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

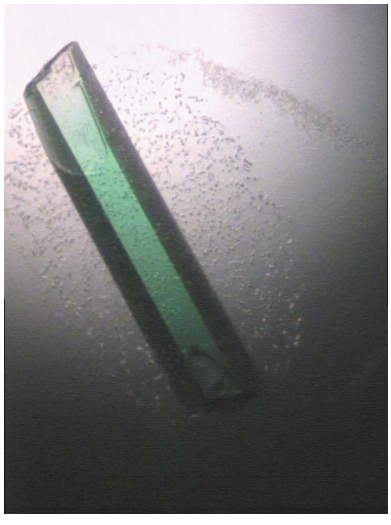
Crystallization and preliminary X-ray analysis of binary and ternary complexes of *Haloferax mediterranei* glucose dehydrogenase

Julia Esclapez,[a] K. Linda Britton,[b] Patrick J. Baker,[b] Martin Fisher,[b] Carmen Pire,[a] Juan Ferrer,[a] María José Bonete[a] and David W. Rice[b]*

[a]Departamento de Agroquímica y Bioquímica, Facultad de Ciencias, Universidad de Alicante, Ap. 99 Alicante 03080, Spain, and [b]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

Received 6 May 2005
Accepted 23 June 2005
Online 8 July 2005



Haloferax mediterranei glucose dehydrogenase (EC 1.1.1.47) belongs to the medium-chain alcohol dehydrogenase superfamily and requires zinc for catalysis. In the majority of these family members, the catalytic zinc is tetrahedrally coordinated by the side chains of a cysteine, a histidine, a cysteine or glutamate and a water molecule. In *H. mediterranei* glucose dehydrogenase, sequence analysis indicates that the zinc coordination is different, with the invariant cysteine replaced by an aspartate residue. In order to analyse the significance of this replacement and to contribute to an understanding of the role of the metal ion in catalysis, a range of binary and ternary complexes of the wild-type and a D38C mutant protein have been crystallized. For most of the complexes, crystals belonging to space group *I*222 were obtained using sodium/potassium citrate as a precipitant. However, for the binary and non-productive ternary complexes with NADPH/Zn, it was necessary to replace the citrate with 2-methyl-2,4-pentanediol. Despite the radical change in conditions, the crystals thus formed were isomorphous.

1. Introduction

Extremely halophilic archaea are found in highly saline environments such as natural salt lakes or saltern pools. These microorganisms require between 2.5 and 5.2 *M* salt for optimal growth (Kamekura, 1998) and they can balance the external salt concentration by accumulating intracellular KCl close to saturation. The biochemical machinery of these microorganisms has therefore been adapted in the course of evolution to be able to function at salt concentrations at which most biochemical systems cease to function.

Comparison of the amino-acid compositions and structures of halophilic proteins and their mesophilic counterparts has shown that a significant difference in the characteristics of the surface of halophilic proteins is an excess of acidic over basic residues (Lanyi, 1974; Böhm & Jaenicke, 1994; Dym *et al.*, 1995; Frolow *et al.*, 1996; Britton *et al.*, 1998, 2005).

The extremely halophilic archaeon *Haloferax mediterranei* (ATCC 33500/R4) is able to grow in a minimal medium with glucose as the sole carbon source, which is catabolized by a modified Enter–Doudoroff pathway. Glucose dehydrogenase (GlcDH) catalyses the first step in this pathway, the oxidation of β-D-glucose to gluconic acid, preferentially using $NADP^+$ as a coenzyme. Sequence analysis has shown that GlcDH belongs to the zinc-dependent medium-chain alcohol dehydrogenase (MDR) superfamily (Bonete *et al.*, 1996; Pire *et al.*, 2001). Biochemical studies have established that the protein is dimeric, with a subunit molecular weight of 39 kDa, and have confirmed the requirement of zinc for catalysis. In previous work, the crystallization of recombinant GlcDH in the presence of $NADP^+$ has been reported under conditions which closely mimic those experienced by the enzyme in the cell of the halophile (Ferrer *et al.*, 2001); more recently, the structure has been determined to 1.6 Å resolution (Britton *et al.*, 2005). This structure has revealed that the surface of the enzyme is not only decorated with acidic residues, but also displays a significant reduction in the fraction of exposed hydrophobic surface compared with non-halophilic glucose dehydrogenases. This reduction in hydrophobic surface predominately arises from the loss of the exposed alkyl component of lysine side chains as a result of the reduction of lysine content in the enzyme (Britton *et*

al., 2005). Moreover, genome comparisons have shown that there appears to be a general reduction in the frequency of lysine in halophilic proteins (Kennedy *et al.*, 2001).

At the active site of the prototypical MDR superfamily member, horse liver alcohol dehydrogenase (HLADH), the essential catalytic zinc ion is coordinated by three protein ligands (Cys46, His67 and Cys174), with the coordination shell of the zinc being completed by a water molecule (Eklund *et al.*, 1982). The mechanism of enzymes of the MDR superfamily is commonly proposed to involve the exchange of the zinc-bound water molecule with the hydroxyl of the substrate, from which a proton is then removed by a base to generate an alkoxide intermediate, which subsequently collapses to a carbonyl with concomitant reduction of $NAD(P)^+$. In this mechanism, the zinc is proposed to remain tetrahedrally coordinated throughout (Eklund *et al.*, 1982; Ehrig *et al.*, 1991; Ramaswamy *et al.*, 1999). In a recent alternative proposal, it has been suggested the zinc ion cycles between different four- and five-coordinate intermediates, the identities of which are not yet clear (Makinen *et al.*, 1983; Kleifeld *et al.*, 2003).

Of the three protein ligands to the zinc in enzymes of the MDR superfamily, the first cysteine (Cys46 in HLADH) and the histidine (His67 in HLADH) are very strongly conserved in the sequence. In some family members, the second cysteine (Cys174 in HLADH) is replaced by a glutamate [for example, Glu155 in *Thermoplasma acidophilum* GlcDH (John *et al.*, 1994) and Glu153 in rat sorbitol dehydrogenase (Johansson *et al.*, 2001)]. This sequence pattern is also seen in the structure of *H. mediterranei* GlcDH, where a glutamate residue (Glu64) occupies the equivalent position to the second cysteine (Britton *et al.*, 2005). However, the structure of *H. mediterranei* GlcDH has shown that there is an additional difference, with an aspartate residue (Asp38) occurring in a structurally equivalent position to the highly conserved first cysteine of this motif (Britton *et al.*, 2005). This is a sequence change rarely seen in the MDR family, but that has also been observed in the sequence of other halophilic glucose dehydrogenases [for example, the *Halobacterium* sp1 (Ng *et al.*, 2000) and *Haloferax volcanii* GlcDHs; http://zdna2.umbi.umd.edu] and raises the intriguing question as to whether the presence of this aspartate residue in the active site is a halophilic adaptation. In order to investigate the role of the metal and its ligands in catalysis and to enhance our understanding of the reaction mechanism in *H. mediterranei* GlcDH, we have explored a range of crystallization conditions of complexes of both the wild-type and the mutant D38C enzymes. In this paper, we report the preliminary crystallographic analysis of these various enzyme–substrate complexes.

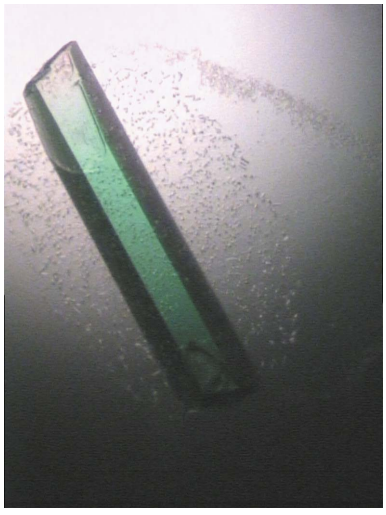


Figure 1
A crystal of the binary complex of D38C GlcDH with NADPH and zinc.

2. Materials and methods

2.1. Site-directed mutagenesis, expression and purification

The gene encoding the halophilic GlcDH was cloned into pGEM-11Zf(+) and site-directed mutagenesis was performed using the GeneEditor *in vitro* site-directed mutagenesis system (Promega). The protocol supplied with the kit was followed except that the length of the DNA-denaturation stage was increased from 5 min at room temperature to 20 min at 310 K. The expression, renaturation and purification of the recombinant wild-type and mutant proteins were performed as described previously (Pire *et al.*, 2001). The purified D38C GlcDH was dialyzed against 50 m*M* phosphate buffer pH 7.3 containing 2 *M* NaCl at 277 K overnight. Prior to crystallization, protein samples were concentrated to approximately 20 mg ml^{-1} using a Vivaspin concentrator (30 kDa molecular-weight cutoff).

2.2. Crystallization and diffraction data collection

Crystals of the wild-type and D38C mutant of GlcDH grew after 4–6 d in the presence of different substrates using the hanging-drop vapour-diffusion method, mixing small volumes (2–3 µl) of protein sample with an equal volume of a precipitant solution at 290 K. For the wild-type protein, crystals of the free enzyme and the binary complex with $NADP^+$, using sodium citrate as the precipitate, have already been reported (Ferrer *et al.*, 2001), but using these conditions crystals of the binary complex with NADPH could not be obtained either in the presence or absence of zinc. However, crystals of the NADPH–Zn complex could be produced by preparing the protein with 1 m*M* NADPH and 1 m*M* $ZnCl_2$ and using 67–72%(*v*/*v*) 2-methyl-2,4-pentanediol (MPD) in 100 m*M* HEPES pH 7.5 as the precipitant. The D38C mutant protein behaved in the same manner as the wild type, with crystals of the apoenzyme and various $NADP^+$ complexes forming in the presence of citrate and the NADPH complexes in the presence of MPD. For the D38C protein, the protein sample was mixed with 1 m*M* $NADP^+$ or 1 m*M* NADPH, 1 m*M* $ZnCl_2$, 10 m*M* glucose and 10 m*M* gluconate as appropriate. For the free enzyme and the complexes containing $NADP^+$ the precipitant used was 1.4–1.6 *M* sodium citrate in 100 m*M* HEPES pH 7.0 and for the complexes containing NADPH the precipitant was 62–72%(*v*/*v*) MPD in 100 m*M* HEPES pH 7.5. Crystals of D38C GlcDH were also grown in the presence of 2 *M* KCl, 1 m*M* $ZnCl_2$ and 1 m*M* $NADP^+$ using 1.4–1.6 *M* potassium citrate as precipitant in 100 m*M* HEPES buffer pH 7.0.

The crystals of the wild-type and D38C free enzyme grown in the presence of sodium citrate showed a hexagonal bipyramidal morphology (maximum dimensions 0.25 × 0.40 × 0.25 mm), whereas crystals with a rod-like morphology (Fig. 1; maximum dimensions 0.6 × 0.6 × 0.4 mm) were obtained for all the wild-type and mutant complexes with $NADP^+$ or NADPH, irrespective of whether the precipitant was sodium citrate or MPD.

Crystals grown with sodium citrate as the precipitant were mounted in X-ray-transparent capillaries. Preliminary data sets were collected at 290 K by the rotation method with 1° rotations per frame using a MAR 345 detector, with double-mirror-focused Cu $K\alpha$ X-rays produced by a Rigaku RU-200 rotating-anode generator. The crystals grown in the presence of MPD were flash-cooled in a cold nitrogen-gas stream and data were collected using the same method at 100 K. Data from the crystals of the D38C GlcDH–$NADP^+$–Zn complex

Table 1
X-ray data-collection statistics.

Values in parentheses refer to the highest resolution shell.

Protein sample	Wild type	D38C					
Complex	NADPH/Zn	Free enzyme	$NADP^+$/Zn	$NADP^+$/Zn/gluconate	NADPH/Zn/glucose	NADPH/Zn	NADPH
Source/wavelength	Cu $K\alpha$	Cu $K\alpha$	SRS† (0.97 Å)	Cu $K\alpha$	Cu $K\alpha$	Cu $K\alpha$	Cu $K\alpha$
Resolution (Å)	2.01	3.55	1.50	2.2	2.0	1.84	2.0
Highest resolution shell (Å)	2.08–2.01	3.63–3.55	1.55–1.50	2.27–2.22	2.05–2.0	1.89–1.84	2.05–2.00
Space group	$I222$	$P6_222$	$I222$	$I222$	$I222$	$I222$	$I222$
Unit-cell parameters (Å)							
a	60.6	89.4	60.5	61.6	60.1	60.5	60.7
b	107.7	89.4	109.3	112.2	106.3	108.9	108.1
c	153.6	211.5	151.9	150.5	151.9	151.2	152.6
Unique reflections	32173	6269	66981	24766	30873	44001	32792
Completeness (%)	94.9 (82.5)	96.7 (100)	94.4 (93.5)	94.5 (92.9)	92.8 (89.6)	94.5 (91.5)	95.7 (99.3)
Multiplicity	3.9 (3.4)	8.74 (3.4)	5.1 (4.6)	9.34 (3.96)	8.03 (2.16)	7.1 (2.23)	5.4 (2.62)
$\langle I/\sigma(I)\rangle$	30.1 (3.7)	11.71 (2.4)	1.37 (1.01)	14.3 (3.3)	10.0 (1.8)	16.3 (1.8)	14.58 (2.41)
$R_{\rm merge}$‡	0.059 (0.354)	0.10 (0.50)	0.051 (0.709)	0.10 (0.42)	0.095 (0.45)	0.047 (0.49)	0.092 (0.447)

† SRS: CCLRC Daresbury Synchrotron Radiation Source. ‡ $R_{\rm merge} = \sum_h \sum_i |I(h,i) - \langle I(h)\rangle| / \sum_h \sum_i I(h,i)$.

grown in the presence of KCl and potassium citrate were collected at the SRS Daresbury synchrotron on station PX14.2 at a wavelength of 0.97 Å with 1° rotations using an ADSC Q4 detector. The data for each crystal were processed and analysed using the *HKL* suite of programs (Otwinowski & Minor, 1997) and subsequently handled using the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

Analysis of the various data sets using the autoindexing routine of *DENZO* showed that D38C GlcDH had crystallized in two different forms, designated I and II, which are the same as those seen for the wild-type enzyme (Ferrer *et al.*, 2001). The crystals of the D38C free enzyme (form I) belong to a hexagonal lattice *P*622, with unit-cell parameters $a = b = 89.4$, $c = 211.5$ Å, $\alpha = \beta = 90$, $\gamma = 120°$ and a unit-cell volume of 1.46×10^6 Å^3. Given the GlcDH subunit molecular weight of 39 kDa, the $V_{\rm M}$ for a monomer in the asymmetric unit is 3.1 Å^3 Da^{-1}, which lies within the normal range for proteins (Matthews, 1977). The crystals of the binary complexes and non-productive ternary complexes (form II) were all isomorphous, whether they were grown using citrate or MPD as the precipitant, and diffracted to up to 1.5 Å resolution. These crystals belong to one of the special pair of space groups $I222$ or $I2_12_12_1$, with unit-cell parameters as shown in Table 1 and a unit-cell volume of 1.0×10^6 Å^3. The asymmetric unit appears to contain a monomer ($V_{\rm M}$ = 3.2 Å^3 Da^{-1}). Data-collection statistics for the form I and form II crystals are given in Table 1.

A cross-rotation function and translation function were calculated in both $I222$ or $I2_12_12_1$ space groups on the form II data for the D38C NADPH–Zn–glucose complex at a resolution of 20–3.0 Å, using a single subunit of the wild-type *H. mediterranei* GlcDH as a search model and the program *AMoRe* (Navaza, 1994). A clear translation-function solution of correlation coefficient 67.6% and *R* factor 36.9% was only seen in space group $I222$, indicating that this is the correct space group. For the data from the form I crystals a similar process was undertaken. In this case, a clear translation-function solution was only seen in space group $P6_222$ (correlation coefficient 71.6%, *R* factor 38.7%), identifying this as the correct space group.

Previously, we have obtained form II crystals from the wild-type enzyme in complex with $NADP^+$ and also from the D38C mutant in complex with $NADP^+$ and zinc (PDB codes 1ss0 and 1ss5, respectively). All the D38C GlcDH complex crystals reported here are isomorphous with the form II crystals. However, it is interesting to note that the crystals of the $NADP^+$ complexes grow using citrate as the precipitant, whereas those complexes containing NADPH only grow when MPD is used as the precipitant. Despite these radically different conditions, the crystals thus produced are isomorphous.

Analysis of these structures is now under way and given the successful crystallization of this wide range of complexes of the wild-type and mutant D38C GlcDH, we will be able to provide new insights into the structure–function relationships of the enzyme and perhaps shed light on the question of whether the aspartate residue at position 38 in the *H. mediterranei* GlcDH is a halophilic adaptation or purely serendipitous.

We thank the support staff at CCLRC Synchrotron Radiation Source, Daresbury for assistance with station alignment. This work was supported by the BBSRC, the Wellcome Trust, The Woolfson Foundation and MCYT (BIO2002-03179) and by FPI fellowships from Generalitat Valenciana (Spain) to JE. The Krebs Institute is a designated BBSRC Biomolecular Sciences Centre and a member of the North of England Structural Biology Centre.

References

Böhm, G. & Jaenicke, R. (1994). *Protein Eng.* **7**, 213–220.
Bonete, M. J., Pire, C., Llorca, F. I. & Camacho, M. L. (1996). *FEBS Lett.* **383**, 227–229.
Britton, K. L., Baker, P. J., Fisher, M., Ruzheinikov, S., Gilmour, J., Bonete, M. J., Ferrer, J., Pire, C., Esclapez, J. & Rice, D. W. (2005). Submitted.
Britton, K. L., Stillman, T. J., Yip, K. S. P., Forterre, P., Engel, P. C. & Rice, D. W. (1998). *J. Biol. Chem.* **273**, 9023–9030.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dym, O., Mevarech, M. & Sussman, J. L. (1995). *Science*, **267**, 1344–1346.
Ehrig, T., Hurley, T. D., Edenberg, H. J. & Bosron, W. F. (1991). *Biochemistry*, **30**, 1062–1068.
Eklund, H., Plapp, B. V., Samama, J. P. & Branden, C. I. (1982). *J. Biol. Chem.* **257**, 14349–14358.
Ferrer, J., Fisher, M., Burke, J., Sedelnikova, S. E., Baker, P. J., Gilmour, D. J., Bonete, M. J., Pire, C., Esclapez, J. & Rice, D. W. (2001). *Acta Cryst.* D**57**, 1887–1889.
Frolow, F., Harel, M., Sussman, J. L., Mevarech, M. & Sholman, M. (1996). *Nature Struct. Biol.* **3**, 452–458.
Johansson, K., El-Ahmad, M., Kaiser, C., Jörnvall, H., Eklund, H., Höög, J. O. & Ramaswamy, S. (2001). *Chem. Biol. Interact.* **130–132**, 351–358.
John, J., Crennell, S. J., Hough, D. W., Danson, M. J. & Taylor, G. L. (1994). *Structure*, **2**, 385–393.
Kamekura, M. (1998). *Extremophiles*, **2**, 289–295.

Kennedy, S. P., Ng, W. V., Salzberg, S. L., Hood, L. & DasSharma, S. (2001). *Genome Res.* **11**, 1641–1650.
Kleifeld, O., Frenkel, A., Martin, J. M. & Sagi, I. (2003). *Nature Struct. Biol.* **10**, 98–103.
Lanyi, J. K. (1974). *Bacteriol. Rev.* **38**, 272–290.
Makinen, M. W., Maret, W. & Yim, M. B. (1983). *Proc. Natl Acad. Sci. USA*, **80**, 2584–2588.
Matthews, B. W. (1977). *The Proteins*, 3rd ed., Vol. 3, edited by H. Neurath & R. Hill, pp. 468–477. New York: Academic Press.
Navaza, J. (1994) *Acta Cryst.* A**50**, 157–163.
Ng, W. V. *et al.* (2000). *Proc. Natl Acad. Sci. USA*, **97**, 12176–12181.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pire, C., Esclapez, J., Ferrer, J. & Bonete, M. J. (2001). *FEMS Microbiol. Lett.* **200**, 221–227.
Ramaswamy, S., Park, D. H. & Plapp, B. V. (1999). *Biochemistry*, **38**, 13951–13959.