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

Cloning, purification and crystallization of *Thermus thermophilus* proline dehydrogenase

Tommi A. White and John J. Tanner*

Departments of Chemistry and Biochemistry, University of Missouri-Columbia, Columbia, Missouri 65211, USA

Correspondence e-mail: tannerjj@missouri.edu

Received 13 April 2005
Accepted 22 June 2005
Online 8 July 2005

© 2005 International Union of Crystallography
All rights reserved

Nature recycles L-proline by converting it to L-glutamate. This four-electron oxidation process is catalyzed by the two enzymes: proline dehydrogenase (PRODH) and Δ^1-pyrroline-5-carboxylate dehydrogenase. This note reports the cloning, purification and crystallization of *Thermus thermophilus* PRODH, which is the prototype of a newly discovered superfamily of bacterial monofunctional PRODHs. The results presented here include production of a monodisperse protein solution through use of the detergent *n*-octyl β-D-glucopyranoside and the growth of native crystals that diffracted to 2.3 Å resolution at Advanced Light Source beamline 4.2.2. The space group is $P2_12_12_1$, with unit-cell parameters a = 82.2, b = 89.6, c = 94.3 Å. The asymmetric unit is predicted to contain two protein molecules and 46% solvent. Molecular-replacement trials using a fragment of the PRODH domain of the multifunctional *Escherichia coli* PutA protein as the search model (24% amino-acid sequence identity) did not produce a satisfactory solution. Therefore, the structure of *T. thermophilus* PRODH will be determined by multiwavelength anomalous dispersion phasing using a selenomethionyl derivative.

1. Introduction

Proline utilization A (PutA) proteins are membrane-associated bifunctional proline-catabolic enzymes that catalyze the two-step oxidation of proline to glutamate (Menzel & Roth, 1981*b*; Brown & Wood, 1993; Surber & Maloy, 1998; Becker & Thomas, 2001; Vinod *et al.*, 2002; Zhu & Becker, 2003). In the first step of proline catabolism, proline is oxidized to Δ^1-pyrroline-5-carboxylate (P5C) by the FAD-dependent PutA proline dehydrogenase (PRODH) domain. P5C is hydrolyzed nonenzymatically to glutamic semialdehyde and the semialdehyde is oxidized to glutamate by the NAD-dependent PutA P5C dehydrogenase (P5CDH) domain. PutA proteins typically contain 1000–1300 amino-acid residues, with the PRODH domain located in the N-terminal half of the polypeptide chain and the P5CDH domain located in the C-terminal half. In addition to their PRODH and P5CDH activities, some PutA proteins, such as *Escherichia coli* PutA, serve as autogenous repressors and therefore contain an N-terminal DNA-binding domain (Gu *et al.*, 2004; Menzel & Roth, 1981*c*; Ostrovsky De Spicer & Maloy; 1993; Becker & Thomas, 2001; Wood, 1981).

Several functional aspects of PutA have been investigated, including enzymatic activity (Menzel & Roth, 1981*a*,*b*; Surber & Maloy, 1998; Zhu *et al.*, 2002; Baban *et al.*, 2004), redox-linked conformational changes (Brown & Wood, 1993; Zhu & Becker, 2003; Zhang, Zhou *et al.*, 2004) and transcriptional repression (Menzel & Roth, 1981*c*; Muro-Pastor & Maloy, 1995; Muro-Pastor *et al.*, 1997; Ostrovsky De Spicer & Maloy, 1993; Gu *et al.*, 2004). In addition, the crystal structure of the PRODH domain of *E. coli* PutA has been determined (Nadaraia *et al.*, 2001; Lee *et al.*, 2003; Zhang, White *et al.*, 2004). The structures of PutA P5CDH and DNA-binding domains are currently not known.

Until recently, it was thought that most bacteria express PutA proteins. However, our analysis of genome-sequence data suggests that PutA proteins are restricted to Gram-negative bacteria, whereas most Gram-positive bacteria express PRODH and P5CDH as separate enzymes encoded by separate genes (data not shown). These newly discovered monofunctional proline-catabolic enzymes have

not been characterized; therefore, we cloned the genes for *Thermus thermophilus* PRODH and P5CDH in preparation for biochemical, biophysical and structural analyses. PRODH from *T. thermophilus* shares 24% amino-acid sequence identity with the *E. coli* PutA PRODH domain. Here, we report the cloning, expression, purification and crystallization of *T. thermophilus* PRODH. Interestingly, PRODH and P5CDH appear as separate enzymes in eukaryotes (Phang, 1985; Adams & Frank, 1980); thus, the study of bacterial monofunctional proline-catabolic enzymes may provide insights into the human enzymes that are not readily obtained from studies of multifunctional PutA proteins.

2. Materials and methods

2.1. Cloning

The *T. thermophilus* PRODH gene was cloned from genomic DNA purchased from the American Type Culture Collection and introduced into the plasmid pKA8H between *Bam*HI and *Nde*I sites. The pKA8H vector codes for an N-terminal 8×His affinity tag and a tobacco etch virus protease site. Since the PRODH gene contains a *Bam*HI site, digestion of the PCR product with *Bam*HI was not possible. Therefore, the staggered reannealing method (Ailenberg & Silverman, 1996) was used with the following three primers: forward, 5′-CCTTGATCATATGAACCTGGACCTGGCTTACCGTTC-3′; reverse 1, 5′-GATCCCTAGCCGGAAACCAGGCTCCTCAGG-3′; reverse 2, 5′-CCTAGCCGGAAACCAGGCTCCTCAGG-3′. Two separate PCR amplification experiments were performed using the forward primer in conjunction with each of the two reverse primers. The two PCR products were purified, mixed in equimolar amounts, denatured at 369 K for 5 min, annealed by slow cooling and finally digested with *Nde*I. The resulting PCR product was ligated into pKA8H, which had been digested with *Bam*HI and *Nde*I. Sequencing confirmed that the gene was successfully cloned into the vector.

2.2. Protein expression and purification

Unless stated otherwise, all chemicals were purchased from Fisher Scientific or Sigma–Aldrich. PRODH was expressed in BL21(DE3)pLysS cells (Novagen) as follows. Small (10 ml) cultures were grown overnight in LB media and used to inoculate 1.5 l LB media. Protein expression was induced with 0.5 m*M* IPTG after the culture reached an optical density of OD_{600} = 0.6. Cells were harvested 3 h after induction, resuspended in 50 m*M* NaH_2PO_4, 300 m*M* NaCl, 10 m*M* imidazole, 5% glycerol pH 8.0 and frozen. Cells were thawed at 277 K and lysed by a French press at 110 MPa in the presence of five protease inhibitors (0.1 m*M* TPCK, 0.05 m*M* AEBSF, 0.1 µ*M* pepstatin, 0.01 m*M* leupeptin, 5 µ*M* E-64). The supernatant was collected after centrifugation at 15 000 rev min^{-1} for 30 min at 277 K, filtered through a 0.45 µm filter (Millipore) and applied to 5 ml of Ni–NTA Superflow resin (Qiagen). The column was washed in two steps with the loading buffer supplemented with 50 m*M* imidazole followed by 75 m*M* imidazole. PRODH was eluted with 250 m*M* imidazole and dialyzed overnight in the dark at 277 K into 4 l of 50 m*M* Tris–HCl, 50 m*M* NaCl, 0.5 m*M* EDTA, 0.5 m*M* DTT and 5% glycerol pH 8.0. Prior to dialysis, flavin adenine dinucleotide (FAD) at a concentration of 0.1 m*M* was added to the dialysis bag containing the eluted PRODH. Excess FAD was removed using a desalting column (Biorad P100). The protein was concentrated to 13 mg ml^{-1} using centrifugal concentrators. Protein concentration was determined using the Bradford method with bovine serum albumin as the standard.

Figure 1
T. thermophilus proline dehydrogenase crystals. The largest dimension of these crystals is 0.2 mm.

The purified protein exhibited the intense yellow color that is characteristic of flavoenzymes. The enzyme displayed PRODH activity as measured by a previously described assay (Zhang, White *et al.*, 2004). The molecular mass as determined by MALDI–TOF mass spectrometry at the University of Missouri-Columbia Proteomics Center was 37968 ± 3. This value was in good agreement with the theoretical weight predicted from the gene sequence (37 923 Da). The 8×His affinity tag was not removed prior to crystallization trials.

2.3. Dynamic light scattering

Gel-filtration chromatography suggested that the protein exhibited a high degree of aggregation. For example, protein injected onto a Superdex-200 column eluted entirely in the void volume. Aggregation of PRODH could indicate that *T. thermophilus* PRODH is a membrane-associated protein, which would be consistent with the fact that PutA proteins are peripheral membrane proteins (Zhang, Zhou *et al.*, 2004; Surber & Maloy, 1999; Brown & Wood, 1993; Wood, 1987). The correlation between monodispersity and crystallizability has been well established (Ferre-D'Amare & Burley, 1997; Zulauf & D'Arcy, 1992); therefore, the protein was exposed to various solvent conditions in an attempt to identify conditions that promoted monodispersity. The monodispersity of each protein solution tested was assessed with a Protein Solutions DynaPro 99 Molecular Sizing Instrument. The parameters varied included pH, ionic strength, addition of Pro and Pro analogues and addition of various detergents, including *n*-octyl β-D-glucopyranoside (BOG). The protein concentration was in the range 1–3 mg ml^{-1}.

Only BOG had a significant effect on protein aggregation. The amount of high-molecular-weight species was dramatically reduced when BOG was added to a final concentration of 20 m*M*. The protein–BOG solution had an apparent protein molecular weight of 35 kDa and the polydispersity index was Cp/Rh = 37%. For comparison, in the absence of BOG, the apparent protein molecular weight was 1700–2600 kDa and the polydispersity index was Cp/Rh = 40–52%. Based on these results, the purification procedure described in §2.2 was modified by the addition of 20 m*M* BOG to the protein after the final dialysis step. Excess detergent and FAD were then removed using a desalting column (Biorad P100). The concentration of BOG that remained in the protein solution after desalting was not determined.

3. Results

3.1. Crystallization

All crystallization experiments were performed at 295 K using the sitting-drop method of vapor diffusion with drops formed by mixing

Table 1
Data-collection statistics.

Values for the outer resolution shell of data are given in parentheses.

Beamline	ALS 4.2.2
Wavelength (Å)	0.9869
Space group	$P2_12_12_1$
Unit-cell parameters	a = 82.16, b = 89.62, c = 94.26
Resolution limits	45–2.30 (2.38–2.30)
Observations	226041
Unique reflections	31569
Average redundancy	7.2 (7.0)
Completeness (%)	99.9 (99.9)
Average $I/\sigma(I)$	14.5 (5.1)
$R_{\rm merge}$	0.076 (0.325)

equal volumes of the reservoir (2 µl) and protein (2 µl) solutions. Commercially available crystal screens (Hampton Research and Decode Genetics) were used to identify initial crystallization conditions. Several conditions in the screens yielded crystals of various size and quantity. The precipitating agent 2-methyl-2,4-pentanediol (MPD) was present in many of the positive conditions. After several rounds of optimization, the best crystals were grown with a reservoir containing 50 m*M* $MgCl_2$, 100 m*M* imidazole pH 7.5, 35% MPD (Fig. 1). Note that BOG was not added to the reservoir and that the only potential source of BOG in the crystallization drop was that which remained in the protein stock solution after desalting. Since the mother liquor provided cryoprotection, the crystals were picked up with Hampton mounting loops and frozen directly in liquid nitrogen.

3.2. Data collection and processing

X-ray diffraction data were collected at beamline 4.2.2 of the Advanced Light Source at Lawrence Berkeley National Laboratory using a NOIR-1 CCD detector. Autoindexing of the data with *d*TREK* (Pflugrath, 1999) suggested a primitive orthorhombic lattice with unit-cell parameters a = 82.2, b = 89.6, c = 94.3 Å. A 2.3 Å native data set consisting of 180 frames was collected with a crystal-to-detector distance of 125 mm, an oscillation angle of 1° and an exposure time of 4 s per frame. Analysis of the data with *d*TREK* confirmed *Pmmm* as the Laue symmetry and suggested $P2_12_12_1$ as the space group. The Matthews coefficient was 2.3 $Å^3$ Da^{-1}, implying a solvent content of 46% with two molecules of PRODH in the asymmetric unit (Matthews, 1968). See Table 1 for data-processing statistics.

Molecular-replacement calculations (high-resolution limit = 4 Å) were performed with *MOLREP* (Vagin & Teplyakov, 2000) using the $\beta_8\alpha_8$ barrel of the PRODH domain of *E. coli* PutA (Lee *et al.*, 2003; Zhang, White *et al.*, 2004) as the search model (residues 264–435 and 457–562 of PDB entry 1tiw). All eight possible primitive orthorhombic lattices were tested. The top solution had an R factor of 0.57 and a correlation coefficient of 0.23, which indicated that molecular replacement was not a suitable phasing method. Therefore, the structure of *T. thermophilus* PRODH will be determined by multi-wavelength anomalous dispersion phasing using a selenomethionyl derivative.

This work was supported by NIH grant GM065546. We thank Mingyi Zhou of the University of Missouri-Columbia DNA Core for performing cloning experiments and Beverly Dague of the University of Missouri-Columbia Proteomics Center for performing mass-spectral analysis. We thank Katherine Ferrell for generously providing the pKA8H vector. We thank Jay Nix and Darren Sherrell of ALS beamline 4.2.2 for help with data collection and processing. The ALS is supported by the Director, Office of Science, Office of Basic Energy Sciences, Materials Sciences Division of the US Department of Energy under Contract No. DE-AC03-76SF00098 at Lawrence Berkeley National Laboratory.

References

Adams, E. & Frank, L. (1980). *Annu. Rev. Biochem.* **49**, 1005–1061.
Ailenberg, M. & Silverman, M. (1996). *Biochem. Mol. Biol. Int.* **39**, 771–779.
Baban, B., Vinod, M. P., Tanner, J. J. & Becker, D. F. (2004). *Biochem. Biophys. Acta*, **1701**, 49–59.
Becker, D. F. & Thomas, E. A. (2001). *Biochemistry*, **40**, 4714–4722.
Brown, E. D. & Wood, J. M. (1993). *J. Biol. Chem.* **268**, 8972–8979.
Ferre-D'Amare, A. R. & Burley, S. K. (1997). *Methods Enzymol.* **276**, 157–166.
Gu, D., Zhou, Y., Kallhoff, V., Baban, B., Tanner, J. J. & Becker, D. F. (2004). *J. Biol. Chem.* **279**, 31171–31176.
Lee, Y. H., Nadaraia, S., Gu, D., Becker, D. F. & Tanner, J. J. (2003). *Nature Struct. Biol.* **10**, 109–114.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Menzel, R. & Roth, J. (1981*a*). *J. Biol. Chem.* **256**, 9755–9761.
Menzel, R. & Roth, J. (1981*b*). *J. Biol. Chem.* **256**, 9762–9766.
Menzel, R. & Roth, J. (1981*c*). *J. Mol. Biol.* **148**, 21–44.
Muro-Pastor, A. M. & Maloy, S. (1995). *J. Biol. Chem.* **270**, 9819–9827.
Muro-Pastor, A. M., Ostrovsky, P. & Maloy, S. (1997). *J. Bacteriol.* **179**, 2788–2791.
Nadaraia, S., Lee, Y. H., Becker, D. F. & Tanner, J. J. (2001). *Acta Cryst.* D**57**, 1925–1927.
Ostrovsky De Spicer, P. & Maloy, S. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 4295–4298.
Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.
Phang, J. M. (1985). *Curr. Top. Cell Regul.* **25**, 92–132.
Surber, M. W. & Maloy, S. (1998). *Arch. Biochem. Biophys.* **354**, 281–287.
Surber, M. W. & Maloy, S. (1999). *Biochim. Biophys. Acta*, **1421**, 5–18.
Vagin, A. & Teplyakov, A. (2000). *Acta Cryst.* D**56**, 1622–1624.
Vinod, M. P., Bellur, P. & Becker, D. F. (2002). *Biochemistry*, **41**, 6525–6532.
Wood, J. (1987). *Proc. Natl Acad. Sci. USA*, **84**, 373–377.
Wood, J. M. (1981). *J. Bacteriol.* **146**, 895–901.
Zhang, M., White, T. A., Schuermann, J. P., Baban, B. A., Becker, D. F. & Tanner, J. J. (2004). *Biochemistry*, **43**, 12539–12548.
Zhang, W., Zhou, Y. & Becker, D. F. (2004). *Biochemistry*, **43**, 13165–13174.
Zhu, W. & Becker, D. F. (2003). *Biochemistry*, **42**, 5469–5477.
Zhu, W., Gincherman, Y., Docherty, P., Spilling, C. D. & Becker, D. F. (2002). *Arch. Biochem. Biophys.* **408**, 131–136.
Zulauf, M. & D'Arcy, A. (1992). *J. Cryst. Growth*, **122**, 102–106.