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

Cloning, overexpression, purification, crystallization and preliminary X-ray diffraction analysis of an inositol monophosphatase family protein (SAS2203) from *Staphylococcus aureus* MSSA476

Sudipta Bhattacharyya, Debajyoti Dutta, Ananta Kumar Ghosh and Amit Kumar Das*

Department of Biotechnology, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721 302, India

Correspondence e-mail: amitk@hijli.iitkgp.ernet.in

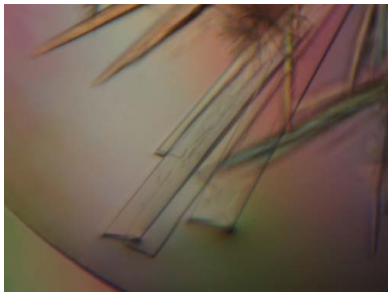
Received 7 December 2010
Accepted 26 January 2011

The gene product of the *sas*2203 ORF of *Staphylococcus aureus* MSSA476 encodes a 30 kDa molecular-weight protein with a high sequence resemblance (29% identity) to tetrameric inositol monophosphatase from *Thermotoga maritima*. The protein was cloned, expressed, purified to homogeneity and crystallized. Crystals appeared in several conditions and good diffraction-quality crystals were obtained from 0.2 *M* Li_2SO_4, 20% PEG 3350, 0.1 *M* HEPES pH 7.0 using the sitting-drop vapour-diffusion method. A complete diffraction data set was collected to 2.6 Å resolution using a Rigaku MicroMax-007 HF Cu $K\alpha$ X-ray generator and a Rigaku R-AXIS IV^{++} detector. The diffraction data were consistent with the orthorhombic space group $P2_12_12_1$, with unit-cell parameters $a = 49.98$, $b = 68.35$, $c = 143.79$ Å, $\alpha = \beta = \gamma = 90^\circ$, and the crystal contained two molecules in the asymmetric unit.

1. Introduction

Inositol monophosphatases are ubiquitous proteins with molecular mass $\simeq$ 30 kDa which are abundant in bacteria, unicellular eukaryotes and plant and animal cells. This group of enzymes belongs to the FIG superfamily of proteins, which also includes fructose-1,6-bisphosphatases (FBPases; both the major and the *glp*X-encoded variants), inositol monophosphatases (IMPases) and inositol polyphosphatases (IPPases). The IMPase group of proteins have two well conserved motifs: motifs A, W-$x_{0,1}$(IV)-D-P-(IL)-D-x-T-x_2-(FYI)-x-(HK), and B, W-D-x_2-(AG)-(AG)-x-(AIL)-(ILV)-(ALV)-x_3-G-(AG) (Neuwald *et al.*, 1991). Most of the members of the IMPase family of proteins exhibit Mg^{2+}-activated Li^+-inhibited phosphatase activity towards many phosphorylated compounds ranging from inositol monophosphate to sugar and nucleotide phosphates. The degree of Li^+ inhibition as well as the substrate specificity of this group of enzymes varies greatly depending on the source organism and the molecular basis of this variation is still a matter of controversy. However, in accordance with recent structural studies, the Li^+-sensitivity of IMPases can be related to the length and the flexibility of the active-site mobile loop containing the residues binding the third Mg^{2+} ion (Johnson *et al.*, 2001; Stieglitz *et al.*, 2002; Li *et al.*, 2010). In IMPases, the participation of the third Mg^{2+} ion in phosphoester hydrolysis is crucial as it can activate the water nucleophile in order to initiate the reaction (Li *et al.*, 2010). Li^+, which has almost the same ionic radius as Mg^{2+}, can effectively replace the latter, but owing to the difference in their preferred coordination number (four *versus* six) and the geometries of their coordination complexes (tetrahedral *versus* octahedral) Li^+ cannot place the water nucleophile in the proper position and hence the phosphatase activity is severely impaired.

IMPases have long been identified as one of the key enzymes involved in the phosphatidyl inositol signalling cascade in eukaryotic systems. The role of this enzyme in human beings is well documented as it plays a pivotal role in manic depressive disorder and has been found to be the primary target of Li^+-based therapies (Berridge *et al.*, 1989). However, the role of IMPase in prokaryotes is not precisely known as inositol and its derivatives are not as abundant in



prokaryotes. Exceptions are the cases of mycobacteria, in which inositol-based phospholipids are important immunomodulatory cell-wall components (Nigou *et al.*, 2002), and high-temperature-dwelling archaea, in which DIP (di-*myo*-inositol-1,1′-phosphate) acts as a compatible solute that is required for cellular osmoregulation at high temperatures (Chen & Roberts, 1998). Despite the absence of a phosphatidyl inositol signalling system and also of other inositol derivatives, it is a most intriguing fact that most eubacteria possess one or more than one copy of IMPase homologues, suggesting an alternative role for this enzyme in bacterial life. Interestingly, the genome sequences of all pathogenic *Staphylococcus aureus* strains show the presence of two conserved ORFs with high sequence similarity to other well known IMPases. It is highly unlikely that a eubacterium such as *S. aureus* would have such an activity, suggesting an alternate role for these enzymes in this organism. Therefore, we have focused our attention on structural and mechanistic studies of these enzymes in order to decipher their role in *S. aureus*. In the present study, the putative identification of the *sas*2203 gene product from *S. aureus* MSSA476 as an IMPase-family protein and its high sequence similarity to IMPase from the high-temperature-dwelling bacterium *Thermotoga maritima* prompted us to clone and express the gene product in *Escherichia coli*. The protein was subsequently purified and crystallized and a complete set of diffraction data was collected. The present work reports the cloning, overexpression, purification, crystallization and preliminary X-ray diffraction studies of SaIMP-1 (SAS2203) from *S. aureus* MSSA476.

2. Materials and methods

2.1. Cloning

The *sas*2203 ORF was amplified by PCR using *S. aureus* MSSA476 genomic DNA as the template with the primer pair 5′-ATCTCTAGA**CTGCAG**CTATCTGTATTTTTG-3′ (forward primer containing a *Pst*I recognition site) and 5′-GAAGTGGTG**GAGCTC**ATGACAGATAAAAC-3′ (reverse primer containing a *Sac*I recognition site). The PCR-amplified DNA was then double digested with *Pst*I and *Sac*I and cloned into the corresponding sites of the expression vector pQE30 (Qiagen, USA) in frame with a sequence encoding six consecutive histidine residues at the N-terminus. The recombinant plasmid was then transformed into *E. coli* M15 (pREP4) cells and subsequently selected on ampicillin/kanamycin plates. The positive clones were verified by DNA sequencing.

Table 1
Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

Wavelength (Å)	1.5418
Space group	$P2_12_12_1$
Unit-cell parameters (Å, °)	a = 49.98, b = 68.35, c = 143.79, $\alpha = \beta = \gamma = 90$
Resolution (Å)	19.67–2.6 (2.7–2.6)
Total No. of observations	108947 (14352)
No. of unique reflections	15728 (2219)
Completeness (%)	99.2 (97.5)
Multiplicity	6.9 (6.5)
Average $I/\sigma(I)$	11.1 (3.2)
$R_{\rm merge}$† (%)	0.13 (0.57)
Monomers per asymmetric unit (Z)	2
Solvent content (%)	37.94

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity of reflection hkl and $\langle I(hkl)\rangle$ is the mean intensity of reflection hkl.

2.2. Overexpression and purification

The positive clone harbouring the desired construct of *sas*2203 was grown in 4 l Luria broth supplemented with 100 µg ml^{-1} ampicillin and 25 µg ml^{-1} kanamycin at 310 K until the A_{600} reached 0.6, induced with 100 µ*M* IPTG and grown for a further 4 h at 310 K in order to obtain the maximum protein yield. Cells were then harvested by centrifugation at 10 000*g* for 10 min and the cell pellet was resuspended in buffer *A* (10 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl and 10 m*M* imidazole) containing 0.1 m*M* each of leupeptin, pepstatin and aprotinin and 0.02 m*M* phenylmethylsulfonyl fluoride (PMSF). The suspension was lysed by ultrasonication on ice and the lysate was centrifuged at 22 000*g* for 40 min. The supernatant was loaded onto Ni-Sepharose High Performance affinity matrix (GE Healthcare Biosciences) pre-equilibrated with buffer *A*. The column was then washed extensively with buffer *A* followed by buffer *B* (10 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl and 100 m*M* imidazole) to remove non-specifically bound contaminants. The protein was finally eluted with buffer *C* (10 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl and 300 m*M* imidazole). The eluted protein was subjected to gel-filtration chromatography using Superdex 200 prep-grade matrix in a 16/70 C column (GE Healthcare Biosciences) on an ÄKTAprime Plus system (GE Healthcare Biosciences) equilibrated with buffer *D* (10 m*M* Tris–HCl pH 8.0, 50 m*M* NaCl, 2 m*M* $MgCl_2$, 5 m*M* DTT). Fractions

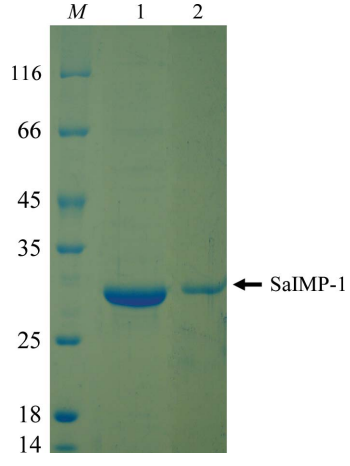


Figure 1
12% SDS–PAGE analysis of SaIMP-1 after Ni–NTA and gel-filtration chromatographic steps. Lane *M*, protein molecular-weight markers (kDa); lane 1, Ni–NTA-purified SaIMP-1; lane 2, purified SaIMP-1 after gel-filtration chromatography.

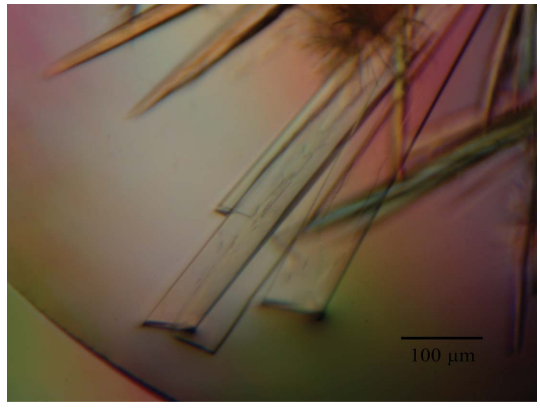


Figure 2
Good diffraction-quality single crystals of SaIMP-1 obtained using 0.2 *M* Li_2SO_4, 20% PEG 3350, 0.1 *M* HEPES pH 7.0 with the sitting-drop vapour-diffusion method.

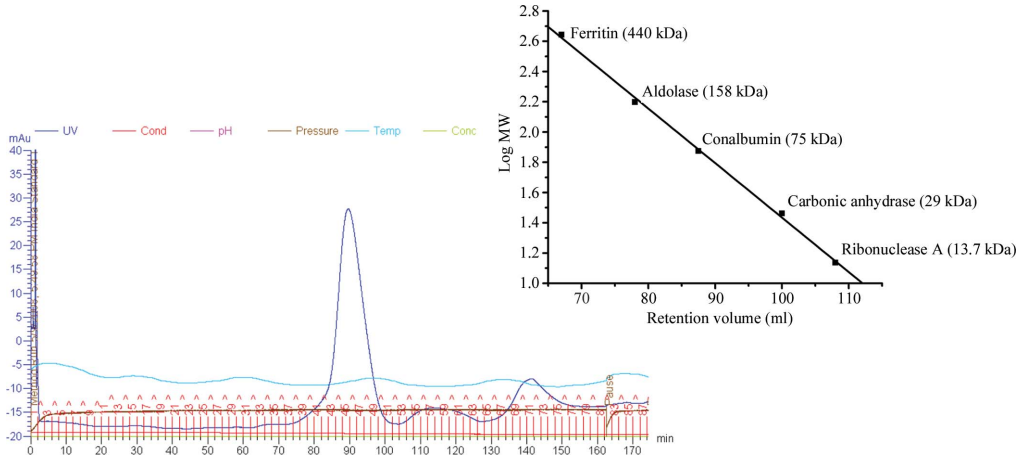


Figure 3
Determination of the oligomeric state of SaIMP-1 in solution by gel-filtration chromatography using a Superdex 200 16/70 C column. SaIMP-1 was eluted as described in §2.2. The gel-filtration column was calibrated with globular protein molecular-weight standards consisting of ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa), carbonic anhydrase (29 kDa) and ribonuclease A (13.7 kDa). The retention volume of SaIMP-1 was 90 ml, corresponding to a native molecular mass of ∼60 kDa.

(2 ml) were collected at a flow rate of 1 ml min^{-1}. To determine the oligomeric state of SaIMP-1 in solution, the gel-filtration column was calibrated with globular protein molecular-weight standards consisting of ferritin (440 kDa), aldolase (158 kDa), conalbumin (75 kDa), carbonic anhydrase (29 kDa) and ribonuclease A (13.7 kDa), which had retention volumes of 67, 78, 88, 100 and 108 ml, respectively. After the gel-filtration chromatographic step the fractions containing the desired protein were pooled together. The protein concentration was estimated by the method of Bradford (1976) and the purity was verified by 12% SDS–PAGE (Fig. 1).

2.3. Crystallization

The purified SaIMP-1 protein was concentrated to 24 mg ml^{-1} using a 10 kDa cutoff Vivaspin 20 concentrator (GE Healthcare). Preliminary screening for initial crystallization conditions was performed using Crystal Screen, Crystal Screen 2 and Index solutions (all from Hampton Research, USA) using the hanging-drop vapour-diffusion method at 298 K; droplets were formed by mixing 2 µl concentrated protein solution with an equal volume of reservoir solution in 24-well Linbro plates. Initial hits were only obtained using Index solutions [0.2 *M* Li_2SO_4, 25%(*w*/*v*) PEG 3350, 0.1 *M* Bis-Tris pH 6.5 and 0.2 *M* Li_2SO_4, 25%(*w*/*v*) PEG 3350, 0.1 *M* HEPES pH 7.5]. The single crystals resulting from the hanging-drop vapour-diffusion method using the Index solutions diffracted very poorly (∼3.5–4 Å). However, from this primary screening PEG 3350 and Li_2SO_4 were identified as successful precipitant and salt conditions, respectively, and were found to be mandatory for single-crystal formation. Conditions were further optimized using different concentrations of Li_2SO_4 and different percentages of PEG 3350 in the presence of 0.1 *M* HEPES pH 7.0 with the sitting-drop vapour-diffusion method. All sitting drops were formed by mixing 4 µl concentrated protein solution with an equal volume of reservoir solution using Micro-Bridges in 24-well Linbro plates. Finally, crystals of good diffraction quality (Fig. 2) appeared in about 24 h from 0.2 *M* Li_2SO_4, 20%(*w*/*v*) PEG 3350, 0.1 *M* HEPES pH 7.0.

2.4. Data collection and processing

Before data collection, a single crystal was picked up using a silicon loop (Hampton Research, USA), quick-soaked in mother liquor and flash-cooled in an N_2 stream at 100 K. X-ray diffraction data were collected on an in-house Rigaku R-AXIS IV^{++} image-plate detector using Cu $K\alpha$ X-rays generated by a Rigaku MicroMax-007 HF rotating-anode generator. The crystal-to-detector distance was maintained at 180 mm and the crystal was rotated through a total of 180° with 0.5° rotation per frame and an exposure time of 2 min per frame. The crystal diffracted to a resolution of 2.6 Å. A total of 360 frames were processed and scaled in space group $P2_12_12_1$ using *XDS* (Kabsch, 2010) and *SCALA* (Evans, 1993), respectively. The presence of screw axes was determined by observing the systematic absences and was confirmed by *POINTLESS* from the *CCP*4 suite (Evans, 2006). The resulting overall R_{merge} value was 0.13. The final statistics of data collection and processing are tabulated in Table 1.

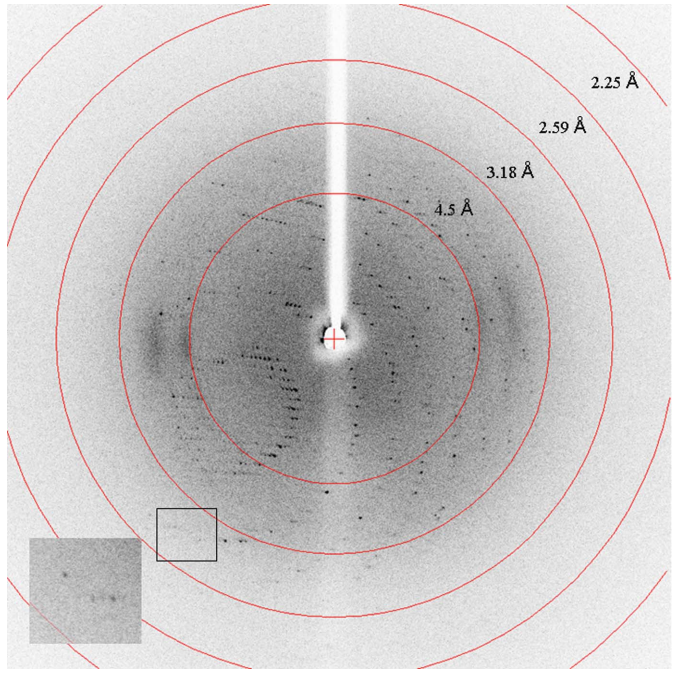


Figure 4
X-ray diffraction pattern of a typical SaIMP-1 crystal.

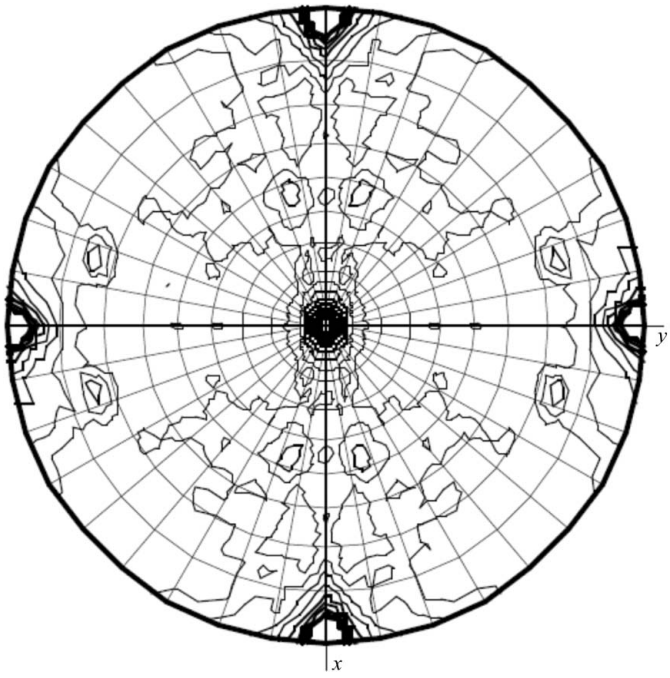


Figure 5
Stereographic projection of the self-rotation function of SaIMP-1 for the $\chi = 180^\circ$ section, showing three perpendicular twofold axes along the *x*, *y*, *z* directions. This figure was generated using the program *MOLREP* (Vagin & Teplyakov, 2010; Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

SaIMP-1 (SAS2203), one of the two inositol monophosphatase family proteins from *S. aureus* MSSA476, has been successfully isolated and cloned in pQE30 for recombinant expression in *E. coli* M15. The overexpressed protein was subsequently purified to homogeneity by immobilized metal-affinity chromatography (IMAC) followed by gel-filtration chromatography. The subunit molecular weight of His_6 SaIMP-1 was confirmed to be approximately 30 kDa by 12% SDS–PAGE (Fig. 1) as predicted from the sequence. Gel-filtration chromatographic analysis indicates that the protein exists as a dimer in solution, with an apparent molecular mass of ∼60 kDa (Fig. 3). Good diffraction-quality protein crystals were obtained from 0.2 *M* Li_2SO_4, 20%(*w*/*v*) PEG 3350, 0.1 *M* HEPES pH 7.0 using the sitting-drop vapour-diffusion method. The SaIMP-1 crystals diffracted to 2.6 Å resolution (Fig. 4) and belonged to an orthorhombic space group, with unit-cell parameters $a = 49.98$, $b = 68.35$, $c = 143.79$ Å, $\alpha = \beta = \gamma = 90^\circ$. The assumption of two chains of SaIMP-1 in the asymmetric unit of the $P2_12_12_1$ unit cell leads to a Matthews coefficient of 1.98 $Å^3\ Da^{-1}$ and a solvent content of 38% (Matthews, 1968). Calculation of the self-rotation function (Fig. 5) shows the presence of crystallographic twofold symmetry axes along the *x*, *y* and *z* axes ($\chi = 180^\circ$), a typical feature of *P*222 symmetry. However, no extra peaks arising from noncrystallographic symmetry could be seen in the self-rotation function. This observation, together with the results of the cell-content analysis (Matthews, 1968), suggest that the noncrystallographic axis runs either parallel or tilted at an angle of less than 10° to a crystallographic axis and hence is not visible as an additional peak in the self-rotation function. This ambiguity can only be resolved once the structure has been solved.

This work was carried out using the Protein X-ray Crystallography Facility funded by the Department of Biotechnology, Government of India and housed at the Central Research Facility (CRF) of the Indian Institute of Technology (IIT) Kharagpur. SB thanks IIT Kharagpur for an individual institute fellowship.

References

Berridge, M. J., Downes, C. P. & Hanley, M. R. (1989). *Cell*, **59**, 411–419.
Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Chen, L. & Roberts, M. F. (1998). *Appl. Environ. Microbiol.* **64**, 2609–2615.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Evans, P. R. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 114–122. Warrington: Daresbury Laboratory.
Johnson, K. A., Chen, L., Yang, H., Roberts, M. F. & Stec, B. (2001). *Biochemistry*, **40**, 618–630.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Li, Z., Stieglitz, K. A., Shrout, A. L., Wei, Y., Weis, R. M., Stec, B. & Roberts, M. F. (2010). *Protein Sci.* **19**, 309–318.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Neuwald, A. F., York, J. D. & Mejerus, P. W. (1991). *FEBS Lett.* **294**, 16–18.
Nigou, J., Dover, L. G. & Besra, G. S. (2002). *Biochemistry*, **41**, 4392–4398.
Stieglitz, K. A., Johnson, K. A., Yang, H., Roberts, M. F., Seaton, B. A., Head, J. F. & Stec, B. (2002). *J. Biol. Chem.* **277**, 22863–22874.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.