during molecular-replacement calculations for VcLMWPTP-1. Before proceeding with molecular-replacement calculations, waters and a long loop preceding the DPY loop were deleted from the coordinates and mismatched residues were truncated to Ala. Using this truncated model, Phaser (McCoy et al., 2007) placed four molecules in the asymmetric unit with RFZ = 3.8, TFZ = 10.3 and LLG = 247 (Table 2). This model was then subjected to several cycles of rigid-body refinement, which removed the clashes between the molecules; the orientations of the four molecules in the asymmetric unit after rigid-body refinement are shown in Fig. 4(a). An electron-density map calculated using the molecular-replacement solution thus obtained showed continuous electron density, and refinement of the structure is in progress to obtain the correct structure. The packing of the VcLMWPTP-1 molecules in the crystal clearly indicates threefold symmetry and large solvent channels (Fig. 4b).

Initial molecular-replacement trials to solve the structure of VcLMWPTP-2 did not produce any clear-cut solution. This might be because of the low sequence identity (29%) of the search models used.

Table 2

<table>
<thead>
<tr>
<th>Improvement of Phaser statistics.</th>
<th>RFZ</th>
<th>TFZ</th>
<th>PAK</th>
<th>LLG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Molecule 1</td>
<td>3.1</td>
<td>5.0</td>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>Molecule 2</td>
<td>3.5</td>
<td>8.4</td>
<td>1</td>
<td>85</td>
</tr>
<tr>
<td>Molecule 3</td>
<td>3.6</td>
<td>10.9</td>
<td>3</td>
<td>159</td>
</tr>
<tr>
<td>Molecule 4</td>
<td>3.8</td>
<td>10.3</td>
<td>8</td>
<td>247</td>
</tr>
</tbody>
</table>

Figure 4

(a) Arrangement of four VcLMWPTP-1 molecules in the asymmetric unit showing no intermolecular clashes. The four molecules are shown in four different colours. (b) Arrangement of the VcLMWPTP-1 molecules in the crystal showing the threefold symmetry and the large solvent channels.
for molecular replacement. The moderate quality of the diffraction data of VcLMWPTP-2 (Table 1) might also be a reason why a clear solution was not obtained, especially when using models with low sequence identity. At present, we are trying to obtain the phases of VcLMWPTP-2 experimentally.

The laboratory of US is partly supported by the SPGHGD project, DAE, Government of India, SINP. RB is a senior research fellow of Council of Scientific and Industrial Research. US thanks Abhijit Bhattacharya of SINP for help in data collection.

References


