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

Eugénie Grigorian, Thomas Roret, Mirjam Czjzek, Catherine Leblanc, Ludovic Delage

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

Eugénie Grigorian[a], Thomas Roret[b], Mirjam Czjzek[a]*, Catherine Leblanc[a] and Ludovic Delage[a]*

[a]Sorbonne Université, CNRS, Integrative Biology of Marine Models (LBI2M), Station Biologique de Roscoff (SBR), 29680 Roscoff, Bretagne, France

[b]Sorbonne Université, CNRS, FR2424, Station Biologique de Roscoff (SBR), 29680 Roscoff, Bretagne, France

*For correspondence: czjzek@sb-roscoff.fr and ludovic.delage@sb-roscoff.fr

running title : Structure and mechanism of ZgHAD

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

50-75 words statement outlining the importance and impact of the work presented

The present work on the first characterized *Bacteroidetes* haloacid dehalogenase provides additional knowledge on marine L-2-HAD structure/function characteristics of the

monophyletic group B. The 3D crystal structure reveals specificities related to unique catalytic features, such as a catalytic His/Glu dyad, as compared to known L-2-HADs from soil organisms of the monophyletic group A. An unprecedented helical arrangement of 13 dimers observed within one crystal highlights the presence of possible movements between monomeric units.

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-HADi act with a configuration-inverting mechanism whereas DL-2-HADr 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 (Wang et al., 2021; Ang *et al.,* 2018).

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), DhlB 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.*, 2021). 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; Schmidberger *et al.,* 2007; Rye *et al.,* 2009).

The complete genome sequence of the marine flavobacteria *Zobellia galactanivorans* Dsij^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 towards 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.

RESULTS

Thermostability and solvent stability

A temperature gradient between 20°C and 95°C was applied to evaluate the stability zone of the protein fold. The experiment showed that ZgHAD was denatured between 60 and 70°C with a melting temperature of 65°C (**Fig. 1A**), as also confirmed by Dynamic Light Scattering (DLS) measurements (**Fig. 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°C and 90°C and after 30 minutes of enzyme incubation. The residual activity was measured and plotted as a percentage of the initial activity (**Fig. 1B**). ZgHAD activity was found to be stable up to 50°C but decreased rapidly between 60°C 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 one hour before evaluating the residual activity (**Fig. 1C**). The effects of ethanol and methanol treatments were very similar. The activity of the enzyme was conserved up to 80%, when 10 to 20% of these organic solvents were added, and decreased to 20% in presence of 20% to 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 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 to 80% of acetonitrile.

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 $P2_12_12_1$ and the mutant ZgHAD_H179N with space group $P2_1$; 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 (**Fig. 2A**). The core domain has a typical "Rossmann fold" which consists of six parallel β-strands surrounded by five α-helices and 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 β3. The cap domain is composed of four α-helices, and a 3_{10} helix is inserted in the core domain between strand β1 and helix α5. The active site is located between the core and the cap domains, right after strand β1. The four helices of the cap domain shield the top of the active site cavity from the solvent (**Fig. 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 rmsd values between matching Cα 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 x 35 x 44 Å (**Fig. 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 two-fold symmetry axis running nearly parallel to the α2 helix. On dimer formation, 16% of each monomeric subunit's accessible surface is buried at the interface as identified by PDBePISA server (Krissinel and Henrick, 2007). In comparison, this value ranges between 13.4% in DehSft to 19% in DhlB. The subunit interface of ZgHAD is mainly

formed by helices α2 and α3. In addition, a salt bridge between glutamate E53 of helix α2 of one monomeric subunit and arginine R188 of helix α8 of the other (and vice versa) reinforces the interaction between the two monomeric subunits.

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 (**Fig. 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 (Kurihara *et al.,* 1995; Pang and Tsang, 2001; Nakamura *et al.,* 2009; Adamu *et al.,* 2016). 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 (**Fig. S2** and **Table S1**). The positively charged arginine that binds and stabilizes the halide ion in the active site of DhlB (R39) and L-DEX YL (R42) (Ridder *et al.,* 1997; Kondo *et al.,* 2014) 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

198 a potential catalytic histidine at position 179 as part of the hydrophobic pocket around the active
199 site. This histidine (H179) has been changed to alanine and asparagine by site directed
200 mutagenesis to generate the ZgHAD_H179A and ZgHAD_H179N mutant enzymes leading to
201 the loss of dehalogenase activity. To analyze the potential structural rearrangements due to these
202 mutations, we have also crystallized and solved the crystal structures of both ZgHAD_H179A
203 and ZgHAD_H179N. No major structural differences were observed for the H179N and H179A
204 mutants, except for the position of glutamate E17. In both cases, the main chain of E17 moves
205 by 1 to 2 Å towards the active site and the side chain displays an alternative conformer that
206 moves the carboxyl group 6 Å closer to the catalytic residues, which may possibly interfere
207 with substrate binding or water activation (**Fig. 3B** and **3C)**.
208 In the case of the mutant ZgHAD_H179N, the crystal structure revealed 13 dimers displaying
209 a helical arrangement within the asymmetric unit, as shown in **Figure 4A**. Notably, the same
210 space group and large unit cell parameters as for ZgHAD_H179N were also observed for certain
211 wild-type ZgHAD crystals, indicating that this spatial arrangement (dependent on the pH of the
212 crystallization condition) is also possible for the native protein. But due to diffraction at low
213 resolution (i.e. 3.5 - 3.2 Å), these crystals were not investigated further, since better diffracting
214 crystals were obtained for the wild-type protein. For ZgHAD_H179N, when all independent
215 dimers of the asymmetric unit are superimposed based on a single monomeric subunit, these
216 subunits match well with a rmsd between 0.288 and 0.445 (**Fig. 4B**), respectively. However,
217 the other monomeric subunits, which are not included in the superimposition calculations,
218 appear to display different relative positions with respect to the first monomeric subunit, with
219 largest main chain distances of up to 2.7 Å (**Fig. 4C**) for the outer structural elements. This
220 indicates that the dimeric arrangement has some flexibility, allowing a rotational freedom at the
221 interface of the two monomeric subunits.
222

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

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 (**Fig. S3A**). The largest entrance has an ellipsoid configuration and was estimated to have an average surface of 74.9 $Å^2$ and the smallest an average surface of 66.5 $Å^2$ (**Fig. S3B** and **Fig. S3C**). The diameters are 2.64 Å and 1.56 Å 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 $Å^3$ and 606.9 $Å^3$. In the inactive H179A mutant, a reduction of the cavity is observed with approximately 564.2 $Å^3$, 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) (**Fig. 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 (**Fig. 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 3 lysine (K67, K94 and K125) and 2 aspartic acid (D61 and D87) residues near the surface (**Fig. 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 (**Fig. S4D**). The equivalent to that zone is covered by K94, N123, L126, L127 and Q130, forming helix α6 in ZgHAD (**Fig. S5**). By contrast, this helix is replaced by an arch-forming structure

and rather divergent residues in the loop ^{126}SAPSPAPSP^{134} in the *Rhodobacteraceae* homolog (**Fig. S5**).

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 (**Fig. 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 (**Fig. 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.

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 (Wang *et al.*, 2021; Novak *et al.,* 2013; Schmidberger *et al.,* 2007; Hisano *et al.*, 1996; Nakamura *et al.,* 2009). A

295 similar His/Glu dyad is well-known to operate in an equivalent manner in haloalkane
296 dehalogenases that are widespread in the marine environment (Janssen, 2004).
297 Among the nine residues, depicted to be the most important for catalytic activity, five residues
298 are fully conserved, while the four other amino acids of L-DEX YL (two serines, an arginine
299 and an asparagine) are less systematically conserved, but present in most bacterial and archaeal
300 sequences (**Table S1** and **Fig. S2**). Other conserved amino acids have been shown to belong to
301 a hydrophobic pocket surrounding the active site and appear to play an important role in
302 determining the stereo-specificity of the enzyme (**Fig. S2**; Novak *et al.*, 2013). F43, L46, L47,
303 F62, H179 and W181 compose this hydrophobic cluster in ZgHAD. Based on the high-
304 resolution structure of ZgHAD in the unbound state, computational analyses by docking
305 substrate molecules into the active site revealed that only two important residues for substrate
306 binding or catalysis are different between ZgHAD and DehRhb (L47 and S120 of ZgHAD vs.
307 I51 and T124 of DehRhb). Since serine and threonine are neutral polar amino acids, the
308 substitution between them is unlikely to introduce a strong significant change of the catalytic
309 activity. The same conservative replacement can be assumed for the exchange of hydrophobic
310 amino acids isoleucine to leucine in ZgHAD. Nevertheless, a mutational study of L-DEX YL
311 was shown to produce a reduction by 20% of the dehalogenation activity towards 2-
312 chloropropionic acid by a single mutant, S175T (Kurihara *et al.*, 1995), suggesting that this
313 substitution might have a similar effect in ZgHAD, as compared to DehRhb. The recombinant
314 ZgHAD is shown to have preferable activity towards short-carbon-chain substrates and most
315 specifically for C2 compared to C3 haloalkanoic acids. Iodoacetic acid and bromoacetic acid
316 are found to be the best substrates and the catalytic turnover rates are similar for both substrates
317 (Grigorian *et al.*, 2021). When compared to DehRhb, we found that ZgHAD removes bromide
318 more efficiently from bromoacetic acid, with a Vmax value of 1.12 $\mu M.sec^{-1}$, in respect to 1.75
319 $\mu M.min^{-1}$ for the *Rhodobacteraceae* enzyme (Grigorian *et al.*, 2021). The Km 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 DhlB (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 minutes (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 minutes. On the other hand, while ZgHAD was shown to be quite stable in low concentrations of ethanol, methanol,

345 acetonitrile and DMSO, it was rapidly inactivated at high concentrations. Similar results were
346 presented for DehRhb (Novak *et al.,* 2013) and DehSft (Rye *et al.,* 2009). ZgHAD and DehRhb
347 appear to be more stable than DehSft, when incubated with the same organic solvents. The
348 optimum pH was not determined since the activity assay is pH dependent, but according to
349 previous studies, L-2-HADs enzymes are generally reported to be alkaline (Van der Ploeg *et*
350 *al,* 1991; Liu *et al.,* 1994). In this respect, although it could only be due to crystal packing
351 artifact, notably the possible flexibility at the interface of the dimer was observed at alkaline
352 pH, as shown by the dimer positional variability of ZgHAD_H190N crystallized at pH 8.5.
353 These differences in the interactions at the dimeric interface might explain the variations
354 observed in thermal and solvent stabilities of the L-2-HADs, if the dimer formation is necessary
355 to conserve active enzymes.
356 The most important differences between DehRhb and ZgHAD are seen at the two entrances
357 leading to the active site (**Fig. S4 and S5**), possibly explaining the observed differences in
358 substrate specificity. While they both exhibit the presence of two openings connected with the
359 catalytic cavity, in DehRhb these are situated on the same side of the monomeric subunit, one
360 giving direct access to the hydrophobic pocket called the "halogen cradle", whereas in ZgHAD
361 the two orifices are disposed on opposite sides of the monomeric subunit. This positions the
362 potential 'halogen cradle" differently with respect to the openings in ZgHAD. The charge
363 distribution at the openings is also different in the two enzymes, where DehRhb presents a less
364 charged environment than ZgHAD. In addition, the electrostatic properties of the catalytic
365 cavities of ZgHAD vs DehRhb, calculated the APBS software appear to be opposite, since a
366 global acidic environment is predicted for DehRhb, while it is basic in ZgHAD (data not
367 shown). This basic environment might be more attractive for small halogenated acids in the
368 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 towards 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 (**Fig. 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.

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.

MATERIALS AND METHODS

Gene cloning and site-directed mutagenesis

The ZgHAD gene sequence (Zobellia_4183) was cloned from the genomic DNA of *Zobellia 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* DH5α 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** from supplementary data.

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.

Thermal unfolding experiments

The proteins were diluted to a final concentration of 10 µM. For each condition, 10 µL 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 °C 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.

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

Crystallization, data collection, structure determination and refinement

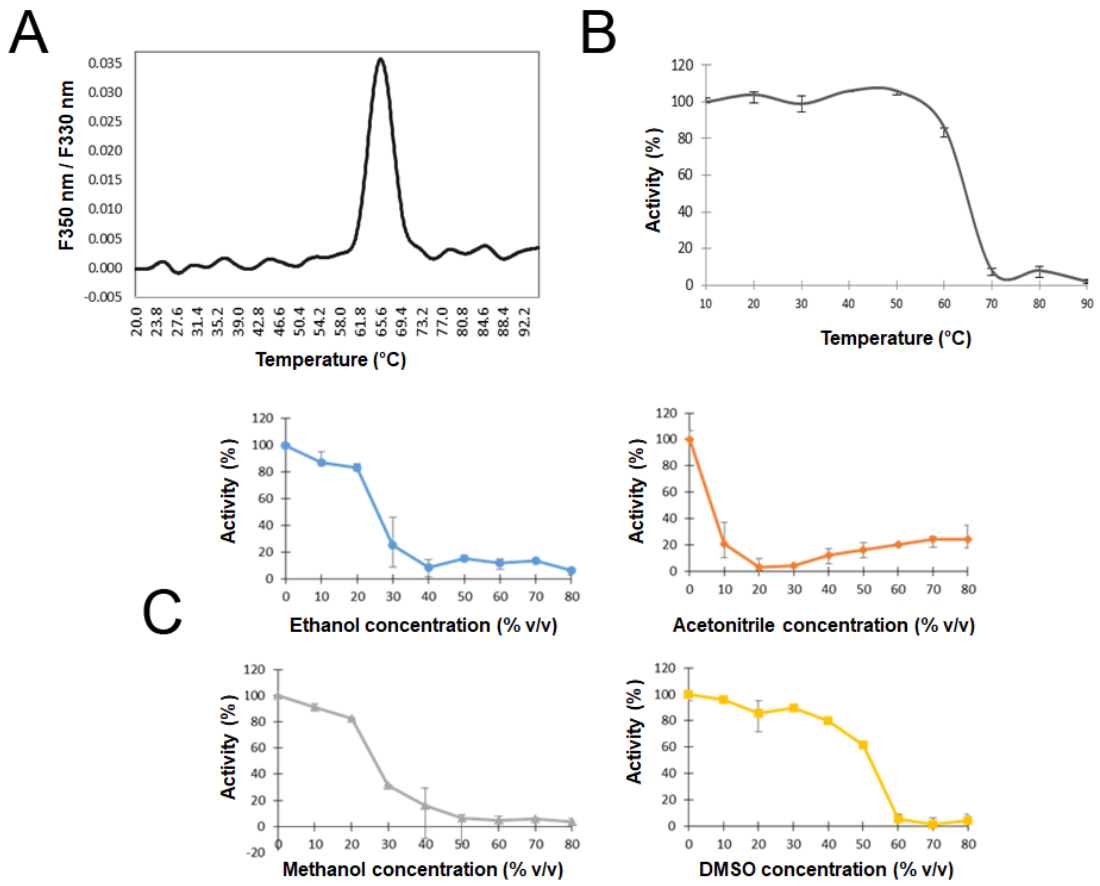
The purified L-haloacid dehalogenase (ZgHAD) and its variants (ZgHAD_H179A and ZgHAD_H179N) were concentrated using a 10 kDa membrane Amicon Ultra-15 centrifugals Filters (Mercks Millipore) at 3600 g and 4°C until a final concentration of 15 mg/mL was reached. Hanging drops were prepared by mixing 2µL of ZgHAD (15 mg/mL) and 1 µL of reservoir, and were equilibrated by vapor diffusion at 20°C. Diffraction-quality crystals appeared after approximately three days in a condition containing 0.33 M potassium thiocyanate and 31% (w/v) PEG 3350 for ZgHAD and ZgHAD_H1179A; and containing 25% (w/v) 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% (v/v) 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).

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 and 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 × 25 Å × 25 Å, encompassing the entire active site, was used. The calculation yielded 9 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).

FIGURES AND LEGENDS



458

459 Figure 1: Thermal and solvent stability of the recombinant L-2-HAD from Z. galactanivorans.
460 (A) Thermal break point of ZgHAD protein as determined by the Prometheus NT.48. (B)
461 Thermal stability of ZgHAD activity as determined after the pre-incubation of the enzyme at
462 varying temperatures for 30 minutes before measuring the residual activity at 20°C for all the
463 points. (C) Solvent stability of ZgHAD activity as determined after pre-exposure to different
464 concentrations of ethanol, methanol, acetonitrile and DMSO for 1 hour before measuring the
465 residual activity in a standard solvent concentration.

466

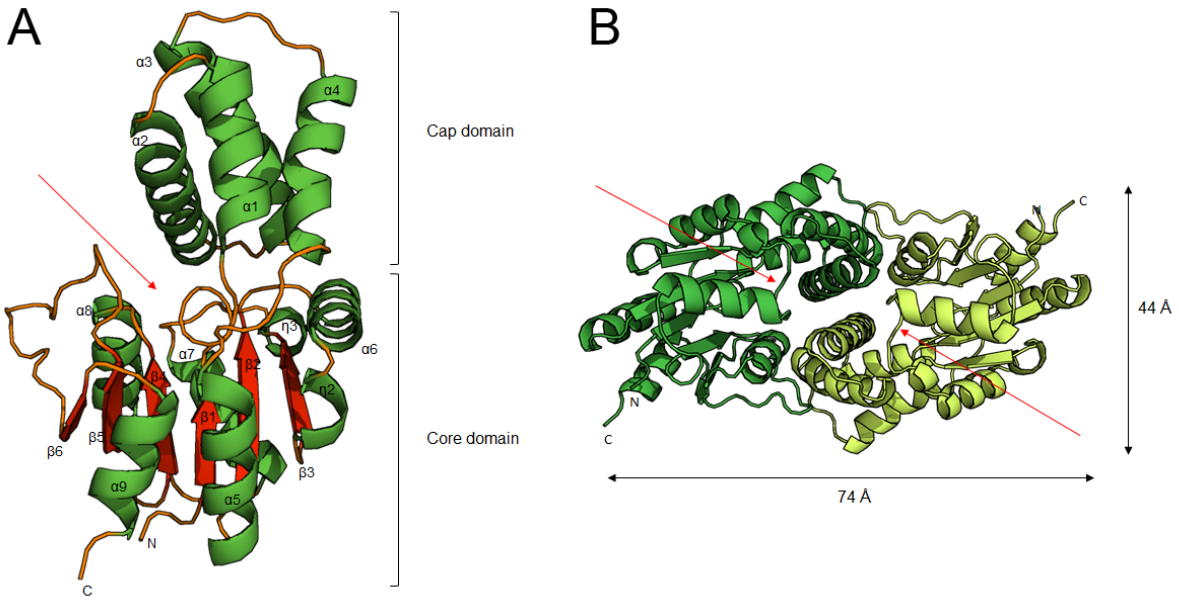


Figure 2: (A) Overall fold of the ZgHAD monomeric subunit presented as a ribbon diagram colored by secondary structure elements. The α helices, β strands and loops are colored in green, red and orange respectively. (B) Ribbon diagram of the ZgHAD dimer viewed along the two-fold horizontal axis. Each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

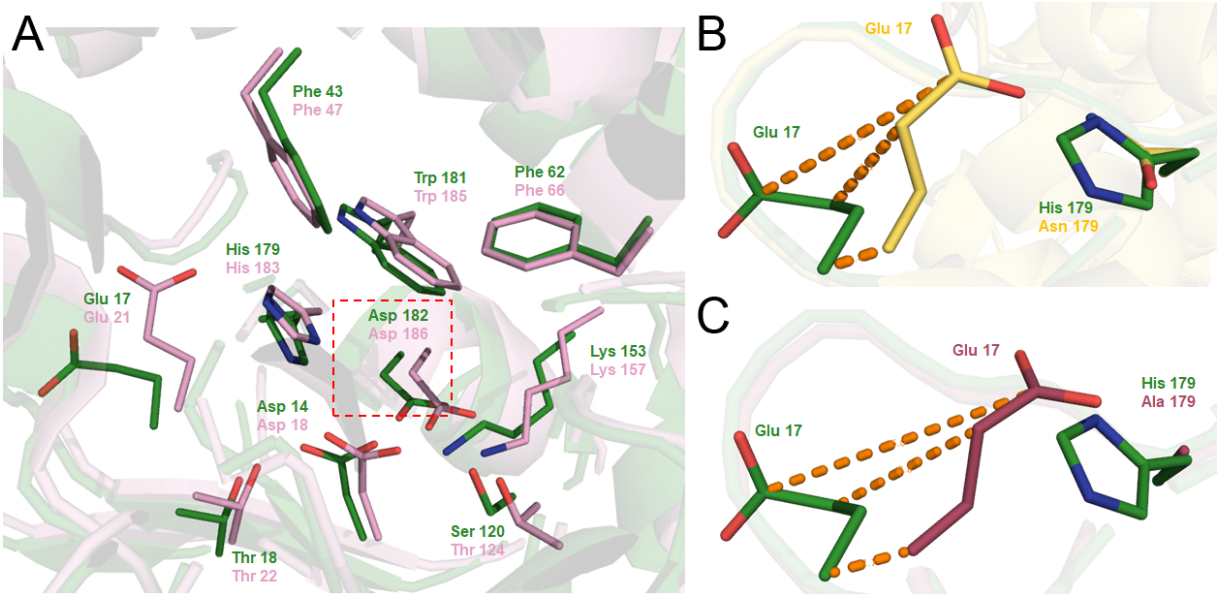


Figure 3: Ribbon representation showing the active site residues in the structures of ZgHAD and DehRhb. The side chains of selected residues are shown as sticks (carbons are colored in

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.

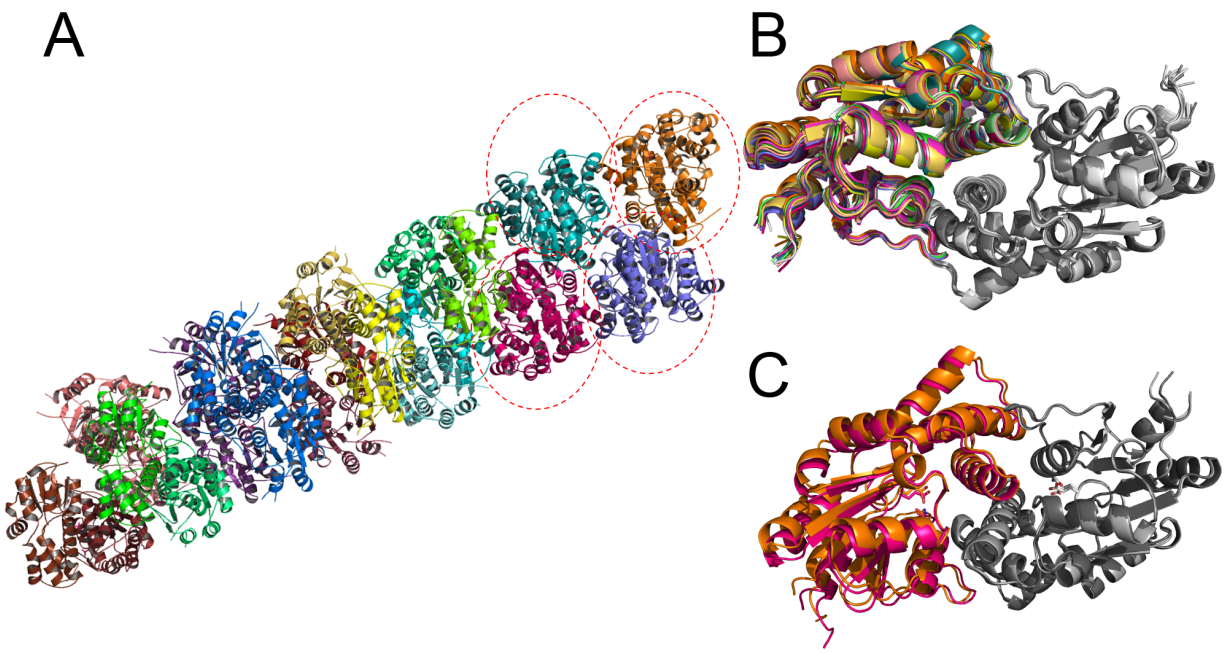


Figure 4: Crystal structure of the mutant ZgHAD_H179N shows an unexpected helical arrangement of 13 dimers within the asymmetric unit. (A) Ribbon representation of the thirteen ZgHAD_H179N dimers present in the asymmetric unit. The individual monomeric subunits of each dimer are colored with similar colors, highlighting the helical arrangement of the dimers. At one end, each circle of doted lines surrounds one dimer. (B) Superimposition of all dimers based on the calculation of a single monomeric subunit of each (grey). 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.

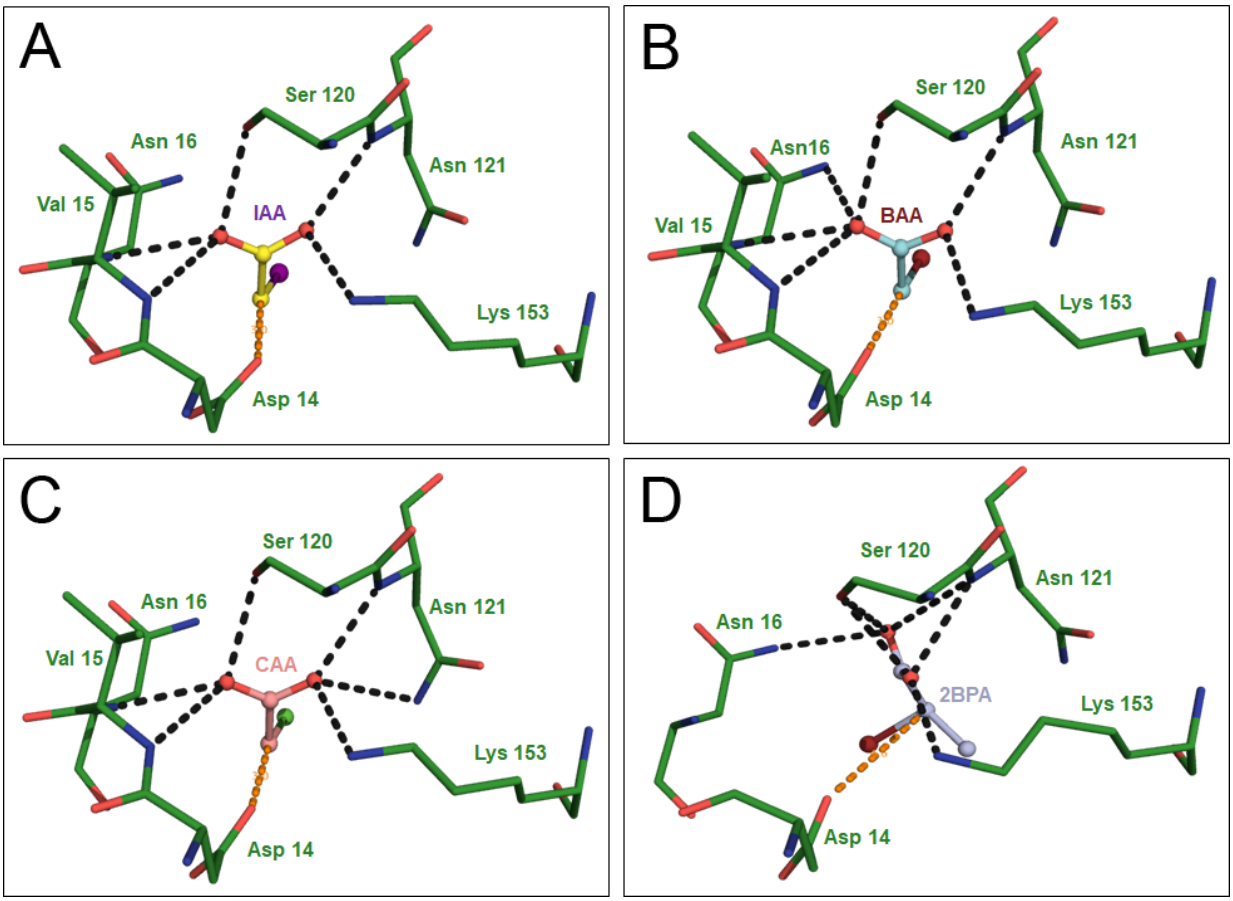


Figure 5: Illustration of the molecular docking calculations of various substrate molecules in the active site of ZgHAD. Selected residues are shown as sticks. (A) Docking of IAA (yellow) in ZgHAD (green) structures. (B) Docking of BAA (cyan) in ZgHAD (green) structure. (C) Docking of CAA (pink) in ZgHAD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

Table 1: Data collection and refinement statistics for ZgHAD. Values in parentheses refer to the outer resolution shell.

	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 68.83 103.74	59.716 71.655 116.009	76.17 132.79 275.70
Wavelength (Å)	0.97986	0.980116	0.980114
Resolution (Å)	45.03 to 1.60	45.874 to 1.716	49.46 to 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	41657 (2364)	41318 (2628)	144,483 (4,343)
Mean I/σI	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
Refinement			
Resolution (Å)	1.78	1.88	2.75
R_{free}/R_{work}	0.1774/0.1552	0.1991/0.1680	0.2543/0.2065
Total number of atoms	7481	7415	44949
Water	473	413	56
Average B factor (Å^2)	24.7	29.6	73.31
Ligands	PO4 ; SCN	PO4 ; SCN	PO4
R.m.s deviations			
Bonds	0.010	0.017	0.009
Angles	1.0	1.4	1.5
MolProbity analysis			
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

SUPPLEMENTARY MATERIAL DESCRIPTION

Figure S1: Thermal stability of the ZgHAD enzyme by Dynamic Light Scattering (DLS) measurement. The peak of denaturation corresponding to a Z-average diameter increase appears at 65°C.

Figure S2: Amino acid sequence alignment of ZgHAD and its homologues DehRhb, DehIVa, L-DEX, DehlB and DehSft, using the programs MultAlin (Corpet, 1988) and ESPript (Robert and Gouet, 2014).

Figure S3: Opening zones, tunnels and determination of ZgHAD dimer obtained by CAVER Analyst program.

Figure S4: Visualization of the possible entrance sites leading to the active site cavity of ZgHAD as compared to DehRhb. The surface view is modelled on the left and the corresponding residues are represented as sticks on the right view.

Figure S5: Visualization of the possible entrance sites leading to the active site cavity in DehRhb. (A) representation of the arch-like structure formed by a loop rich in proline residues that liberates two orifices connected with the active site.

Figure S6: Representations of the dimeric interface of ZgHAD and DehRhb, analyzed using PDBe PISA v1.52. (A) ZgHAD dimer and (B) dimer of DehRhb.

Figure S7: Surface view of the position of Glu17 at the putative substrate entrance site in (A) ZgHAD wild type enzyme, (B) ZgHAD mutant H190A and (C) ZgHAD mutant H190N.

Table S1: Residues found to be essential for catalysis in L-DEX YL (Kurihara *et al.*, 1995), and the corresponding amino acids in DhlB, DehIVa, DehSft, DehRhb and ZgHAD.

Table S2: sequences of primers used for PCR-based cloning for ZgHAD and and site-directed mutants H179A and H179N.

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 EG was also supported by Region Bretagne (ARED 2017, projet MHALIN).

AUTHOR CONTRIBUTIONS

EG and LD conceived the experiments. EG contributed to the majority of the experiments and analyses. MC contributed to crystal preparation, data collection and refinement and TR contributed to structure resolution, refinement and docking. EG wrote the manuscript with the support of LD, CL and MC. All authors corrected and approved the final manuscript.

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;66:213-221.

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;15:373-377.

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;51:985–997.

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;361:1003–1034.

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 Nov 25;16:10881-10890.

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 Dsij 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, Soda K. 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;8:150–159.

Jurcik A, Bednar D, Byska J, Marques SM, Furmanova K, Daniel L, Kokkonen P, Brezovsky J, Strnad O, Stourac J, Pavelka A, Manak M, Damborsky J, Kozlikova B. CAVER Analyst 2.0: analysis and visualization of channels and tunnels in protein structures and molecular dynamics trajectories. Bioinformatics. 2018;34:3586-3588.

Jurrus E, Engel D, Star K, Monson K, Brandi J, Felberg LE, Brookes DH, Wilson L, Chen J, Liles K, Chun M, Li P, Gohara DW, Dolinsky T, Konecny R, Koes DR, Nielsen JE, Head-Gordon T, Geng W, Krasny R, Wei GW, Holst MJ, McCammon JA, Baker NA. Improvements to the APBS biomolecular solvation software suite. Protein Sci. 2018;27:112-128.

Kabsch W. XDS. Acta Crystallogr D Biol Crystallogr. 2010;66:125-132.

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, Tyka M, Baker D, Karplus K. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. Proteins. 2009;9:114-122.

Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J Mol Biol. 2007;372:774-797.

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;117:1317–1322.

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;43:2812–2820.

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;75:861-877.

Liu JQ, Kurihara T, Hasan AK, Nardi-Dei V, Koshikawa H, Esaki, N, Soda K. Purification and characterization of thermostable and nonthermostable 2-haloacid dehalogenases with different stereospecificities from Pseudomonas sp. strain YL. Appl Environ Microbiol. 1994;60:2389–2393.

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;270:18309–18312.

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: 2625–2634.

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;179:4232–4238.

Novak HR, Sayer C, Isupov MN, Paszkiewicz K, Gotz D, Spragg AM, Littlechild JA. Marine Rhodobacteraceae L -haloacid dehalogenase contains a novel His / Glu dyad that could activate the catalytic water. FEBS J. 2013;280:1664–1680.

Pang BCM and Tsang JSH. Mutagenic analysis of the conserved residues in dehalogenase IVa of Burkholderia cepacia MBA4. FEMS Microbiol Lett. 2001;204:135–140.

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;272:33015–33022.

Robert X and Gouet P. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. 2014;42:W320-324.

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;13:179–190.

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;368:706–717.

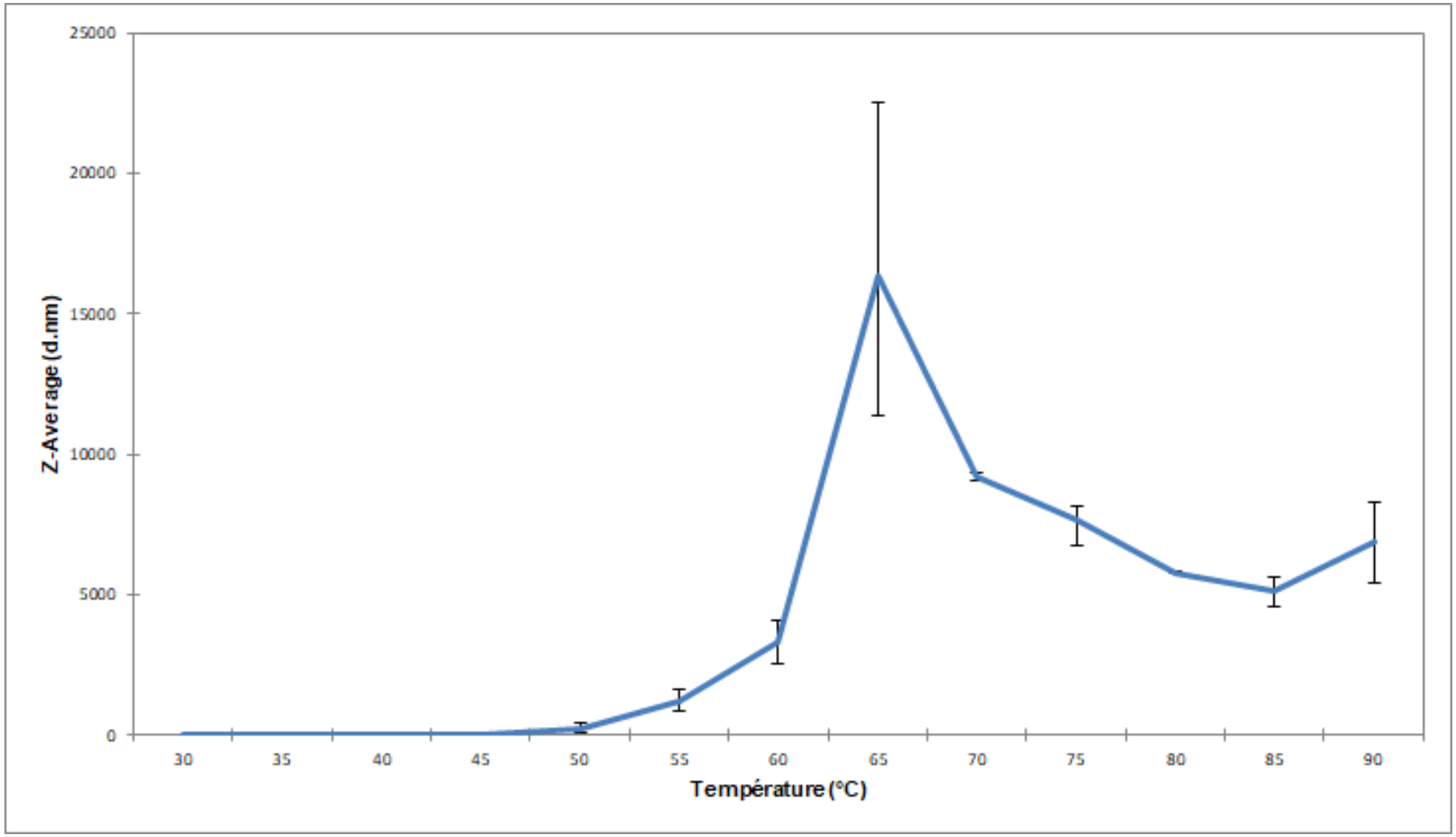
Trott O. and Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31:455-461.

Vagin A. and Teplyakov A. Molecular replacement with MOLREP. Acta Crystallogr D Biol Crystallogr. 2010;66:22-25.

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;173:7925–7933.

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, Keegan RM, *et al.* Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011;67:235-242.



<u>Figure S1:</u> Thermal stability of the ZgHAD enzyme by Dynamic Light Scattering (DLS) measurement.
The peak of denaturation corresponding to a Z-average diameter increase appears at 65°C.

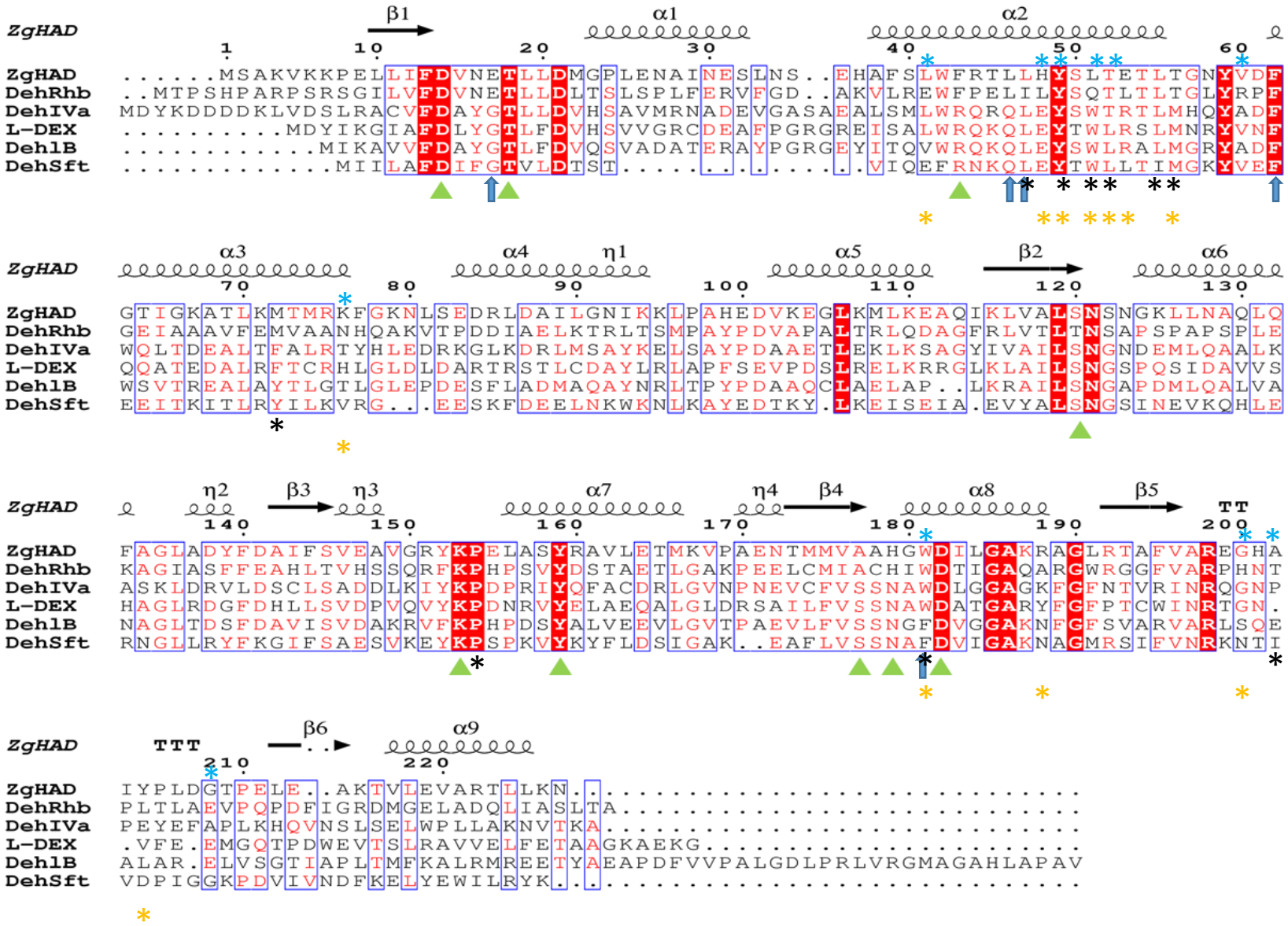
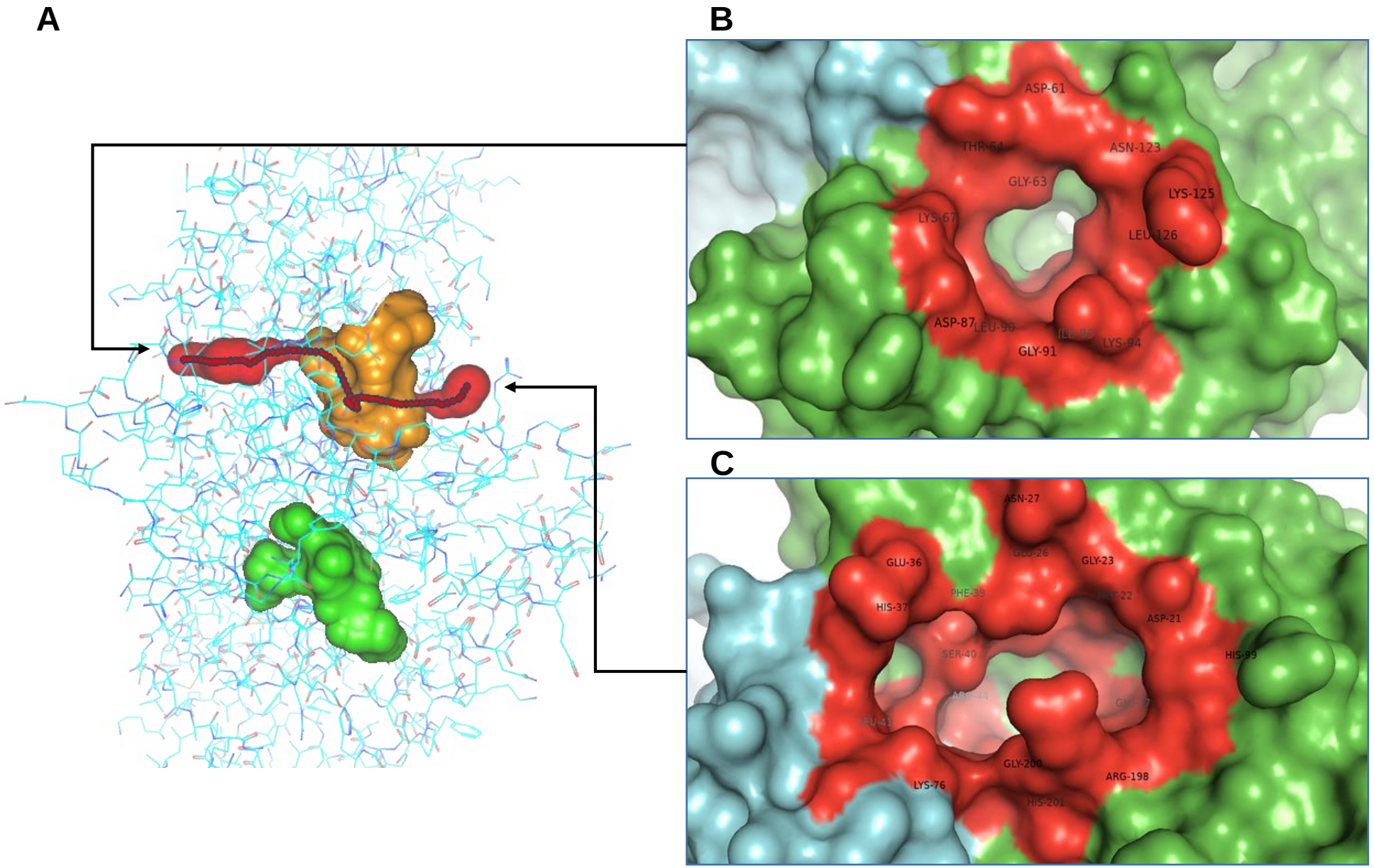
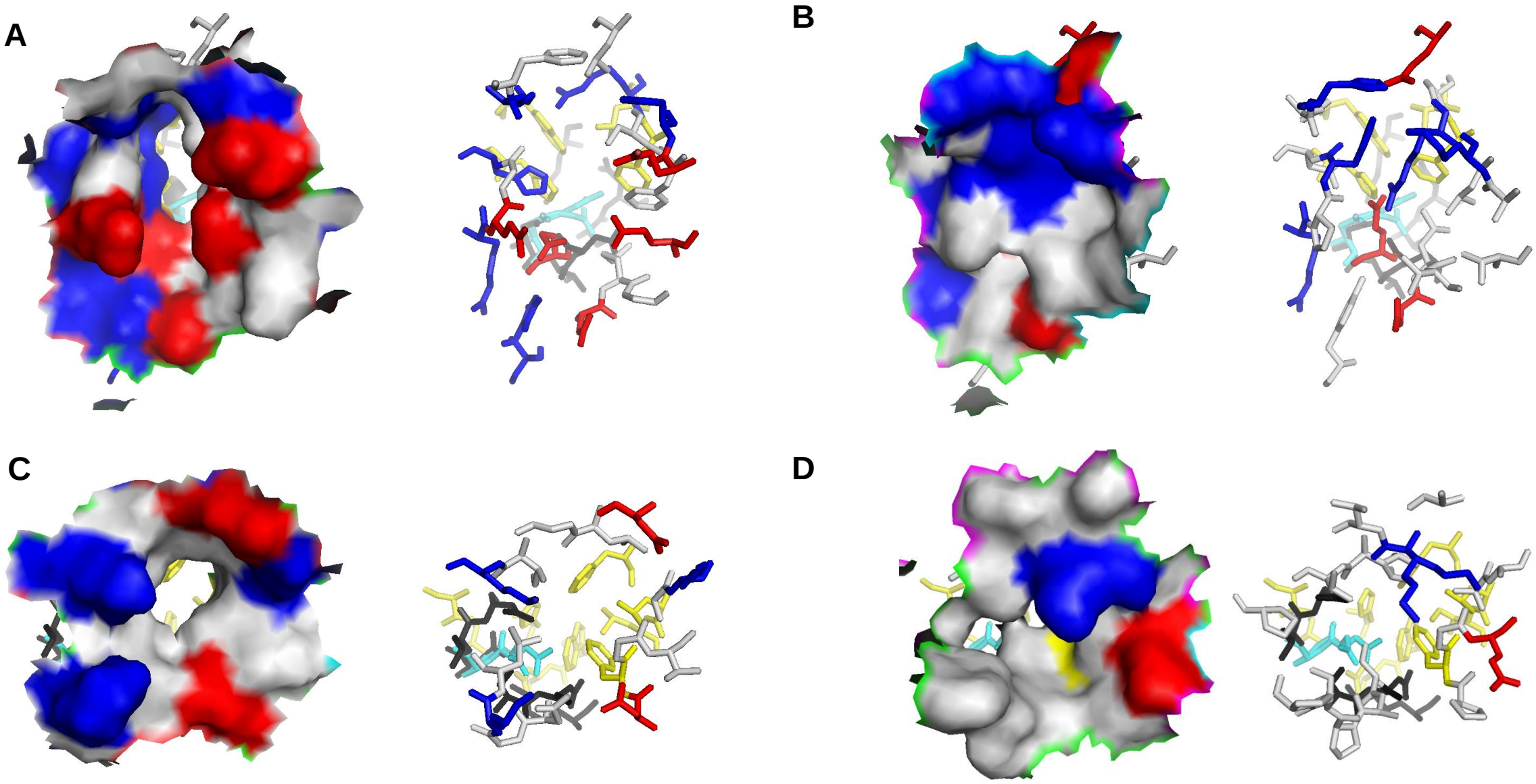


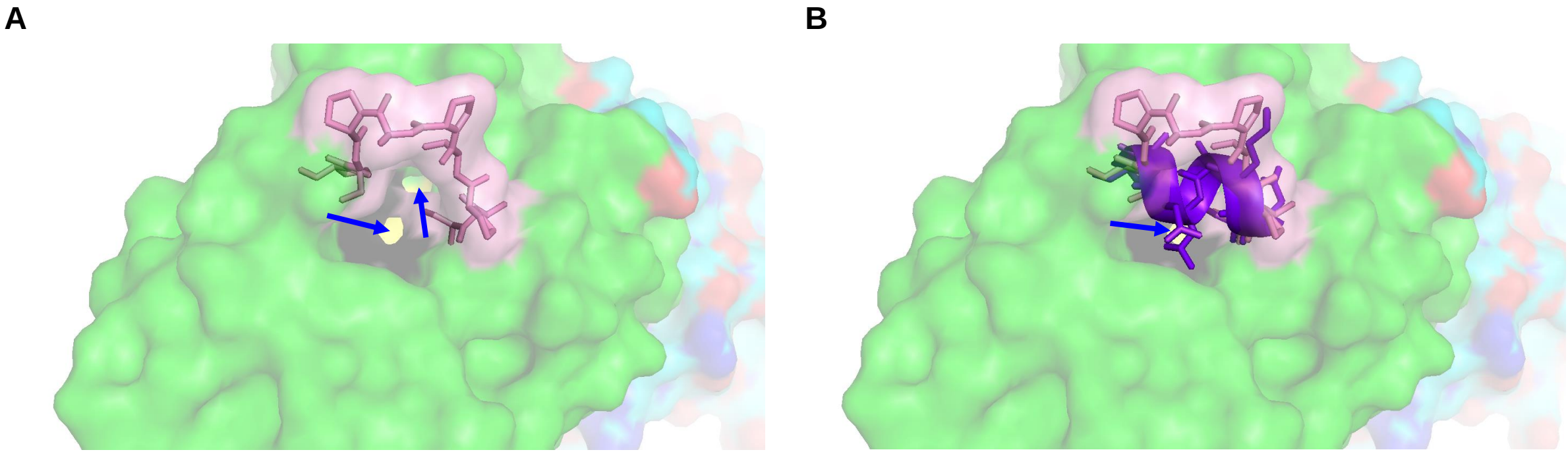
Figure S2: Amino acid sequence alignment of ZgHAD and its homologues DehRhb, DehIVa, L-DEX, DehlB and DehSft, using the programs MultAlin (Corpet, 1988) and ESPript (Robert & Gouet 2014). Identical residues are shown in red boxes and residues with similar properties are in blue boxes. The green triangles highlight the residues that have been shown to be important for catalytic activity and substrate recognition. The blue arrows highlight additional residues belonging to the hydrophobic cluster around the active site in comparison with DehRhb. Black, blue and orange stars indicate the conserved residues respectively involved in dimerization of DehSft according to Rye *et al.*, 2009 and DehRhb according to Novak *et al.*, 2013 and in ZgHAD.



<u>Figure S3:</u> Opening zones, tunnels and determination of ZgHAD dimer obtained by CAVER Analyst program.
(A) Global view of the cavities and tunnels present in the dimer. The 2 catalytic cavities of each subunit are orange and green coloured. The tunnels of one monomer are shown in red and the pathway from one extremity to the other is marked with red dots. (B) Zoom of the smaller entrance site. (C) Zoom on the larger entrance site.

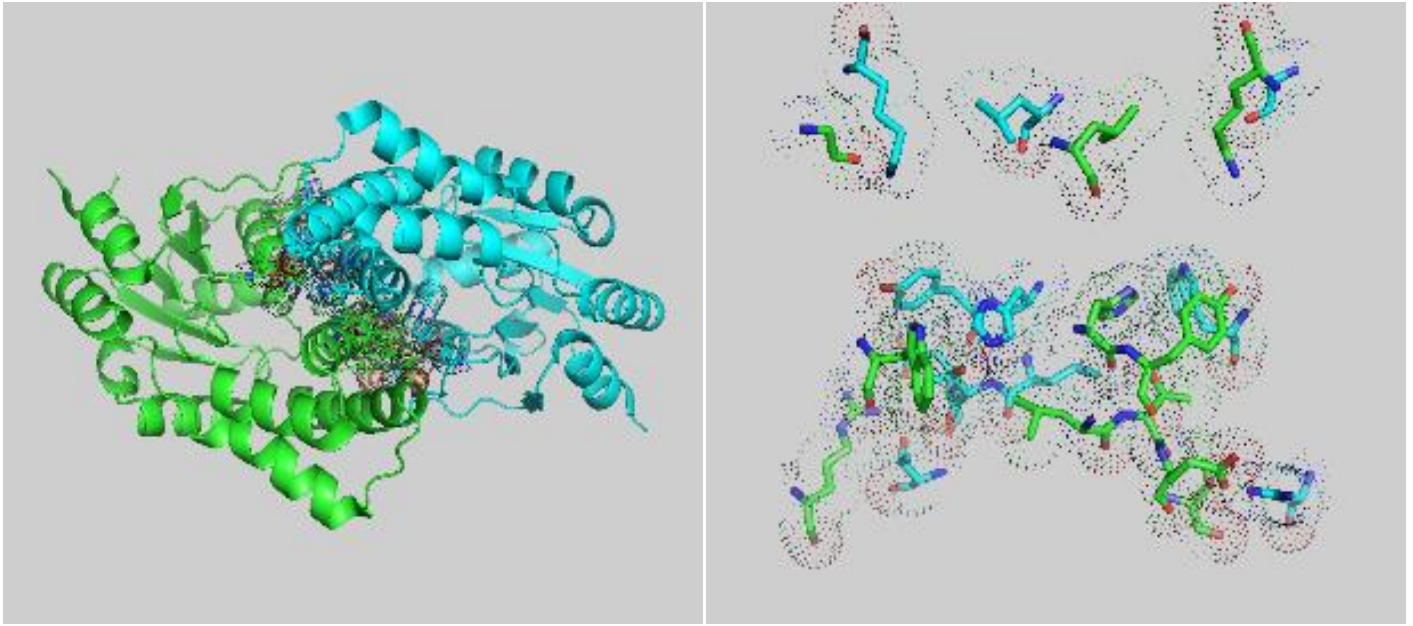


<u>Figure S4:</u> Visualization of the possible entrance sites leading to the active site cavity of ZgHAD compared to DehRhb.
The surface view is modelled on the left and the corresponding residues are represented on the right via stick view. (A) Large entrance accession zone in ZgHAD and (B) corresponding obstructed zone in DehRhb. (C) Small entrance accession zone in ZgHAD and (D) corresponding obstructed zone in DehRhb.
Acidic negatively charged, basic positively charged and neutral amino acids are respectively in red, blue and light grey. Residues implicated in substrate binding and in the hydrophobic pocket are in black and in yellow. Aspartic acid-4-carboxymethyl ester representing the intermediate of chloroacetic acid linked to D14 in DehRhb is shown in cyan. This intermediate has been artificially introduced in the catalytic site of ZgHAD by overimposition of the DehRhb structure (PDB entry: 2YMP).

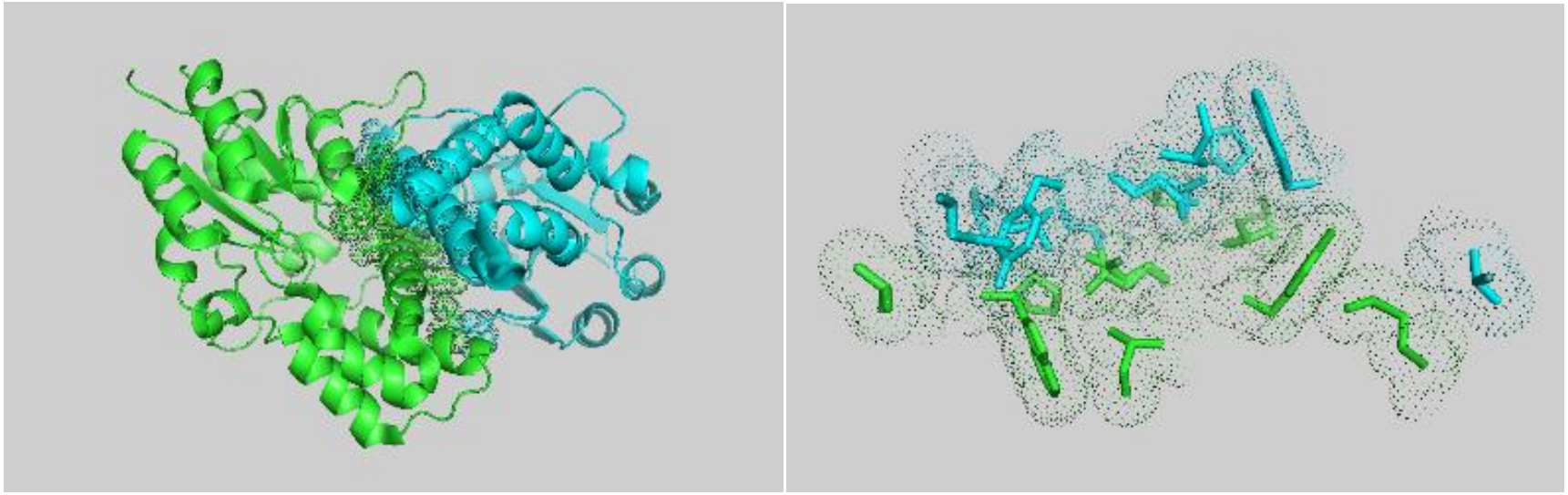


<u>Figure S5:</u> Visualization of the possible entrance sites leading to active site cavity in DehRhb.
(A) Ark structure forming by a loop rich in proline residues liberates 2 orifices connected to the active site. (B) Corresponding ark zone in ZgHAD forms a alpha helix masking the accession to the active site.
Arch structure is represented in pink over the 2 orifices composed respectively of the substrate binding residues in black and the hydrophobic pocket residues in yellow indicated by blue arrows. The overimposed alpha helix structure of ZgHAD is shown in violet.

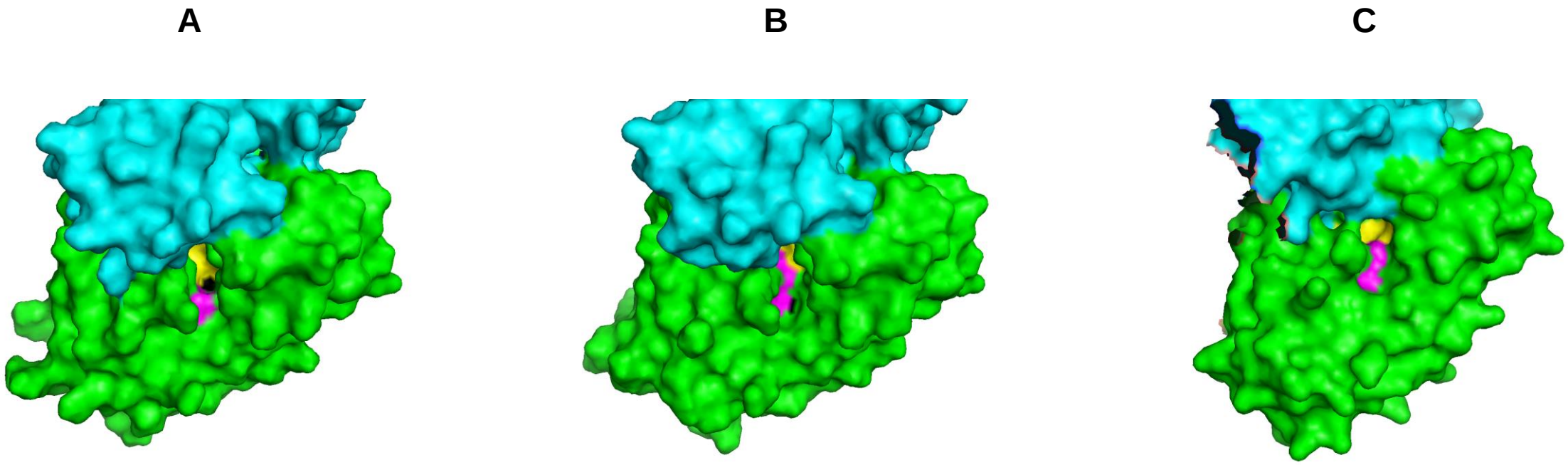
A



B



<u>Figure S6:</u> Dimeric interface representation of ZgHAD compared to DehRhb anlysed using PDBe PISA v1.52.
Every left panel represents the overall view and right panel shows the zoom on dimeric interface.
(A) Pictures of the ZgHAD dimer and (B) Pictures of DehRhb. In each enzyme, one monomer is coloured in green and the other coloured in cyan. Amino acids are surrounded by dots edited in pymol to visualize the interaction zone between the residues.



<u>Figure S7:</u> Surface view of E17 amino acid position at the putative substrate entrance site in (A) ZgHAD wild type enzyme, (B) ZgHAD mutant H190A and (C) ZgHAD mutant H190N.
Each monomer is represented in green and cyan. Putative substrate binding residues are in black and hydrophobic pocket residues of the active site are in yellow. E17 amino acid is shown in magenta.

L-DEX YL	DhlB	DehIVa	DehSft	DehRhb	ZgHAD
D10	D8	D11	D7	D18	D14
T14	T12	T15	T11	T22	T18
R41	R39	R42	R23	F47	F43
S118	S114	S119	S95	T124	S120
K151	K147	K152	K128	K157	K153
Y157	Y135	Y158	Y134	Y163	Y159
S175	S171	S176	S150	A181	A177
N177	N173	N178	N152	H183	H179
D180	D176	D181	D155	D186	D182

Table S1: Comparison of residues found to be essential for catalysis in L-DEX YL (Kurihara *et al.*, 1995), and the corresponding amino acids in DhlB, DehIVa, DehSft, DehRhb and ZgHAD.
Differences between enzymes are highlighted in red.

H179N_Fwd	CCATGATGGTCGCTGCCAACGGATGGGATATTTTAGG
H179N_Rv	CCTAAAATATCCCATCCGTTGGCAGCGACCATCATGG
H179A_Fwd	CCATGATGGTCGCTGCCGCCGGATGGGATATTTTAGG
H179A_Rv	CCTAAAATATCCCATCCGGCGGCAGCGACCATCATGG
ZgHAD_Fwd	GGATCCTCGGCAAAAGTAAAGAAACCCGAG
ZgHAD_Rv	GAATTCTTAGTTCTTCAATAAGGTTCGGGCTAC

Table S2: sequences of primers used for PCR-based cloning for ZgHAD and and site-directed mutants H179A and H179N.