


X-ray structure and mechanism of ZgHAD, a L-2-haloacid dehalogenase from the marine *Flavobacterium Zobellia galactanivorans*

Eugénie Grigorian¹ | Thomas Roret² | Mirjam Czjzek¹ | Catherine Leblanc¹ | Ludovic Delage¹ 

¹Station Biologique de Roscoff (SBR), Sorbonne Université, CNRS, Integrative Biology of Marine Models (LBI2M), Roscoff, France

²Station Biologique de Roscoff (SBR), CNRS FR2424, Sorbonne Université, Roscoff, France

Correspondence

Mirjam Czjzek and Ludovic Delage, Station Biologique de Roscoff (SBR), Sorbonne Université, CNRS, Integrative Biology of Marine Models (LBI2M), 29680 Roscoff, Bretagne, France.

Email: czjzek@sb-roscoff.fr and ludovic.delage@sb-roscoff.fr

Funding information

Agence Nationale de la Recherche, Grant/Award Number: ANR-10-BTBR-04; Région Bretagne (projet MHALIN), Grant/Award Number: ARED 2017; Centre National de la Recherche Scientifique

Review Editor: Jeanine Amacher

Abstract

Haloacid dehalogenases are potentially involved in bioremediation of contaminated environments and few have been biochemically characterized from marine organisms. The L-2-haloacid dehalogenase (L-2-HAD) from the marine *Bacteroidetes Zobellia galactanivorans* Dsij^T (ZgHAD) has been shown to catalyze the dehalogenation of C2 and C3 short-chain L-2-haloalkanoic acids. To better understand its catalytic properties, its enzymatic stability, active site, and 3D structure were analyzed. ZgHAD demonstrates high stability to solvents and a conserved catalytic activity when heated up to 60°C, its melting temperature being at 65°C. The X-ray structure of the recombinant enzyme was solved by molecular replacement. The enzyme folds as a homodimer and its active site is very similar to DehRhb, the other known L-2-HAD from a marine *Rhodobacteraceae*. Marked differences are present in the putative substrate entrance sites of the two enzymes. The H179 amino acid potentially involved in the activation of a catalytic water molecule was confirmed as catalytic amino acid through the production of two inactive site-directed mutants. The crystal packing of 13 dimers in the asymmetric unit of an active-site mutant, ZgHAD-H179N, reveals domain movements of the monomeric subunits relative to each other. The involvement of a catalytic His/Glu dyad and substrate binding amino acids was further confirmed by computational docking. All together our results give new insights into the catalytic mechanism of the group of marine L-2-HAD.

KEYWORDS

catalytic mechanism, computational docking, crystal structure, His/Glu dyad, L-2-haloacid dehalogenase, marine *Bacteroidetes*, *Zobellia galactanivorans*

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *Protein Science* published by Wiley Periodicals LLC on behalf of The Protein Society.

1 | INTRODUCTION

The haloacid dehalogenases (HAD) superfamily includes dehalogenating enzymes together with diverse enzymes that hydrolyze carbon-phosphorus bonds, such as epoxide hydrolases, phosphatases, phosphomutases, or nucleotidases. Widely present among living organisms, they are involved in a variety of cellular processes ranging from amino acid biosynthesis to detoxification (Burroughs et al., 2006). Since the industrial boom, halogenated xenobiotic pollutants are contaminating soils and aquatic environments. The accumulation of those toxic compounds led to the research of new tools for detoxification and bioremediation.

The “true” dehalogenases of the HAD superfamily are classified into four types relative to the substrate specificity and stereoselectivity. D-2-haloacid dehalogenases (D-2-HADs) and L-2-haloacid dehalogenases (L-2-HADs) have a strict enantioselective dehalogenating activity on D-2-haloacids and L-2-haloacids, respectively, to produce the corresponding alcohols with an inverted chirality. The two other types are DL-2-haloacid dehalogenases (DL-2-HADs) which accept both D- or L-2-haloacids as substrates: DL-2-HAD_i act with a configuration-inverting mechanism whereas DL-2-HAD_r retain the configuration. The 2-haloacid dehalogenases are also categorized into two groups according to their amino acid sequence homology. D-2-HADs (EC 3.8.1.9) and DL-2-HADs (EC 3.8.1.10 and EC 3.8.1.11) are part of Group I and L-2-HADs (E.C. 3.8.1.2) belong to Group II (Ang et al., 2018; Wang et al., 2021).

In addition, L-2-HADs were recently classified on the basis of phylogenetic/environment analyses as two monophyletic groups, where the group A contains a mix of terrestrial and marine sequences and the group B includes mostly marine sequences (Grigorian et al., 2021). There are currently nine available crystal structures of L-2-HADs from diverse organisms, mainly bacteria. Four of these 3D structures belong to the group A and were solved before 2020. They correspond to L-DEX YL from *Pseudomonas* sp. YL (Hisano et al., 1996), Dh1B from *Xanthobacter autotrophicus* GJ10 (Ridder et al., 1997), DehIVa from *Burkholderia cepacia* (Schmidberger et al., 2007), and DehSft from *Sulfolobus tokodaii* (Rye et al., 2009). Four other crystal structures were determined very recently and they led to the discovery of the first defluorinating L-2-HAD enzymes for two of them, named Bpro0530 from *Polaromonas* sp. JS666 and Rha0230 from *Rhodococcus* sp. RHA1 isolated from polluted environments (Chan et al., 2022). All these eight characterized L-2-HADs originate from terrestrial bacteria. The last L-2-HAD characterized structure was DehRhB from a marine *Rhodobacteraceae* sp. (Novak et al., 2013) and the only one relative to the phylogenetic group B (Grigorian et al., 2021).

All the biochemically and structurally described L-2-HAD enzymes to date are dimers with two domains in each subunit. The L-2-HADs consist in a characteristic core domain with a conserved alpha/beta hydrolase fold, similar to the “Rossmann-fold,” and a second small cap domain exhibiting varying folds and functions. This cap domain is responsible for the biochemical diversification within the HAD superfamily (Lahiri et al., 2004). These enzymes transform the substrate(s) molecule(s) according to a conserved nucleophilic substitution involving a conserved aspartic acid that forms an intermediate ester bond with the substrate. In the case of L-2-HADs, the enzyme-substrate ester bond is then hydrolyzed by another nucleophilic attack with an activated water molecule (Liu et al., 1995; Nardi-Dei et al., 1997). In contrast to other HAD superfamily enzymes, the cap domain is similar in all L-2-HADs and is composed of a four-helix bundle where the active site is flanked by a hydrophobic cavity situated in between the core and cap domains (Ridder et al., 1997; Rye et al., 2009; Schmidberger et al., 2007).

The complete genome sequence of the marine flavobacteria *Zobellia galactanivorans* Dsjj^T revealed the presence of a L-2-HAD enzyme, further on named ZgHAD. Subsequent to cloning and purification, the biochemical characterization of the recombinant ZgHAD enzyme has been described recently (Grigorian et al., 2021). The enzyme is specific toward L-2-enantiomer substrates having a short carbon chain (C2 and C3) and it can deiodinate, debrominate, or dechlorinate the α -carbon position. The highest activity was observed with iodoacetic and bromoacetic acids, while the reactions with chloroacetic and L-2-bromopropionic acids were much lower, suggesting catalytic specificities when compared to the other characterized marine L-2-HAD (Grigorian et al., 2021).

Here, we describe the 3D crystal structure of wild-type ZgHAD, as well as that of two point-mutated enzymes, ZgHAD-H179A and ZgHAD-H179N, and, together with computational docking we compare the observed structural details with those of related enzymes, previously described. The thermal and solvent stabilities of this enzyme have also been assessed in view of potential biotechnological applications.

2 | RESULTS

2.1 | Thermostability and solvent stability

A temperature gradient between 20 and 95°C was applied to evaluate the stability zone of the protein fold. The experiment showed that ZgHAD was denatured between

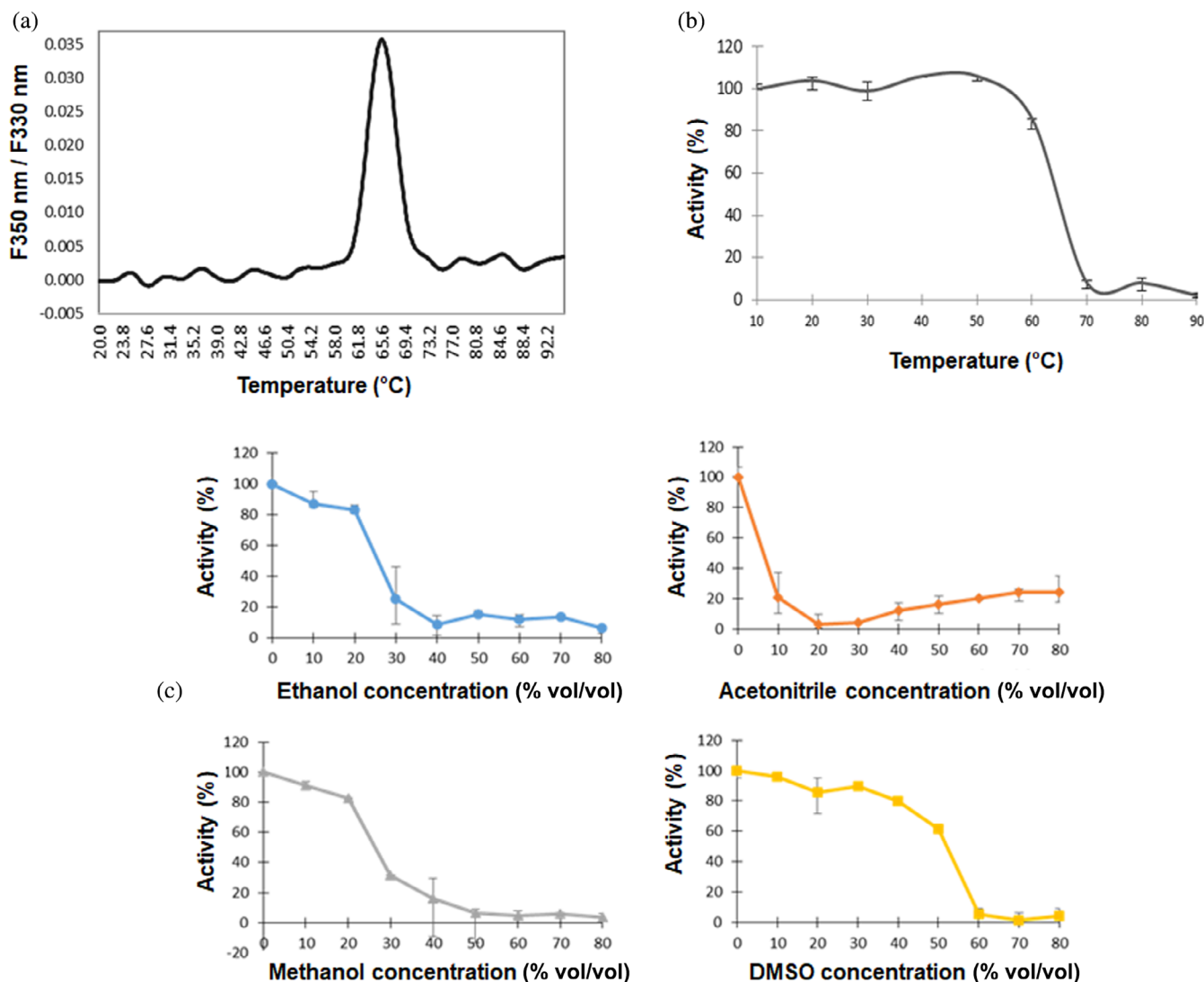


FIGURE 1 Thermal and solvent stability of the recombinant L-2-HAD from *Zobellia galactanivorans*. (a) Thermal break point of ZgHAD protein as determined by the Prometheus NT.48. (b) Thermal stability of ZgHAD activity as determined after the pre-incubation of the enzyme at varying temperatures for 30 min before measuring the residual activity at 20°C for all the points. (c) Solvent stability of ZgHAD activity as determined after pre-exposure to different concentrations of ethanol, methanol, acetonitrile and DMSO for 1 h before measuring the residual activity in a standard solvent concentration

60 and 70°C with a melting temperature of 65°C (Figure 1a), as also confirmed by dynamic light scattering measurements (Figure S1). These results suggest that ZgHAD is thermostable up to the maximal temperature of 65°C. The thermal stability of the enzymatic activity was also investigated between 10 and 90°C and after 30 min of enzyme incubation. The residual activity was measured and plotted as a percentage of the initial activity (Figure 1b). ZgHAD activity was found to be stable up to 50°C but decreased rapidly between 60 and 70°C, from 80% residual activity to a complete loss of activity. The results on protein denaturation and enzymatic activity are strongly correlated and show that ZgHAD turns completely unfolded and therefore inactive at 70°C.

The solvent stability of ZgHAD was assessed by incubating the enzyme with different concentrations of ethanol, methanol, acetonitrile and dimethyl sulfoxide (DMSO) for 1 h before evaluating the residual activity (Figure 1c). The effects of ethanol and methanol treatments were very similar. The activity of the enzyme was conserved up to 80%, when 10%–20% of these organic solvents were added, and decreased to 20% in presence of 20%–40% of these alcohols. Above 40% of methanol or ethanol, the activity was almost completely inhibited. DMSO had the least inhibitory effect of all the solvents tested. Only 20% loss of activity was observed after addition of up to 40% DMSO. The activity decreased to 60% with 50% of DMSO, and higher concentrations drastically reduced the

TABLE 1 Data collection and refinement statistics for ZgHAD

| Data collection | ZgHAD wt | ZgHAD H179A | ZgHAD H179N |
|------------------------------------|--|--|----------------------|
| Beam line | PROXIMA-2A | PROXIMA-2A | PROXIMA-2A |
| Space group | P 2 ₁ 2 ₁ 2 ₁ | P 2 ₁ 2 ₁ 2 ₁ | P 2 ₁ |
| Average unit cell (Å) | 59.53 68.83103.74 | 59.716 71.655116.009 | 76.17132.79275.70 |
| Wavelength (Å) | 0.97986 | 0.980116 | 0.980114 |
| Resolution (Å) | 45.03–1.60 | 45.874–1.716 | 49.46–2.75 |
| R_{merge} | 0.098 (0.745) | 0.080 (0.872) | 0.132 (1.471) |
| R_{meas} | 0.102 (0.774) | 0.083 (0.908) | 0.143 (1.612) |
| R_{pim} | 0.028 (0.208) | 0.023 (0.252) | 0.054 (0.643) |
| No. unique reflections | 41,657 (2364) | 41,318 (2628) | 144,483 (4343) |
| Mean $I/\sigma I$ | 15.2 (3.1) | 21.6 (3.5) | 9.6 (0.9) |
| $CC_{1/2}$ | 0.998 (0.896) | 0.999 (0.925) | 0.998 (0.447) |
| Completeness (%) | 100 (100) | 100 (100) | 98.0 (60.1) |
| Average redundancy | 13.5 (13.7) | 13.0 (13.1) | 6.9 (5.6) |
| Wilson B-factor (Å ²) | 21.70 | 25.06 | 83.7 |
| <i>Refinement</i> | | | |
| Resolution (Å) | 1.78 | 1.88 | 2.75 |
| $R_{\text{free}}/R_{\text{work}}$ | 0.1774/0.1552 | 0.1991/0.1680 | 0.2543/0.2065 |
| Total number of atoms | 7481 | 7415 | 44,949 |
| Water | 473 | 413 | 56 |
| Average B factor (Å ²) | 24.7 | 29.6 | 73.31 |
| Ligands | PO4; SCN | PO4; SCN | PO4 |
| <i>RMS deviations</i> | | | |
| Bonds | 0.010 | 0.017 | 0.009 |
| Angles | 1.0 | 1.4 | 1.5 |
| <i>MolProbity analysis</i> | | | |
| Clashscore, all atoms | 3.14 | 2.14 | 4 |
| MolProbity score | 1.21 | 1.02 | 1.21 |
| Ramachandran outliers | 0.0% | 0.0% | 0.12% (7) |
| Ramachandran allowed | 0.46% | 0.68% | 1.96% |
| Ramachandran favored | 99.54% | 99.32% | 97.91% |
| PDB entry | 7ARP | 7ASZ | 7QNM |

Note: Values in parentheses refer to the outer resolution shell.

activity to 10% of the control without any solvent. In acetonitrile, the activity of ZgHAD was drastically reduced to 20% in the presence of 10% of this solvent. It was completely inhibited between 20% and 30% of acetonitrile, and surprisingly it increased again for higher concentrations tested, with a stabilization around 20% of residual activity in the presence of 60%–80% of acetonitrile.

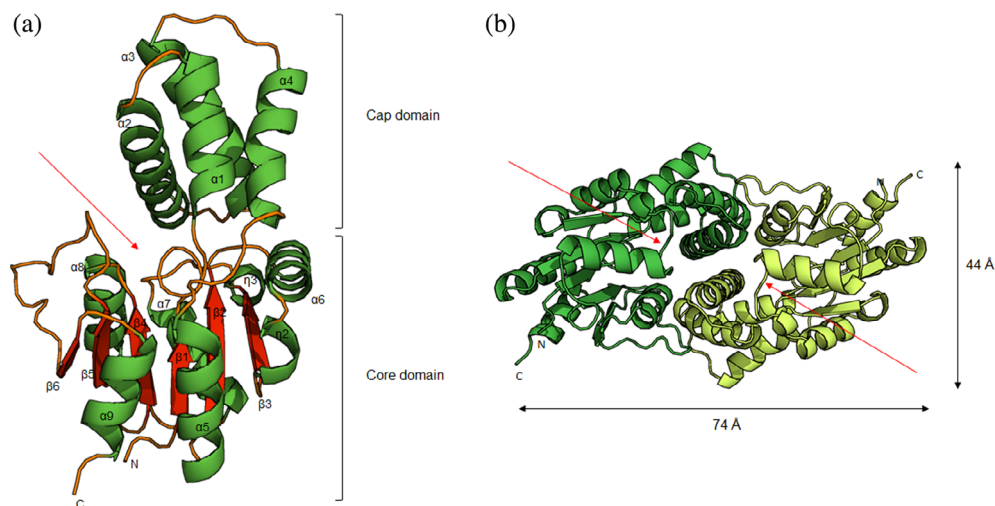
2.2 | Overall structure of ZgHAD

The crystal structure of ZgHAD was solved by the molecular replacement method at 1.6 Å, using the closest

structural representative, DehRhb (PDB accession: [2YML](#)) from a marine *Rhodobacteraceae*. The two proteins share 31% identity and 50% similarity of amino acid sequences. ZgHAD crystallized with the space group P2₁2₁2₁ and the mutant ZgHAD_H179N with space group P2₁; unit cell parameters are reported in Table 1.

ZgHAD crystallized as a homodimer and, likewise all reported HAD enzymes, the monomeric subunit is composed of two domains, comprising a core domain, formed by the residues 18–30 and 110–238, and a cap domain, formed by the residues 31–109 (Figure 2a). The core domain has a typical “Rossmann fold” which consists of six parallel β-strands surrounded by five α-helices and

FIGURE 2 (a) Overall fold of the ZgHAD monomeric subunit presented as a ribbon diagram colored by secondary structure elements. The α helices, β strands, and loops are colored in green, red, and orange respectively. (b) Ribbon diagram of the ZgHAD dimer viewed along the twofold horizontal axis. Each subunit is shown in a different color. The red arrows indicate the position of the catalytic site



three 3_{10} helices. Overall, the connectivity follows the pattern β -strand– α -helix– β -strand, except for β -strands 5 and 6, which are connected by a β -turn. 3_{10} helices are found before and after strand $\beta 3$. The cap domain is composed of four α -helices, and a 3_{10} helix is inserted in the core domain between strand $\beta 1$ and helix $\alpha 5$. The active site is located between the core and the cap domains, right after strand $\beta 1$. The four helices of the cap domain shield the top of the active site cavity from the solvent (Figure 2a). The core and the cap domains of ZgHAD are similar to those of DehRhB (2YML), DehSft (2W43), DehIVa (2NO4), and L-DEX YL (1ZRM) with root-mean-square deviation (RMSD) values between matching C α positions of 1.21 Å over 212, 1.65 Å over 194, 1.77 Å over 213, and 1.30 Å over 203 residues, respectively.

The homodimer of ZgHAD has the dimensions of $74 \times 35 \times 44$ Å (Figure 2b) and this oligomeric state is in agreement with the estimated size of the protein in solution determined by gel filtration (as referred in Grigorian et al., 2021). Up to date, all other structurally-characterized L-2-HADs have been reported to occur as homodimers, except for one putative L-2-HAD named PH0459 that has been crystallized in a monomeric state, but no dehalogenase activity was yet reported for this enzyme (Arai et al., 2006). Similar to these other L-2-HADs, the two subunits are related by a twofold symmetry axis running nearly parallel to the $\alpha 2$ helix. On dimer formation, 16% of each monomeric subunit's accessible surface is buried at the interface as identified by PDBePISA server (Krissinel & Henrick, 2007). In comparison, this value ranges between 13.4% in DehSft to 19% in Dh1B. The subunit interface of ZgHAD is mainly formed by helices $\alpha 2$ and $\alpha 3$. In addition, a salt bridge between glutamate E53 of helix $\alpha 2$ of one monomeric subunit and arginine R188 of helix $\alpha 8$ of the other (and vice versa) reinforces the interaction between the two monomeric subunits.

2.3 | Active site of ZgHAD

While 69% of the primary amino acid sequence differ between DehRhB and ZgHAD, all the amino acids of their active sites are conserved, except for one amino acid that changes from a serine (S120) in ZgHAD to a threonine (T124) in DehRhB (Figure 3a). This corresponds to a minor difference, as both amino acids are exchangeable and conservative with respect to their functional group. Previous site-directed mutagenesis performed on other L-2-HADs has shown that nine conserved amino acids are essential for catalytic activity (Adamu et al., 2016; Kurihara et al., 1995; Nakamura et al., 2009; Pang & Tsang, 2001). Among those nine essential amino acids, three are different in ZgHAD compared to the terrestrial L-2-HADs, as shown by the sequence alignment of structurally characterized L-2-HADs (Figure S2 and Table S1). The positively charged arginine that binds and stabilizes the halide ion in the active site of Dh1B (R39) and L-DEX YL (R42) (Kondo et al., 2014; Ridder et al., 1997) is replaced by a non-polar phenylalanine (F43) in ZgHAD. In L-DEX YL and DehIVa, a serine residue (S175 and S176, respectively) is described to form a hydrogen bond with the catalytic aspartate (D10 and D11, respectively) to maintain a suitable orientation of its carboxyl group for the nucleophilic attack on the substrate (Hisano et al., 1996; Schmidberger et al., 2007). In ZgHAD and DehRhB, this serine is replaced by an alanine (A177) that cannot bind with the catalytic D14. In contrast, a threonine (T18) and a lysine (K153) form hydrogen bonds with the carboxyl group of the catalytic aspartate (D14), respectively. In DehRhB, a histidine (H183) was proposed to participate to the activation of the catalytic water molecule instead of the conserved asparagine in other characterized L-2-HADs (Novak et al., 2013). Similar to DehRhB, ZgHAD possesses a potential catalytic histidine

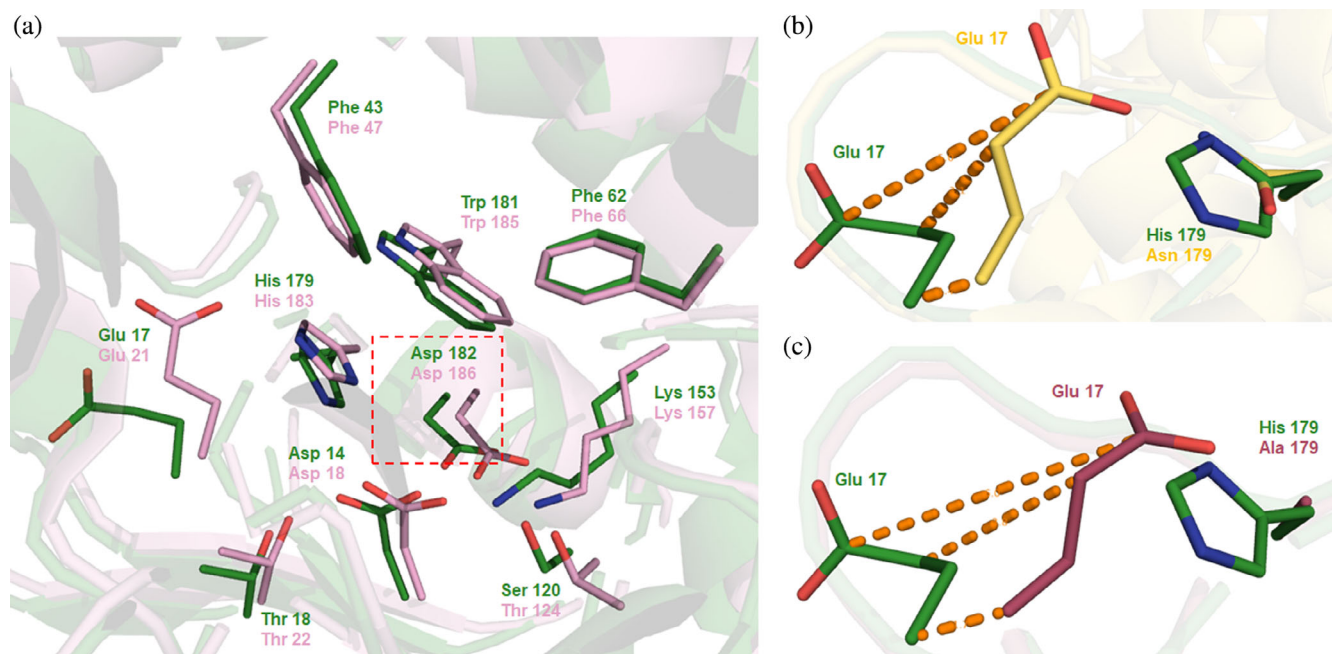


FIGURE 3 Ribbon representation showing the active site residues in the structures of ZgHAD and DehRhb. The side chains of selected residues are shown as sticks (carbons are colored in green, pink, yellow, or purple, oxygen red, and nitrogen blue). (a) Superimposition of the structures of ZgHAD (green) and DehRhb (pink). The red square shows the fixation site of the substrate. (b) Superimposition of ZgHAD (green) and mutant H179N (yellow). (c) Superimposition of ZgHAD (green) and mutant H179A (purple). Orange dotted lines show difference of positions of side chain carbons of the movement of Glu17 between wild-type and mutant proteins

at position 179 as part of the hydrophobic pocket around the active site. This histidine (H179) has been changed to alanine and asparagine by site directed mutagenesis to generate the ZgHAD_H179A and ZgHAD_H179N mutant enzymes leading to the loss of dehalogenase activity. To analyze the potential structural rearrangements due to these mutations, we have also crystallized and solved the crystal structures of both ZgHAD_H179A and ZgHAD_H179N. No major structural differences were observed for the H179N and H179A mutants, except for the position of glutamate E17. In both cases, the main chain of E17 moves by 1–2 Å toward the active site and the side chain displays an alternative conformer that moves the carboxyl group 6 Å closer to the catalytic residues, which may possibly interfere with substrate binding or water activation (Figure 3b,c).

In the case of the mutant ZgHAD_H179N, the crystal structure revealed 13 dimers displaying a helical arrangement within the asymmetric unit, as shown in Figure 4a. Notably, the same space group and large unit cell parameters as for ZgHAD_H179N were also observed for certain wild-type ZgHAD crystals, indicating that this spatial arrangement (dependent on the pH of the crystallization condition) is also possible for the native protein. But due to diffraction at low resolution (i.e., 3.5–3.2 Å), these crystals were not investigated further, since better diffracting crystals were obtained for the wild-type

protein. For ZgHAD_H179N, when all independent dimers of the asymmetric unit are superimposed based on a single monomeric subunit, these subunits match well with a RMSD between 0.288 and 0.445 (Figure 4b), respectively. However, the other monomeric subunits, which are not included in the superimposition calculations, appear to display different relative positions with respect to the first monomeric subunit, with largest main chain distances of up to 2.7 Å (Figure 4c) for the outer structural elements. This indicates that the dimeric arrangement has some flexibility, allowing a rotational freedom at the interface of the two monomeric subunits.

2.4 | Putative substrate binding residues and docking analysis

Molecular docking studies of ZgHAD with iodoacetic acid (IAA), bromoacetic acid (BAA), chloroacetic acid (CAA), or L-2-bromopropionic acid (2-BPA) substrate analogs were undertaken using AutoDock Vina (Trott & Olson, 2010). For each docking result, the top ranked position based on affinity score (kcal/mol) was selected as the most likely solution. The binding energy calculated for these substrates is comprised between –5.6 and –6.1 kcal/mol, indicating that they can all be considered as potential substrates for ZgHAD. In these models,

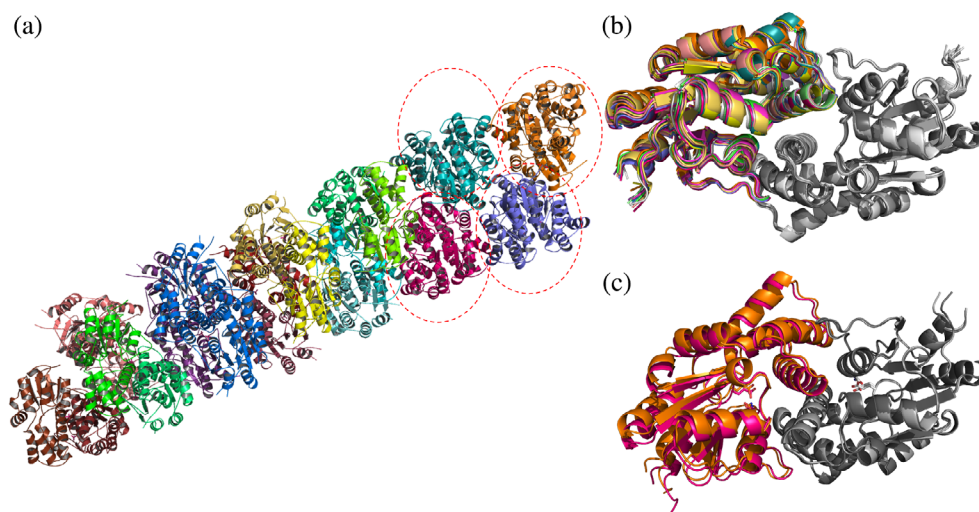


FIGURE 4 Crystal structure of the mutant ZgHAD_H179N shows an unexpected helical arrangement of 13 dimers within the asymmetric unit. (a) Ribbon representation of the 13 ZgHAD_H179N dimers present in the asymmetric unit. The individual monomeric subunits of each dimer are colored with similar colors, highlighting the helical arrangement of the dimers. At one end, each circle of dotted lines surrounds one dimer. (b) Superimposition of all dimers based on the calculation of a single monomeric subunit of each (gray). The colored monomeric subunits highlight the variability of relative positions to the gray monomeric subunits, which in contrast are almost identical. (c) Two of the dimers from (b), displaying the most distant relative orientations, are presented as ribbons and the amino acids E17 and N179, represented as sticks, highlight the position of the respective catalytic active sites

residues V15, N16, S120, N121, and K153 are potentially interacting with each docked substrate molecule through hydrogen bonds (Figure 5).

According to Novak et al., 2013, a “halogen cradle” is formed by the side chains of residues F47, I51, F66, N125, and W185 in DehRhB. Four of the corresponding amino acids in ZgHAD (F43, F62, N121, and W181) are conserved and could be similarly involved in stabilizing the halogen atom, as well as the fifth residue, L47, that replaces I51 in a conservative manner.

2.5 | Active site entrances, tunnels, and cavities

The presence of cavities and potential tunnels/channels in ZgHAD was analyzed using CAVER Analyst 2.0 software (Jurcik et al., 2018). Two putative entry sites were determined on each monomeric subunit, a large and a smaller one, both leading to the catalytic cavity by a short and tight tunnel. The two orifices are on opposite sides of the monomeric subunit but connected (Figure S3A). The largest entrance has an ellipsoid configuration and was estimated to have an average surface of 74.9 \AA^2 and the smallest an average surface of 66.5 \AA^2 (Figure S3B,C). The diameters are 2.64 and 1.56 \AA for the larger and smaller tunnel, respectively. The catalytic cavities present on each monomeric subunit are shown to have a very similar volume of 606.8 and 606.9 \AA^3 . In the inactive

H179A mutant, a reduction of the cavity is observed with $\sim 564.2 \text{ \AA}^3$, corresponding to a decrease of the volume by 7%.

The analysis of the amino acid composition around the cavity and along the tunnel allowed to identify two patches of basic (H37, R44, H99, R198, H201 on monomeric subunit A and K76, F77 on monomeric subunit B) and acidic residues (E17, D21, E26, E36, E199), complemented by hydrophobic, neutral and small residues (M22, G23, N27, F39, S40, G200 on monomeric subunit A and L41 on monomeric subunit B) (Figure S4A). Deeper in the tunnel and near the catalytic site, a patch of hydrophobic amino acids (F43, L47, W181) is present and the putative substrate binding residues (V15, N16, S120, N121, K153) are more buried in the core of the cavity. The electrostatic potential in proximity of the haloacid binding site was estimated to be positive as calculated by APBS in PyMOL version 2.4.1 (Jurrus et al., 2018). In the closest homologous protein DehRhB, the amino acid composition in the same area is quite different (Figure S4B). Among the basic residues at the surface, two major changes are the replacement of the small residues S40 and G200 in ZgHAD by cumbersome R44 and H204 in DehRhB. These modifications lead to a closed conformation in DehRhB, compared to the potential entrance site for the substrates in ZgHAD.

The second, smaller opening in the structure of ZgHAD is circular and globally more basic than acidic, with three lysine (K67, K94, and K125) and two aspartic

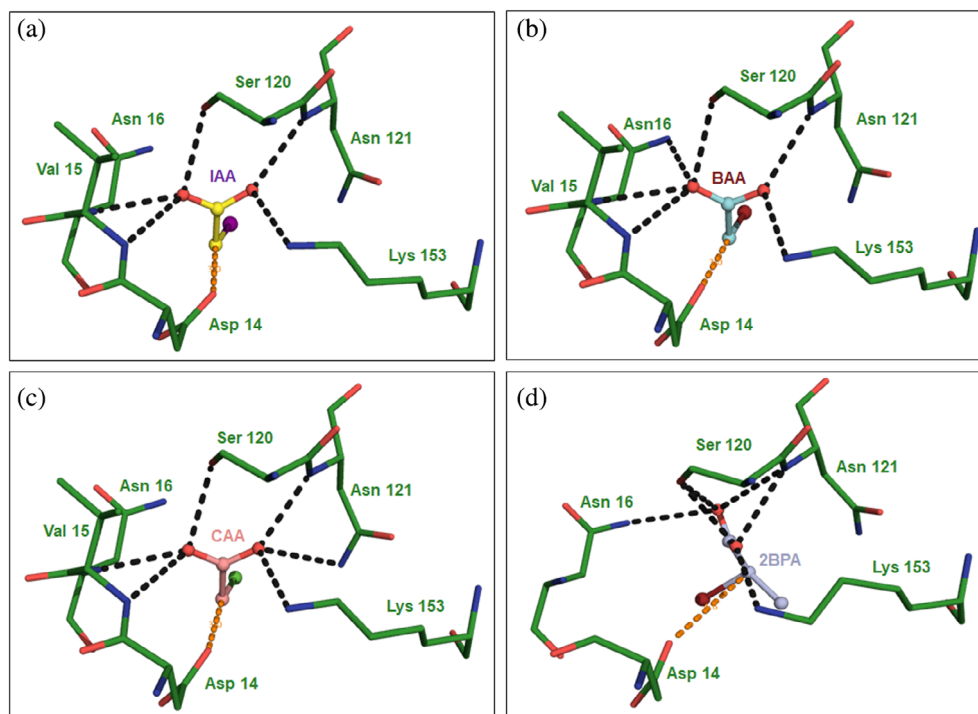


FIGURE 5 Illustration of the molecular docking calculations of various substrate molecules in the active site of ZgHAD. Selected residues are shown as sticks. (a) Docking of IAA (yellow) in ZgHAD (green) structures. (b) Docking of BAA (cyan) in ZgHAD (green) structure. (c) Docking of CAA (pink) in ZgHAD (green) structure. (d) Docking of 2BPA (purple) in ZgHAD (green) structure. Orange dotted lines represent the distance between Asp14 and the carbon 2 of the substrate. Black dotted lines represent hydrogen bonds of the substrates with surrounding amino acids

acid (D61 and D87) residues near the surface (Figure S4C). Neutral and hydrophobic residues (G63, T64, L90, G91, I93, N123, L126) complete this patch of amino acids at the entrance site. Several hydrophobic residues (L25, W42, L46, F62, S122) constitute the tunnel leading to the active site and in particular to the substrate binding residues. The corresponding zone of this smaller opening is cluttered in DehRhB by the presence of K94 and A70, as well as L97 although to a lesser extent.

Nevertheless, two channels are also present, side by side, in a nearby zone in DehRhB (Figure S4D). The equivalent to that zone is covered by K94, N123, L126, L127, and Q130, forming helix α_6 in ZgHAD (Figure S5). By contrast, this helix is replaced by an arch-forming structure and rather divergent residues in the loop $^{126}\text{SAPSPAPSP}^{134}$ in the *Rhodobacteraceae* homolog (Figure S5).

2.6 | Dimeric interface

There are 12 amino acids, hydrophobic in majority, that are involved at the dimeric interface in ZgHAD: L41, H48, Y49, L51, T52, E53, T56, K76, W181, R188, G200, Y204 (Figure S6). Eight of these 12 positions are also at the dimer interface in DehRhB, while only 5 equivalent positions, in the alignment with DehSft, are also involved in dimerization (Figure S2). Among these corresponding residues, 3 out of 8 are identical in DehRhB (Y53, T56, and W185) and only one (Y52) is conserved in DehSft. Furthermore, R188 is implicated in hydrogen bonds with E53 and T56 and W181 forms a salt bridge with Y49.

3 | DISCUSSION

Expectedly, ZgHAD displays the common fold and dimeric arrangement shared with the other L-2-HADs, where the catalytic site is located in a cavity between the core and cap domain. Despite this similarity, major differences regarding lengths of loops and interface interactions are reflected in the low sequence identity (between 19% and 31%) that ZgHAD shares with other characterized L-2-HADs. The highest similarity is shared with DehRhB the closest homolog, which is also reflected in similar catalytic properties. In all enzymes, a conserved aspartate residue (D14 in ZgHAD) is the main catalytic amino acid, performing the first step of the reaction. The activation of the water molecule is done by a His/Glu dyad for both ZgHAD and DehRhB, whereas various other amino acid dyads, such as Asp/Asn, Asp/Lys, or Lys/Tyr, have been proposed to be responsible of this step in other L-2-HAD enzymes (Hisano et al., 1996; Nakamura et al., 2009; Novak et al., 2013; Schmidberger et al., 2007; Wang et al., 2021). A similar His/Glu dyad is well-known to operate in an equivalent manner in haloalkane dehalogenases that are widespread in the marine environment (Janssen, 2004).

Among the nine residues, depicted to be the most important for catalytic activity, five residues are fully conserved, while the four other amino acids of L-DEX YL (two serines, an arginine and an asparagine) are less systematically conserved, but present in most bacterial and archaeal sequences (Table S1 and Figure S2). Other conserved amino acids have been shown to belong to a

hydrophobic pocket surrounding the active site and appear to play an important role in determining the stereo-specificity of the enzyme (Figure S2; Novak et al., 2013). F43, L46, L47, F62, H179, and W181 compose this hydrophobic cluster in ZgHAD. Based on the high-resolution structure of ZgHAD in the unbound state, computational analyses by docking substrate molecules into the active site revealed that only two important residues for substrate binding or catalysis are different between ZgHAD and DehRhb (L47 and S120 of ZgHAD vs. I51 and T124 of DehRhb). Since serine and threonine are neutral polar amino acids, the substitution between them is unlikely to introduce a strong significant change of the catalytic activity. The same conservative replacement can be assumed for the exchange of hydrophobic amino acids isoleucine to leucine in ZgHAD. Nevertheless, a mutational study of L-DEX YL was shown to produce a reduction by 20% of the dehalogenation activity toward 2-chloropropionic acid by a single mutant, S175T (Kurihara et al., 1995), suggesting that this substitution might have a similar effect in ZgHAD, as compared to DehRhb. The recombinant ZgHAD is shown to have preferable activity toward short-carbon-chain substrates and most specifically for C2 compared to C3 haloalkanoic acids. Iodoacetic acid and bromoacetic acid are found to be the best substrates and the catalytic turnover rates are similar for both substrates (Grigorian et al., 2021). When compared to DehRhb, we found that ZgHAD removes bromide more efficiently from bromoacetic acid, with a V_{\max} value of $1.12 \mu\text{M s}^{-1}$, in respect to $1.75 \mu\text{M min}^{-1}$ for the *Rhodobacteraceae* enzyme (Grigorian et al., 2021). The K_m value of ZgHAD for bromoacetic acid was also found to be less than that of DehRhb (0.46 mM against 6.72 mM respectively), suggesting a better affinity for this substrate. Consequently, the replacement of these two amino acids might be associated to an increase of substrate reactivity and/or affinity in ZgHAD, as compared to that in DehRhb. L47 is located in the cap domain, which is assumed to be responsible for substrate recognition and binding. However, this hypothesis requires further experimental examination, such as site-directed mutagenesis. The high structural conservation of active site residues between ZgHAD and DehRhb, suggests that the equivalent His/Glu dyad might be responsible for water activation. Domain movements, such as that of the cap domain relative to the core domain but also of the two monomeric subunits relative to each other, have been described to be important for the activity of DehIVa (Schmidberger et al., 2007). Despite the difference of residues involved in the mechanism between DehIVa and ZgHAD, it is interesting to note that the crystal structure of ZgHAD_H179N highlights the possibility of such movements, which potentially open the access to the

active site pocket, allowing the entrance of substrate molecules and the outward diffusion of the reaction products. The importance of these movements for activity might be the basis for the dimeric assemblage of the enzymes of this class.

In the same line, interactions involved in dimerization of ZgHAD are most similar to those of DehRhb and different from DehSft (Rye et al., 2009), L-DEX YL (Hisano et al., 1996), and DhIB (Ridder et al., 1997), as described by Novak et al. (2013). L-2-HADs were shown to be robust enzymes as they display significant thermostability and resistance to organic solvents. L-DEX YL retained 100% of its activity when incubated at 60°C for 30 min (Liu et al., 1994) and DehRhb from *Rhodobacteraceae* retained 90% activity when incubated at 55°C (Novak et al., 2013). Similarly, we also observe a high thermal stability for ZgHAD as it conserved 100% of its activity after incubation at 55°C during 30 min. On the other hand, while ZgHAD was shown to be quite stable in low concentrations of ethanol, methanol, acetonitrile and DMSO, it was rapidly inactivated at high concentrations. Similar results were presented for DehRhb (Novak et al., 2013) and DehSft (Rye et al., 2009). ZgHAD and DehRhb appear to be more stable than DehSft, when incubated with the same organic solvents. The optimum pH was not determined since the activity assay is pH dependent, but according to previous studies, L-2-HADs enzymes are generally reported to be alkaline (Liu et al., 1994; Van der Ploeg et al., 1991). In this respect, although it could only be due to crystal packing artifact, notably the possible flexibility at the interface of the dimer was observed at alkaline pH, as shown by the dimer positional variability of ZgHAD_H190N crystallized at pH 8.5. These differences in the interactions at the dimeric interface might explain the variations observed in thermal and solvent stabilities of the L-2-HADs, if the dimer formation is necessary to conserve active enzymes.

The most important differences between DehRhb and ZgHAD are seen at the two entrances leading to the active site (Figures S4 and S5), possibly explaining the observed differences in substrate specificity. While they both exhibit the presence of two openings connected with the catalytic cavity, in DehRhb these are situated on the same side of the monomeric subunit, one giving direct access to the hydrophobic pocket called the “halogen cradle,” whereas in ZgHAD the two orifices are disposed on opposite sides of the monomeric subunit. This positions the potential ‘halogen cradle’ differently with respect to the openings in ZgHAD. The charge distribution at the openings is also different in the two enzymes, where DehRhb presents a less charged environment than ZgHAD. In addition, the electrostatic properties of the

catalytic cavities of ZgHAD versus DehRhb, calculated the APBS software appear to be opposite, since a global acidic environment is predicted for DehRhb, while it is basic in ZgHAD (data not shown). This basic environment might be more attractive for small halogenated acids in the enzyme from *Z. galactanivorans*, in agreement with its substrate preference (Grigorian et al., 2021). These differences might also explain the higher catalytic efficiency of ZgHAD toward bromoacetic acid than DehRhb.

Another interesting structural difference between ZgHAD and DehRhb that could explain substrate specificity is the position of the glutamic acid at the beginning of the tunnel plunging into the larger entrance of ZgHAD (E17) whereas it is located inside the core of the catalytic cavity of DehRhb (E13). When H179 was mutated to alanine or asparagine in ZgHAD_H179A and ZgHAD_H179N mutant enzymes, this led to the movement of E17 and an obstruction of the larger entrance that potentially blocked the access to substrate binding residues (Figure S7). As the catalytic H179 was affected in both mutant enzymes it was not possible to observe the effect of the tunnel closing by E17 on dehalogenation activity or substrate affinity but it might be interesting to study in future.

4 | CONCLUSION

In conclusion, our structural study by comparison to all available HAD structures allows pinpointing the subtle differences on a same overall quaternary arrangement that lead to variations of the catalytic activity and/or substrate specificity. In particular, the position and charge distribution at the entrance to the active site cavity appear to vary among homologous enzymes. Our data also confirm the possibility that domain movements, occurring between the two monomeric subunits of the dimer, may play a key role in substrate tunneling to and from the active site. Future work using site-directed mutagenesis and methods to analyze the dynamics will help confirm these findings.

5 | MATERIALS AND METHODS

5.1 | Gene cloning and site-directed mutagenesis

The ZgHAD gene sequence (Zobellia_4183) was cloned from the genomic DNA of *Z. galactanivorans* as described by Barbeyron et al. (2001); using primers Zgal_4183fw and Zgal_4183rv (Table S2) The PCR product was ligated into pFO4 vector using *Bam*HI and *Eco*RI restriction sites and the T4 DNA ligase protocol (New England Biolabs).

The recombinant vector was transformed first into *Escherichia coli* DH5 α for sequence verification and subsequently into *E. coli* BL21(DE3) expression strain.

ZgHAD mutants H179A and H179N were produced using QuickChange Lightning Site-directed Mutagenesis (Agilent Technologies). Primers used are listed in Table S2.

5.2 | Gene overexpression and protein purification

The procedure for heterologous gene expression with subsequent production and protein purification was performed as described in Grigorian et al. (2021).

5.3 | Thermal unfolding experiments

The proteins were diluted to a final concentration of 10 μ M. For each condition, 10 μ l of sample per capillary were prepared. The samples were loaded into UV capillaries and experiments were carried out using the Prometheus NT.48 (NanoTemper Technologies) that can detect changes in the fluorescence of tryptophan (Trp) residues in the proteins. The temperature gradient was set to an increase of 2°C/min in a range from 20 to 95°C. Protein unfolding was measured by detecting the temperature-dependent change in tryptophan fluorescence at 330 and 350 nm emission wavelengths. The increase of the ratio of Trp fluorescence emission between 350 and 330 nm indicates the thermal unfolding transition midpoint of the protein.

5.4 | Thermostability measurements

The thermostability of the protein was determined by incubation in the presence of monochloroacetic acid (MCA) at different temperatures between 10 and 90°C for 30 min. The assay solution was added and then incubated on ice for 1 h. The solvent stability was investigated by incubating the enzyme in the presence of MCA and with ethanol, methanol, acetonitrile, and dimethylsulfoxide (DMSO) at concentrations between 10% and 80% for 1 h at room temperature. The assay solution was added and then incubated on ice for 1 h. The activity was determined by measuring the absorbance at 560 nm as previously described.

5.5 | Crystallization, data collection, structure determination and refinement

The purified L-haloacid dehalogenase (ZgHAD) and its variants (ZgHAD_H179A and ZgHAD_H179N) were

concentrated using a 10 kDa membrane Amicon Ultra-15 centrifugals Filters (Mercks Millipore) at 3600 g and 4°C until a final concentration of 15 mg/ml was reached. Hanging drops were prepared by mixing 2 μ l of ZgHAD (15 mg/ml) and 1 μ l of reservoir, and were equilibrated by vapor diffusion at 20°C. Diffraction-quality crystals appeared after \sim 3 days in a condition containing 0.33 M potassium thiocyanate and 31% (wt/vol) PEG 3350 for ZgHAD and ZgHAD_H1179A; and containing 25% (wt/vol) PEG3350, 0.1 M Tris-HCl pH 8.5 and 0.2 M NaCl for ZgHAD_H179N. Crystals were soaked in their reservoir solutions supplemented with 10% (vol/vol) glycerol before flash-freezing in liquid nitrogen. Diffraction data were collected at 100 K at microfocus beamline Proxima 2-A (Soleil, France). The data were processed using XDS (Kabsch, 2010) and scaled with Aimless from the CCP4 program package (Winn et al., 2011). The structure of ZgHAD was solved by molecular replacement with the CCP4 suite program MolRep (Vagin & Teplyakov, 2010) using the marine *Rhodobacteraceae* L-Haloacid Dehalogenase as the starting model (PDB code: 2YML). Iterative rounds of model building and refinement were carried out using Coot (Emsley et al., 2010) and the Phenix.Refine module of PHENIX (Adams et al., 2010). The validation of the crystal structures was performed with MolProbity (Chen et al., 2010).

5.6 | Computational docking

Computational docking of haloacetic acids (Cl-, Br-, and I-) and of the 2-bromopropionic acid to ZgHAD X-ray structure was performed using AutoDock Vina (Trott & Olson, 2010). The initial coordinates of these molecules were generated from the SMILES string using PHENIX.eLBOW (Liebschner et al., 2019). The ZgHAD protein was kept rigid during docking. A docking grid with dimensions 25 Å \times 25 Å \times 25 Å, encompassing the entire active site, was used. The calculation yielded nine possible models, of which the one with the highest affinity in kcal/mol was selected as the most likely. Then the complexes were energy minimized using the Yasara energy minimization server (Krieger et al., 2009).

AUTHOR CONTRIBUTIONS

Eugénie GRIGORIAN: Conceptualization (equal); formal analysis (lead); methodology (lead); validation (equal); writing – original draft (lead). **Thomas RORET:** Data curation (supporting); writing – original draft (supporting); writing – review and editing (supporting). **Mirjam Czjzek:** Data curation (lead); funding acquisition (supporting); validation (equal); writing – review and editing (supporting). **Catherine LEBLANC:** Funding

acquisition (lead); supervision (equal); validation (supporting); writing – review and editing (supporting). **Ludovic DELAGE:** Conceptualization (equal); data curation (supporting); formal analysis (supporting); methodology (supporting); supervision (lead); validation (equal); writing – review and editing (lead).

ACKNOWLEDGMENTS

We would like to thank the synchrotron SOLEIL for access to beamtime through the BAG (beamtime allocation group) MX-20181002. We are also thankful to the beamline staff at Proxima 2A of the French synchrotron SOLEIL for their help during X-ray data collection and treatment. The authors appreciated the access to the CristalO platform (FR2424, Station Biologique de Roscoff), which is part of the Biogenouest core facility network. This work benefited from the support of the French Government via the National Research Agency investment expenditure program IDEALG (ANR-10-BTBR-04) and via the Centre National de la Recherche Scientifique (CNRS). The PhD project of Eugénie Grigorian was also supported by Region Bretagne (ARED 2017, projet MHALIN).

ORCID

Ludovic Delage  <https://orcid.org/0000-0001-6198-2771>

REFERENCES

- Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*. 2010;66:213–21.
- Adamu A, Wahab RA, Huyop F. L-2-Haloacid dehalogenase (DehL) from *Rhizobium* sp. RC1. *Springerplus*. 2016;5:695.
- Ang TF, Maiangwa J, Salleh AB, Normi YM, Leow TC. Dehalogenases: from improved performance to potential microbial dehalogenation applications. *Molecules*. 2018;23:1–40.
- Arai R, Kukimoto-Niino M, Kuroishi C, Bessho Y, Shirouzu M, Yokoyama S. Crystal structure of the probable haloacid dehalogenase PH0459 from *Pyrococcus horikoshii* OT3. *Protein Sci*. 2006;15:373–7.
- Barbeyron T, L'Haridon S, Corre E, Kloareg B, Potin P. *Zobellia galactanovorans* gen. nov., sp. nov., a marine species of Flavobacteriaceae isolated from a red alga, and classification of [Cytophaga] uliginosa (ZoBell and Upham 1944) Reichenbach 1989 as *Zobellia uliginosa* gen. nov., comb. nov. *Int J Syst Evol Microbiol*. 2001;51:985–97.
- Burroughs AM, Allen KN, Dunaway-Mariano D, Aravind L. Evolutionary genomics of the HAD superfamily: understanding the structural adaptations and catalytic diversity in a superfamily of phosphoesterases and allied enzymes. *J Mol Biol*. 2006;361:1003–34.
- Chan PWY, Chakrabarti N, Ing C, Halgas O, To TKW, Wälti M, et al. Defluorination Capability of l-2-Haloacid Dehalogenases in the HAD-Like Hydrolase Superfamily Correlates with Active Site Compactness. *ChemBiochem*. 2022;23(1):e202100414.

- Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*. 2010;66:12–21.
- Corpet F. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res*. 1988;25(16):10881–90.
- Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. 2010;66:486–501.
- Grigorian E, Thomas F, Groisillier A, Leblanc C, Delage L. Functional characterization of a L-2-haloacid dehalogenase from *Zobellia galactanivorans* Dsj T suggests a role in haloacetic acid catabolism and a wide distribution in marine environments. *Front Microbiol*. 2021;12:725997.
- Hisano T, Hata Y, Fujii T, Liu JQ, Kurihara T, Esaki N, et al. Crystal structure of a 2-haloacid dehalogenase from *Pseudomonas* sp. YL. *J Biol Chem*. 1996;271:20322–30230.
- Janssen DB. Evolving haloalkane dehalogenases. *Curr Opin Chem Biol*. 2004;8:150–9.
- Jurcik A, Bednar D, Byska J, Marques SM, Furmanova K, Daniel L, et al. CAVER Analyst 2.0: analysis and visualization of channels and tunnels in protein structures and molecular dynamics trajectories. *Bioinformatics*. 2018;34:3586–8.
- Jurrus E, Engel D, Star K, Monson K, Brandi J, Felberg LE, et al. Improvements to the APBS biomolecular solvation software suite. *Protein Sci*. 2018;27:112–28.
- Kabsch W. XDS. *Acta Crystallogr D Biol Crystallogr*. 2010;66:125–32.
- Kondo H, Nakamura T, Tanaka S. A significant role of Arg41 residue in the enzymatic reaction of haloacid dehalogenase L-DEX YL studied by QM/MM method. *J Mol Catal B: Enzym*. 2014;110:23–31.
- Krieger E, Joo K, Lee J, Lee J, Raman S, Thompson J, et al. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. *Proteins*. 2009;9:114–22.
- Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol*. 2007;372:774–97.
- Kurihara T, Liu JQ, Nardi-dei V, Koshikawa H, Esaki N, Soda K. Comprehensive site-directed mutagenesis of L-2-halo acid dehalogenase to probe catalytic amino acid residues. *J Biochem*. 1995;117:1317–22.
- Lahiri SD, Zhang G, Dai J, Dunaway-Mariano D, Allen KN. Analysis of the substrate specificity loop of the HAD superfamily cap domain. *Biochemistry*. 2004;43:2812–20.
- Liebschner D, Afonine PV, Baker ML, Bunkóczi G, Chen VB, Croll TI, et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol*. 2019;75:861–77.
- Liu JQ, Kurihara T, Hasan AK, Nardi-Dei V, Koshikawa H, Esaki N, et al. Purification and characterization of thermostable and nonthermostable 2-haloacid dehalogenases with different stereospecificities from *Pseudomonas* sp. strain YL. *Appl Environ Microbiol*. 1994;60:2389–93.
- Liu JQ, Kurihara T, Miyagi M, Esaki N, Soda K. Reaction mechanism of L-2-haloacid dehalogenase of *Pseudomonas* sp. YL: identification of Asp10 as the active site nucleophile by 18O incorporation experiments. *J Biol Chem*. 1995;270:18309–12.
- Nakamura T, Yamaguchi A, Kondo H, Watanabe H, Kurihara T, Esaki N, et al. Roles of K151 and D180 in L-2-haloacid dehalogenase from *Pseudomonas* sp. YL: analysis by molecular dynamics and ab initio fragment molecular orbital calculations. *J Comput Chem*. 2009;30:2625–34.
- Nardi-Dei V, Kurihara T, Park C, Esaki N, Soda K. Bacterial DL-2-haloacid dehalogenase from *Pseudomonas* sp. strain 113: gene cloning and structural comparison with D- and L-2-haloacid dehalogenases. *J Bacteriol*. 1997;179:4232–8.
- Novak HR, Sayer C, Isupov MN, Paszkiewicz K, Gotz D, Spragg AM, et al. Marine Rhodobacteraceae L-haloacid dehalogenase contains a novel His/Glu dyad that could activate the catalytic water. *FEBS J*. 2013;280:1664–80.
- Pang BCM, Tsang JSH. Mutagenic analysis of the conserved residues in dehalogenase IVa of *Burkholderia cepacia* MBA4. *FEMS Microbiol Lett*. 2001;204:135–40.
- Ridder IS, Rozeboom HJ, Kalk KH, Janssen DB, Dijkstra BW. Three-dimensional structure of L-2-haloacid dehalogenase from *Xanthobacter autotrophicus* GJ10 complexed with the substrate-analogue formate. *J Biol Chem*. 1997;272:33015–22.
- Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*. 2014;42:W320–4.
- Rye CA, Isupov MN, Lebedev AA, Littlechild JA. Biochemical and structural studies of a l-haloacid dehalogenase from the thermophilic archaeon *Sulfolobus tokodaii*. *Extremophiles*. 2009;13:179–90.
- Schmidberger JW, Wilce JA, Tsang JSH, Wilce MCJ. Crystal structures of the substrate free-enzyme, and reaction intermediate of the HAD superfamily member, haloacid dehalogenase DehIVa from *Burkholderia cepacia* MBA4. *J Mol Biol*. 2007;368:706–17.
- Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem*. 2010;31:455–61.
- Vagin A, Teplyakov A. Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr*. 2010;66:22–5.
- Van der Ploeg J, Van Hall G, Janssen DB. Characterization of the haloacid dehalogenase from *Xanthobacter autotrophicus* GJ10 and sequencing of the dhlB gene. *J Bacteriol*. 1991;173:7925–33.
- Wang Y, Xiang Q, Zhou Q, Xu J, Pei D. Mini review: advances in 2-haloacid dehalogenases. *Front Microbiol*. 2021;12:758886.
- Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr*. 2011;67:235–42.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Grigorian E, Roret T, Czjzek M, Leblanc C, Delage L. X-ray structure and mechanism of ZgHAD, a L-2-haloacid dehalogenase from the marine Flavobacterium *Zobellia galactanivorans*. *Protein Science*. 2023; 32(1):e4540. <https://doi.org/10.1002/pro.4540>