Extremophiles (2009) 13:179–190
DOI 10.1007/s00792-008-0208-0

ORIGINAL PAPER

# Biochemical and structural studies of a L-haloacid dehalogenase from the thermophilic archaeon *Sulfolobus tokodaii*

**Carrie A. Rye · Michail N. Isupov · Andrey A. Lebedev · Jennifer A. Littlechild**

Received: 28 July 2008 / Accepted: 29 October 2008 / Published online: 29 November 2008


**Abstract** Haloacid dehalogenases have potential applications in the pharmaceutical and fine chemical industry as well as in the remediation of contaminated land. The L-2-haloacid dehalogenase from the thermophilic archaeon *Sulfolobus tokodaii* has been cloned and over-expressed in *Escherichia coli* and successfully purified to homogeneity. Here we report the structure of the recombinant dehalogenase solved by molecular replacement in two different crystal forms. The enzyme is a homodimer with each monomer being composed of a core-domain of a $\beta$-sheet bundle surrounded by $\alpha$-helices and an $\alpha$-helical subdomain. This fold is similar to previously solved mesophilic L-haloacid dehalogenase structures. The monoclinic crystal form contains a putative inhibitor L-lactate in the active site. The enzyme displays haloacid dehalogenase activity towards carboxylic acids with the halide attached at the C2 position with the highest activity towards chloropropionic acid. The enzyme is thermostable with maximum activity at 60°C and a half-life of over 1 h at 70°C. The enzyme is relatively stable to solvents with 25% activity lost when incubated for 1 h in 20% v/v DMSO.

Communicated by A. Driessen.

A. A. Lebedev
Structural Biology Laboratory,
Department of Chemistry, University of York,
Heslington, York YO10 5YW, UK

C. A. Rye · M. N. Isupov · A. A. Lebedev · J. A. Littlechild (✉)
Henry Wellcome Building for Biocatalysis,
School of Biosciences, University of Exeter,
Stocker Road, Exeter EX4 4QD, UK
e-mail: J.A.Littlechild@exeter.ac.uk

**Keywords** Haloacid dehalogenase · *Sulfolobus tokodaii* · X-ray structure

## Introduction

There are many different types of dehalogenase enzyme which are often involved in pathways to degrade complex halogenated compounds. Dehalogenases have important applications for the biodegradation of environmental toxic halogenated compounds. Many of these are produced synthetically for use as herbicides and growth regulators (Allpress and Gowland 1998). Over 60% of herbicides contain at least one chlorine atom (Slater 1982).

The 2-haloacid dehalogenases (EC 3.8.1; halidohydrolases) catalyse the hydrolytic dehalogenation of 2-haloalkanoic acids to produce 2-hydroxyalkanoic acids. As well as having potential for the degradation of toxic halogenated compounds from the environment haloacid dehalogenases also have industrial applications in the production of chiral haloalkanoic acids and hydroxyalkanoic acids.

Based on their substrate specificity three different types of haloacid dehalogenase have been identified depending on the enantiomer that they show activity against, DL, D and L. The DL-haloacid dehalogenases work equally well on both enantiomers of the haloacid with either retention or inversion of the stereochemistry at the C2 atom position. D and L-2-haloacid dehalogenases are specific to only one enantiomer and cause an inversion in the C2 configuration of the hydroxyalkanoic acid produced (Slater et al. 1997). L-2-haloacid dehalogenases belong to a large superfamily, the haloacid dehalogenase (HAD, Pfam PF00702) superfamily. Alongside haloacid dehalogenases this family also includes some ATPases, epoxide hydrolases and a number of different phosphatases. All members of this family have

a completely conserved aspartic acid residue at the start of a five amino acid N-terminal motif. The middle amino acid in this motif is often conserved in the different groups of this superfamily. For example a threonine residue is commonly found in ATPases and a tyrosine residue found in dehalogenases.

There are currently two X-ray structures available for mesophilic L-2-haloacid dehalogenases. These structures are L-DEX-YL from *Pseudomonas* sp. YL (Hisano et al. 1996) and DhlB from *Xanthobacter autotrophicus* GJ10 (Ridder et al. 1997). Both of these enzymes are homodimers with each subunit having a core domain of a Rossmann fold like six-stranded parallel $\beta$-sheet flanked by five $\alpha$-helices and a four-helix bundle sub-domain. The monomeric structure of a putative haloacid dehalogenase (PH0459) from the thermophilic archaeon *Pyrococcus horikoshii* OT3 has also been solved (Arai et al. 2006).

The enzymes isolated from thermophilic organisms often show stability to a number of extrinsic factors such as temperature and organic solvents. These properties make them particularly useful for industrial biocatalysis. The genome of *Sulfolobus tokodaii* has been previously sequenced (Kawarabayasi et al. 2001). This organism was isolated from Beppu hot springs in Kyushu, Japan in 1983 and is a member of the Crenarchaeota which is able to convert hydrogen sulfide to sulphate growing optimally at 80°C in an aerobic, acidic, sulphur-rich environment. A L-2-haloacid dehalogenase sequence identified from the *S. tokodaii* genome sequence, has 31% sequence identity to L-DEX YL dehalogenase, 28% to DhlB dehalogenase and 29% to PH0459 dehalogenase.

We have cloned, characterised and solved the X-ray structure of the *S. tokodaii* strain 7 L-2-haloacid dehalogenase (Rye et al. 2007). Here we present the high-resolution three-dimensional structure of this thermophilic dehalogenase together with is biochemical characterisation and comparison with other related enzymes previously described.

## Materials and methods

The *S. tokodaii* L-2-haloacid dehalogenase enzyme was over-expressed in *E. coli* then purified using nickel affinity and gel filtration column chromatography and crystallised as previously described (Rye et al. 2007). Two different buffers were used for the elution of the gel filtration column; 100 mM Tris–HCl, 100 mM NaCl, pH 8.5 or 50 mM $KH_2PO_4/K_2HPO_4$, 100 mM NaCl, pH 6.5. Enzyme prepared using each buffer was used for crystallisation trials. The enzyme used for all other studies was prepared in phosphate buffer.

### Enzyme activity

The L-2-haloacid dehalogenase activity was detected by a modified colorimetric assay as described in Rye et al. 2007. For the standard assay, substrate was added to a final concentration of 10 mM and the assay mixture was incubated at 50°C for 5 min before the addition of 20 µl of the purified L-2-haloacid dehalogenase enzyme. The activity was detected by measuring the decrease in absorbance at 540 nm.

### Stability trials

The temperature at which *S. tokodaii* haloacid dehalogenase showed maximum activity was determined by carrying out assays between 20 and 70°C. The assay mixture was pre-incubated at the temperature of interest for 5 min and the reaction started by the addition of enzyme. Standard curves of HCl concentration against change in absorbance were produced at each temperature tested to take into account the effect of temperature on the buffering capacity of the assay solution.

Temperature stability was assayed by incubating the enzyme at different temperatures (50–90°C) for 30, 60 and 90 min in triplicate followed by cooling on ice. The standard activity assay was then carried out.

To determine the solvent stability of *S. tokodaii* haloacid dehalogenase, the enzyme was incubated with 10, 20, 30 and 50% (v/v) ethanol, methanol, DMSO and acetonitrile. Samples were then taken in triplicate after 1 h and assayed for enzyme activity. As a control, reactions were carried out without enzyme to determine the effect of solvent on the substrate and the assay.

### Crystallisation and data collection

Two crystal forms were grown using the microbatch method as described earlier (Rye et al. 2007). The monoclinic crystals were grown from the first batch of protein purified using Tris–HCl buffer; the drop contained 0.1 M HEPES, 20% PEG 6,000, pH 7.0. The data were collected at the Daresbury Synchrotron as described earlier (Rye et al. 2007). The orthorhombic crystals were grown from the second batch of protein purified using phosphate buffer; the drop contained 150 mM MES, 16% (w/v) PEG 6000 pH 5.5. These crystals were frozen under silicon oil prior to data collection. The data were collected 'in house' at 100 K using a MAR Research 345 Image Plate and copper K$\alpha$ radiation from a rotating anode generator (Bruker AXS) operated at 100 mA and 35 KV (3.5 kW) with XENOCS FOX2D CU25_25P mirrors. Data were processed using the programs DENZO and SCALEPACK (Otwinowski and Minor 1997).

### Model building and refinement

Molecular replacement using the program MOLREP (Vagin and Teplyakov 1997) was used for phase determination. The structure of the haloacid dehalogenase from *X. autotrophicus* (PDB code 1QQ5) had previously been used as a model to solve the 1.9 Å resolution monoclinic structure of this enzyme to an R-factor of 0.21 ($R_{free}$ 0.27) (Rye et al. 2007). The monoclinic data was found to be twinned by a reticular merohedry. The data were detwinned by an original algorithm (Rye et al. 2007) and the model was subjected to further refinement (Table 1). The refined monoclinic structure was used as the model for molecular replacement for the orthorhombic crystal data set. Maximum likelihood refinement was carried out in REFMAC 5.2 (Murshudov et al. 1997) and the manual model rebuilding was carried out in the visualisation program COOT (Emsley and Cowtan 2004). Water molecules were added in COOT and refined and checked manually using the 2Fo–Fc map. After refinement the quality of the model was checked using the program PROCHECK (Laskowski et al. 1993). Ionic bonds, hydrogen bonds and hydrophobic interactions were detected using the program WHATIF (Vriend 1990) and then visually checked in the program COOT with bond length cut offs of 3.0, 3.2 and 4.5 Å, respectively. Superimpositions were carried out using the programs PyMol (Delano Scientific) and ccp4MG (Potterton et al. 2004). Accessible surfaces were calculated using the AREAIMOL program in the CCP4 suite of programs (CCP4 1994).

## Results and discussion

### Quality of the model

The crystals grown from the protein purified in Tris–HCl buffer belong to the monoclinic space group C2 with unit cell parameters $a = 127.6$ Å, $b = 58.1$ Å, $c = 51.2$ Å, $\beta = 97.2°$. The electron density obtained allowed the positioning of the entire 201 amino acids of the protein as well as part of N-terminal His-tag sequence. This resulted in the model being made of 206 amino acids in subunit A and 204 amino acids is subunit B. The final model contained 336 water molecules. The *G* factor as calculated by PROCHECK (Laskowski et al. 1993) is 0.2, which is better than expected for a structure of this resolution. The Ramachandran plot shows that 94.1% of the residues fall within the most favoured regions. The properties of the refined structure are given in Table 1. One residue from each monomer (Lys128) is in the cis conformation. Additional density was observed in the active site that has been modelled as L-lactate.

The crystals grown from the protein purified in phosphate buffer belong to the orthorhombic space group P212121 with unit cell parameters $a = 69.8$ Å, $b = 71.9$

**Table 1** Summary of data collection and refinement statistics for both crystal forms

| | Monoclinic | Orthorhombic |
|---|---|---|
| Space group | *C*2 | $P2_12_12_1$ |
| Unit cell *a*/*b*/*c* (Å) | 127.6, 58.1, 51.2 $\beta = 97.2°$ | 69.8, 71.9, 85.6 |
| Wavelength (Å) | 1.729 | 1.5418 |
| Resolution range (Å) | 25–1.9 | 15–1.66 |
| No. measured reflections | 99871 | 324359 |
| No. unique reflections | 28692 | 51594 |
| Completeness | 97.6 (75.0)[a] | 99.7 (99.8)[a] |
| I > 3σ(I) (%) | 90.0 (51.0)[a] | 93.3 (74.1)[a] |
| (I)/σ (I) | 16 (5.7)[a] | 28.8 (7.3)[a] |
| $R_{sym}$ (%) | 9.6 (16.1)[a] | 5.1 (22.6)[a] |
| Wilson B-factor (Å$^2$) | 31.6 | 30.6 |
| Final $R_{cryst}$ | 0.16 | 0.19 |
| $R_{free}$ | 0.21 | 0.24 |
| No. of water molecules | 336 | 606 |
| Average B-factor (protein) (Å$^2$) | 22.5 | 19.7 |
| Average B-factor (solvent) (Å$^2$) | 33.1 | 24.4 |
| Ligand | L-Lactate | $PO_4$, MES |
| Average B-factor (ligand) (Å$^2$) | 23.8 | 17.0 |
| Rms deviation from ideality | | |
| Bond lengths (Å) | 0.012 (0.022) | 0.018 (0.022) |
| Bond angles (°) | 1.254 (1.964) | 1.630 (1.967) |

[a] Values in parentheses are given for the outer resolution shell. $R_{sym} = \Sigma_h\Sigma_J|<I_h>-I_J(h)|/\Sigma_h\Sigma_J I(h)$, where $I(h)$ is the intensity of reflection h. $\Sigma_h$ is the sum over all reflections and $\Sigma_J$ is the sum over *J* measurements of the reflection. $R_{cryst} = \sum||F_o|-|F_c||/\sum|F_o|$

Å, $c = 85.6$ Å. The electron density obtained allowed the positioning of all 201 amino acids of subunit A and 200 amino acids of subunit B (excluding the N-terminal methionine). The final model contains 401 amino acid residues and 606 water molecules and was refined to an R-factor of 0.19 and $R_{free}$ of 0.24 (Table 1). The program PROCHECK (Laskowski et al. 1993) shows that 95.3% of the residues fall within the most favoured region of the Ramachandran plot. One residue from each monomer (Lys128) is in the cis conformation. The *G* factor as calculated by PROCHECK is 0.0. Additional density was observed in the active site and on the dimer interface. In this case an inorganic phosphate molecule was found in the active site and a MES molecule at the interface.

The secondary structure composition of the *S. tokodaii* L-haloacid dehalogenase is approximately 12% β-sheet, 44% α-helix and 10% $3_{10}$ helix. The two subunits of the monoclinic form superpose with an r.m.s deviation of 0.121 Å. and of the orthorhombic form 0.304 Å. Due to the high similarity of the monoclinic and orthorhombic structures the main discussion below is limited to the orthorhombic structure, with the exception of discussions about the observed density in the active site.

## Monomer structure

The monomer of *S. tokodaii* L-haloacid dehalogenase (Fig. 1) consists of 201 amino acid residues and has dimensions of approximately 50 × 36 × 32 Å. The monomer folds into two domains, a core domain and a subdomain with the active site located between the two. The overall fold is common to all members of the HAD enzyme superfamily.

The core domain has a Rossmann fold like six-stranded parallel β-strand bundle surrounded by five α-helices and three $3_{10}$ helices and is composed of residues 1–13 and 73–201 making up approximately 70% of all the amino acid residues of the protein. The β-sheet has $-1x$, $-1x$, $3x$, $1x$, $1x$ connectivity. The strands all show β-strand α-helix β-strand connectivity, except for β-strands 5 and 6, which are connected via a $3_{10}$ helix. Additional $3_{10}$ helices are found before and after β-strand 3. The subdomain has a three α-helix bundle and a $3_{10}$ helix and is composed of residues 14–72, making up approximately 30% of all the amino acid residues of the protein. The subdomain is composed of α-helices 1, 2 and 3 and $3_{10}$ helix 1 and is inserted into the core domain between β-strand 1 and α-helix 4. In the subdomain, α-helix 2 and α-helix 3 run approximately side-by-side and α-helix 1 crosses the other two.

The *S. tokodaii* L-haloacid dehalogenase shares 29% sequence identity to the 2.0 Å structure of the putative L-haloacid dehalogenase (PH0459) from *P. horikoshii* OT3

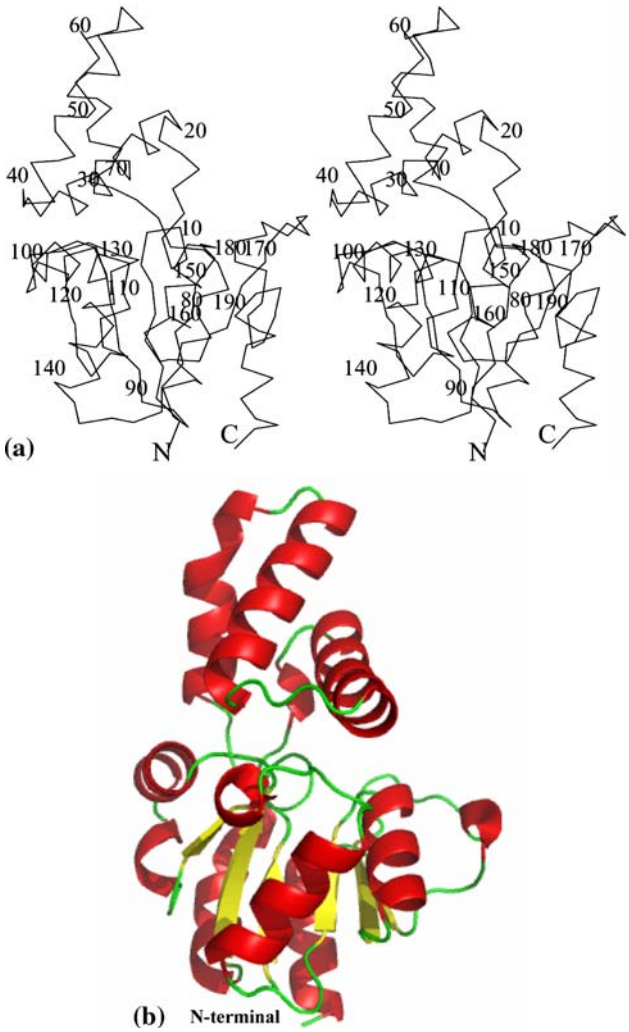


**Fig. 1** **a** The overall fold of the *S. tokodaii* haloacid dehalogenase monomer shown as a Cα trace, with every tenth residue numbered. This stereo picture was produced using BOBSCRIPT (Esnouf 1997). **b** A ribbon diagram coloured by secondary structure elements showing the overall secondary structure of the *S. tokodaii* dehalogenase. The lower portion is the core domain and the upper the smaller subdomain. The figure was produced using PyMol (Delano Scientific)

(1 × 42, Arai et al. 2006). The 2.5 Å structure of L-DEX-YL from *Pseudomonas* sp. YL (1JUD, Hisano et al. 1996) shares 31% identity with the *S. tokodaii* dehalogenase. Both of these enzymes contain a core domain and a subdomain, however, the 1.5 Å structure of DhlB from *X. autotrophicus* (1QQ5), which shares 28% sequence identity to the *S. tokodaii* enzyme has been shown to be made of three domains (a core domain, subdomain and a dimerisation domain consisting of two loosely packed anti parallel α-helices (Ridder et al. 1997).

The superimposition of the *S. tokodaii* monomer with the three other structures (Fig. 2) as calculated using CCP4MG (Potterton et al. 2004) gives an RMS deviation

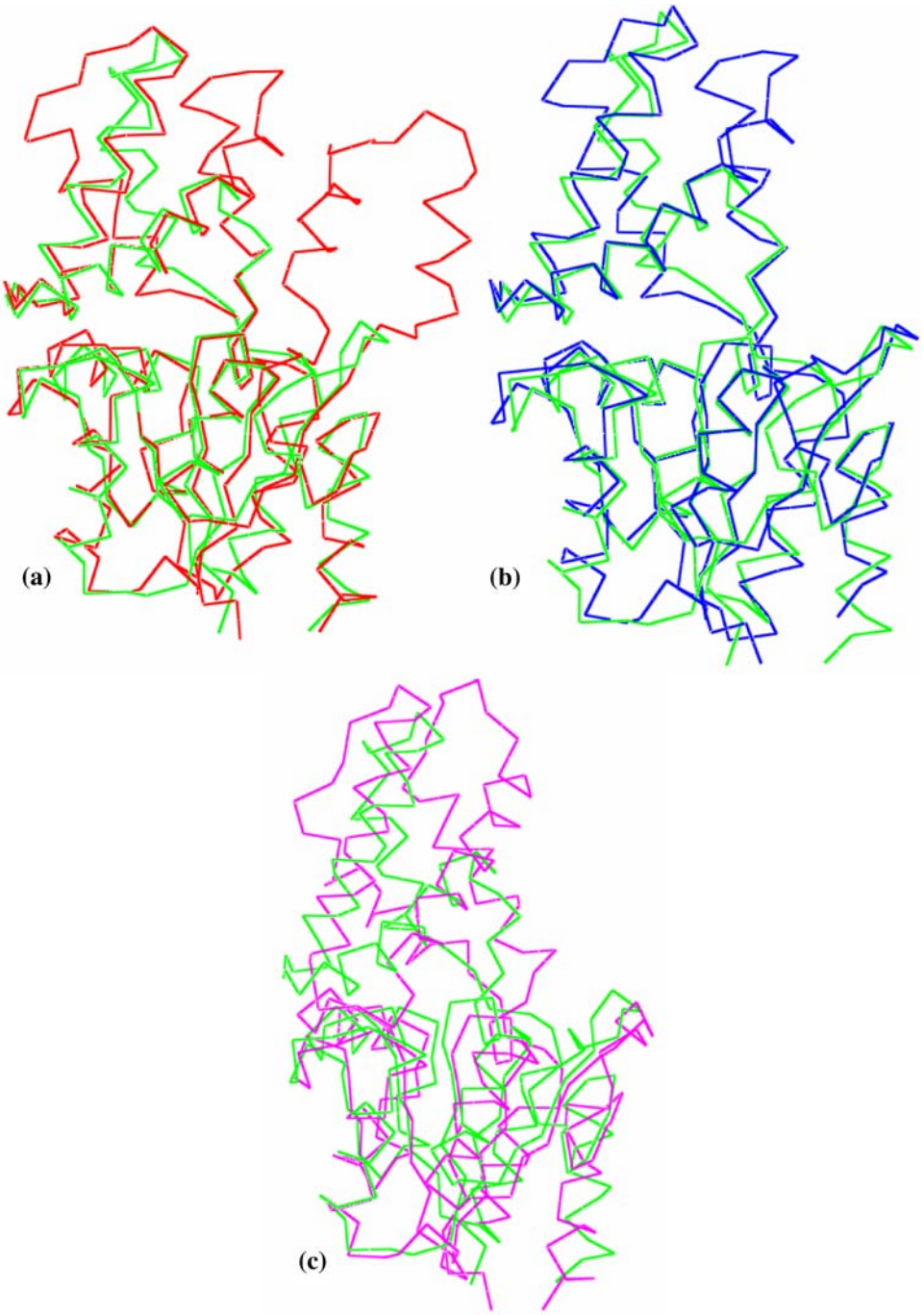


**Fig. 2** Superimposition of the monomers of *S. tokodaii* dehalogenase (*green*) and **a** DhlB, **b** L-DEX-YL, **c** PH0459

of 1.61 Å to DhlB over 190 Cα atoms, 1.48 Å to L-DEX-YL over 190 Cα atoms and 2.49 Å to PH0459 over 168 Cα atoms. The most noticeable differences are within the subdomain in the *S. tokodaii* dehalogenase where this domain is more compact and has one less α-helix. The greatest difference is seen when compared to PH0459 where the difference is over both domains, with PH0459 having an additional α-helix in its core domain as well as an additional α-helix in the subdomain.

## Quaternary structure

The *S. tokodaii* haloacid dehalogenase crystallises as a homodimer, which is in agreement with gel filtration analysis of the protein in solution. This same oligomeric state has been seen in both of the mesophilic haloacid dehalogenases with solved structures. The only other thermophilic archaeal dehalogenase (PHO459) structure was found to be a monomer and there is no biochemical

characterisation for this enzyme. The two subunits of the *S. tokodaii* haloacid dehalogenase are related by a twofold axis that is approximately parallel to α-helix 1 of each subunit. The dimer (Fig. 3) is compact with approximate dimensions of 75 × 42 × 44 Å. The monomer has an accessible surface area of 10,401 $Å^2$ and the dimer has an accessible surface area of 18,022 $Å^2$, therefore 1,390 $Å^2$ or 13.4% of each monomer's accessible surface is buried on dimer formation. This compares to a buried surface area of 13.5% in L-DEX-YL and 19% in DhlB.

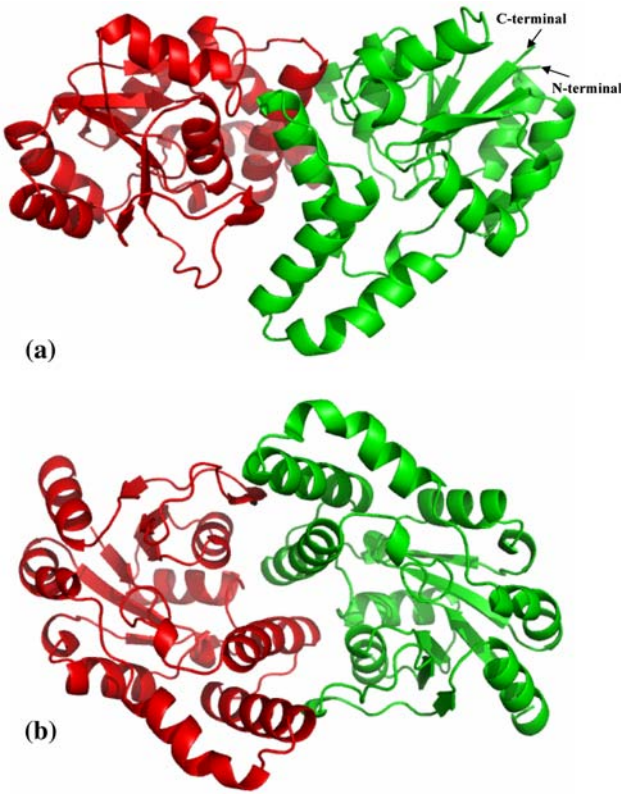


**Fig. 3** The L-haloacid dehalogenase dimer shown as a ribbon representation. Each subunit is shown in a different colour. **a** View with the twofold axis vertical, **b** rotation of 90° on the horizontal axis. The figure was produced using the program PyMol (Delano Scientific)

The main area of contact on dimer formation is along α-helix 1 of the subdomain. The α-helices from each subunit are closer together at their C-terminal ends than at the N-terminal ends, resulting in a depression that leads to the active site. There are five main areas of contact between the two subunits, which are stabilised mainly by a hydrophobic cluster involving the side chains of Leu27, Tyr29, Trp31, Leu32, Ile35, Met36, Phe52, Pro129, Phe154, Gly158 and Ile171 (Fig. 4). The first area (A1) is on α-helix 1; the second area (A2) is on α-helix 2; the third area (A3) is the turn region between $3_{10}$ helix 3 and α-helix 6; the fourth area (A4) is α-helix 7 and the fifth area (A5) is on the loop between β-strand 5 and $3_{10}$ helix 4. The largest area of contact is the interaction between A1 of both subunits. A1 also makes contacts with A3 (Tyr29 of one subunit to Pro129 of the other subunit), A4 (hydrophobic interaction between helices involving mainly Leu32 on one subunit and Phe154/G158 on the other) and A5 (Tyr29/Leu32 on one subunit and Ile179 on the other). A5 makes an additional contact with A2 (Ile175 to Tyr52). Similar to that seen in L-DEX-YL structure (Hisano et al. 1996) there are hydrogen bond interactions between the residues Lys25, Tyr29, Pro179 and Asp178 of each subunit, including a salt bridge between Asp178 of one subunit and Lys25 of the other.

The interactions between the two monomers of *S. tokodaii* haloacid dehalogenase are very similar to those seen in L-DEX-YL and DhlB enzymes, with a number of hydrophobic residues being conserved between all three enzymes (Fig. 5). There are however additional hydrogen bonds

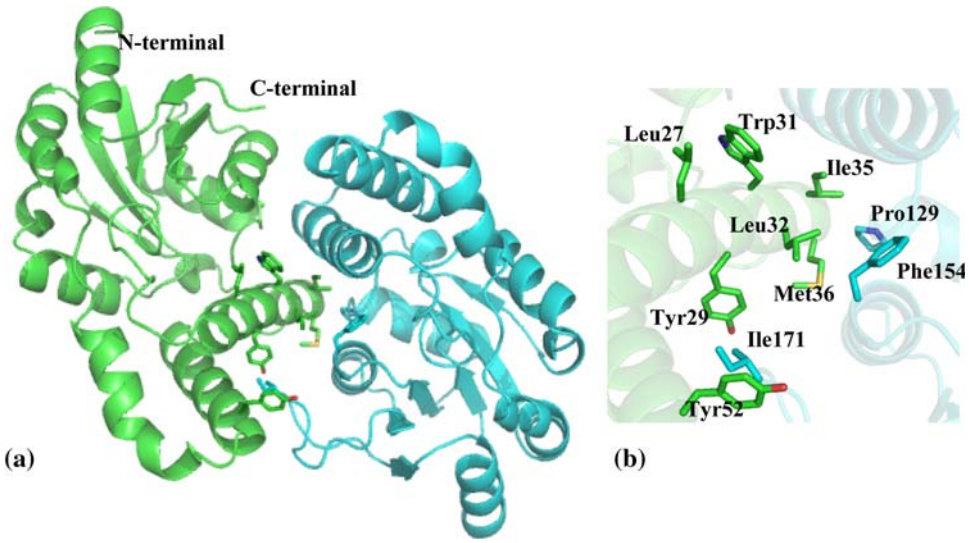


**Fig. 4** **a** View looking down the dimer interface showing that the α-helices from each subunit are closer together at one end than at the other, **b** the residues on the dimer interface that make up a hydrophobic cluster. The figure was produced using PyMOL (Delano Scientific)

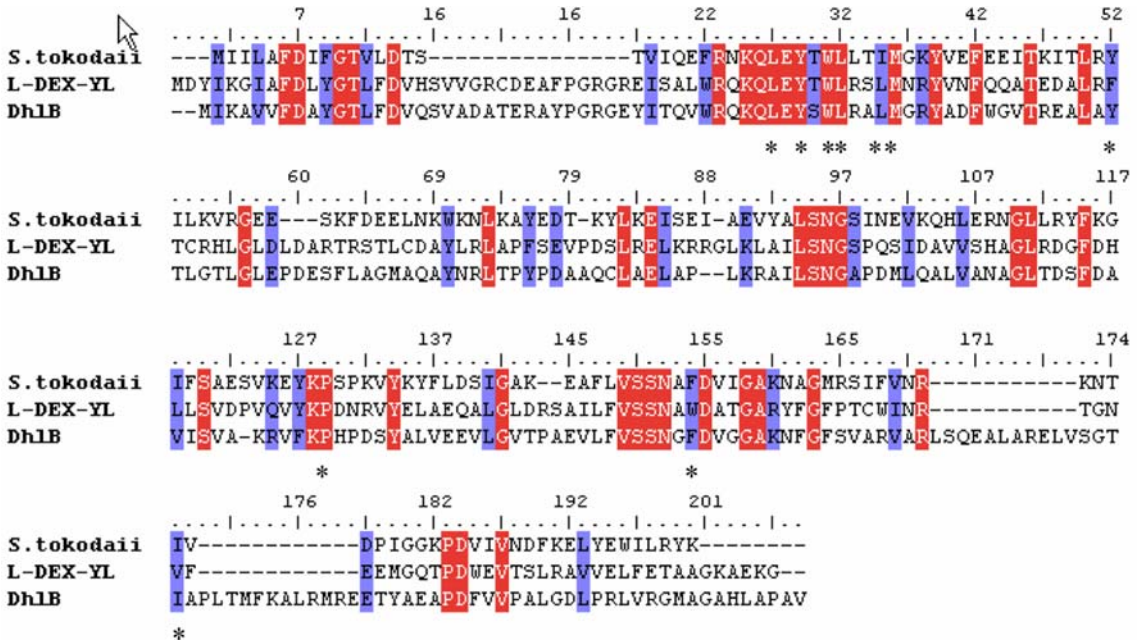


**Fig. 5** Sequence alignment of *S. tokodaii* haloacid dehalogenase (Accession Number: BAB67681) with L-DEX-YL (Accession Number: AAB32245) and DhlB (Accession Number: CAA60039). Identical residues shown in *red* and similar in *blue*, the conserved hydrophobic residues involved in dimerisation are highlighted by an *asterisk* (*)

between the α-helix 2 of each subunit in L-DEX-YL. DhlB has a similar interface but contains an extra dimerisation domain of two anti parallel α-helices. On dimerisation, the two α-helices in this extra domain form hydrophobic interactions with α-helices 2 and 3 of the subdomain (Ridder et al. 1997).

## Active site

The enzymes active site is located between the two domains of the protein. Amino acids from the core domain make up the bottom of this cavity and the subdomain shields the top. The proposed active site of *S. tokodaii* dehalogenase is composed of atoms from the conserved amino acid residues Asp7, Ala8, Phe9, Arg23, Leu27, Phe42, Leu94, Ser95, Asn96, Gly97, Lys128, Asn152, Phe154 and Asp155. The only difference in these residues between *S. tokodaii* dehalogenase and DhlB/L-DEX-YL is that Phe9 of *S. tokodaii* dehalogenase is replaced with a tyrosine.

Mutagenesis studies on L-DEX-YL have identified a number of important residues; the mutation of these residues results in reduced activities (Kurihara et al. 1995). Based on sequence comparisons, these residues in *S. tokodaii* dehalogenase are Asp7, Thr11, Arg23, Ser95, Lys128, Tyr134, Ser150, Asn152 and Asp155. These residues are completely conserved in DhlB and the majority are conserved in PH0459.

Studies of the ester intermediates of L-DEX-YL have previously identified Ser118 as the binding residue for the carboxylic acid group of the substrate as well as the main chain nitrogens of Leu11 and Tyr12 (Li et al. 1998). In DhlB, the residues Ser114, Ala9, and Tyr10 are performing these roles (Ridder et al. 1999). In *S. tokodaii* dehalogenase Ser95 is in the same position as the respective serine residues in L-DEX-YL and DhlB enzymes so is likely to perform this role, as are the main chains of residues Ile8 and Phe9.

The substrate carboxylic group binding site of the protein crystallised in the orthorhombic form is occupied by an inorganic phosphate molecule. Phosphate was present at 50 mM concentration in the gel filtration column buffer. However, the active site of the protein crystallised in the monoclinic form, which was purified in the Tris–HCl buffer, has additional density, which corresponds to a molecule with six non-hydrogen atoms. The carboxylate group of this putative compound has the same interactions as phosphate in the other crystal form and as the covalent intermediates in other known haloacid dehalogenase structures (PDB codes 1QQ7, 1QQ5; Ridder et al. 1999). No additional compound with such properties was present in the purification or crystallisation buffers. It is therefore proposed that it has come from the *E. coli* host cells and its high affinity for the active site of the enzyme has allowed it to remain bound during the enzyme purification. Several candidate compounds—isobutyrate, pyruvate, L-lactate and

D-lactate which is a product of the haloacid dehalogenase reaction, were modelled into the additional electron density in several possible conformations which were subsequently refined. One conformation of L-lactate and one conformation of pyruvate had better fits to the electron density and more consistent B-factors than the other candidate molecules/conformations. However the pyruvate geometry was distorted after refinement whereas the L-lactate retained a nearly ideal geometry. Therefore, with the present data, the L-lactate is proposed to be the single molecule occupying this density (Fig. 6). Inhibition studies were carried using the standard colorimetric assay to determine whether lactate or pyruvate had any inhibitory effect on the activity of *S. tokadaii* haloacid dehalogenase. No inhibition was observed with the addition of up to 100 mM of D/L-lactate or pyruvate.

A hydrophobic pocket in L-DEX-YL composed of the residues Tyr12, Gln44, Leu45, Phe60, Lys151, Asn177 and Trp179 was shown to fit around the alkyl group of the substrate and play a role in stabilising the substrate as well as having roles in the stereospecificity. If the alkyl group and H atom of the substrate are swapped, the substrate will not fit in the active site as the alkyl group would be sterically hindered by the main chain and side chains of the amino acids Leu11, Tyr12 and nearby residues. In *S. tokodaii* dehalogenase the conserved residues likely to form the hydrophobic pocket are Phe9, Gln26, Leu27, Phe42, Lys128, Asn152 and Phe154.

The residue in the position of Phe154 is often a tryptophan in other haloacid dehalogenases including L-DEX-YL. Additionally, the majority of haloacid dehalogenase (including L-DEX-YL and DhlB) have a tyrosine residue in place of Phe9. The Phe9 residue in the proposed *S. tokodaii* hydrophobic pocket is the middle amino acid residue in the first HAD superfamily motif. It is this residue that is largely conserved amongst the different groups in the HAD superfamily. Mutations produced in this residue as well as to the other aromatic residue in the hydrophobic pocket of L-DEX-YL led to a large decrease in activity compared to the native enzyme (Kurihara et al. 1995). The mutation Y12F led to a 60% drop in activity, whilst the mutation W179F led to a 90% decrease in activity. It would therefore appear that the presence of these residues is important in forming hydrophobic interactions with the substrate.

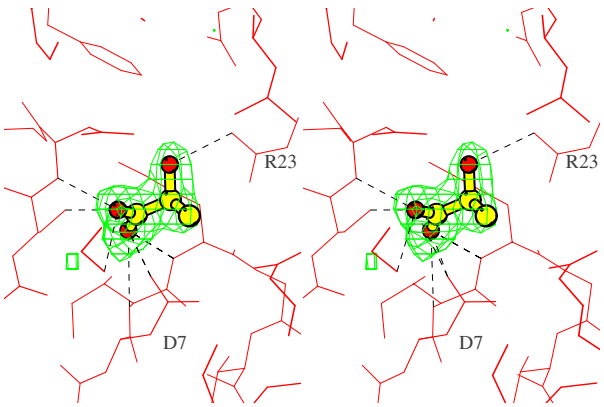


**Fig. 6** Stereo diagram showing the additional electron density in the haloacid dehalogenase active site in the monoclinic crystal form. The Fo–Fc map was calculated using phases from the model in which the ligand was omitted and the model was subjected to 20 cycles of refinement with REFMAC to remove model bias. The map was contoured at a level of 3$\sigma$. The best fitting ligand, L-lactate is shown as a ball-and-stick model. The figure was prepared using BOBSCRIPT (Esnouf 1997)

These interactions with the substrate produce the correct orientation for the conserved aspartic acid residue of the dehalogenase to nucleophilically attack the substrate leading to the formation of the ester intermediate. The ester intermediate is then hydrolysed to produce the hydroxyacid product.

## Residues involved in ester formation and hydrolysis

In all of the substrate bound structures of L-DEX-YL Asp10 was esterified with the dechlorinated moiety of the substrate. The formation of the ester intermediate results in conformational changes in this region, with the side chain of Asp10 adopting different conformations depending of the size of the substrate bound (Li et al. 1998). The residue Asp7 is the completely conserved active site nucleophile of *S. tokodaii* dehalogenase, which will perform the same function.

In both L-DEX-YL and PH0459 there is a salt bridge present between Lys155 and Asp8, which will reduce the pKa of Asp8 increasing its nucleophilicity (Arai et al. 2006). A similar salt bridge is present between Asp7 and Lys128 in the *S. tokodaii* dehalogenase. This combined with the fact that the remainder of the conserved amino acid residues implicated in ester formation and hydrolysis in L-DEX-YL and DhlB are in similar orientations in the *S. tokodaii* dehalogenase suggests the *S. tokodaii* enzyme is likely to carry out ester formation in a similar way.

## Halide stabilisation and abstraction

The residue Arg23 from the subdomain of *S. tokodaii* dehalogenase points into the active site. In L-DEX-YL the equivalent residue has been shown to be the acceptor of the halide (Ridder et al. 1999). Studies on DhlB by Ridder et al. (1997) showed that the halide stabilisation cradle was made of Arg39, Asn115 and Phe175. The arginine and asparagine residues are conserved in L-DEX-YL and *S. tokodaii* haloacid dehalogenase, however, the phenylalanine is replaced with a tryptophan residue in L-DEX-YL. As mentioned earlier, mutations to this tryptophan residue

in L-DEX-YL resulted in a decreased activity. Superimpositions of the halide stabilising cradles of *S. tokodaii* dehalogenase, L-DEX-YL and DhlB reveal that the residues are in very similar orientations.

## Biochemical characterisation

### Substrate specificity

The dehalogenases have different substrate specificities, with some being able to accept much larger substrates than others. The PH0459 dehalogenase is a monomer and its active site is very open to solvent. Whereas the DhlB dehalogenase has a much more enclosed active site and has been shown to only accept small substrates up to the size of chloropropionic acid (van der Ploeg et al. 1991). Much of the active site hindrance in DhlB dehalogenase is due to the presence of the extra dimerisation domain that is not present in the other haloacid dehalogenase structures. The L-DEX-YL dehalogenase active site is much more exposed and this enzyme has been shown to have activity towards substrates with longer chain lengths, especially in organic solvents (Liu et al. 1995). The *S. tokodaii* dehalogenase active site appears less exposed than that of L-DEX-YL dehalogenase but more exposed than the DhlB dehalogenase which is consistent with the activity of this enzyme towards 2-chlorobutyric acid. Surface representations of the four dehalogenases with known structures are shown in Fig. 7.

### Solvent stability

Many industrial processes are carried out in the presence of organic solvents, therefore an enzymes stability in organic solvent can be an important factor in deciding whether it is suitable to use in an industrial process. Organic solvents can lead to the stabilisation of 2-bromoalkanoic acids and so it can be very useful for a haloacid dehalogenase to be stable in organic solvents. The L-2-haloacid dehalogenase from *Pseudomonas* sp. YL can use longer chain 2-haloalkanoic acids (including 2-bromohexadecanoate) when the reaction is carried out in n-heptane (Liu et al. 1995). The activity in n-heptane is however, greatly reduced with the activity towards 2-chloropropionic acid reduced by over 95%. Previous studies on haloacid dehalogenases show that activities are often greatly reduced by a number of organic solvents. By incubating *S. tokodaii* haloacid dehalogenase with different concentrations of ethanol, methanol, DMSO and acetonitrile for 1 h, the solvent stability was assessed (Fig. 8). DMSO had the least inhibitory

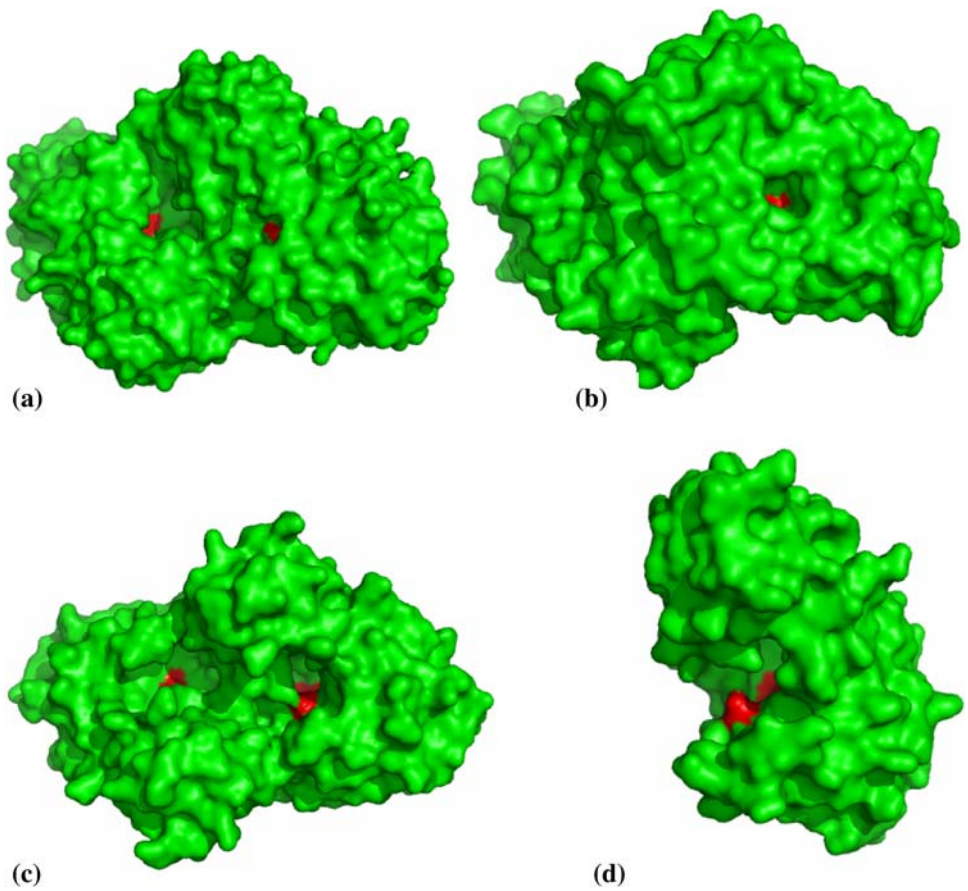


**Fig. 7** Surface representations of **a** *S. tokodaii* dehalogenase, **b** DhlB, **c** L-DEX-YL dehalogenase and **d** PH0459. All of the structures are oriented in the same direction and the active site residues are shown in *red*. Figure produced using PyMOL (Delano Scientific)

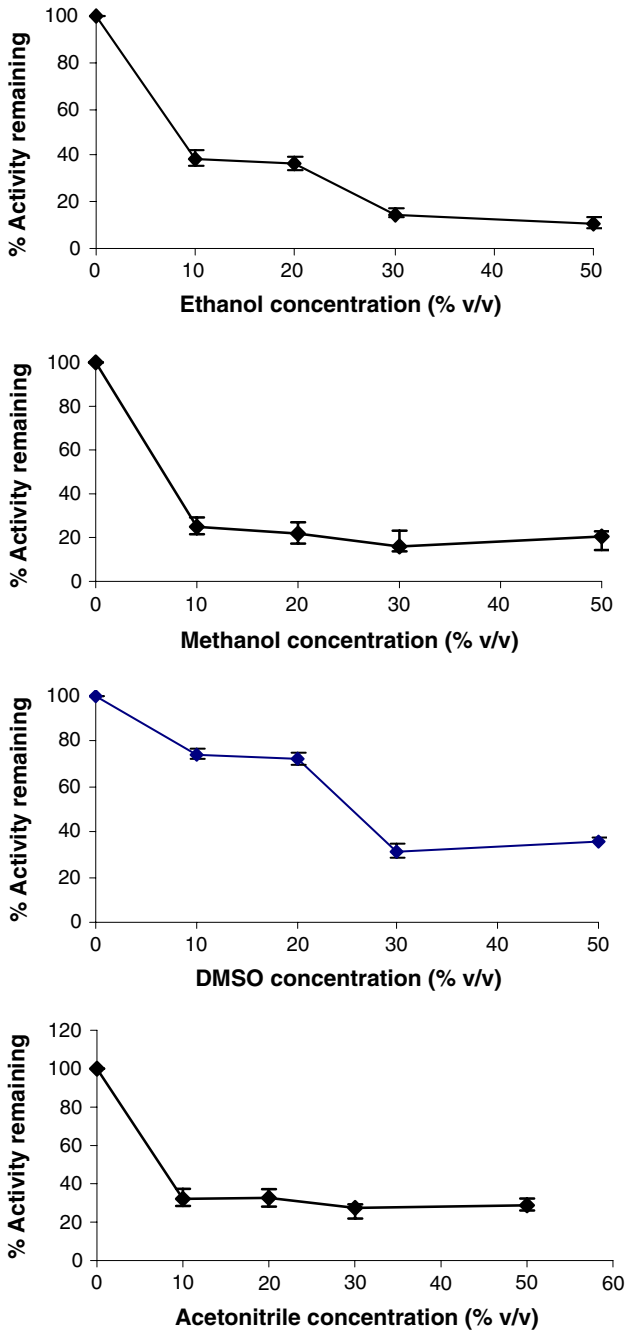


**Fig. 8** The activity of *S. tokodaii* haloacid dehalogenase after exposure to different concentrations of solvent

effect of all the solvents tested, with only 25% activity lost with the addition of up to 20% v/v DMSO. However, at higher concentrations of DMSO the activity was reduced to below 40% of normal. The activity in 10% v/v methanol was only 25% of normal; however, similar activity was observed for the higher concentrations of methanol studied. Surprisingly in 10% v/v ethanol activity was still 40% of normal. However, further increases in ethanol concentration reduced the activity until at 50%(v/v) ethanol less than

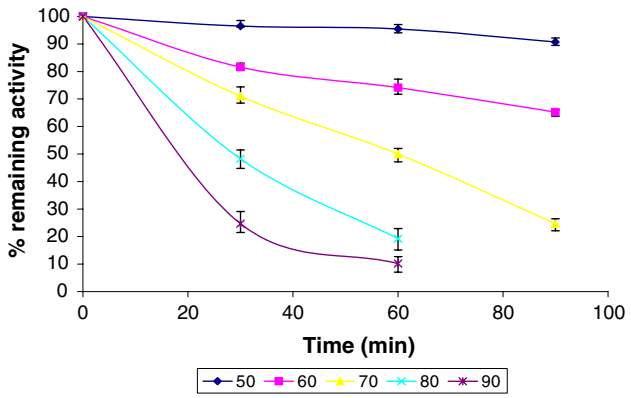


**Fig. 9** Thermal deactivation of *S. tokodaii* haloacid dehalogenase shown as a percentage of the activity remaining at different incubation temperatures

10% activity was observed. In acetonitrile the activity was reduced by over 70% for any concentration tested.

## Thermostability

The optimal growth temperature for *S. tokodaii* is 75–80°C and therefore its enzymes, including the haloacid dehalogenase, are likely to have maximum activities at elevated temperatures. Experiments were carried out by assaying the enzyme at temperatures between 20 and 70°C. The activity of the dehalogenase increases steadily up to 60°C after which it begins to fall. There are no other reports for thermophilic dehalogenases currently published, however, a number of dehalogenases from mesophilic sources have been characterised. Interestingly L-DEX-YL from *Pseudomonas* sp. YL displayed most activity at a temperature of 65°C even though the optimal growth temperature of the organism is 30°C. The *Azotobacter* sp. RC26 haloacid dehalogenase has maximum activity between 30 and 40°C but by immobilisation of the enzyme on ion-exchange media this increases to 60°C (Diez et al. 1996).

The reaction rate and enzyme stability will determine the temperature at which an industrial process is carried out. Sometimes it is cost effective to run a process at a high temperature where the enzyme displays maximum activity and substrates are soluble but in other cases reactions are carried out at moderate temperatures to reduce running costs. The *S. tokodaii* haloacid dehalogenase is still active at 20°C but displays approximately 20% of the maximum activity and at 40°C, 85% of the maximum activity is observed.

Temperature stability studies were carried out between 50 and 90°C for 30, 60 and 90 min and the remaining activity plotted as a percentage of original activity (Fig. 9). The *S. tokodaii* dehalogenase was found to be stable to prolonged exposure to 50°C, losing only 8% activity after 90 min incubation. As the temperature was increased the stability

decreased. The optimal growth temperature of *S. tokodaii* is between 75 and 80°C so as expected at temperatures greater than this the half-life was greatly reduced. This temperature stability is comparable to L-DEX-YL from *Pseudomonas* sp. YL which is not inactivated by 30 min incubation at 60°C (Liu et al. 1995). In contrast the L-haloacid dehalogenase from *P. putida* 109 loses 75% activity after a 30 min 60°C heat treatment (Motosugi et al. 1982) and the haloacid dehalogenase from *B. cepacia* loses 25% activity when heated to 35°C for 15 min (Tsang and Sam 1999).

As structures are available for two mesophilic L-haloacid dehalogenase enzymes as well as the structure from a hyperthermophile dehalogenase a detailed comparison can be carried out with the structure of the *S. tokodaii* enzyme to help determine potential mechanisms of thermostability.

The α-helices 2 and 3 of the *S. tokodaii* haloacid dehalogenase subdomain run side by side and both of these helices contain negatively charged residues towards their N-termini and positively charged residues towards the C-termini. This could play a role in stabilising the protein. Higher oligomeric states are often observed in thermophilic proteins. Both of the mesophilic dehalogenase structures are dimers, with each subunit containing an active site. The thermophilic *S. tokodaii* dehalogenase adopts the same oligomeric state. In contrast the hyperthermophilic PH0459 enzyme is a monomer stabilised by a disulfide bond (Arai et al. 2006). A number of X-ray crystallographic structures of hyperthermophilic proteins (Mallick et al. 2002) have shown the presence of disulfide bonds. The *S. tokodaii* enzyme does not contain any cysteine residues so is not capable of using this mechanism of stabilisation. Putative haloacid dehalogenase sequences from other *Sulfolobus* species also lack the cysteine residues needed for disulfide formation. The same is true for the putative dehalogenases from the thermophiles *Metallosphaera sedula* and *Haloarcula marismortui*. The two cysteines residues in the *P. horikoshii* dehalogenase are however conserved amongst other putative haloacid dehalogenases from hyperthermophilic archaea (proteins PF0463 from *P. furiosus* and TK0058 from *T. kodakarensis*) (Arai et al. 2006). There is no information available about the oligomeric state that these putative dehalogenases adopt.

The amino acid composition of an enzyme has been widely implicated as an important factor for thermostability. The preference for different amino acids in thermostable proteins include an increase in charged amino acid residues and an increase in proline residues. Within the haloacid dehalogenases, there is a general increase in the percentage of charged amino acid residues as the optimal growth temperature increases (Table 2). These can be involved in the formation of an increased number of ionic interactions to stabilise the protein structure. There is almost a 10% increase in charged residues between the mesophilic and thermophilic dehalogenase sequences. An increase in glutamic acid accounts for most of the charged amino acid content in the dehalogenases from thermophilic organisms. Glutamic acid will stabilise helix formation (Pace and Scholtz 1998) and its content nearly doubles between the dehalogenases from mesophilic and thermophilic organisms. The majority of glutamic acid residues in the four dehalogenases are in helices, with PH0459 having the highest by number (18) and percentage (87%) in helices. In the *S. tokodaii* haloacid dehalogenase there are salt bridges present between glutamic acid and lysine residues in α-helices 3 and 4. These interactions are not present in the dehalogenases from mesophilic organisms, as the amino acid residues involved are not conserved. Similar interactions are seen in the α-helices of PH0459. These interactions are likely to play a role in stabilising the α-helices and the overall protein structure.

Of the amino acid substitutions commonly seen in thermophilic proteins, the most common are glycine substituted by alanine and lysine by arginine. However, amongst the six dehalogenases, no trends are seen in the amino acid content of these four residues. There are actually more alanine residues in the mesophilic dehalogenases. There are also more lysine residues in the thermophilic structures and more arginine residues in the mesophilic structures.

Determination of the structure of the *S. tokodaii* dehalogenase has given some insight into its substrate specificity and mechanisms of thermostability when compared with other L-haloacid dehalogenases. Its biochemical properties with a view to its potential for commercial applications demonstrate that it can be used at both high and moderate

**Table 2** Amino acid composition of the four L-haloacid dehalogenases for which there are available structures and two with no known structure

| | PH0459 | *S. tokodaii* | L-DEX-YL | DhlB | *P. putida* | *B. cepacia* |
|---|---|---|---|---|---|---|
| Optimal growth temp. | 90–105°C | 75–80°C | 30°C (optimal activity 65°C) | 28°C | 28°C | 28°C |
| Chain length | 232 | 201 | 232 | 253 | 224 | 280 |
| Charged residues | 32.3 | 29.4 | 25.0 | 21.0 | 24.9 | 19.2 |
| Uncharged polar residues | 16.0 | 23.4 | 21.5 | 19.0 | 18.7 | 17.5 |
| Aromatic residues | 10.6 | 13.4 | 10.8 | 9.9 | 9.3 | 6.8 |

Two are from thermophilic archaeal sources (*S. tokodaii* and PH0459) and four are from mesophilic bacterial sources (L-DEX-YL, DhlB, *P. putida* 109 and *B. cepacia*)

temperatures as well as having increased stability to temperature and in organic solvents.

The structure factors and refined coordinates for each crystal form of the *S. tokodaii* dehalogenase have been deposited with the Protein Data Bank; the access codes are 2W11 and 2W43.

**Acknowledgments** The Engineering and Physical Sciences Research Council, UK and TMO Renewables are acknowledged for PhD studentship funding to CAR. We thank Prof. T. Oshima for providing the *S. tokodaii* genomic DNA and we thank Dr. Roger Cripps for useful discussions. AAL was supported by a BBSRC grant. The Synchrotron Radiation Source, Daresbury, UK is acknowledged for provision of Data Collection Facilities on the NWSGC MAD10 beamline funded by BBSRC grants (719/B15474 and 719/REI20571) and an NWDA project award (N0002170).

## References

Allpress JD, Gowland PC (1998) Dehalogenases: environmental defence mechanism and model of enzyme evolution. Biochem Education 26:267–276

Arai R, Kukimoto-Niino M, Kuroishi C, Bessho Y, Shirouzu M, Yokoyama S (2006) Crystal structure of the probable haloacid dehalogenase PH0459 from *Pyrococcus horikoshii* OT3. Protein Sci 15:373–377

CCP4. Collaborative Computational Project, Number 4 (1994) Acta Cryst D 50:760–763

Diez A, Prieto MI, Alvarez MJ, Bautista JM, Garrido J, Puyet A (1996) Improved catalytic performance of a 2-haloacid dehalogenase from *Azotobacter* sp. by ion-exchange immobilisation. Biochem Biophys Res Comm 220:828–833

Emsley P, Cowtan K (2004) COOT: model-building tools for molecular graphics. Acta Cryst D 60:2184–2195

Esnouf RM (1997) An extensively modified version of MOLSCRIPT that includes greatly enhanced colouring capabilities. J Mol Graph Model 15:132–134

Hisano T, Hata Y, Fujii T, Liu JQ, Kurihara T, Esaki N, Soda K (1996) Crystal structure of L-2-haloacid dehalogenase from *Pseudomonas* sp. YL. J Biol Chem 271:20322–20330

Kawarabayasi Y, Hino Y, Horikawa H, Jin-no K, Takahashi M, Sekine M, Baba S, Ankai A, Kosugi H, Hosoyama A, Fukui S, Nagai Y, Nishijima K, Otsuka R, Nakazawa H, Takamiya M, Kato Y, Yoshizawa T, Tanaka T, Kudoh Y, Yamazaki J, Kushida N, Oguchi A, Aoki K, Masuda S, Yanagii M, Nishimura M, Yamagishi A, Oshima T, Kikuchi H (2001) Complete genome sequence of an aerobic thermoacidophilic crenarchaeon *Sulfolobus tokodaii* strain 7. DNA Res 8:123–140

Kurihara T, Liu JQ, Nardi-Dee V (1995) Comprehensive site directed mutagenesis of L-2-haloacid dehalogenases to probe catalytic amino acid residues. J Biochem 117:1317–1322

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291

Li YF, Hato Y, Fujii T, Hisano T, Nishihara M, Kurihara T, Esaki N (1998) Crystal structures of reaction intermediates of L-2-haloacid dehalogenase and implications for the reaction mechanism. J Biol Chem 273:15035–15044

Liu JQ, Kurihara T, Hasan KMQ, Nardi-dee V, Koshikawa N, Esaki N, Soda K (1995) Purification of thermostable and nonthermostable 2-haloacid dehalogenases with different stereospecificities from *Pseudomonas* sp. strain YL. J Biol Chem 270:18309–18312

Mallick P, Boutz DR, Eisenberg D, Yeates TO (2002) Genomic evidence that the intracellular proteins of archaeal microbes contain disulfide bonds. Proc Natl Acad Sci USA 99:9679–9684

Motosugi K, Esaki N, Soda K (1982) Purification and properties of a new enzyme, DL-2-haloacid dehalogenase from *Pseudomonas* sp. J Bacteriol 150:522–527

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Cryst D 53:240–255

Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscullation mode. Methods Enzymol 276:307–326

Pace CN, Scholtz JM (1998) A helix propensity scale based on experimental studies of peptides and proteins. Biophys J 75:422–427

Potterton P, McNicholas S, Krissinel E, Gruber J, Cowtan K, Emsley P, Murshudov GN, Cohen S, Perrakis A, Noble M (2004) Developments in the CCP4 molecular-graphics project. Acta Cryst D 60:2288–2294

Ridder IS, Rozeboom HJ, Kalk KH, Dijkstra BW (1997) Three-dimensional structure of L-2-haloacid dehalogenases from *Xanthobacter autotrophicus* GJ10 complexed with the substrate analogue formate. J Biol Chem 272:33015–33022

Ridder IS, Rozeboom HJ, Kalk KH, Dijkstra BW (1999) Crystal structures of intermediates in the dehalogenation of haloalkanoates by L-2-haloacid dehalogenase. J Biol Chem 274:30672–30678

Rye CA, Isupov MN, Lebedev AA, Littlechild JA (2007) An order-disorder twin crystal of L-2-haloacid dehalogenase from *Sulfolobus tokodaii*. Acta Cryst D 63:926–930

Slater JH (1982) New microbes to tackle toxic compounds. South Af J Sci 78:101–104

Slater JH, Bull AT, Hardman DJ (1997) Microbial dehalogenation of halogenated alkanoic acids, alcohols and alkanes. Adv Microbiol Physiol 38:133–176

Tsang JSH, Sam L (1999) Cloning and characterisation of a cryptic haloacid dehalogenases from *Burkholderia cepacia* MBA4. J Bacteriol 181:6003–6009

Vagin AA, Teplyakov A (1997) MOLREP: an automated program for molecular replacement. J Appl Crystallogr 30:1022–1025

van der Ploeg J, van Hall G, Janssen DB (1991) Characterization of the haloacid dehalogenase from *Xanthobacter autotrophicus* GJ10 and sequencing of the DhlB gene. J Bacteriol 173:7925–7933

Vriend G (1990) WHAT IF: a molecular modeling and drug design program. J Mol Graph 8:52–56