



Crystallization and X-ray diffraction analysis of an L-arabinonate dehydratase from *Rhizobium leguminosarum* bv. *trifolii* and a D-xylonate dehydratase from *Caulobacter crescentus*

Mohammad Mubinur Rahman,^a Martina Andberg,^b Anu Koivula,^b Juha Rouvinen^a and Nina Hakulinen^{a*}

Received 4 May 2016

Accepted 25 June 2016

Edited by J. Newman, Bio21 Collaborative Crystallisation Centre, Australia

Keywords: L-arabinonate dehydratase; D-xylonate dehydratase; IlvD/EDD enzymes; [Fe–S] cluster; *Rhizobium leguminosarum* bv. *trifolii*; *Caulobacter crescentus*.

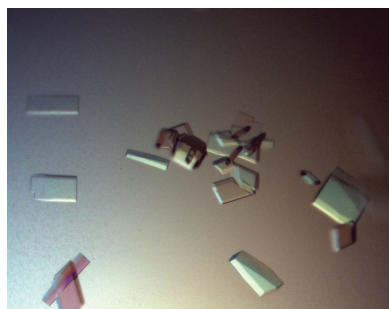
Supporting information: this article has supporting information at journals.iucr.org/f

^aDepartment of Chemistry, University of Eastern Finland, Joensuu Campus, PO Box 111, FIN-80101 Joensuu, Finland, and ^bVTT Technical Research Centre of Finland Ltd, PO Box 1000, FIN-02044 VTT Espoo, Finland. *Correspondence e-mail: nina.hakulinen@uef.fi

L-Arabinonate dehydratase (EC 4.2.1.25) and D-xylonate dehydratase (EC 4.2.1.82) are two enzymes that are involved in a nonphosphorylative oxidation pathway of pentose sugars. L-Arabinonate dehydratase converts L-arabinonate into 2-dehydro-3-deoxy-L-arabinonate, and D-xylonate dehydratase catalyzes the dehydration of D-xylonate to 2-dehydro-3-deoxy-D-xylonate. L-Arabinonate and D-xylonate dehydratases belong to the IlvD/EDD family, together with 6-phosphogluconate dehydratases and dihydroxyacid dehydratases. No crystal structure of any L-arabinonate or D-xylonate dehydratase is available in the PDB. In this study, recombinant L-arabinonate dehydratase from *Rhizobium leguminosarum* bv. *trifolii* (RlArDHT) and D-xylonate dehydratase from *Caulobacter crescentus* (CcXyDHT) were heterologously expressed in *Escherichia coli* and purified by the use of affinity chromatography followed by gel-filtration chromatography. The purified proteins were crystallized using the hanging-drop vapour-diffusion method at 293 K. Crystals of RlArDHT that diffracted to 2.40 Å resolution were obtained using sodium formate as a precipitating agent. They belonged to space group $P2_1$, with unit-cell parameters $a = 106.07$, $b = 208.61$, $c = 147.09$ Å, $\beta = 90.43^\circ$. Eight RlArDHT molecules (two tetramers) in the asymmetric unit give a V_M value of $3.2 \text{ \AA}^3 \text{ Da}^{-1}$ and a solvent content of 62%. Crystals of CcXyDHT that diffracted to 2.66 Å resolution were obtained using sodium formate and polyethylene glycol 3350. They belonged to space group $C2$, with unit-cell parameters $a = 270.42$, $b = 236.13$, $c = 65.17$ Å, $\beta = 97.38^\circ$. Four CcXyDHT molecules (a tetramer) in the asymmetric unit give a V_M value of $4.0 \text{ \AA}^3 \text{ Da}^{-1}$ and a solvent content of 69%.

1. Introduction

Lignocellulosic biomass is an important source of hexose and pentose sugars for the biorefinery industry. Apart from biofuels, the bioconversion of lignocellulosic hexose and pentose sugars can provide various platform chemicals which can be used as precursors for polymeric materials (Menon & Rao, 2012). The hemicellulose part of the lignocellulose biomass contains a significant amount of D-xylose and L-arabinose (Schädel *et al.*, 2010). In nature, some archaea, bacteria and fungi are known to metabolize L-arabinose (Boguta *et al.*, 2014; Brouns *et al.*, 2006; Dien *et al.*, 1996) and D-xylose (Okamoto *et al.*, 2014; Gu *et al.*, 2010) by using different nonphosphorylated oxidative pathways. In the Weimberg pathway, L-arabinose or D-xylose is converted to α -ketoglutarate, which is an intermediate metabolite in the tricarboxylic acid (TCA) cycle (Weimberg, 1961), or, in the



OPEN  ACCESS

Table 1
Macromolecule-production information.

	<i>RIArDHT</i>	<i>CcXyDHT</i>
Source organism	<i>R. leguminosarum</i> bv. <i>trifolii</i> (GenBank accession No. KT260159)	<i>C. crescentus</i> (GenBank accession No. KT260154)
DNA source	Synthetically made	Synthetically made
Forward primer†	CACGTCCATGGACTGGTCTCATCCACAATTCGAGAAGAAAAAAGCT-GAATGGCCG	ACGTCCATGGACTGGTCTCATCCACAATTCGAGAAGTCTAATCGCACCC-CGGCTC
Reverse primer‡	GACGAATTCTCAATG	GTCGACGAATTCTCAATG
Cloning vector	pBAT4	pBAT4
Expression vector	pBAT4	pBAT4
Expression host	<i>E. coli</i> BL21(DE3)	<i>E. coli</i> BL21(DE3)
Complete amino-acid sequence of the construct produced§	MDW [§] SH [‡] P [‡] Q [‡] FEK [‡] KK [‡] AEW [‡] PR [‡] KL [‡] RS [‡] Q [‡] E [‡] W [‡] Y [‡] G [‡] T [‡] S [‡] R [‡] D [‡] V [‡] I [‡] Y [‡] H [‡] R [‡] G [‡] W [‡] L [‡] K [‡] N [‡] Q [‡] G [‡] Y [‡] P [‡] H [‡] D [‡] L [‡] - FDGRPVIGILNTWSDMTPCNGHLRELAEKVKAGVWEAGGFLEVPVF- SASENTFRPTAMMYRNLAALAVEEAIRGQPMDCVLLVGCDKTTPSL- LMGAASCDLPSIVVTGGPMLNGYFRGERVSGTHLWKFSEMVKAGEM- TQAEFLEAEASMSRSGTCNTMGTA [‡] STMASMAEALGMALSGNAAI [‡] PG- VDSRRKVM [‡] Q [‡] L [‡] T [‡] GRR [‡] IV [‡] Q [‡] M [‡] V [‡] K [‡] D [‡] L [‡] K [‡] P [‡] SE [‡] IM [‡] T [‡] K [‡] Q [‡] A [‡] FENAI [‡] R [‡] T [‡] NAI [‡] GG- STNAVIHLLAI [‡] AGRV [‡] GID [‡] LSL [‡] DD [‡] W [‡] RC [‡] GR [‡] D [‡] V [‡] PT [‡] IV [‡] N [‡] LP [‡] SG [‡] K [‡] Y [‡] L [‡] MEE- FFYAGGLPV [‡] L [‡] K [‡] RL [‡] GEAG [‡] LL [‡] H [‡] K [‡] D [‡] AL [‡] T [‡] V [‡] S [‡] GET [‡] V [‡] W [‡] DEV [‡] K [‡] D [‡] V [‡] N [‡] W [‡] NED [‡] V [‡] I- LPAEKAL [‡] T [‡] SSGG [‡] I [‡] V [‡] LR [‡] GN [‡] L [‡] PK [‡] G [‡] AV [‡] L [‡] K [‡] PSA [‡] AS [‡] PH [‡] LV [‡] H [‡] K [‡] GRA [‡] V [‡] VF [‡] E- DIDDYK [‡] AK [‡] IND [‡] N [‡] LD [‡] IDEN [‡] CM [‡] VM [‡] K [‡] NG [‡] CP [‡] K [‡] YP [‡] GMA [‡] EV [‡] GN [‡] M [‡] GL [‡] PP [‡] K [‡] V- LKK [‡] ILD [‡] M [‡] VR [‡] IS [‡] D [‡] AR [‡] MS [‡] T [‡] AY [‡] GT [‡] V [‡] VL [‡] HT [‡] SP [‡] EAA [‡] V [‡] GG [‡] PL [‡] AV [‡] V [‡] K [‡] NG [‡] DM [‡] I- ELD [‡] V [‡] PN [‡] RL [‡] H [‡] LD [‡] IS [‡] DE [‡] EL [‡] ARR [‡] LA [‡] EW [‡] Q [‡] PN [‡] HD [‡] L [‡] PT [‡] SG [‡] Y [‡] AF [‡] L [‡] H [‡] Q [‡] H [‡] VEGA [‡] - DTGAD [‡] LD [‡] FL [‡] K [‡] GR [‡] NAV [‡] G [‡] K [‡] DSH	MDW [§] SH [‡] P [‡] Q [‡] FEK [‡] SN [‡] R [‡] T [‡] PR [‡] R [‡] FR [‡] S [‡] R [‡] D [‡] W [‡] F [‡] D [‡] N [‡] P [‡] D [‡] H [‡] D [‡] M [‡] T [‡] AL [‡] Y [‡] LF [‡] RF [‡] M [‡] NY [‡] G [‡] IT [‡] PEE- L [‡] RS [‡] G [‡] P [‡] I [‡] I [‡] G [‡] I [‡] A [‡] Q [‡] T [‡] G [‡] S [‡] D [‡] I [‡] SP [‡] C [‡] N [‡] R [‡] I [‡] H [‡] L [‡] D [‡] L [‡] V [‡] Q [‡] R [‡] V [‡] R [‡] D [‡] G [‡] I [‡] R [‡] D [‡] A [‡] G [‡] G [‡] I [‡] P [‡] M [‡] E [‡] F [‡] P [‡] V- H [‡] P [‡] I [‡] F [‡] E [‡] N [‡] C [‡] R [‡] R [‡] P [‡] T [‡] A [‡] L [‡] D [‡] R [‡] N [‡] L [‡] S [‡] Y [‡] L [‡] G [‡] L [‡] V [‡] E [‡] T [‡] H [‡] G [‡] Y [‡] P [‡] I [‡] D [‡] A [‡] V [‡] L [‡] T [‡] T [‡] G [‡] C [‡] K [‡] T [‡] T [‡] P [‡] A- G [‡] I [‡] M [‡] A [‡] A [‡] T [‡] V [‡] N [‡] I [‡] P [‡] A [‡] I [‡] V [‡] L [‡] S [‡] G [‡] G [‡] P [‡] M [‡] L [‡] D [‡] G [‡] W [‡] H [‡] E [‡] N [‡] E [‡] L [‡] V [‡] G [‡] S [‡] G [‡] T [‡] V [‡] I [‡] W [‡] S [‡] R [‡] R [‡] K [‡] L [‡] A [‡] A [‡] G [‡] E- I [‡] T [‡] E [‡] E [‡] F [‡] I [‡] D [‡] R [‡] A [‡] A [‡] S [‡] A [‡] P [‡] S [‡] A [‡] G [‡] H [‡] C [‡] N [‡] T [‡] M [‡] G [‡] T [‡] A [‡] S [‡] T [‡] M [‡] N [‡] A [‡] V [‡] A [‡] E [‡] L [‡] G [‡] L [‡] S [‡] L [‡] T [‡] G [‡] C [‡] A [‡] I [‡] P- A [‡] P [‡] Y [‡] R [‡] E [‡] R [‡] G [‡] Q [‡] M [‡] A [‡] Y [‡] K [‡] T [‡] G [‡] Q [‡] R [‡] I [‡] V [‡] D [‡] L [‡] A [‡] Y [‡] D [‡] V [‡] K [‡] P [‡] L [‡] D [‡] I [‡] L [‡] T [‡] K [‡] Q [‡] A [‡] FENAI [‡] AL [‡] V [‡] AA [‡] A [‡] G- G [‡] S [‡] T [‡] NA [‡] Q [‡] P [‡] H [‡] IV [‡] AM [‡] AR [‡] H [‡] AG [‡] VEI [‡] T [‡] AD [‡] W [‡] RA [‡] A [‡] Y [‡] D [‡] I [‡] PL [‡] IV [‡] N [‡] M [‡] Q [‡] P [‡] AG [‡] Y [‡] L [‡] GER- F [‡] HR [‡] AG [‡] AP [‡] AV [‡] L [‡] W [‡] E [‡] L [‡] L [‡] Q [‡] Q [‡] R [‡] L [‡] H [‡] G [‡] D [‡] V [‡] L [‡] T [‡] V [‡] T [‡] G [‡] K [‡] T [‡] M [‡] SE [‡] N [‡] L [‡] Q [‡] GRE [‡] T [‡] S [‡] D [‡] RE [‡] VI- F [‡] P [‡] Y [‡] HE [‡] LA [‡] E [‡] K [‡] AG [‡] F [‡] L [‡] V [‡] L [‡] K [‡] N [‡] L [‡] F [‡] D [‡] FA [‡] IM [‡] K [‡] SS [‡] VI [‡] GE [‡] EF [‡] R [‡] K [‡] RY [‡] LS [‡] Q [‡] PG [‡] Q [‡] E [‡] - V [‡] FE [‡] ARA [‡] IV [‡] FD [‡] G [‡] S [‡] DD [‡] Y [‡] H [‡] K [‡] R [‡] IND [‡] PA [‡] L [‡] E [‡] ID [‡] ERC [‡] IL [‡] VI [‡] R [‡] G [‡] AG [‡] P [‡] IG [‡] W [‡] PG [‡] SAE- V [‡] V [‡] N [‡] Q [‡] P [‡] D [‡] H [‡] L [‡] L [‡] K [‡] G [‡] I [‡] MS [‡] L [‡] P [‡] T [‡] L [‡] G [‡] D [‡] GR [‡] Q [‡] S [‡] T [‡] AD [‡] SP [‡] IL [‡] N [‡] AS [‡] PESA [‡] I [‡] G [‡] GG- L [‡] SW [‡] L [‡] R [‡] T [‡] G [‡] D [‡] T [‡] IR [‡] IDL [‡] T [‡] GR [‡] CD [‡] AL [‡] V [‡] DE [‡] AT [‡] IA [‡] ARK [‡] Q [‡] D [‡] G [‡] IP [‡] AV [‡] P [‡] AT [‡] M [‡] T [‡] P [‡] W [‡] Q- E [‡] I [‡] Y [‡] RA [‡] H [‡] AS [‡] QL [‡] D [‡] T [‡] G [‡] V [‡] LE [‡] FA [‡] V [‡] K [‡] Y [‡] Q [‡] DL [‡] AA [‡] KL [‡] PR [‡] H [‡] NH

† The NcoI site is underlined. ‡ The EcoRI site is underlined. § The *Strep*-Tag II is underlined.

Dahms pathway, the pentose sugars are converted to glycolaldehyde and pyruvate (Dahms, 1974). In both pathways, L-arabinonate dehydratase (EC 4.2.1.25) catalyzes the removal of one water molecule from L-arabinonate to produce 2-dehydro-3-deoxy-L-arabinonate (Fig. 1a). Similarly, D-xylonate dehydratase (EC 4.2.1.82) catalyzes the removal of one water molecule from D-xylonate to produce 2-dehydro-3-deoxy-D-xylonate (Fig. 1b) (Stephens *et al.*, 2007; Dilworth *et al.*, 1986; Weimberg, 1961). Several bacterial and yeast strains have been engineered to metabolize L-arabinose or D-xylose through the oxidative pathway and to produce biofuels, in addition to other value-added industrial chemicals (Kurosawa *et al.*, 2015; Jin *et al.*, 2014; Demeke *et al.*, 2013; Xiao *et al.*,

2012; Bettiga *et al.*, 2009; Meijnen *et al.*, 2009; Matsushika *et al.*, 2009).

L-Arabinonate dehydratase from *Rhizobium leguminosarum* bv. *trifolii* and D-xylonate dehydratase from *Caulobacter crescentus* belong to the IlvD/EDD enzyme family. Enzymes in this family contain iron–sulfur [Fe–S] clusters as a prosthetic group (Andberg *et al.*, 2016; Nunn *et al.*, 2010; Stephens *et al.*, 2007) and, using spectroscopic methods, active enzymes from this family have been found to contain either a [4Fe–4S] cluster (Watanabe *et al.*, 2006; Rodriguez *et al.*, 1996; Flint *et al.*, 1993) or a [2Fe–2S] cluster (Flint & Emptage, 1988) in the active sites. There is only one crystal structure in the PDB from the IlvD/EDD family (PDB entry 2gp4; Southeast

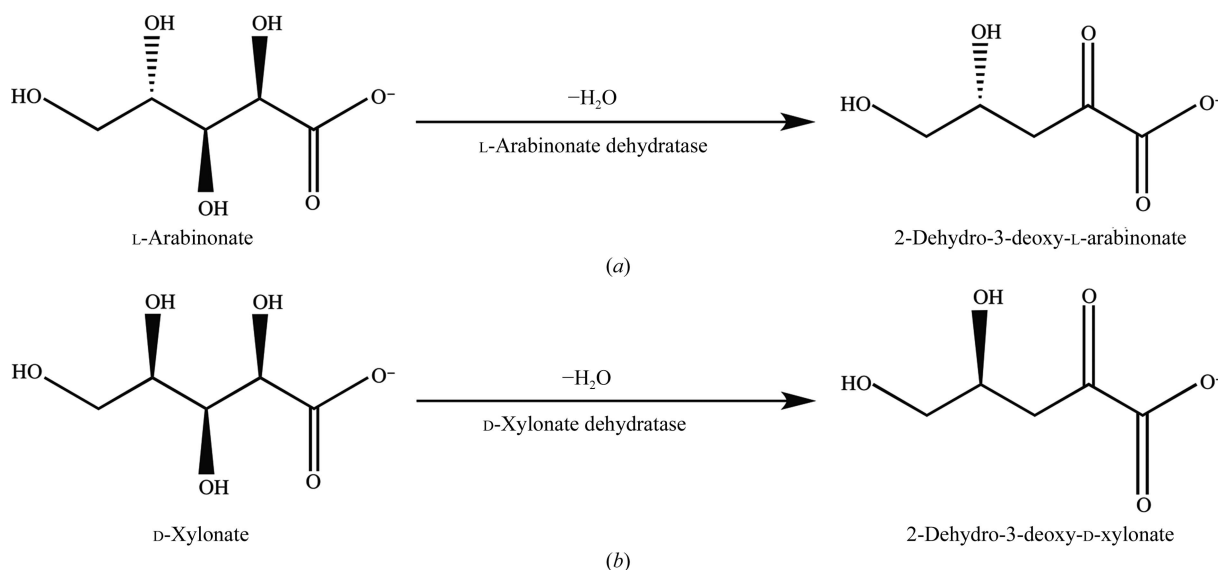


Figure 1
Catalytic reactions of (a) *RIArDHT* and (b) *CcXyDHT*.

Table 2
Crystallization.

	<i>RIArDHT</i>	<i>CcXyDHT</i>
Method	Hanging-drop vapour diffusion	Hanging-drop vapour diffusion
Plate type	24-well cell-culture plate	24-well cell-culture plate
Temperature (K)	293	293
Protein concentration (mg ml ⁻¹)	9.0	7.5
Buffer composition of protein solution	50 mM Tris-HCl pH 7.5, 5 mM MgCl ₂	50 mM Tris-HCl pH 7.5, 5 mM MgCl ₂
Composition of reservoir solution	4.0 M sodium formate, 0.1 M MES pH 7.0	3.8 M sodium formate pH 7.0, 0.1 M TES pH 7.0, 5 mM magnesium formate, 4% PEG 3350
Volume and ratio of drop	4 µl, 1:1	4 µl, 1:1
Volume of reservoir (µl)	500	500

Collaboratory for Structural Genomics, unpublished work) and it lacks the [Fe-S] cluster. Therefore, structural studies of enzymes belonging to this family are needed in order to understand their structure-function relationship and the role of the [Fe-S] cluster. Here, we describe the overproduction, purification, crystallization and X-ray diffraction analysis of two enzymes from the IlvD/EDD family.

2. Materials and methods

2.1. Macromolecule production

The *araD* gene encoding *R. leguminosarum* L-arabinonate dehydratase (*RIArDHT*) and the *xylD* gene encoding *C. crescentus* D-xylonate dehydratase (*CcXyDHT*) were purchased as codon-optimized synthetic genes from GeneArt, Germany and a *Strep-Tag II* (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys) was added at the N-terminus (the deposited GenBank accession numbers for *araD* and *xylD* are KT260159 and KT260154, respectively). The genes were cloned into a pBAT4 expression vector, as described in Andberg *et al.* (2016). Other information on macromolecule production is given in Table 1.

For protein production, *Escherichia coli* BL21(DE3) expression cells were transformed with the recombinant plasmids. Large-scale protein production was performed in 2 l shake flasks containing 500 ml Luria-Bertani (LB) culture medium with 100 µg ml⁻¹ ampicillin. The culture flasks were incubated at 303 K in a shaker incubator at 180 rev min⁻¹. When the OD_{600 nm} reached 0.5–0.6, protein expression was induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM, and cultivation was continued under identical conditions overnight (Andberg *et al.*, 2016). Cells were harvested from the culture medium by centrifugation at 4000 rev min⁻¹ at 277 K for 25 min. The cell pellets were suspended in an extraction buffer consisting of 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM MgCl₂, 1 mM DTT, protease inhibitor (EDTA-free, Roche) and 100 µg ml⁻¹ lysozyme (Sigma-Aldrich, Germany). Intracellular proteins were isolated by centrifugation at 18 500 rev min⁻¹, followed by a one-step freeze-thawing of the suspended cells at 193 K and sonication.

Crude protein samples were loaded onto a StrepTrap HP 5 ml column (GE Healthcare, Sweden) equilibrated with binding buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM MgCl₂). *Strep-Tag II*-bound proteins were eluted with elution buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM MgCl₂,

2.5 mM D-desthiobiotin). A final polishing step was conducted by gel-filtration chromatography on a Superdex 200 HR 10/30 column (GE Healthcare, Sweden) equilibrated with a buffer consisting of 50 mM Tris-HCl pH 7.5, 5 mM MgCl₂. Purified samples were concentrated using an Amicon Ultra-4 centrifugal filter device with a molecular-weight cutoff of 30 kDa (Merck Millipore, Germany). The molecular weights and the purity of the enzymes were checked by SDS-PAGE with silver staining in a PhastSystem (GE Healthcare, Sweden). The enzyme activities were checked by the thiobarbituric acid (TBA) assay and the semicarbazide assay (Andberg *et al.*, 2016).

2.2. Crystallization

Crystallization of the purified enzymes was performed by the hanging-drop vapour-diffusion method in a 24-well cell-culture plate (Greiner Bio-One, Germany) at 293 K. The following commercial crystallization kits were used: Crystal Screen, Crystal Screen 2, PEG/Ion, SaltRx 1, SaltRx 2, Index HT, PEGRx 1 and PEGRx 2 (Hampton Research). Droplets were prepared by mixing protein sample and crystallization reagent in a 1:1 ratio on a siliconized glass cover slide. The cover slides were then turned drop-side down against 500 µl crystallization reservoir reagent and sealed with grease to ensure that they were airtight. The plates were stored at 293 K. A microseeding technique (Bergfors, 2003) was applied to improve the quality of the *CcXyDHT* crystals. An initial condition that consisted of 3.6 M sodium formate, 5% PEG 3350 gave clusters of *CcXyDHT* crystals. A small cluster of crystals was removed in 4 µl crystallization reagent on a glass cover slide, crushed with a micro needle, diluted with 46 µl crystallization reagent to a final volume of 50 µl and sonicated for 3 min in a sonication bath (FinnSonic, Finland). A dog hair was dipped into the seed-stock solution and streaked through pre-equilibrated crystallization drops. Crystallization information is shown in Table 2.

2.3. Data collection and processing

RIArDHT crystals were mounted on a nylon loop and soaked in a cryoprotectant solution consisting of 4.0 M sodium formate, 0.1 M MES pH 7.0, 15% glycerol. *CcXyDHT* crystals were soaked with 3.9 M sodium formate pH 7.0, 0.1 M TES pH 7.0, 5 mM magnesium formate, 4% PEG 3350, 40 mM calcium D-xylonate. Data collection for *RIArDHT* crystals was carried

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

	<i>RIArDHT</i>	<i>CcXyDHT</i>
Diffraction source	ID14-2, ESRF, France	I04, DLS, England
Wavelength (Å)	0.97957	0.97949
Temperature (K)	100	100
Detector	CCD	Pilatus 6M
Crystal-to-detector distance (mm)	335.51	519.95
Rotation range per image (°)	0.25	0.5
Total rotation range (°)	180	180
Exposure time per image (s)	0.1	0.1
Space group	$P2_1$	$C2$
a, b, c (Å)	106.07, 208.61, 147.09	270.42, 236.13, 65.17
α, β, γ (°)	90, 90.43, 90	90, 97.38, 90
Mosaicity (°)	0.65	0.7
Resolution range (Å)	40.00–2.40 (2.50–2.40)	59.69–2.66 (2.73–2.66)
Total No. of reflections	779327 (88849)	393801 (60419)
No. of unique reflections	246940 (28307)	114638 (17945)
Completeness (%)	99.3 (99.0)	99.0 (95.8)
Multiplicity	3.2 (3.1)	3.4 (3.4)
$\langle I/\sigma(I) \rangle$	12.3 (2.0)	17.2 (2.1)
$CC_{1/2}$	99.5 (57.8)	99.9 (75.1)
R_{meas} (%)	10.8 (81.0)	7.0 (82.1)
Overall B factor from Wilson plot (Å ²)	43	63

out at 100 K on beamline ID14-2 at the ESRF in France at a wavelength of 0.97957 Å. Data from *CcXyDHT* crystals were collected remotely on beamline I04 at Diamond Light Source (DLS) in England at a wavelength of 0.97949 Å. Data processing was carried out by *XDS* (Kabsch, 2010). Detailed data-collection and processing statistics are shown in Table 3.

3. Results and discussion

Purified dehydratase enzymes contain an [Fe–S] cluster, which is necessary for enzyme activity (Andberg *et al.*, 2016). Both the *RIArDHT* and the *CcXyDHT* protein solutions were dark brown in colour, which was a good indication that the [Fe–S] clusters had been maintained. In addition, the purified proteins were found to be active by an enzyme-activity assay as described previously (Andberg *et al.*, 2016). Based on SDS–PAGE analysis (Supplementary Fig. S1), the molecular weights of the purified *RIArDHT* and *CcXyDHT* were

approximately 64 kDa, and are in good correspondence with the theoretical molecular weights. The calculated molecular weights are 63.738 and 64.314 kDa (without the [Fe–S] cluster) for *RIArDHT* and *CcXyDHT*, respectively.

RIArDHT was crystallized in 4.0 M sodium formate, 0.1 M MES pH 7.0, which was a modification of condition No. 33 (4 M sodium formate) from Crystal Screen. Crystals appeared within two weeks (Fig. 2a) and diffracted to 2.4 Å resolution (Fig. 3a). The space group was $P2_1$, with unit-cell parameters $a = 106.07$, $b = 208.61$, $c = 147.09$ Å, $\beta = 97.43^\circ$. Assuming eight molecules in the asymmetric unit, the calculated Matthews coefficient (V_M) was $3.2 \text{ \AA}^3 \text{ Da}^{-1}$, which corresponds to a solvent content of 62% (Matthews, 1968).

In the initial crystallization screening of *CcXyDHT*, hexagonal bipyramidal crystals were obtained in various conditions. Most of the conditions contained salts, while others contained PEGs as a precipitating agent. However, the crystals showed only poor diffraction to 6 Å resolution. Additive and detergent screens were also tried, but no real improvement was observed. A modified condition that consisted of 3.6 M sodium formate pH 7.0, 5% PEG 3350 gave a cluster of bar-shaped crystals. Microseeding, together with further modification of the crystallization condition (3.8 M sodium formate pH 7.0, 0.1 M TES pH 7.0, 5 mM magnesium formate, 4% PEG 3350) gave crystals within four weeks (Fig. 2b) which diffracted to 2.66 Å resolution (Fig. 3b). The crystals belonged to space group $C2$, with unit-cell parameters $a = 270.42$, $b = 236.13$, $c = 65.17$ Å, $\beta = 97.38^\circ$. Assuming four molecules in the asymmetric unit, the calculated Matthews coefficient (V_M) was $4.0 \text{ \AA}^3 \text{ Da}^{-1}$, which corresponds to a solvent content of 69% (Matthews, 1968).

The structures of *RIArDHT* and *CcXyDHT* were both solved by molecular replacement using *Phaser* (McCoy *et al.*, 2007). The structure of *RIArDHT* was solved using a modified model from PDB entry 2gp4 as a template. This resulted in a clear solution with two tetramers (Z -score = 12) in the asymmetric unit, with initial R and R_{free} values of 37.7 and 46.7%, respectively. The structure of *CcXyDHT* was subsequently solved by molecular replacement using the coordinates of *RIArDHT*. A clear solution with a tetramer in the asymmetric unit (Z -score = 9) was found and resulted in initial R and R_{free} values of 31.7 and 39.7%, respectively. Three-

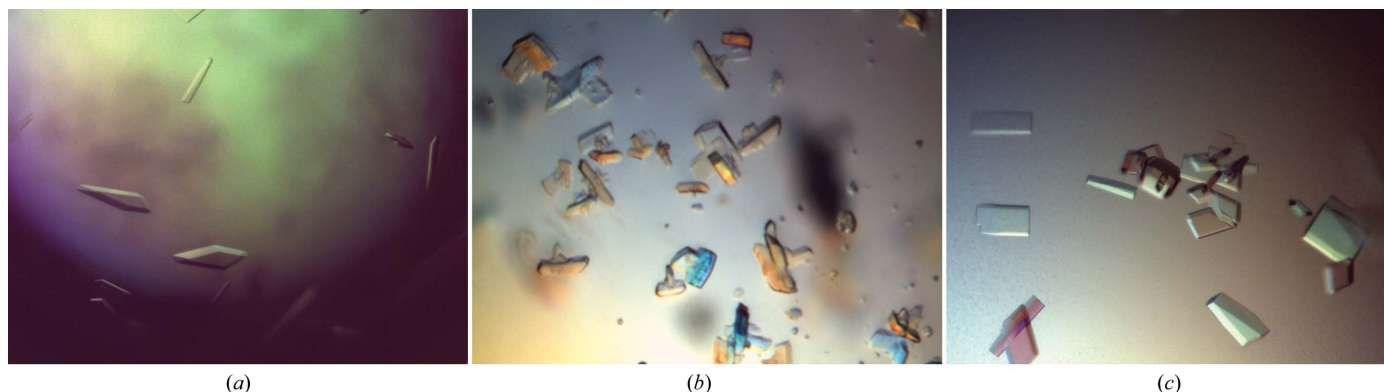


Figure 2
Crystals of (a) *RIArDHT*, (b) *CcXyDHT* before seeding and (c) *CcXyDHT* after seeding.

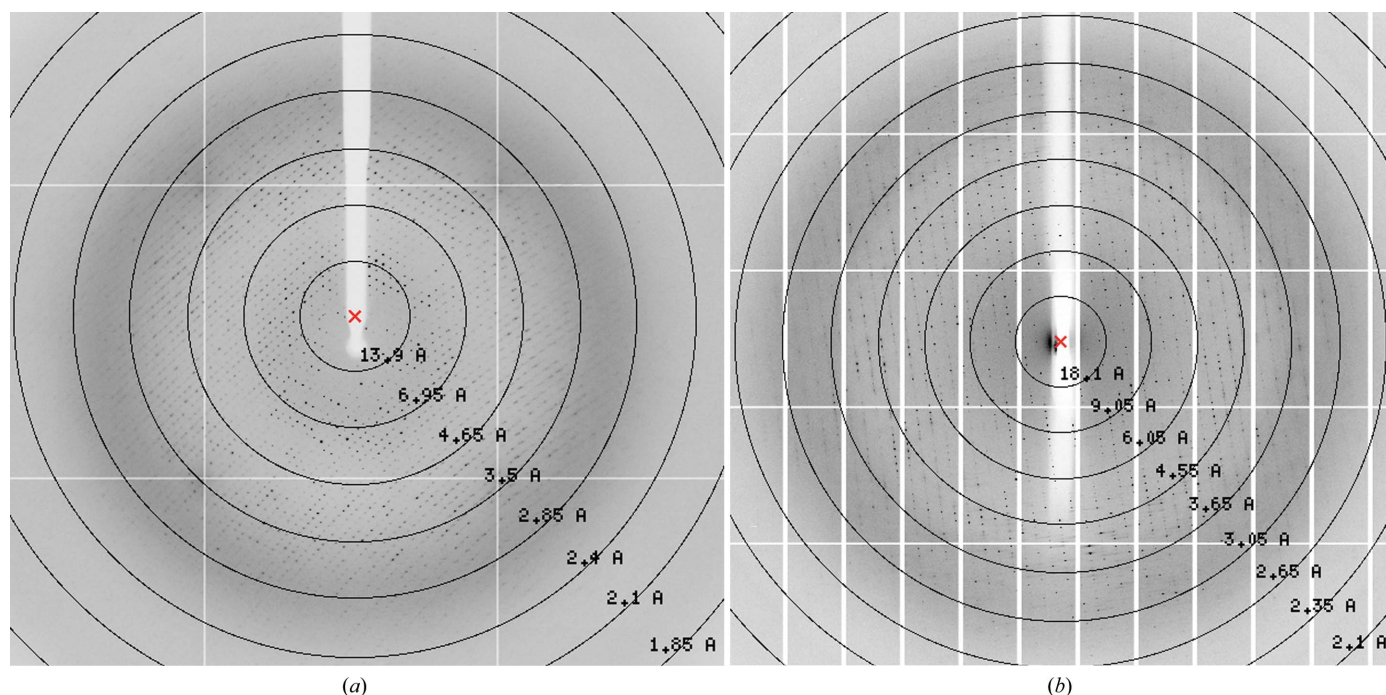


Figure 3
Diffraction images of (a) *RlArDHT* and (b) *CcXyDHT* crystals. The black circles correspond to the resolution limit in Å.

dimensional structure determination and structure refinement of both *RlArDHT* and *CcXyDHT* are ongoing.

Acknowledgements

We would like to thank the ESRF and ID14-2 beamline staff for their assistance during data collection and the I04 beamline staff at DLS for the provision of remote data collection. We would also like to thank all members of the DNA Laboratory at the Biology Department of the University of Eastern Finland, Joensuu Campus for allowing us to use their protein-expression facilities. This project was supported by the Academy of Finland (decisions Nos. 256937 and 263931). The work at VTT was supported by the Finnish Centre of Excellence in the White Biotechnology–Green Chemistry programme (Academy of Finland, decision No. 118573).

References

Andberg, M., Aro-Kärkkäinen, N., Carlson, P., Oja, M., Bozonnet, S., Toivari, M., Hakulinen, N., O'Donohue, M., Penttilä, M. & Koivula, A. (2016). *Appl. Microbiol. Biotechnol.*, doi:10.1007/s00253-016-7530-8.

Bergfors, T. (2003). *J. Struct. Biol.* **142**, 66–76.

Bettiga, M., Bengtsson, O., Hahn-Hägerdal, B. & Gorwa-Grauslund, M. F. (2009). *Microb. Cell Fact.* **8**, 40.

Boguta, A. M., Bringel, F., Martinussen, J. & Jensen, P. R. (2014). *Microb. Cell Fact.* **13**, 97.

Brouns, S. J. *et al.* (2006). *J. Biol. Chem.* **281**, 27378–27388.

Dahms, A. S. (1974). *Biochem. Biophys. Res. Commun.* **60**, 1433–1439.

Demeke, M. M., Dumortier, F., Li, Y., Broeckx, T., Foulquié-Moreno, M. R. & Thevelein, J. M. (2013). *Biotechnol. Biofuels*, **6**, 120.

Dien, B. S., Kurtzman, C. P., Saha, B. C. & Bothast, R. J. (1996). *Appl. Biochem. Biotechnol.* **57–58**, 233–242.

Dilworth, M. J., Arwas, R., McKay, I. A., Saroso, S. & Glenn, A. R. (1986). *J. Gen. Microbiol.* **132**, 2733–2742.

Flint, D. H. & Emptage, M. H. (1988). *J. Biol. Chem.* **263**, 3558–3564.

Flint, D. H., Emptage, M. H., Finnegan, M. G., Fu, W. & Johnson, M. K. (1993). *J. Biol. Chem.* **268**, 14732–14742.

Gu, Y., Ding, Y., Ren, C., Sun, Z., Rodionov, D. A., Zhang, W., Yang, S., Yang, C. & Jiang, W. (2010). *BMC Genomics*, **11**, 255.

Jin, L., Zhang, H., Chen, L., Yang, C., Yang, S., Jiang, W. & Gu, Y. (2014). *J. Biotechnol.* **173**, 7–9.

Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.

Kurosawa, K., Plassmeier, J., Kalinowski, J., Rückert, C. & Sinskey, A. J. (2015). *Metab. Eng.* **30**, 89–95.

Matsushika, A., Inoue, H., Murakami, K., Takimura, O. & Sawayama, S. (2009). *Bioresour. Technol.* **100**, 2392–2398.

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.

Meijnen, J.-P., de Winde, J. H. & Ruijsseenaars, H. J. (2009). *Appl. Environ. Microbiol.* **75**, 2784–2791.

Menon, V. & Rao, M. (2012). *Prog. Energy Combust. Sci.* **38**, 522–550.

Nunn, C. E., Johnsen, U., Schönheit, P., Fuhrer, T., Sauer, U., Hough, D. W. & Danson, M. J. (2010). *J. Biol. Chem.* **285**, 33701–33709.

Okamoto, K., Uchii, A., Kanawaku, R. & Yanase, H. (2014). *Springerplus*, **3**, 121.

Rodriguez, M., Wedd, A. G. & Scopes, R. K. (1996). *Biochem. Mol. Biol. Int.* **38**, 783–789.

Schädel, C., Blöchl, A., Richter, A. & Hoch, G. (2010). *Plant Physiol. Biochem.* **48**, 1–8.

Stephens, C., Christen, B., Fuchs, T., Sundaram, V., Watanabe, K. & Jenal, U. (2007). *J. Bacteriol.* **189**, 2181–2185.

Watanabe, S., Shimada, N., Tajima, K., Kodaki, T. & Makino, K. (2006). *J. Biol. Chem.* **281**, 33521–33536.

Weimberg, R. (1961). *J. Biol. Chem.* **236**, 629–635.

Xiao, H., Li, Z., Jiang, Y., Yang, Y., Jiang, W., Gu, Y. & Yang, S. (2012). *Metab. Eng.* **14**, 569–578.