crystallization communications

Cloning, purification and preliminary crystallographic analysis of a putative pyridoxal kinase from *Bacillus subtilis*

Joseph A. Newman, Sanjan K. Das, Svetlana E. Sedelnikova and David W. Rice*

Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, The University of Sheffield, Sheffield S10 2TN, England

Correspondence e-mail: d.rice@sheffield.ac.uk

Pyridoxal kinases (PdxK) are able to catalyse the phosphorylation of three vitamin B_6 precursors, pyridoxal, pyridoxine and pyridoxamine, to their $5'$-phosphates and play an important role in the vitamin B_6 salvage pathway. Recently, the *thiD* gene of *Bacillus subtilis* was found to encode an enzyme which has the activity expected of a pyridoxal kinase despite its previous assignment as an HMPP kinase owing to higher sequence similarity. As such, this enzyme would appear to represent a new class of 'HMPP kinase-like' pyridoxal kinases. *B. subtilis thiD* has been cloned and the protein has been overexpressed in *Escherichia coli*, purified and subsequently crystallized in a binary complex with ADP and Mg^{2+}. X-ray diffraction data have been collected from crystals to 2.8 Å resolution at 100 K. The crystals belong to a primitive tetragonal system, point group 422, and analysis of the systematic absences suggest that they belong to one of the enantiomorphic pair of space groups $P4_12_12$ or $P4_32_12$. Consideration of the space-group symmetry and unit-cell parameters ($a = b = 102.9$, $c = 252.6$ Å, $\alpha = \beta = \gamma = 90°$) suggest that the crystals contain between three and six molecules in the asymmetric unit. A full structure determination is under way to provide insights into aspects of the enzyme mechanism and substrate specificity.

1. Introduction

Pyridoxal kinase (PdxK) is an enzyme that functions in the salvage pathway of pyridoxal 5′-phosphate (PLP). PdxK catalyses the phosphorylation of pyridoxal, pyridoxine and pyridoxamine to their corresponding 5′-phosphates, which are precursors for the active form of vitamin B_6 (Fig. 1). PLP is required by both prokaryotes and eukaryotes for the synthesis, conversion and catabolism of amino acids and can be synthesized *de novo* by many bacteria as well as salvaged from the growth medium. In *Escherichia coli*, two pyridoxal kinase genes have been identified and named PdxY (Yang *et al.*, 1998; Safo *et al.*, 2004) and PdxK (Yang *et al.*, 1996). PdxY is a pyridoxal kinase which like PdxK is involved in the PLP-salvage pathway, but is not able to catalyse the phosphorylation of the additional substrates pyridoxine and pyridoxamine. PdxK has also been shown to play a role in the thiamine phosphate-salvage pathway and possesses 4-amino-5-hydroxymethyl-2-methylpyrimidine (HMP) kinase activity (Begley *et al.*, 1999; Mizote & Nakayama, 1989).

The *thiD* gene has been previously classified in studies from a number of different species as encoding a 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate kinase (HMPP kinase), which is an enzyme that functions in the thiamine phosphate (vitamin B_1) salvage pathway and is essential for the survival and virulence of a number of pathogens (Hava & Camilli, 2002; Akerley *et al.*, 2002). HMPP kinases are able to catalyse the phosphorylation of the hydroxymethyl group on HMP and also the subsequent phosphorylation of a phosphomethyl group on HMPP, but do not have any known activity for phosphorylating pyridoxal, pyridoxine or pyridoxamine. The product of the *Bacillus subtilis thiD* gene comprises a 29 kDa protein of 271 amino acids which is unable to catalyse the phosphorylation of HMPP as suggested by its genome annotation, but instead possesses the enzyme activity of a pyridoxal kinase (Park *et al.*, 2004). This confusion arose as a result of its higher sequence identities to HMPP kinases than to pyridoxal kinases (35 and 25%, respectively, for the
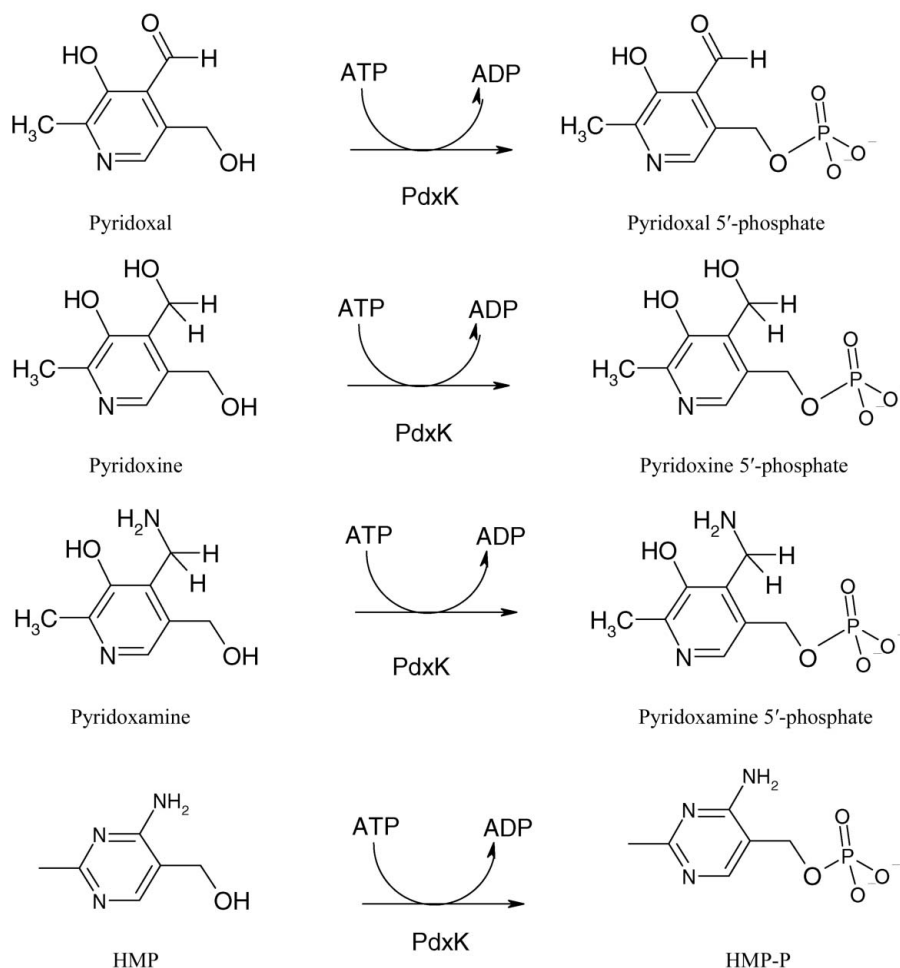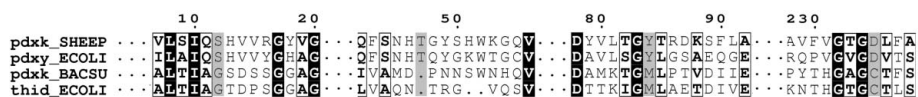
Figure 1
The four kinase reactions catalysed by PdxK.

E. coli enzymes). The *B. subtilis thiD* gene product will therefore be referred to as PdxK for the remainder of this paper.

Analysis of the amino-acid sequence of PdxK has suggested that it is a member of the ribokinase superfamily and this has been confirmed by the determination of the structure of PdxKs from sheep brain (Li *et al.*, 2002, 2004), human (PDB code 2ajp) and *E. coli* (Safo *et al.*, 2006) and the PdxY gene product in *E. coli* (Safo *et al.*, 2004). However, comparing the sequences of these enzymes to that of *B. subtilis* PdxK, the well aligned regions flanking the substrate-binding site reveal that few of the key residues involved in substrate recognition or catalysis are conserved (Fig. 2). Specifically, three residues involved in forming interactions with the substrate (Ser12, Thr47 and Tyr84 in the sheep brain enzyme) are completely conserved in all PdxKs studied to date but are not present in *B. subtilis* PdxK (Gly11, Pro46, Met82). Furthermore, the catalytic machinery of these enzymes is also different, with the sheep, human and *E. coli* enzymes having a catalytic Asp, whereas the *B. subtilis*

enzyme has a Cys at this position. Interestingly, at each of these positions the sequence of *B. subtilis* PdxK is very similar to that found in the family of HMPP kinases (for example, *E. coli* ThiD). This suggests that *B. subtilis* PdxK may be a representative member of a new class of 'HMPP kinase-like' pyridoxal kinases for which no structural data currently exists. The closer sequence similarities found between *B. subtilis* PdxK and other HMPP kinases may indicate that pyridoxal kinase activity has evolved independently on multiple occasions during evolution.

To date, there is no structural data on the mode of nucleotide binding to any HMPP kinase and analysis of the apoenzyme structures [from *Salmonella typhimurium* (Cheng *et al.*, 2002) and *Thermus thermophilus* (PDB code 1ub0)] has revealed that two loops which lie close to the nucleotide-binding site are disordered. The structure of a nucleotide complex may shed light on the role of these loops, which are thought to become ordered upon nucleotide binding and play a role in nucleotide recognition.
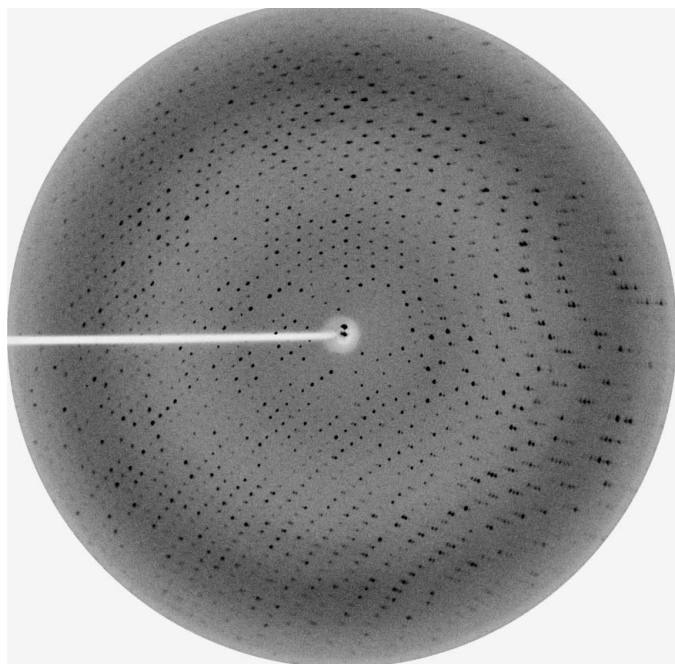

Figure 2
Multiple sequence alignment showing the sequence conservation in four different regions for PdxK from sheep, *E. coli* and *B. subtilis* and *E. coli* ThiD. Identical residues across all sequences are shown in reverse type on a dark background. Strongly conserved residues are shown in bold. Residues that have been identified from previous structural studies on sheep and *E. coli* PdxK to form interactions with the substrate are highlighted by grey boxes and can be seen to be different in *B. subtilis* PdxK, where they are more similar to *E. coli* ThiD. The sequences were aligned using *ClustalW* (Thompson *et al.*, 1994) and the figure was prepared using *ESPript* (Gouet *et al.*, 1999).

As a contribution towards understanding the structure–function relationship of the enzyme and the molecular basis of substrate specificity, we have initiated the structure determination of *B. subtilis* PdxK. In this paper, we describe the cloning, overexpression, purification, crystallization and preliminary X-ray analysis of this enzyme in the presence of ADP and Mg^{2+}.

2. Cloning, overexpression and purification of *B. subtilis* PdxK

The *B. subtilis thiD* gene was PCR-amplified directly from the genomic DNA of *B. subtilis* strain 168 with the primers TCG-GCGGCATATGTCTATGCATAAAGCACT (forward) and ACA-CAGALTCTTATGATTGTTGATTCAGGCG (reverse). The PCR product (813 bp) was purified (Quiagen) and a parallel restriction digest (310 K overnight) of the PCR product and plasmid vector pTB361 was performed to produce linear fragments with sticky ends suitable for ligation. The restriction products were ligated together using a gene:plasmid DNA ratio of 2:1 at 287 K and subsequently transformed into *E. coli* DH5α for verification by PCR screening and restriction-enzyme digestion. For overexpression, the plasmid was transformed into *E. coli* strain BL21 (DE3). A 250 ml flask containing 50 ml LB medium with 50 μg ml^{-1} tetracycline was inoculated with a single colony of the transformed strain and grown overnight at 310 K on a shaking tray at 250 rev min^{-1}. 10 ml of this culture was used to inoculate five 2 l flasks each containing 450 ml of LB medium supplemented with tetracycline as above. Cultures were grown until the absorbance at 600 nm reached 0.6, at which point overexpression was induced by the addition of 0.1 mM IPTG. Cultures were harvested after a further 5 h by centrifugation at 5000g for 20 min at 277 K and the pellets were frozen immediately. Analysis of the soluble fraction by SDS–PAGE showed a large overexpression band corresponding to the expected molecular weight of the enzyme (∼30 kDa).

Cells containing overexpressed PdxK were thawed on ice, suspended in 50 mM Tris–HCl pH 8.0 and disrupted by ultrasonication (3 × 20 s pulses). Cell debris was removed by centrifugation at 20 000g for 20 min. The supernatant fraction was applied onto a column packed with DEAE-Sepharose Fast Flow (Amersham Biosciences) and proteins were eluted with a linear gradient of NaCl from 0 to 1 M in 50 mM Tris–HCl pH 8.0. Fractions containing PdxK were combined and 4.0 M NH$_4$SO$_4$ was added to give a final concentration of 1.5 M. Precipitated proteins were removed by ultracentrifugation at 20 000 g for 5 min. The supernatant fraction was then loaded onto a column packed with Phenyl ToyoPearl 650S (Tosoh) and eluted with a linear gradient of NH$_4$SO$_4$ from 1.5 to 0 M in 50 mM Tris–HCl pH 8.0 and 200 mM NaCl. Fractions containing PdxK were pooled and concentrated using a VivaSpin 10 000 kDa molecular-weight cutoff concentrator to a final volume of 0.5 ml and a protein concentration of 13 mg ml^{-1}. The sample was then loaded onto a Hi-Load Superdex 200 column (Amersham Biosciences) equilibrated with 200 mM NaCl in 50 mM Tris–HCl pH 8.0 and eluted with the same buffer. The elution volume was compared to a standard curve and PdxK ran with an approximate molecular weight of 40 kDa. Peak fractions containing PdxK were collected and concentrated to 10–15 mg ml^{-1} as above, exchanging to buffer containing 10 mM Tris–HCl pH 8.0 and 50 mM NaCl. The purity of the final preparation was estimated to be over 90% by SDS–PAGE, with a yield of approximately 20 mg per gram of cell paste.

3. Crystallization and preliminary X-ray analysis

An initial crystallization screen was attempted using the hanging-drop vapour-diffusion method at 290 K and crystallization kits from Hampton Research in the presence and absence of the nucleotides ADP, ATP, AMP-PNP and ATPγS with and without Mg^{2+}. Initial rod-like crystals of the free enzyme with approximate dimensions 5 × 5 × 50 μm were observed to have grown in 30% PEG 4000, 0.2 M sodium acetate trihydrate, 0.1 M Tris–HCl pH 8.0. Crystals of the PdxK–ADP–Mg^{2+} complex were grown using 28% PEG 4000, 0.17 M sodium acetate trihydrate, 0.1 M Tris–HCl pH 8.5 as the precipitant with the addition of 10 mM ADP, 10 mM MgCl$_2$. Crystals from the latter condition were bipyramidal with approximate dimensions 60 × 25 × 25 μm.

Single crystals of both the free enzyme and the PdxK–ADP–Mg^{2+} complex were loop-mounted and placed into a cryosolution containing 27% PEG 4000, 0.17 M sodium acetate trihydrate, 0.08 M Tris–HCl pH 8.0, 23% glycerol for approximately 2 min. Diffraction images were collected at 100 K with 0.5° rotations using a Rigaku MM007 rotating-anode generator with a MAR Research image-plate detector. Preliminary analysis of the X-ray diffraction data using *MOSFLM* (Leslie, 1992) indicated that the free-enzyme crystals belong to a primitive trigonal system, point group 312 (either $P3_212$ or $P3_112$), with unit-cell parameters $a = b = 50.4, c = 177.9$ Å, $\alpha = \beta = 90$, $\gamma = 120°$. However, the highly anisotropic nature of the diffraction and the poor resolution limit (3.5 Å) made these crystals unsuitable for detailed structural analysis.

Crystals of the PdxK–ADP–Mg^{2+} complex diffracted to 2.8 Å (Fig. 3). Processing of the X-ray diffraction data using *MOSFLM* and *SCALA* (Evans, 1997) from the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994) and analysis of the merging statistics indicated that the crystals belong the a primitive tetragonal space group, point group 422. Further analysis of the pattern of systematic absences suggest that the crystals belong to one of the enantiomorphic pair of space groups $P4_12_12$ or $P4_32_12$, with

Table 1
Data-processing statistics for *B. subtilis* PdxK–ADP–Mg^{2+} complex.

Values in parentheses are for data in the highest resolution shell.

Resolution	20–2.8 (2.95–2.8)
Independent reflections	33604 (4844)
Completeness	99.4 (99.9)
Multiplicity	3.7 (3.8)
$I/\sigma(I)$	9.3 (2.4)
R_{merge}†	0.131 (0.475)

† $R_{merge} = \sum |I - \langle I \rangle| / \sum I$, where I is the integrated intensity of a given reflection.

unit-cell parameters $a = b = 102.9$, $c = 252.6$ Å, $\alpha = \beta = \gamma = 90°$ (data-collection statistics are shown in Table 1). Consideration of the unit-cell volume and space-group symmetry suggests that the crystal contains between three and six monomers in the asymmetric unit, with a corresponding V_M range of 1.8–3.8 Å^3 Da^{-1} (Matthews, 1977). A full structure determination is under way in order to provide insights into the catalytic activity, substrate specificity and evolution of this enzyme.

References

Akerley, B. J., Rubin, E. J., Novick, V. L., Amaya, K., Judson, N. & Mekalanos, J. J. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 966–971.

Begley, T. P., Downs, D. M., Ealick, S. E., McLafferty, F., Van Loon, D., Taylor, S., Chiu, H., Kinsland, C., Reddick, J., Xi, J. & Campobasso, N. (1999). *Arch. Microbiol.* **171**, 293–300.

Cheng, G., Bennett, E. M., Begley, T. P. & Ealick, S. E. (2002). *Structure*, **10**, 225–235.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Evans, P. R. (1997). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **33**, 22–24.

Gouet, P., Courcelle, E., Stuart, D. I. & Metoz, F. (1999). *Bioinformatics*, **15**, 305–308.

Hava, D. L. & Camilli, A. (2002). *Mol. Microbiol.* **45**, 1389–1405.

Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EAMCB Newsl. Protein Crystallogr.* **26**.

Li, M. H., Kwok, F., Chang, W. R., Lau, C. K., Zhang, J. P., Lo, S. C., Jiang, T. & Liang, D. C. (2002). *J. Biol. Chem.* **277**, 46385–46390.

Li, M. H., Kwok, F., Chang, W. R., Liu, S. Q., Lo, S. C. L., Zhang, J. P., Jiang, T. & Liang, D. C. (2004). *J. Biol. Chem.* **279**, 17459–17465.

Matthews, B. W. (1977). *The Proteins*, 3rd ed., edited by H. Neurath & R. L. Hill, Vol. 3, pp. 468–477. New York: Academic Press.

Mizote, T. & Nakayama, H. (1989). *Biochim. Biophys. Acta*, **991**, 109–113.

Park, J. H., Burns, K., Kinsland, C. & Begley, T. P. (2004). *J. Bacteriol.* **186**, 1571–1573.

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.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673-4680.

Yang, Y., Tsui, H. 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.