G. Deka,^a J. N. Kalyani,^b J. F. Benazir,^a H. S. Savithri^{b*} and M. R. N. Murthy^{a*}

^aMolecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India, and

^bBiochemistry Department, Indian Institute of Science, Bangalore 560 012, India

Correspondence e-mail:
bchss@biochem.iisc.ernet.in,
mrn@mbu.iisc.ernet.in

Received 24 December 2013

Accepted 8 March 2014

Successful data recovery from oscillation photographs containing strong polycrystalline diffraction rings from an unknown small-molecule contaminant: preliminary structure solution of *Salmonella typhimurium* pyridoxal kinase (PdxK)

Pyridoxal kinase (PdxK; EC 2.7.1.35) belongs to the phosphotransferase family of enzymes and catalyzes the conversion of the three active forms of vitamin B₆, pyridoxine, pyridoxal and pyridoxamine, to their phosphorylated forms and thereby plays a key role in pyridoxal 5'-phosphate salvage. In the present study, pyridoxal kinase from *Salmonella typhimurium* was cloned and overexpressed in *Escherichia coli*, purified using Ni-NTA affinity chromatography and crystallized. X-ray diffraction data were collected to 2.6 Å resolution at 100 K. The crystal belonged to the primitive orthorhombic space group $P2_12_12_1$, with unit-cell parameters $a = 65.11$, $b = 72.89$, $c = 107.52$ Å. The data quality obtained by routine processing was poor owing to the presence of strong diffraction rings caused by a polycrystalline material of an unknown small molecule in all oscillation images. Excluding the reflections close to powder/polycrystalline rings provided data of sufficient quality for structure determination. A preliminary structure solution has been obtained by molecular replacement with the *Phaser* program in the *CCP4* suite using *E. coli* pyridoxal kinase (PDB entry 2ddm) as the phasing model. Further refinement and analysis of the structure are likely to provide valuable insights into catalysis by pyridoxal kinases.

1. Introduction

Pyridoxal 5'-phosphate (PLP) is the most active form of vitamin B₆ and functions as a cofactor for a large number of enzymes involved in amino-acid metabolism (Toney, 2005). In the PLP salvage pathway, pyridoxal kinase (PK) catalyzes the ATP-dependent conversion of three vitamin B₆ precursors, pyridoxal (PL), pyridoxamine (PMN) and pyridoxine (PN), to their phosphorylated forms (di Salvo *et al.*, 2011). Crystal structures of pyridoxal kinase are known from human (Musayev *et al.*, 2007), *Trypanosoma brucei* (Jones *et al.*, 2012), sheep brain (Li *et al.*, 2002), *Escherichia coli* (Safo *et al.*, 2004), *Bacillus subtilis* (Newman *et al.*, 2006) and *Lactobacillus plantarum* (PDB entry 3h74; New York SGX Research Center for Structural Genomics, unpublished work). In *E. coli*, two genes, *pdxK* and *pdxY*, code for pyridoxal kinase (Yang *et al.*, 1996, 1998; Safo *et al.*, 2006).

In the present study, we have initiated structural and biochemical investigations on *Salmonella typhimurium* pyridoxal kinase (*StPK*). In this paper, we describe the cloning, overexpression, crystallization and preliminary structure determination of *StPK*. X-ray diffraction images obtained by oscillation photography contained nearly continuous rings of strong spots resulting from an unidentified contaminating polycrystalline material. Processing these images with *iMosflm* v.1.0.7 from the *CCP4* suite led to poor data statistics and the data obtained were not optimal for structure determination. Exclusion of reflections on or near the unwanted rings could be accomplished in a number of ways, either before or after integration of the images. For example, *MOSFLM* allows exclusion of all data within user-specified resolution ranges. We used a pre-release version of *iMosflm* that makes use of this *MOSFLM* feature to avoid integrating reflections close to the unwanted diffraction rings. The quality of the

data thus obtained was significantly better and led to successful structure determination.

2. Materials and methods

2.1. Cloning, overexpression and purification

The pyridoxal kinase gene (*pdxK*; STM2435) coding for 288 amino-acid residues was PCR-amplified from the *Salmonella enterica* serovar *typhimurium* strain LT2 genomic DNA using specific sense (CCCCATATGATGGGACAAGAGAGTGATATTCAGTC) and antisense (CCCCTCGAGTCGCGCCTCCCCTGCCGCGCAG) primers. The bold sequences in the sense and antisense primers represent restriction sites for *NdeI* and *XhoI*, respectively. PCR involved initial denaturation at 371 K (1 min), 30 cycles of denaturation at 371 K (30 s), annealing at 329 K (30 s) and elongation at 345 K (1 min) and the final extension was at 345 K (10 min). The PCR-amplified product was digested with *NdeI* and *XhoI* and ligated into the vector pET-22b(+) (Novagen) previously digested with the same restriction enzymes. The positive clone (pET22b-PK) obtained was further confirmed by DNA sequencing. The recombinant construct thus obtained had an extra methionine at the N-terminus and eight residues (LEHHHHHH) at the C-terminus that included a hexahistidine tag. The presence of the hexahistidine tag facilitated purification of *StPK* by Ni-NTA affinity chromatography.

The recombinant pET22b-*StPK* plasmids were transformed into *E. coli* BL21 (DE3) pLysS competent cells for overexpression. An LB agar plate containing 50 µg ml⁻¹ ampicillin was used to select the transformants. A single colony was used to prepare a pre-inoculum and 5 ml of it was transferred to 500 ml of LB medium containing 50 µg ml⁻¹ ampicillin. The culture was grown at 310 K until the OD₆₀₀ reached 0.6. Overexpression of *StPK* was induced by the addition of 0.3 mM IPTG. The culture was incubated at 303 K for 5–6 h post-induction. SDS-PAGE examination revealed the overexpression of a protein corresponding to the expected molecular weight of ~32 kDa. The cells were pelleted by centrifugation at 5855g in a GSA rotor for 10 min at 277 K. The recovered cells were suspended in 50 mM potassium phosphate pH 7.5, 0.1 M NaCl, 10% (v/v) glycerol and

lysed by sonication at 277 K. Cell debris was removed by centrifugation at 12 857g in an SS34 rotor at 277 K for 40 min. The supernatant was collected and the expressed protein was allowed to bind Ni-NTA beads equilibrated with extraction buffer (50 mM potassium phosphate pH 7.5, 0.1 M NaCl) on an end-to-end rotor for 3 h at 277 K. The Ni-NTA beads were loaded onto a column. Nonspecifically bound proteins were washed from the column using extraction buffer containing 20 mM imidazole. *StPK* was then eluted from the column using 200 mM imidazole in extraction buffer at 277 K. The eluted protein was concentrated using a 30 kDa molecular-weight cutoff Centricon (Amicon). 1 ml concentrated protein was loaded onto a Superdex S-200 preparative column previously equilibrated with 20 mM potassium phosphate pH 7.5, 1 mM EDTA and eluted with the same buffer at 277 K. Peak fractions which corresponded to the dimeric form of the protein were pooled, concentrated to 15 mg ml⁻¹ and used for crystallization. The protein concentration was estimated by measuring the OD at 280 nm and using the calculated molar extinction coefficient obtained from the *ProtParam* server (Wilkins *et al.*, 1999). This method of purification provided 15–20 mg ml⁻¹ of protein per litre of cell culture. The purity of the eluted protein was also ascertained by SDS-PAGE (Fig. 1).

2.2. Crystallization and data collection

Initial trials of *StPK* crystallization were carried out using microbatch as well as hanging-drop vapour-diffusion methods at 293 K. Various conditions were screened using commercially available kits from Hampton Research and Jena Bioscience. Each crystallization droplet consisted of 2 µl protein solution at 15 mg ml⁻¹ and 2 µl crystallization condition. In the hanging-drop vapour-diffusion method, crystallization droplets were placed over 400 µl reservoir solution. In the microbatch method, a 1:2 mixture of silicone and paraffin oils was layered over droplets consisting of 2 µl protein solution and 2 µl crystallization cocktail. Crystals (Fig. 2) were obtained at 293 K by the microbatch method from a condition (Jena Bioscience Basic 2/A3) consisting of 20% (w/v) PEG 4000, 10% (v/v) 2-propanol, 100 mM HEPES pH 7.5 in 10–15 d after setting up crystallization. These crystals were used for data collection. No hits were obtained in the hanging-drop method of crystallization.

Cu K α radiation ($\lambda = 1.5418$ Å) generated by a rotating-anode X-ray generator (Rigaku RU200) was used for X-ray diffraction data collection. The X-ray beam was focused using Osmic mirror optics and the images were recorded using a MAR Research image-plate (MAR345) detector system. Data collection was carried out at 100 K

Figure 1
Purification profile of *StPK*. Lane 1, molecular-weight markers (labelled in kDa); lane 2, Ni-NTA eluted protein; lane 3, protein after gel filtration.

Figure 2
A typical crystal of *S. typhimurium* pyridoxal kinase.

Table 1
Data-collection, processing and initial refinement statistics.
Values in parentheses are for the highest resolution shell.

	Data set obtained with <i>iMosflm</i> 1.0.7	Data set obtained with the pre-release version of <i>iMosflm</i> (now <i>iMosflm</i> 7.1.0)
Wavelength (Å)	1.54	1.54
Crystal-to-detector distance (mm)	200.3	200.3
Rotation range per image (°)	1	1
Exposure time per image (s)	300	300
Resolution range (Å)	72.73–2.60 (2.74–2.60)	72.89–2.60 (2.74–2.60)
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 65.25$, $b = 72.73$, $c = 107.56$	$a = 65.11$, $b = 72.89$, $c = 107.52$
Mosaicity (°)	1.5	1.5
Total No. of measured intensities	86565 (11999)	81619 (12715)
Unique reflections	16029 (2276)	14688 (2278)
Multiplicity	5.4 (5.3)	5.6 (5.6)
Mean $I/\sigma(I)$	6.0 (3.0)	9.9 (3.5)
Completeness (%)	98.6 (97.9)	90.1 (98.1)
$R_{\text{merge}}^{\dagger}$ (%)	23.7 (50.6)	13.6 (48.5)
$R_{\text{meas}}^{\ddagger}$ (%)	26.4 (84.0)	15.1 (53.4)
Rigid-body refinement		
$R_{\text{work}}^{\S}$ (%)	50.32	44.01
$R_{\text{free}}^{\P}$ (%)	51.20	44.28
Initial restrained refinement		
$R_{\text{work}}^{\S}$ (%)	47.66	30.74
$R_{\text{free}}^{\P}$ (%)	51.90	36.99

$\dagger R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, $\ddagger R_{\text{meas}} = \frac{\sum_{hkl} [N(hkl)/[N(hkl) - 1]]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, where $I_i(hkl)$ is the intensity of the i th observation of reflection hkl , $\langle I(hkl) \rangle$ is its mean intensity and $N(hkl)$ is the number of measurements (redundancy). $\S R_{\text{work}}^{\S}$ (%) = $\frac{\sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum_{hkl} |F_{\text{obs}}|}$, where F_{obs} and F_{calc} are observed and calculated structure factors, respectively. $\P R_{\text{free}}^{\P}$ (%) is calculated as for R_{work} but for a randomly selected 5% of the data set excluded from the refinement process.

with a 1.0° oscillation angle and an exposure time of 300 s per frame. The crystal-to-detector distance was maintained at 200 mm. The crystal was soaked in crystallization condition containing 20% ethylene glycol as the cryoprotectant for 5 min prior to data collection. The crystal diffracted to 2.6 Å resolution as shown in Fig. 3.

Figure 3
An oscillation image obtained from a crystal of *StPK* illustrating strong polycrystalline noise rings at ~3.14, 2.74 and 2.43 Å resolutions.

3. Results
3.1. Data processing and structure determination

The diffraction data were processed using *iMosflm* 1.0.7 (Battye *et al.*, 2011; Leslie, 2006) and scaled using *SCALA* (Evans, 2006) from the *CCP4* program suite (Winn *et al.*, 2011). Autoindexing indicated that the crystal belonged to the orthorhombic system, with unit-cell parameters $a = 65.11$, $b = 72.89$, $c = 107.52$ Å. Further analysis of systematic absences using the program *POINTLESS* in the *CCP4* suite (Evans, 2006; Winn *et al.*, 2011) suggested that the crystal belonged to space group $P2_12_12_1$. The crystal asymmetric unit was compatible with two monomers (calculated Matthews coefficient of $1.99 \text{ Å}^3 \text{ Da}^{-1}$; Matthews, 1968). The data quality was poor as indicated by the R_{merge} and other parameters (Table 1). Attempts to determine the structure by molecular replacement using *E. coli* PK (PDB entry 2ddm, chain A; Safo *et al.*, 2006) as the phasing model, with which *StPK* shares 68% amino-acid sequence identity, were unsuccessful. It was observed that the oscillation images had a few rings of strong spots owing to diffraction by some polycrystalline material. The resolution of these rings corresponded to ~3.14, 2.74 and 2.43 Å. Ice or crystals of salts that were present in the crystallization cocktail do not cause rings at these resolutions. A pre-release version of the *iMosflm* graphical user interface (GUI) became available for processing the data in a *CCP4* workshop held at the Okinawa Institute of Science and Technology, Okinawa, Japan on 4–8 November 2013. This new version of the *iMosflm* GUI can make use of a feature long available in *MOSFLM* to exclude reflections close to polycrystalline rings from integration. This version of *iMosflm* has now been released as *iMosflm* 7.1.0. The quality of the data obtained after the exclusion of rings was substantially higher, although the data completeness was reduced by almost 10%. Table 1 provides a comparison of the results of processing with the old and the pre-release versions of *iMosflm*. The decrease in overall completeness (90.1%) obtained with the pre-release version of *iMosflm* is due to the deletion of a large number of reflections in shells around 3.14 and 2.74 Å corresponding to the observed polycrystalline rings. However,

Figure 4

A cartoon representation of the preliminary structure of *StPK* (orange) superposed on *E. coli* PK (green).

the highest resolution shell is not close to any polycrystalline ring and hence its completeness is nearly 100%. With the new data, a reasonable initial structure could be determined by molecular replacement using *E. coli* PK as the phasing model in *Phaser* (McCoy *et al.*, 2007) from the *CCP4* suite, with RFZ, TFZ and LLG scores of 4.6, 18.1 and 868, respectively. Rigid-body refinement followed by restrained refinement of positional parameters with *REFMAC5* (Murshudov *et al.*, 2011) from the *CCP4* suite resulted in *R* and *R*_{free} values of 30.74 and 36.99%, respectively. The corresponding values for the data obtained using the earlier version of *iMosflm* were 47.66 and 51.90%, respectively. A cartoon representation of the structure superposed on *E. coli* PK is depicted in Fig. 4. Further refinement and model building are in progress. A detailed analysis of the structure and its implications for the mechanism of enzymatic function will be published elsewhere.

4. Discussion

In this work, PK from *S. typhimurium* was purified to homogeneity and crystallized, and X-ray diffraction data were collected using X-rays from an in-house rotating-anode generator. The crystals belonged to space group *P*₂₁₂₁ with two protomers in the asymmetric unit. Structure determination by molecular replacement using the corresponding *E. coli* enzyme as the phasing model should have been straightforward because of the high (68%) amino-acid sequence identity between the model and the target structures. However, an

acceptable solution to the structure could not be obtained. It was apparent that rings of strong spots appearing on all oscillation images owing to a polycrystalline contaminant were the reason for the poor quality of the scaled diffraction data and perhaps also for the difficulty in structure determination. A pre-release version of *iMosflm* that can avoid spots close to specified diffraction rings was used for integration of the images. This helped to improve the data quality in a significant way. Structure determination using the data obtained with the new version of *iMosflm* was straightforward. Interestingly, it was possible to process the oscillation frames with the old version of *iMosflm* to a limited resolution of 3.2 Å so that integration of reflection intensities is avoided in the resolution rings containing polycrystalline spots and obtain a data set that could be used successfully for structure determination and refinement. The results presented in this paper illustrate in an interesting way the well recognized importance of the processing strategy and data quality in X-ray crystal structure analysis.

The X-ray facility at the Molecular Biophysics Unit, Indian Institute of Science was used to collect the intensity data. MRNM and HSS acknowledge the Department of Science and Technology (DST) and the Department of Biotechnology (DBT), Government of India, for financial support. FB was supported by a Kothari post-doctoral fellowship. GD thanks Fadel Samatey for inviting her to the *CCP4* School held at Okinawa Institute of Science and Technology, Japan and Phil Evans for helping her in processing the data set.

References

- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- di Salvo, M. L., Contestabile, R. & Safo, M. K. (2011). *Biochim. Biophys. Acta*, **1814**, 1597–1608.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Jones, D. C., Alpey, M. S., Wyllie, S. & Fairlamb, A. H. (2012). *Mol. Microbiol.* **86**, 51–64.
- Leslie, A. G. W. (2006). *Acta Cryst.* **D62**, 48–57.
- Li, M.-H., Kwok, F., Chang, W.-R., Lau, C.-K., Zhang, J.-P., Lo, S. C. L., Jiang, T. & Liang, D.-C. (2002). *J. Biol. Chem.* **277**, 46385–46390.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Musayev, F. N., di Salvo, M. L., Ko, T. P., Gandhi, A. K., Goswami, A., Schirch, V. & Safo, M. K. (2007). *Protein Sci.* **16**, 2184–2194.
- Newman, J. A., Das, S. K., Sedelnikova, S. E. & Rice, D. W. (2006). *J. Mol. Biol.* **363**, 520–530.
- Safo, M. K., Musayev, F. N., di Salvo, M. L., Hunt, S., Claude, J. B. & Schirch, V. (2006). *J. Bacteriol.* **188**, 4542–4552.
- Safo, M. K., Musayev, F. N., Hunt, S., di Salvo, M. L., Scarsdale, N. & Schirch, V. (2004). *J. Bacteriol.* **186**, 8074–8082.
- Toney, M. D. (2005). *Arch. Biochem. Biophys.* **433**, 279–287.
- Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D. & Hochstrasser, D. F. (1999). *Methods Mol. Biol.* **112**, 531–552.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Yang, Y., Tsui, H.-C. T., Man, T.-K. & Winkler, M. E. (1998). *J. Bacteriol.* **180**, 1814–1821.
- Yang, Y., Zhao, G. & Winkler, M. E. (1996). *FEMS Microbiol. Lett.* **141**, 89–95.