Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Cloning, expression and preliminary X-ray analysis of the dihydroorotase from the hyperthermophilic eubacterium *Aquifex aeolicus*

Cristina Purcarea, Phillip Martin,† John F. Vickrey,† Hedeel I. Guy, Brian F. P. Edwards and David R. Evans*

Department of Biochemistry and Molecular Biology, Wayne State University School of Medicine, 540 East Canfield Street, Detroit, Michigan 48201, USA

† These authors contributed equivalently to the work.

Correspondence e-mail: devans@cmb.biosci.wayne.edu

Received 30 July 2001
Accepted 16 October 2001

Dihydroorotase (DHOase) catalyzes the formation of dihydroorotate in the *de novo* pyrimidine biosynthetic pathway. The gene encoding the type I DHOase from the hyperthermophilic bacterium *Aquifex aeolicus* has been cloned in *Escherichia coli* with a polyhistidine affinity tag appended to the amino-terminal end and sequenced. The recombinant protein was expressed at high levels and could be purified readily in a single step by Ni^{2+} affinity chromatography. Both native and selenomethionine-labeled proteins were crystallized using the hanging-drop vapor-diffusion technique. Screens of the purified protein identified several conditions that yielded crystals; however, the best crystals were obtained using 1 *M* Li_2SO_4, 10 m*M* $NiCl_2$, 100 m*M* Tris acetate pH 8.5 as the precipitant. Well formed diamond-shaped crystals appeared within 1 d and continued to grow over several weeks to about 0.5 mm in the largest dimension. The crystals diffract to 1.7 Å and belong to space group *C*2, with unit-cell parameters a = 119.8, b = 88.0, c = 55.2 Å, β = 99.0° and a mosaic spread of 0.6°. There is one DHOase monomer in the asymmetric unit.

1. Introduction

Dihydroorotase (DHOase; E.C. 3.5.2.3) is a zinc metalloenzyme that catalyzes the reversible cyclization of carbamoyl aspartate to form dihydroorotate (Fig. 1) in the *de novo* pyrimidine biosynthetic pathway. Collectively, dihydroorotases exhibit a remarkably diverse structural organization. *E. coli* DHOase is a homodimer (Washabaugh & Collins, 1984) of 38 kDa polypeptides. In mammalian cells, the DHOase activity is associated with a 44 kDa domain of CAD (Coleman *et al.*, 1977), a multifunctional protein that also has carbamoyl phosphate synthetase (CPSase) and aspartate transcarbamoylase (ATCase) activities. A 44 kDa fragment of purified CAD containing the active DHO domain has been isolated from controlled proteolytic digests (Kelly *et al.*, 1986), but initial attempts to clone the mammalian domain based on homology to the 38 kDa *E. coli* DHOase were unsuccessful. It was subsequently found (Williams *et al.*, 1990; Musmanno *et al.*, 1991; Zimmermann & Evans, 1993) that additional residues at the carboxyl end of the polypeptide are required for catalytic activity, indicating that the mammalian domain is larger than its bacterial counterpart. Dihydroorotase is also associated with ATCase in a multienzyme complex in *Thermus aquaticus* (Van de Casteele *et al.*, 1997) and *Streptomyces griseus* (Hughes *et al.*, 1999). Interestingly, inactive homologues of DHOase, designated pseudo-DHOs, have been discovered in the yeast bifunctional protein encoded by the *ura*2 locus (Souciet *et al.*, 1989) and in *Pseudomonas* ATCase (Schurr *et al.*, 1995; Vickrey *et al.*, 2002). Extensive sequence comparisons (Kim & Kim, 1998) revealed that the dihydroorotases are closely related to the allantoinases, D-hydantoinases and dihydropyrimidinases.

Holm & Sander (1997), using several powerful search algorithms, proposed that DHOase belongs to a superfamily of amidohydrolases that catalyze a diverse series of hydrolytic reactions. The superfamily consists of two subgroups that have a common catalytic mechanism in which one or two metal ions deprotonate a water molecule for nucleophilic attack on the substrate. The first subgroup, which includes urease and phosphotriesterase, are enzymes with a binuclear metal center and a bridging carboxyl-lysine. Adenosine deaminase (Wilson *et al.*, 1991), a representative of the second subclass, has a similar active-site configuration but lacks the carboxylated lysine and has a single metal ion at the active site. DHOase was originally thought to belong to the first subgroup (Holm & Sander, 1997), although recent structural studies (Thoden *et al.*, 2001), described below, place the *E. coli* enzyme in the second subgroup. All members of this superfamily were predicted to have a common structural core consisting of eight alternating β-sheet/α-helices similar to the

© 2002 International Union of Crystallography
Printed in Denmark – all rights reserved

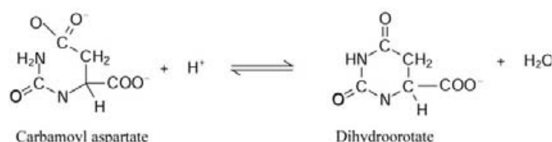

Figure 1
Reaction catalyzed by *A. aeolicus* dihydroorotase. DHOase catalyzes the reversible interconversion of carbamoyl aspartate and dihydroorotate in the *de novo* pyrimidine biosynthetic pathway.

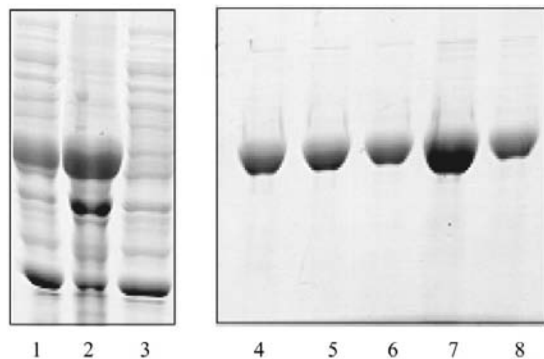

Figure 2
Expression and purification of *A. aeolicus* dihydroorotase. SDS–polyacrylamide gel electrophoresis of supernatant (lane 1), pellet obtained by high-speed centrifugation of the cell extract (lane 2), the Ni^{2+} affinity column flowthrough (lane 3) and the fractions obtained by elution with 100 m*M* (lanes 4–5) and 200 m*M* (lanes 6–8) imidazole in 50 m*M* Tris–HCl pH 8, 200 m*M* NaCl.

TIM-barrel motif found in urease and a signature sequence that includes four histidine residues and an aspartate.

An extensive phylogenetic analysis (Fields *et al.*, 1999) resulted in a classification of the dihydroorotases into two major classes thought to have arisen by an ancestral gene duplication. Type I DHOases are the most ancient and widely divergent form of the enzyme and are found in multifunctional proteins (*e.g.* mammalian CAD), multienzyme complexes (*e.g. T. aquaticus*) and as monofunctional enzymes (*e.g. Bacillus subtilis*). Type II dihydroorotases, such as those found in *E. coli*, are a more recent evolutionary development that have undergone significant changes in sequence. They are smaller (38 kDa), containing only the catalytic core, and are 50 and ten residues shorter on the amino and carboxyl ends, respectively. The type II proteins also have several internal insertions and deletions when compared with the type I enzymes.

E. coli DHOase, recently solved to a resolution of 1.7 Å (Thoden *et al.*, 2001), is the first structure of any dihydroorotase of either type. In accord with the predictions (Holm & Sander, 1997), the architecture of the enzyme resembles the catalytic core of urease. Although the *E. coli* and mammalian DHOases have been reported to have one catalytic zinc per monomer (Washabaugh & Collins, 1984, 1986; Kelly *et al.*, 1986; Brown & Collins, 1991; Huang *et al.*, 1999), the structure clearly showed that the *E. coli* enzyme has a carboxylated lysine and a binuclear center with two zinc ions at the active site. A novel catalytic mechanism has been proposed based on the juxtaposition of residues near the bound substrate.

The DHOase from *A. aeolicus*, an extreme hyperthermophile occupying the deepest branch of the eubacterial phylogenetic tree, has been classified (Fields *et al.*, 1999) on the basis of its size and sequence as a type I DHOase. Compared with the *E. coli* enzyme, *A. aeolicus* DHOase is 52 residues longer on the amino end, seven residues longer on the carboxyl end, has only 20% sequence identity and contains several large insertions and deletions. *A. aeolicus* DHOase is more closely related to the type I mammalian enzyme in size and sequence. It is likely that the type I and II enzymes have a similar catalytic core, active-site configuration and catalytic mechanism, but otherwise may be quite different.

2. Cloning, expression and purification

The *pyrC* gene, identified in the recently completed *A. aeolicus* genome project (Deckert *et al.*, 1998), encodes a 422-residue protein with a molecular mass of 46 379. A *BLAST* search of the entire genome failed to retrieve any other DHOase homologues. The zinc-signature sequence and all the other residues implicated in catalysis were conserved in the putative *A. aeolicus* enzyme (Williams *et al.*, 1995; Zimmermann *et al.*, 1995; Thoden *et al.*, 2001). The 1.3 kbp *pyrC* gene was amplified by PCR and inserted into the *Bam*HI/*Nco*I sites of pRSETC (Invitrogen), an expression vector that incorporates a His tag on the amino end of the recombinant protein to facilitate purification. After sequencing the resulting construct, pAapyrC, to verify the fidelity of the PCR amplification, the plasmid was cotransformed with pSJS1240 into *E. coli* strain BL21(DE3) (Novagen). The helper plasmid pSJS1240 (Kim *et al.*, 1998) encodes two tRNA synthetases that recognize codons that are common in *A. aeolicus* DHOase but occur rarely in *E. coli*.

While a significant fraction of the protein was aggregated, the soluble protein was expressed at high levels, 100 mg per liter of culture, and could be purified to homogeneity (Fig. 2) in a single step by Ni^{2+} affinity chromatography. The molecular mass of the protein, determined by SDS–PAGE, was 49 kDa, close to the value predicted from the deduced amino-acid sequence when the size of the His tag is taken into consideration. Zinc analysis by atomic absorption spectroscopy (University of Georgia Chemical Analysis Facility, Athens, Georgia, USA) showed that there was 0.76 mole zinc per mole of monomer. It is possible that a second loosely bound metal ion is removed by dialysis prior to analysis. Other metals, including Co^{2+}, Ni^{2+}, Mn^{2+}, Mg^{2+}, Cd^{2+} and Pb^{2+}, represented less than 0.01 mol per mole of monomer. Interestingly, the enzyme was found to contains 1.5 mole of potassium ion per mole of monomer despite extensive pre-dialysis.

The selenomethionine DHOase derivative was prepared using the procedure devised by Hendrickson *et al.* (1990). The plasmids pAapyrC and pSJS1240 were cotransformed into the *E. coli* methionine auxotroph strain B834 (DE3) (Novagen) and the cells were grown in minimal media in which methionine was replaced with selenomethionine (Hendrickson *et al.*, 1990). The same isolation protocol was followed and the yield of pure protein was comparable with that of the wild-type enzyme. An X-ray fluorescence scan confirmed that

Figure 3
Crystal of *A. aeolicus* dihydroorotase. An *A. aeolicus* DHOase crystal, 0.2 mm in the longest dimension, was obtained by the hanging-drop method at 295 K using 1 *M* lithium sulfate, 100 m*M* Tris acetate pH 8.5 and 10 m*M* nickel chloride hexahydrate in the well. The protein (6.7 mg ml^{-1}, 0.002 ml) in 50 m*M* HEPES pH 7.4 was diluted 1:2 with the precipitating solution. The crystal was grown for ten weeks.

selenomethionine was present and unoxidized.

3. Crystallization

The crystals were obtained by the hanging-drop method, using 2–6 µl of 6–7 mg ml^{-1} DHOase in 50 mM HEPES pH 7.4. Initially, crystallization conditions were sought using the sparse-matrix Crystal Screens 1 and 2 form Hampton Research (Jancarik & Kim, 1991). Small crystals were observed under 11 different conditions, but the best quality crystals were obtained at 295 K using 1 M lithium sulfate, 100 mM Tris acetate pH 8.5 and 10 mM nickel chloride hexahydrate. Well formed diamond-shaped crystals (Fig. 3) appeared within 1 d and continued to grow to 0.5 mm in the longest dimension over a period of 10–12 weeks. The selenomethionine DHOase derivative was crystallized under the same conditions.

4. Data collection and processing

A crystal of native DHOase (0.2 × 0.3 × 0.4 mm) was soaked in 20% trehalose, 6.4 M lithium sulfate, 100 mM Tris acetate pH 8.5 for 1 min and mounted on the goniometer in the Cryostream. The data set was collected with an R-AXIS IV image-plate detector mounted on a Rigaku R200 X-ray generator equipped with MSC Osmic blue optics and running at 50 kV and 100 mA. Diffraction data were collected over 1° increments in φ for a total rotation of 194°. The data were indexed, integrated and scaled with *XGEN* (Howard *et al.*, 1987). The crystal belongs to the monoclinic space group $C2$, with unit-cell parameters a = 119.83, b = 88.00, c = 55.24 Å, β = 99.0° and a mosaic spread of 0.6°. The packing density suggests there is a monomer (V_M = 2.8 Å^3 Da^{-1}) in the asymmetric unit as opposed to a dimer (V_M = 1.5 Å^3 Da^{-1}). The calculated solvent content of the crystals is 54%. The data-collection statistics are summarized in Table 1. Data were also collected as described above on crystals soaked for 20 h in the substrate carbamoyl aspartate (0.1 mM) in 8.2 M lithium sulfate, 100 mM Tris acetate pH 8.5. While at this point there is no conclusive evidence for substrate binding to the crystals, the mean isomorphous difference is 14% compared with 10% for two native data sets. The DHOase reaction is reversible and both carbamoyl aspartate and dihydroorotate bind with high affinity to the active site of mammalian DHOase (Christopherson & Jones, 1980; Christopherson *et al.*, 1981). At the pH at which the crystals are grown, the ratio of carbamoyl aspartate to dihydroorotate in the equilibrium mixture is 211. Thus, it is likely that the crystals soaked in carbamoyl aspartate represent the Michaelis complex.

Table 1
Diffraction data statistics.

Values in parentheses are for the outer 0.1 Å shell.

No. of measured reflections	100568
No. of unique reflections	37582
Completeness (%)	97 (85)
Resolution (Å)	1.99
$I/\sigma(I)$	12.35 (1.4)
R_{sym} (%)	9.3 (56.2)

Preliminary diffraction data were collected on small native and selenomethionine crystals to 2.5 Å at the IMCA-CAT beamline at the Advanced Photon Source (APS) at the Argonne National Laboratory. The resolution limit of the diffraction pattern of the Se-Met crystals is significantly more asymmetric in certain directions than that of the native crystals, although annealing the crystals prior to data collection in saturated lithium sulfate make the diffraction pattern appreciably more isometric. It should be possible to collect high-resolution data for MAD phasing from these annealed crystals once additional synchrotron time is available. Attempts to solve the structure by molecular replacement using the recently published coordinates of *E. coli* DHOase (Thoden *et al.*, 2001), which has only 20% sequence identity with the *A. aeolicus* DHOase, were unsuccessful, reinforcing the idea that the structures are different.

We would like to acknowledge Drs Karl O. Stetter and Robert Huber (Regensburg University, D-93053 Regensburg, Germany) for the generous gift of *A. aeolicus* chromosomal DNA and Dr Andrew Howard (APS) for help and advice in X-ray data collection. This research was supported by an NSF grant MCB-9810325 and NIH grants GM 47399 and GM/CA60371.

References

Brown, D. C. & Collins, K. D. (1991). *J. Biol. Chem.* **266**, 1597–1604.
Christopherson, R. I. & Jones, M. E. (1980). *J. Biol. Chem.* **255**, 3358–3370.
Christopherson, R. I., Traut, T. W. & Jones, M. E. (1981). *Curr. Topics Cell. Reg.* **18**, 59–77.
Coleman, P., Suttle, D. & Stark, G. (1977). *J. Biol. Chem.* **252**, 6379–6385.
Deckert, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L., Graham, D. E., Overbeek, R., Snead, M. A., Keller, M., Aujay, M., Huber, R., Feldman, R. A., Short, J. M., Olsen, G. J. & Swanson, R. V. (1998). *Nature (London)*, **392**, 353–358.
Fields, C., Brichta, D., Shephardson, M., Farinha, M. & O'Donovan, G. (1999). *Paths Pyrimidines*, **7**, 49–63.
Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.
Holm, L. & Sander, C. (1997). *Proteins*, **28**, 72–82.
Howard, A. J., Gilliland, G. L., Finzel, B., Poulos, T. L., Ohlendorf, D. H. & Salemme, F. R. (1987). *J. Appl. Cryst.* **20**, 383–387.
Huang, D. T., Thomas, M. A. & Christopherson, R. I. (1999). *Biochemistry*, **38**, 9964–9970.
Hughes, L. E., Hooshdaran, M. Z. & O'Donovan, G. A. (1999). *Curr. Microbiol.* **39**, 175–179.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Kelly, R. E., Mally, M. I. & Evans, D. R. (1986). *J. Biol. Chem.* **261**, 6073–6083.
Kim, G.-J. & Kim, H.-S. (1998). *Biochem. J.* **330**, 295–302.
Kim, R., Sandler, J., Goldman, S., Yokota, H., Clarke, A. J. & Kim, S. H. (1998). *Biochem. J.* **330**, 295–302.
Musmanno, L. A., Maley, J. A. & Davidson, J. N. (1991). *Gene*, **99**, 211–216.
Schurr, M. J., Vickrey, J. F., Kumar, A. P., Campbell, A. L., Cunin, R., Benjamin, R. C., Shanley, M. S. & O'Donovan, G. A. (1995). *J. Bacteriol.* **177**, 1751–1759.
Souciet, J.-L., Nagy, M., Legouar, M., Lacroute, F. & Potier, S. (1989). *Gene*, **79**, 59–70.
Thoden, J. B., Phillips, G. N. Jr, Neal, T. M., Raushel, F. M. & Holden, H. M. (2001). *Biochemistry*, **40**, 6989–6997.
Van de Casteele, M., Chen, P., Roovers, M., Legrain, C. & Glansdorff, N. (1997). *J. Bacteriol.* **179**, 3470–3481.
Vickrey, J. F., Hervé, G. & Evans, D. R. (2002). Submitted.
Washabaugh, M. W. & Collins, K. D. (1984). *J. Biol. Chem.* **259**, 3293–3298.
Washabaugh, M. W. & Collins, K. D. (1986). *J. Biol. Chem.* **261**, 5920–5929.
Williams, N. K., Manthey, M. K., Hambley, T. W., O'Donoghue, S. I., Keegan, M., Chapman, B. E. & Christopherson, R. I. (1995). *Biochemistry*, **34**, 11344–11352.
Williams, N. K., Simpson, R. J., Moritz, R. L., Peide, Y., Crofts, L., Minasian, E., Leach, S. J., Wake, R. G. & Christopherson, R. I. (1990). *Gene*, **94**, 283–288.
Wilson, D. K., Rudolph, F. B. & Quiocho, F. A. (1991). *Science*, **252**, 1278–1284.
Zimmermann, B. H. & Evans, D. R. (1993). *Biochemistry*, **32**, 1519–1527.
Zimmermann, B. H., Kemling, N. M. & Evans, D. R. (1995). *Biochemistry*, **34**, 7038–7046.