Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

A unique dye-decolorizing peroxidase, DyP, from *Thanatephorus cucumeris* Dec 1: heterologous expression, crystallization and preliminary X-ray analysis

Takao Sato,[a] Shusaku Hara,[b] Takuro Matsui,[c] Gen Sazaki,[c] Shinya Saijo,[a] Tadashi Ganbe,[a] Nobuo Tanaka,[a] Yasushi Sugano[b] and Makoto Shoda[b]*

[a]Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama 226-8501, Japan, [b]Chemical Resources Laboratory, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama 226-8503, Japan, and [c]Center for Interdisciplinary Research, Tohoku University, Aramaki, Aoba-ku, Sendai 980-8578, Japan

Correspondence e-mail: mshoda@res.titech.ac.jp

Received 15 September 2003
Accepted 1 November 2003

The dye-decolorizing peroxidase DyP is a key enzyme in the decolorizing fungus *Thanatephorus cucumeris* Dec 1 that degrades azo and antraquinone dyes. The gene *dyp* from *T. cucumeris* Dec 1, which has low homology to other peroxidase genes, was cloned and transformed into *Aspergillus oryzae* and glycosylated DyP was expressed at high levels. Purified DyP was deglycosylated using GST Endo F1 and then crystallized in a strong magnetic field (10 T) at 283 K using ammonium sulfate as precipitant. X-ray diffraction data to 2.96 Å resolution collected from a native crystal at the Photon Factory (Tsukuba, Japan) showed that the crystal belonged to the hexagonal space group $P6_522$, with unit-cell parameters $a = b = 136.15$, $c = 363.46$ Å. The asymmetric unit of the crystal contained four DyP molecules, with a corresponding Matthews coefficient (V_M) of 2.50 Å^3 Da^{-1} and a solvent content of 51%. Heavy-atom derivatives of DyP have been obtained and electron-density maps have been calculated. The haem is visible and continuous electron density between the haem and protein clearly indicates the location of the proximal histidine ligand.

1. Introduction

A basidiomycete, *Thanatephorus cucumeris* Dec 1 (formerly called *Geotrichum candidum* Dec 1), isolated from soil, shows broad degrading activity toward poorly biodegradable materials such as lignin, molasses and synthetic dyes (Kim *et al.*, 1995; Kim & Shoda, 1998, 1999*a*; Shintani *et al.*, 2002). This fungus has potential for removing these wastes as an energy-saving treatment method. A peroxidase (DyP) has been associated with the broad degradation spectrum of Dec 1 (Kim & Shoda, 1999).

Although DyP has been proposed to belong to the fungus peroxidase superfamily, its amino-acid sequence homology shows low similarity to any other protein registered in the DNA Data Bank of Japan. The highest identity to DyP was 55% for *Termitomyces albuminosus* peroxidase (Johjima *et al.*, 2003) and *Polyporaceae* sp. hypothetical peroxidase (unpublished work) by *PSI-BLAST* search (Stephen *et al.*, 1997). However, DyP and these two proteins showed no homology to other fungal peroxidases such as lignin peroxidase (LiP) and manganese peroxidase (MnP) (Renganathan & Gold, 1986; Johjima *et al.*, 1999; Wariishi *et al.*, 1988; Heinfling *et al.*, 1998) by *PSI-BLAST* search. Moreover, no typical haem-binding region common to the superfamily, which has conserved proximal histidines and aspartic acid plus distal histidines and arginine residues as proposed by Welinder (1992), was found using any other homology-search programs.

In order to provide structural information on DyP from *T. cucumeris* Dec 1, we initiated structure determination of DyP. DyP is a 58 kDa monomeric enzyme (pI = 4.2) that exhibits a unusually low pH optimum (pH 3.2). It contains one haem *b* and its polypeptide chain consists of 442 amino-acid residues. Its composition is typical of secreted glycoproteins: it includes 17% saccharides and has three probable N-glycosylation sites (Kim & Shoda, 1999*b*; Sugano *et al.*, 1999). It shows both dye-decolorizing and lignin-degrading activities when H_2O_2 is added. As the first step toward structural elucidation of the degradation mechanism, we report here the crystallization and preliminary X-ray crystallographic analysis of this enzyme. Furthermore, structure analysis will provide opportunities to apply structure-based protein-engineering principles in next-generation efforts to rationally design novel peroxidases.

2. Experimental

2.1. Protein preparation

With glycoproteins, much time and many experiments may be required to find useful crystallization conditions and success is never assured. We experienced just such difficulties in crystallization trials using our original DyP samples in glycosylated form. Therefore, we

© 2004 International Union of Crystallography
Printed in Denmark – all rights reserved

decided to deglycosylate DyP using the endoglycosidase GST-Endo F1 (Hampton Research). When combined with heterologous expression of the secreted glycoprotein by *Aspergillus oryzae* (Sugano *et al.*, 2000), the deglycosylation process (283 K, 2 d) provided highly pure deglycosylated DyP in sufficient quantities for crystallization trials. The purification process and enzyme assay were carried out by modification of methods described in previous reports (Kim & Shoda, 1999*b*; Sugano *et al.*, 2000). DyP purity was checked by assaying its specific activity and determining the Reinheitzahl (RZ) value (Dunford, 1999). RZ is defined as the ratio of the absorbance at 407 nm to that at 280 nm and indicates the purity of the enzyme. The deglycosylated protein was characterized using the periodic acid and Schiff reagent (PAS) staining assay (Zacharius *et al.*, 1969) and SDS–PAGE (data not shown). Glycosylated DyP is stained purple with PAS. In contrast, non-glycosylated DyP is not stained. A molecular mass of 47 kDa was also determined by gel-filtration chromatography using a Superdex 75 column (Amersham Pharmacia) and by dynamic light scattering (DLS; ALV-5000, Langen, Germany). There was no significant change in the RZ value or specific activity between deglycosylated DyP and native DyP; the deglycosylated DyP maintained the same conformation as DyP. Based on these data, the protein was found to be a monomer in the solution.

2.2. Crystallization

The crystallization experiments were continued with the purified deglycosylated DyP.

The protein solution was concentrated to 30 mg ml^{-1} in 0.1 *M* citrate buffer pH 4.2 and 0.5 *M* NaCl. Initially, DyP was concentrated to 30 mg ml^{-1}, but dynamic light scattering indicated that the protein solution was not monodisperse. A reduction in the protein concentration and addition of NaCl improved the monodispersity. At this protein concentration (twofold supersaturation; the solubility of DyP in 0.5 *M* NaCl and 0.1 *M* sodium citrate buffer pH 4.2 at 283 K is 15 mg ml^{-1}), DLS measurements indicated the formation of aggregates, indicating these conditions to be in the protein nucleation and crystal-growth zone of the precipitation diagram, and crystal optimization was carried out by varying three factors (gradual reduction of the starting protein concentration and the ammonium sulfate precipitant concentration and the the addition of NaCl) and exploring factor space during the screening step.

Crystals were grown by the batch method. 20 µl of protein and 20 µl of precipitating agent solution were mixed in a small flat-bottomed glass tube of 4 mm inner diameter. Subsequently, pH 4.2 was selected at 283 K in order to obtain reproducible crystallization-assay results. Optimum crystallization results were obtained using 0.89 *M* $(NH_4)_2SO_4$, 0.92 *M* NaCl and 15 mg ml^{-1} protein. To improve the crystal quality, crystallization experiments were conducted in a continuous homogeneous magnetic field (10 T) produced by a liquid helium-free superconducting magnet (JMTD-10T100M, Japan Magnet Technology Inc.) over a three-week period. The crystals obtained were of larger volume or deeper colour than those obtained at 0 T, resulting in diffraction of sufficient intensity to determine the position of the haem iron. Improvement using a magnetic field has been reported in the crystallization of lysozyme (Sato *et al.*, 2000). Controls were not run, however, that would distinguish the effects of growth in the magnetic field from those arising from deglycosylation.

2.3. X-ray diffraction experiment

To provide cryoprotection of the crystals for flash-cooling, crystals were briefly soaked in a cryoprotectant solution consisting of 15%(*v*/*v*) glycerol in artificial mother liquor [2.0 *M* $(NH_4)_2SO_4$, 0.5 *M* NaCl and 0.1 *M* citrate buffer pH 4.2]. All data were collected on an ADSC Quantum 4R (Area Detector Systems Corporation, California, USA) at 95 K using synchrotron radiation (beamlines BL6A and BL18B, Photon Factory, Tsukuba, Japan) to enhance the anomalous scattering effects of the Hg and haem iron. These wavelengths were chosen from inspection of the fluorescence spectrum. Data were processed and reduced with the *Crystal Clear* program (Rigaku/MSC, Texas, USA).

Table 1
Summary of data collection for native and heavy-atom derivative crystals.

Values in parentheses correspond to the highest resolution shell.

Data set	Native	K_2PtCl_4	$Hg(OAc)_2$	$KAu(CN)_2$	Native anomalous
Beamline	BL-6A	BL-18B	BL-6A	BL-6A	BL-18B
Wavelength (Å)	0.9770	0.9780	1.0089	0.9770	1.7435
Target scatterers		Pt	Hg	Au	Fe
Resolution (Å)	2.96 (3.10–2.96)	2.90 (3.0–2.9)	3.4 (3.52–3.4)	3.0 (3.11–3.0)	4.0 (4.14–4.0)
$I/\sigma(I)$†	8.1	9.2	5.1	9.8	4.0
Total observations	173683	142923	173128	137395	78281
Unique data	47508	40101	27103	38882	17284
Completeness (%)	98.4 (97.1)	84.2 (41.5)	94.9 (76.9)	91.7 (63.2)	99.3 (99.4)
$R_{\rm merge}$‡ (%)	5.9 (12.2)	4.9 (8.1)	14.2 (26.8)	5.6 (8.5)	11.9 (13.8)
Mosaicity (°)	0.126	0.174	0.215	0.170	0.148

† The maximum resolution is set to cut $\langle I\rangle/\sigma(\langle I\rangle)$ above 2.0. ‡ $R_{\rm merge} = \sum_h \sum_i |\langle I(h)\rangle - I(h)_i| / \sum_h \sum_i I(h)_i$, where $I(h)$ is the intensity of reflection h, $\sum_h$ is the sum over all reflections and $\sum_i$ is the sum over all i measurements of reflection h.

3. Results and discussion

3.1. Data collection

The data-collection images were indexed in a hexagonal space group, either $P6_122$ or $P6_522$. A total of 30° of data were collected (120 frames) from a single crystal; a total of 186 923 reflections were measured to a resolution limit of 2.96 Å. The native data set is 98.4% complete and the overall $R_{\rm merge}$ for 47 508 unique reflections is 0.059. The crystal mosaic spread is 0.126° (Table 1). The space group was identified to be $P6_522$ as this space group produced the correct hand using anomalous dispersion from a Pt derivative. The unit-cell parameters are $a = b = 136.15$, $c = 363.46$ Å and the asymmetric unit contains four DyP molecules. The unit-cell volume is 58.0×10^5 Å^3 which, on the basis of a tetrameric MW of 190 kDa, gives a $V_{\rm M}$ value of 2.50 Å^3 Da^{-1}, which is within the range given as typical for protein crystals by Matthews (1968). This was confirmed by measuring the crystal density in a linear gradient produced by mixing *m*-xylene and bromobenzene (Low & Richards, 1952). The results show $V_{\rm prot}$ to be 0.49 and the solvent content to be 51%.

3.2. Multiple isomorphous replacement (MIR) phasing

Heavy-atom derivatives of the deglycosylated DyP crystals were prepared by soaking in 10 µl aliquots of 1–5 m*M* solutions of heavy-metal compounds in artificial mother liquor for 2 d. The data-collection details are provided in Table 1. After scaling with *SCALEIT* (Howell & Smith, 1992) from the *CCP*4 program package (Collaborative Computational Project, Number 4, 1994), the program *FFT* (Ten Eyck, 1973) was used to calculate difference Patterson

Table 2
Summary of phase calculations with the resolution range 15–2.96 Å limit before density modification.

Data set	K_2PtCl_4	Hg $(OAc)_2$	$KAu(CN)_2$	Native anomalous
No. of sites	4	2	2	2
R_{iso}†	0.213	0.200	0.079	0.070
R_{cullis}‡				
Iso	0.88/098	0.88/0.99	0.99/1.00	
Ano	0.80	0.81	0.95	
Phasing power§				
Acentric	0.96	0.80	0.34	
Centric	0.70	0.67	0.32	

† $R_{iso} = \sum_h |F_{PH} - F_P| / \sum_h F_P$, where F_{PH} and F_P are the native and derivative structure-factor amplitudes, respectively. ‡ $R_{cullis} = \sum ||F_{PH} \pm F_P| - |F_H|| / \sum |F_{PH}| \pm |F_P|$, where F_H is the calculated structure-factor amplitude of the heavy-atom structure. § Phasing power $= \sum |F_H(\text{calc})|^2 / \sum |E|^2)^{1/2}$, where F_H(calc) is the calculated scattering amplitude for the heavy-atom structure and E is the estimated lack-of-closure error.

maps from reflection data and the positions of heavy atoms were confirmed as significant peaks on four Harker sections. For the platinum derivative, the program *MLPHARE* (Otwinowski, 1991) was used to refine positional and scattering parameters for the four independent Pt atoms and calculate initial phases. An electron-density map based on data in the resolution range 20–3.5 Å revealed the approximate positions of two Au and Hg atoms, respectively. *MLPHARE* was again used to refine the parameters of the two-Hg and two-Au scattering models and to calculate phases to 3.2 Å resolution. The phasing statistics are provided in Table 2. The program *DM* (Cowtan & Main, 1996) was used for non-crystallographic symmetry (NCS) averaging and solvent flattening of the electron-density maps, as well as phase extension from 12 to 2.96 Å resolution. The overall figure of merit for 51 911 reflections from 12 to 2.96 Å is 0.937. On the resulting electron-density map, a number of secondary-structure elements were recognizable and the space group was also confirmed to be $P6_522$ by the appearance of right-handed helices. Manual model building using the program *O* (Jones *et al.*, 1991) is under way.

3.3. Haem environment

One of the most important targets of this work is to determine the haem-binding region. The haem iron positions were located by isomorphous difference Patterson maps and cross-difference Fourier maps. Their positions were confirmed with the native anomalous Fe difference Patterson map. The position of one of the haem groups is shown in the solvent-flattened MIR map (Fig. 1). From analysis of this result, we conclude that the proximal histidine residue is His164 from the density map surrounding the haem. The geometry of the proximal histidine (His164) in DyP is partially similar to that in LiP and MnP from the basidiomycete white-rot fungus *Phanerchaete chrysosporium* (Poulos *et al.*, 1993; Sundaramoorthy *et al.*, 1994). This similarity was limited to too narrow a region to be detected using a homology-search program such as *PSI-BLAST*.

On the other hand, unlike LiP and MnP, DyP contains four-stranded antiparallel β-sheets, indicating that the arrangement on the distal side of haem pocket is significantly different from the known structures of the peroxidases LiP and MnP. Completion of model building and structure refinement will reveal the extent of structural similarity between DyP and the LiP and MnP structures. The final structure will form the starting point for protein engineering of DyP in order to improve its suitability for bioremediation by altering its substrate specificity and increasing its stability.

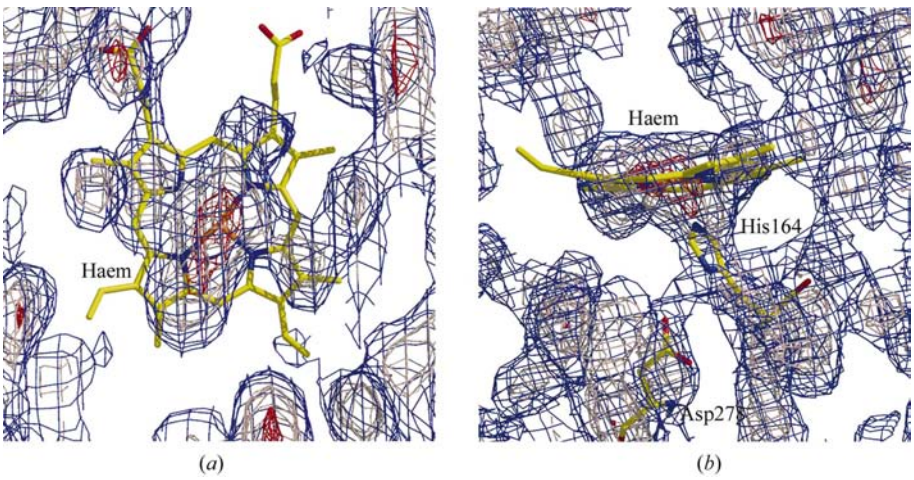

Figure 1
MIR map contoured at 1σ in the vicinity of the haem at 2.96 Å resolution. Top (*a*) and side (*b*) views are shown. In the side view, the continuous electron density between the protein and haem clearly indicates the location of the proximal haem ligands His164 and Asp278 [temporary assignment of the assumed residue numbers where structural homology may exist to ARP (*Arthromyces ramosus* peroxidase; Kunishima *et al.*, 1996) in the proximal regions]. The distance of His164 $N^{\varepsilon 2}$ from the Fe atom is 2.0 Å, which is the same as in LiP. The imidazole ring of His164 forms a hydrogen bond with a carboxyl side chain of Asp278 in DyP. The distance is 2.9 Å and is the same as in LiP.

This research was undertaken with the approval of the Photon Factory Advisory Committee, Japan. The authors wish to express their sincere thanks to the staff at the Photon Factory for their help with the data collection. This work was partly supported by a Grant from the 21st Century COE Program and a Grant-in-Aid for Scientific Research on Priority Areas (A) from the Japanese Ministry of Education, Culture, Sports, Science and Technology.

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cowtan, K. D. & Main, P. (1996). *Acta Cryst.* D**52**, 43–48.
Dunford, H. B. (1999). *Heme Peroxidase*, pp. 21. New York: Wiley–VCH.
Heinfling, A., Martinez, M. J., Martinez, A. T., Bergbauer, M. & Szewzyk, U. (1998). *FEMS Microbiol. Lett.* **165**, 43–50.
Howell, P. L. & Smith, G. D. (1992). *J. Appl. Cryst.* **25**, 81–86.
Johjima, T., Itoh, N., Kabuto, M., Tokimura, F., Nakagawa, T., Wariishi, H. & Tanaka, H. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 1989–1994.
Johjima, T., Ohkuma, M. & Kudo, T. (2003). *Appl. Microbiol. Biotechnol.* **61**, 220–225.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kim, S. J., Ishikawa, K., Hirai, M. & Shoda, M. (1995). *J. Ferment. Bioeng.* **79**, 601–607.
Kim, S. J. & Shoda, M. (1998). *Biotechnol. Tech.* **12**, 497–499.
Kim, S. J. & Shoda, M. (1999*a*). *Biotechnol. Bioeng.* **62**, 114–119.
Kim, S. J. & Shoda, M. (1999*b*). *Appl. Environ. Microbiol.* **65**, 1029–1035.

Kunishima, N., Amada, F., Fukuyama, K., Kawamoto, M., Matsunaga, T. & Matsubara, H. (1996). *FEBS Lett.* **378**, 291–294.
Low, B. W. & Richards, F. M. (1952). *J. Am. Chem. Soc.* **74**, 1660–1666.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. (1991). *Proceedings of the CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 80–86. Warrington: Daresbury Laboratory.
Poulos, T. L., Edward, S. L., Wariishi, H. & Gold, M. H. (1993). *J. Biol. Chem.* **268**, 4429–4440.
Renganathan, V. & Gold, M. H. (1986). *Biochemistry*, **25**, 1626–1631.
Sato, T., Yamada, Y., Saijo, S., Hori, T., Hirose, R., Tanaka, N., Sazaki, G., Nakajima, K., Igarashi, N., Tanaka, M. & Matsuura, Y. (2000). *Acta Cryst.* D**56**, 1079–1083.
Shintani, S., Sugano, Y. & Shoda, M. (2002). *J. Wood Sci.* **48**, 402–408.
Stephen, A. S., Madden, T. L., Shaffer, A. A., Zang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acid. Res.* **25**, 3389–3402.
Sugano, Y., Nakano, R., Sasaki, K. & Shoda, M. (2000). *Appl. Environ. Microbiol.* **66**, 1754–1758.
Sugano, Y., Sasaki, K. & Shoda, M. (1999). *J. Biosci. Bioeng.* **87**, 411–417.
Sundaramoorthy, M., Kishi, K., Gold, M. H. & Poulos, T. L. (1994). *J. Biol. Chem.* **269**, 32759–32767.
Ten Eyck, L. F. (1973). *Acta Cryst.* A**29**, 183–191.
Wariishi, H., Akileswaran, L. & Gold, M. H. (1988). *Biochemistry*, **27**, 5365–5370.
Welinder, K. G. (1992). *Curr. Opin. Struct. Biol.* **2**, 388–393.
Zacharius, R. M., Zell, T. E., Morrison, J. H. & Woodlock, J. J. (1969). *Anal. Biochem.* **30**, 148–152.