

## Introduction

Microorganisms use different chemical compounds as energy source including environmental pollutants. Such organisms have specifically developed a variety of enzyme systems to degrade these compounds (Janssen *et al.*, 2005). Some of these compounds are generated by natural processes in the oceans and in the atmosphere. However, intensive industrial development over the past 100 years has generated many new environmental pollutants such as halogenated aliphatic compounds (Swanson, 1999; Janssen & Schanstra, 1994). At the present, the actual problems are cleaning of contaminated sites and the creation of clean production technologies.

The computation design, protein engineering and X-ray protein crystallography are the basis, which reveals protein structure-function relationships. Haloalkane dehalogenase enzymes catalyze the hydrolytic cleavage of carbon-halogen bonds, which is a key step in aerobic mineralization of many environmental pollutants (Janssen *et al.*, 2005). Due to this property haloalkane dehalogenases can be used as biocatalysts, selective reagents and to be applied as environmental protectant (Verschueren *et al.*, 1993c; Pries *et al.*, 1994, Stucki & Thuer, 1995). The substrate specificity and efficiency of haloalkane dehalogenases reactions depend on the size and geometry of an active site and its entrances.

In this thesis protein X-ray crystallography was used to determine the structures of mutant variants DhaA31 and DhaA12 from *Rhodococcus rhodochrous* NCINB13064 (Kulakova *et al.*, 1997).

Objectives of the first project included:

- a) Finding the most suitable crystallization conditions of DhA31 mutant protein and further optimization process.

- b) Soaking experiments with *1,2,3-trichloropropane*.
- c) Solving structure of DhaA31 variant in two forms: substrate-free form and protein-substrate complex with *1,2,3-trichloropropane*.
- d) Detailed crystallographic analysis of two experimental structures.
- e) Comparison of different stages of DhaA31 enzyme along the reaction pathway with previously described reaction mechanism of haloalkane dehalogenase Dh1A from *Xanthobacter autotrophicus* GJ10 with *1,2-dichloroethane*.
- f) Explanation of dehalogenation mechanism inside the active site of DhaA31.

The second project was focused on creation the new mutant DhaA12 with loop insertion from DbjA in DhaA wild-type haloalkane dehalogenase. The main goal of this work was determination of loop insertion and possible influence on the exchange process between active site and bulk solvent.

## **Haloalkane dehalogenases**

Haloalkane dehalogenases are enzymes that catalyze the cleavage of carbon–halogen bonds by a hydrolytic mechanism. These proteins belong to the most frequently found in nature  $\alpha/\beta$  domain fold. In 1992, a subclass of this fold was called the  $\alpha/\beta$  hydrolase fold (Ollis *et al.*, 1992; Nardini & Dijkstra, 1999). The  $\alpha/\beta$  hydrolase structural fold was found in lipases, acetylcholinesterases, esterases, lactonases, epoxide hydrolases and other enzymes, showing that the haloalkane dehalogenases belong to the protein superfamily of which the members carry out diverse reactions. Several haloalkane dehalogenases are involved in biodegradation pathways of important environmental pollutants (Verschueren *et al.*, 1993; Bosma *et al.*, 2002; Janssen *et al.*, 2005). These enzymes have also been found to be present in pathogenic bacteria and rhizobial bacteria, where their function remains unknown (Chovancova *et al.*, 2007).

### **2.1. Haloalkane dehalogenases classification system**

Many proteins have been suggested to belong to Haloalkane dehalogenases (*HLDs*) family based mainly on sequence similarities. However, it is unclear whether they represent true members of the family in either an evolutionary or functional sense. The evolutionary classification scheme should also serve as a useful platform to identify and classify new family members and to guide predictions concerning their catalytic, biochemical, and structural properties. This phylogenetic analysis of *HLDs* have been published by Chovancova and colleagues (2007), revealing three subfamilies denoted as *HLD-I*, *HLD-II* and *HLD-III*. Haloalkane dehalogenases substrate specificity and the catalytic pentad amino acids are also important factors for classification system (Janssen *et al.*, 2005). The substrate specificity can be function of several structural features such as size and geometry of an active site, the size and

geometry of the entrance to the active site and the efficiency of the transition state and halide stabilization by the active site residues (Damborsky & Koca, 1999). All three subfamilies include experimentally characterized *HLDs* (Table 1). The most of the biochemically characterized *HLDs* belong to the *HLD-II* subfamily. The *HLD-III* subfamily is not as well-defined as *HLD-I* and *HLD-II*.

## 2.2. Structures of haloalkane dehalogenases


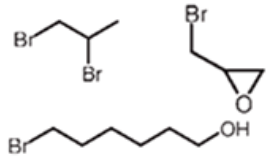

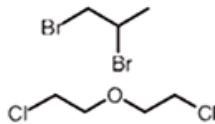
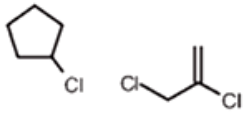
Structurally, haloalkane dehalogenases belong to the  $\alpha/\beta$  hydrolase superfamily (Chovancova *et al.*, 2007). The proteins consist of two domains. The main  $\alpha/\beta$ -hydrolase domain composing a  $\beta$ -sheet surrounded by  $\alpha$ -helices is conserved in members of the  $\alpha/\beta$ -hydrolase superfamily (Fig.1). This domain serves as a scaffold for the catalytic residues. The second helical cap domain laying on the top of the main domain and composing of five helices connected by loops is structurally more variable. The cap domains are much more flexible than the main domains and display concerted motions. It is known that this fact influence the substrate specificity of these enzymes (Koudelakova *et al.*, 2011).

Two domains contribute to the formation of an internal cavity. This cavity is predominantly composed of hydrophobic residues (Fig.1). The only charged residues in the cavity are two of five catalytic residues: a nucleophilic aspartate and a conserved histidine base close to the C-terminus. The third member of the catalytic pentad is a catalytic acid. Another two residues involved in binding of the halide (Janssen *et al.*, 1994; Otyepka & Damborsky, 2002; Janssen *et al.*, 2005).

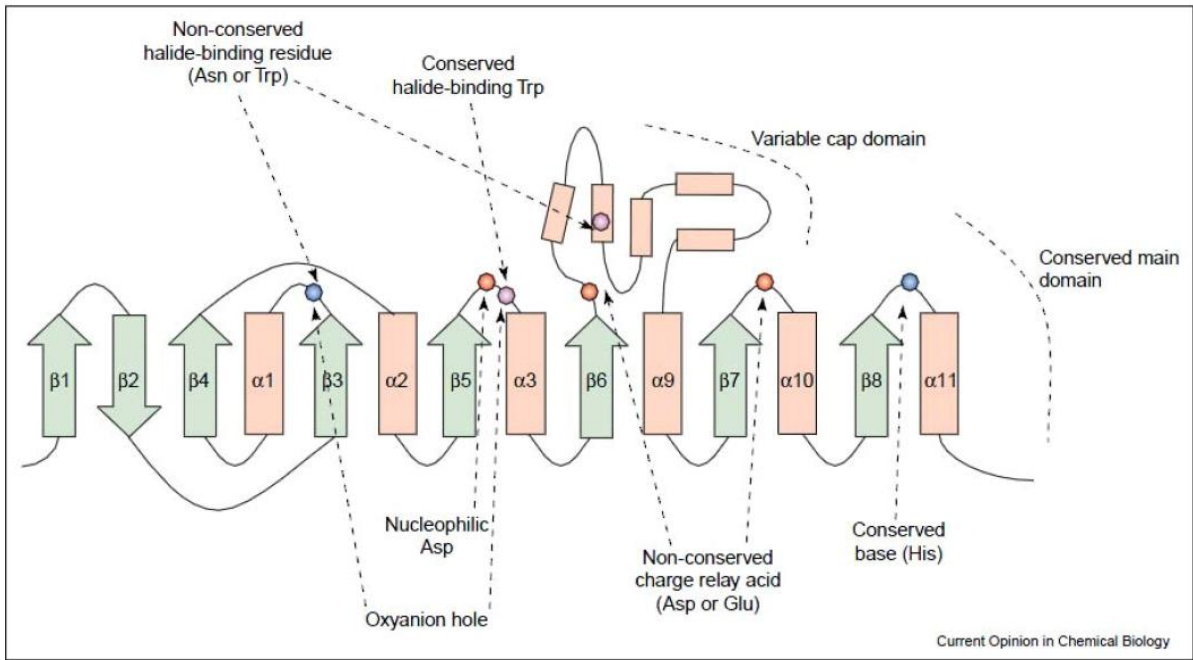
Three available experimental structures of HLDs were compared: Dh1A (PDB ID 1EDE), DhaA (PDBID 1BN6), and LinB (PDB ID 1IZ7). These enzymes differ mainly in the cap domain: 1)  $\alpha 5'$  helix in the cap domain lost

in the *HLD-I* protein Dh1A, but it is common to *HLD-II* subfamily members DhaA and LinB (Fig.2); 2) the spatial arrangement of helices  $\alpha_4$  and  $\alpha_5$  is different in the cap domains of Dh1A and *HLD-II* enzymes; 3) members from different *HLD* subfamilies exhibit structural divergence in the loop regions that connect the cap to the main domain, and in the C-terminal part of the main domain.

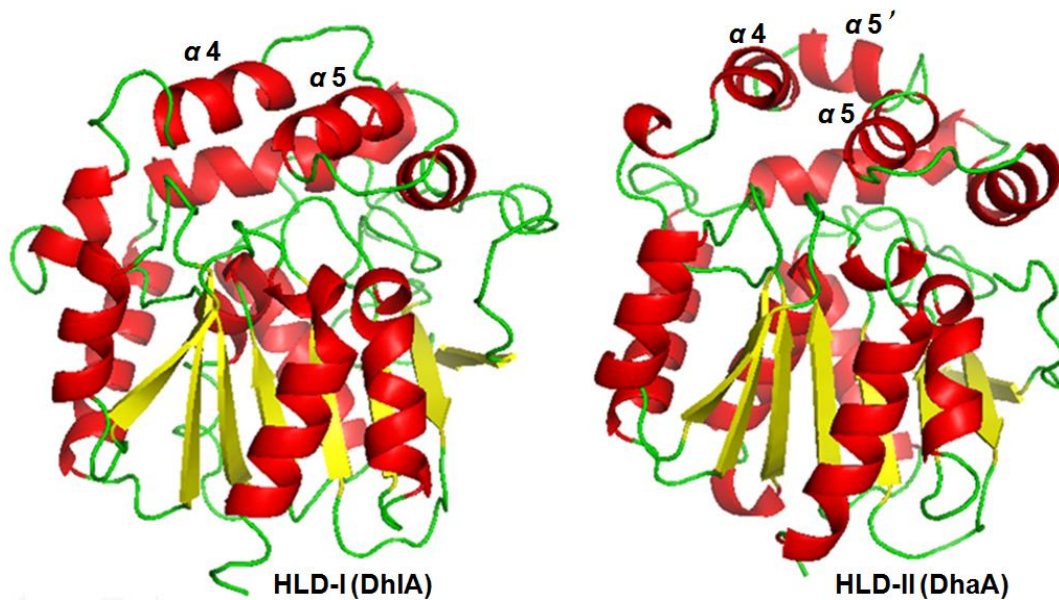
**Table 1. List of experimentally characterized Haloalkane dehalogenases.**

subfamily	protein	organism	substrates	cat.pentad
<i>HLD-I</i>	Dh1A	Xanthobacter autotrophicus	small, terminally halogenated	Asp-His- Asp + Trp-Trp
	DhmA	Mycobacterium avium		
	DmbB	Mycobacterium tuberculosis		
<i>HLD-II</i>	LinB	Sphingobium japonicum	larger, $\beta$ -substituted	Asp-His- Glu + Asn-Trp
	DmbA DmsA	Mycobacterium tuberculosis Mycobacterium smegmatis		
	DhaA	Rhodococcus sp.		
	DbjA DmlA DatA	Bradyrhizobium japonicum Mesorhizobium loti Agrobacterium tumefaciens		
<i>HLD-III</i>	DrbA	Rhodopirellula baltica	unknown	Asp-His- Asp + Asn-Trp
	DmbC	Mycobacterium tuberculosis		

Source: Chovancova et al., 2007.



**Figure 1.** *The topological arrangement of secondary structure elements of haloalkane dehalogenases (adapted from Janssen, 2004).*



**Figure 2.** *Structures of two members haloalkane dehalogenases determined by protein crystallography (adapted from Chovancova, 2007).*

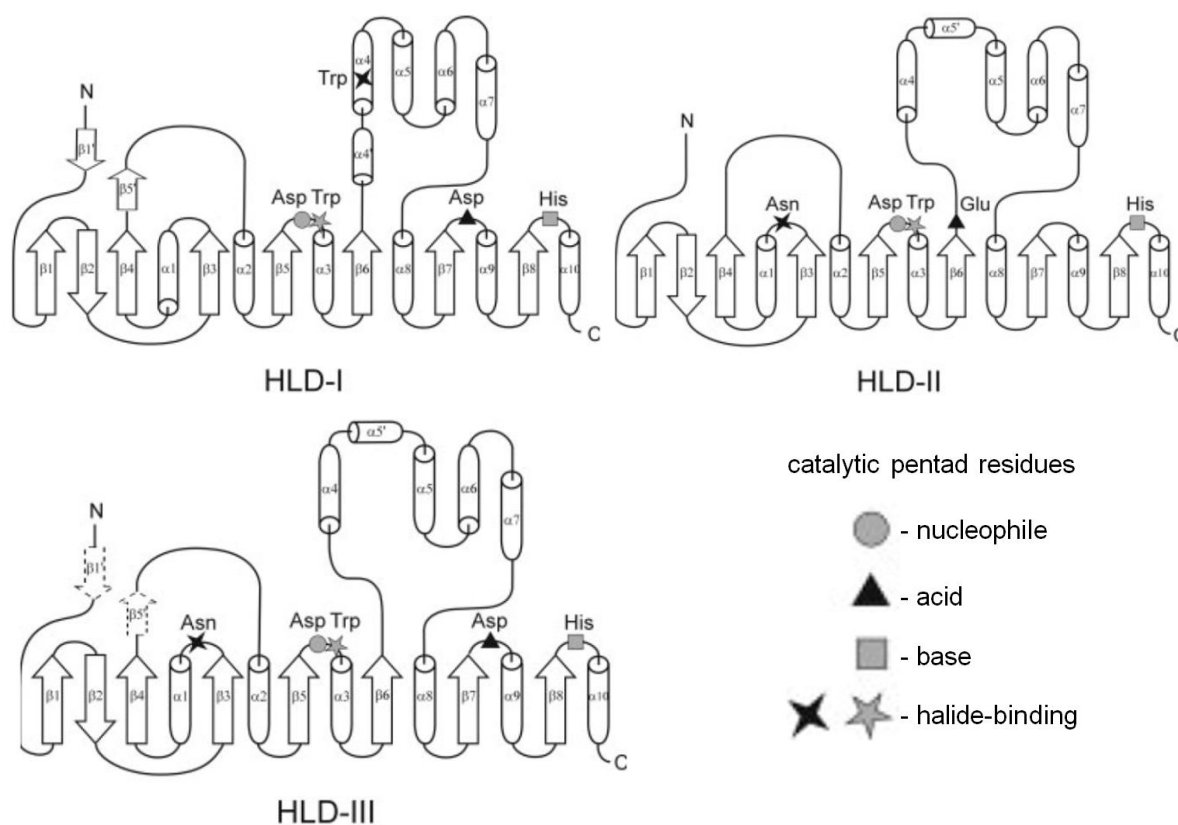
Analysis of the multiple sequence alignment reveals differences within subfamilies. For example, DbjA exhibits an 11 residue-long insertion in the terminus of the cap domain that is not present in other proteins from the *HLD-II* subfamily. In the sequence of Dh1A, a long insertion of 10 amino acids is situated in an analogous region.

In addition, Dh1A active site is small ( $112 \text{ \AA}^3$ ) completely buried inside the protein, the substrate molecules entrance active site through the narrow tunnel. This anatomy of Dh1A determines the dehalogenating activity for small terminally halogenated aliphatic hydrocarbons. On the other hand, representatives of *HLD-II* subfamily have comparably long cavity, bigger active site ( $276 \text{ \AA}^3$  for LinB,  $246 \text{ \AA}^3$  for DhaA, and  $264 \text{ \AA}^3$  for DbjA) as well as wide and open tunnels. The optimum activity of these enzymes is reached with halogenated compounds containing seven, five or six carbon atoms (Verschueren *et al.*, 1993c; Damborsky & Koca, 1999; Sato *et al.*, 2005).

Different compositions of the catalytic pentad also show differences between *HLD* subfamilies (Fig.3). Three residues of the pentad are identical in all subfamilies and thus would be expected to fulfill identical functions: nucleophile (Asp), catalytic base (His), and one of the halide stabilizing residues (Trp). *HLDs* differ by two types of catalytic acid: aspartic acid for *HLD-I* and glutamic acid for *HLD-II*. Moreover, in the *HLD-I* subfamily a Asp, is located in the loop following  $\beta$ -strand 7 (Dh1A), whereas *HLD-II* subfamily members contain a Glu in the loop following  $\beta$ -strand 6 (LinB).

The halide ion released during the dehalogenation reaction is mainly stabilized by the hydrogen bonds provided by two tryptophan residues (*HLD-I*) or by tryptophan and asparagines (*HLD-II*). *HLD-I* members employ Trp (Dh1A) located in helix  $\alpha 4$ , whereas *HLD-II* members use Asn (LinB) located in the loop following  $\beta$ -strand 3.

So, the  $\alpha/\beta$  core is the most conserved region of the *HLDs*, and surface residues are the least conserved. Regions having the highest sequence variability were found in the N-terminal part of the cap domain and in all three helices within the C-terminal region of the main domain. Nine residues are fully conserved in all analyzed



**Figure 3.** The topological arrangement of secondary structure elements in individual haloalkane dehalogenase subfamilies (HLD-I, HLD-II, and HLD-III). (adapted from Chovancova *et al.*, 2007).

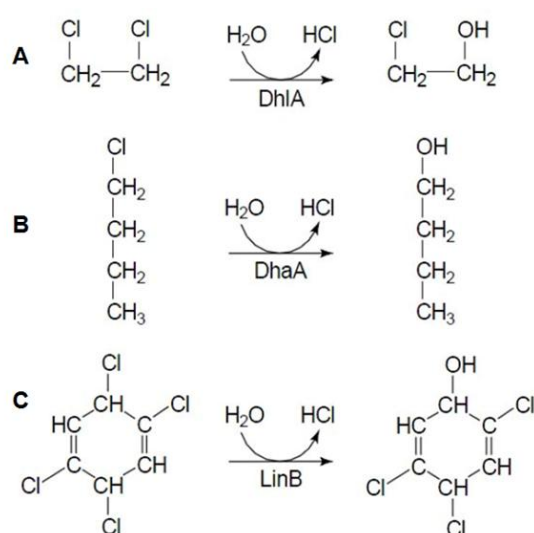
sequences: the catalytic base (His), the catalytic nucleophile (Asp) and one of the halide stabilizing residues (Chovancova *et al.*, 2007).

### 2.3. Chemical reactions of haloalkane dehalogenases

A common nature of dehalogenases is that they carry out nucleophilic displacement reactions on a  $sp^3$  hybridized carbon atom (Janssen *et al.*, 1994). In the first the catalytic mechanism of haloalkane dehalogenase was



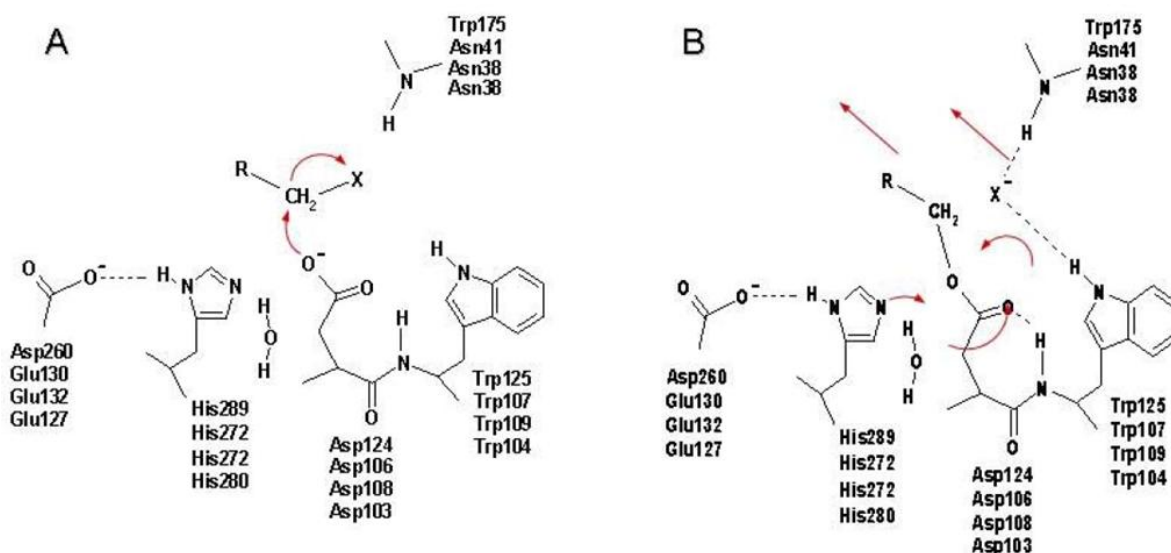
described by Verschueren and colleagues in 1993. It was based on the reaction of Dh1A protein from *Xanthobacter autotrophicus* GJ10 with the chemical compound *1,2-dichloroethane*, mainly used to produce vinyl chloride monomer and as an intermediate for other organic chemical compounds. It forms azeotropes with many other solvents, including water and other chlorocarbons. Dh1A transforms *1,2-dichloroethane* to *2-chloroethanol*, which can be then mineralized. Dh1A shows significant structural similarity to other haloalkane dehalogenase LinB. The dehalogenation activity of LinB occurs with wide range of substrates such as monochloroalkanes, dichloroalkanes, bromoalkanes, and chlorinated aliphatic alcohols. Moreover, the active-site cavities of LinB as well as of DbjA are the largest of all HLDs and this allow both of enzymes to convert bigger substrates. For example LinB can convert laboratory chemical, rare earth compound, *1,3,4,6-tetrachloro-1,4-cyclohexadiene* (1,4-TCDN) to *2,4,5-trichloro-2,5-cyclohexadiene-1-ol* (Fig. 4). Finally, subproduct can be completely degraded with participation of dehydrogenase LinC (Nagata *et al.*, 1997).



**Figure 4.** Dehalogenation reactions in bacterial cultures from (A) *Xanthobacter autotrophicus* (Dh1A), (B) *Rhodococcus erythropolis* (DhaA), (C) *Spingomonas paucimobilis* (LinB) (adapted from Janssen, 2004).

The dehalogenation reaction of haloalkane dehalogenases proceeds in the enzyme active site. At the beginning the one chlorine atom of the halogenated substrate lies in halide binding site between nitrogen atoms of two tryptophan rings (Trp104, 107, 109, 125 and Trp175) or nitrogen atoms of tryptophan ring (Trp104, 107, 109, 125) and asparagines (Asn38, 41) (Verschueren *et al.*, 1993*a,b,c*; Stsiapanava *et al.*, 2010). Next step is nucleophilic attack of nucleophile aspartate (Asp103, 106, 108, 124) on a substrates' carbon atom carrying a halogen. This leads to the cleavage of the carbon-halogen bond, displacement of the halide and formation of a covalent alkyl-enzyme intermediate. Hydrolysis of the covalent intermediate can only occur if an ester intermediate is formed. The hydrolysis of this ester is facilitated by the presence of an oxyanion hole for the carbonyl oxygen of the nucleophilic aspartate (Asp103, 106, 108, 124) with water molecule activated by a catalytic base (His272, 280, 289) (Fig. 5). The last step in the catalytic cycle is releasing of products: alcohol and chloride. Finally, the enzyme is regenerated before the start of next reaction cycle (Verschueren *et al.*, 1993*c*; Janssen *et al.*, 1994; Bosma *et al.*, 2003; Janssen *et al.*, 2004; Otyepka *et al.*, 2007; Pavlova., 2009).

The simplicity of this catalytic mechanism makes this enzyme possible for studying the fundamental principles of enzymatic function of haloalkane dehalogenases (Klvana *et al.*, 2009).



**Figure 5.** General catalytic mechanism of haloalkane dehalogenases.

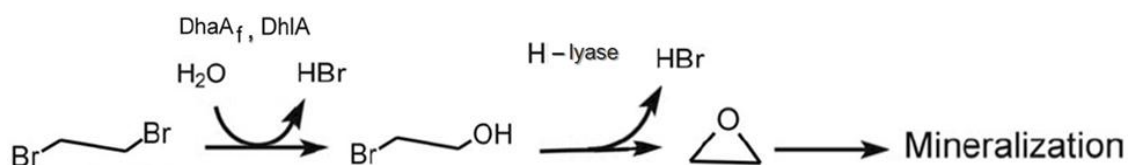
(A) Formation of the covalent intermediate by SN2 substitution.

(B) Hydrolysis of the intermediate. Residue numbers refer to DhIA, DhaA, LinB, and DbjA, respectively (adapted from Janssen, 2004).

### Applications of haloalkane dehalogenases

The main applications of *HLDs* are associated with their participation in biodegradation activity. One of the most important groups of contaminants is halogenated aliphatic compounds. They violate the balance of ecosystems and have adversely affects to the living organisms. Haloalkanes are widely used as solvents, degreasing agents, intermediates in chemical synthesis, and pesticides (Swanson, 1999; Janssen & Schanstra, 1994). The cleavage of the carbon-halogen bonds in these compounds is a critical step in their biodegradation. For example, degradation of *1,2-dibromoethane* (DBE), synthetic pesticide, although trace amounts occur naturally in the ocean. It is a known carcinogen. The effects on animals' organism with short-term exposures to high levels caused depression and collapse, indicating effects on the brain. The first step in *1,2-dibromoethane* metabolism is catalyzed by a hydrolytic haloalkane dehalogenase (DhIA, DhaA<sub>f</sub>). Conversion of the

hydrolytic product, *2-bromoethanol*, is realized by haloalkohol dehalogenase (H-lyase) that converts haloalcohols to the corresponding epoxides and halide ions (Poelarends *et al.*, 1999) (Fig. 6). Another known domain of the use of haloalkane dehalogenases is their application in decontamination of chemical warfare agents (Prokop *et al.*, 2005; Prokop *et al.*, 2006).



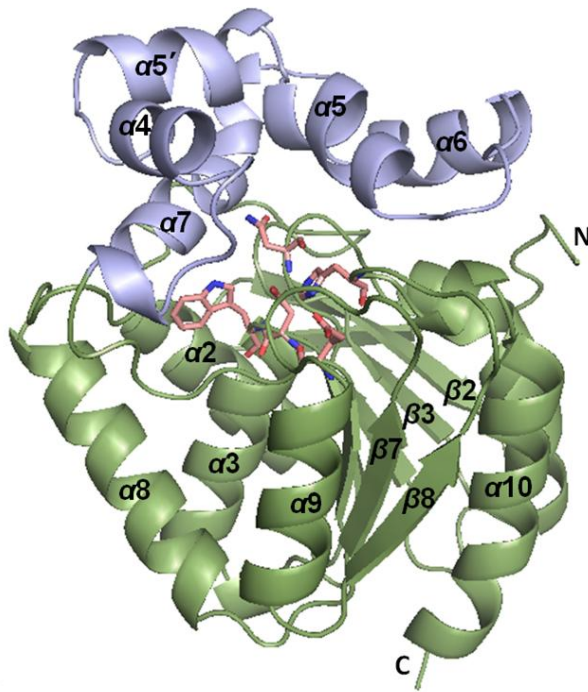
**Figure 6.** Proposed route of the metabolism of 1,2-dibromoethane (adapted from Poelarends, 1999).

New applications of *HLDs* were found in biocatalysis of such toxic compound as *1,2,3-trichloropropane* (Pavlova *et al.*, 2009; Lahoda *et al.*, 2011). The enantioselective properties of haloalkane dehalogenases are useful in pharmacology for preparation of optically pure medicinal components (Prokop *et al.*, 2010). Moreover, *HLDs* have potential in phytoremediation and as active components of biosensors (Campbell *et al.*, 2006; Bidmanova *et al.*, 2010). The haloalkane dehalogenases genes were transferred into plants organisms. The introduction of DhIA and DhIB in *Nicotiana tabacum* shows clean-up activity of halogenated organic pollutant *1,2-dichloroethane* from contaminated soil and groundwater (Mena-Benitez *et al.*, 2008). The transgenic plant *Arabidopsis thaliana* expressing DhIA was engineered as biomarker of *1,2-dichloroethane* contamination (Naested *et al.*, 1999). This is low-cost and fast method of biomonitoring in comparison with traditional analytic techniques. All of these applications are promising approach for their future development in science and their using by people of society.

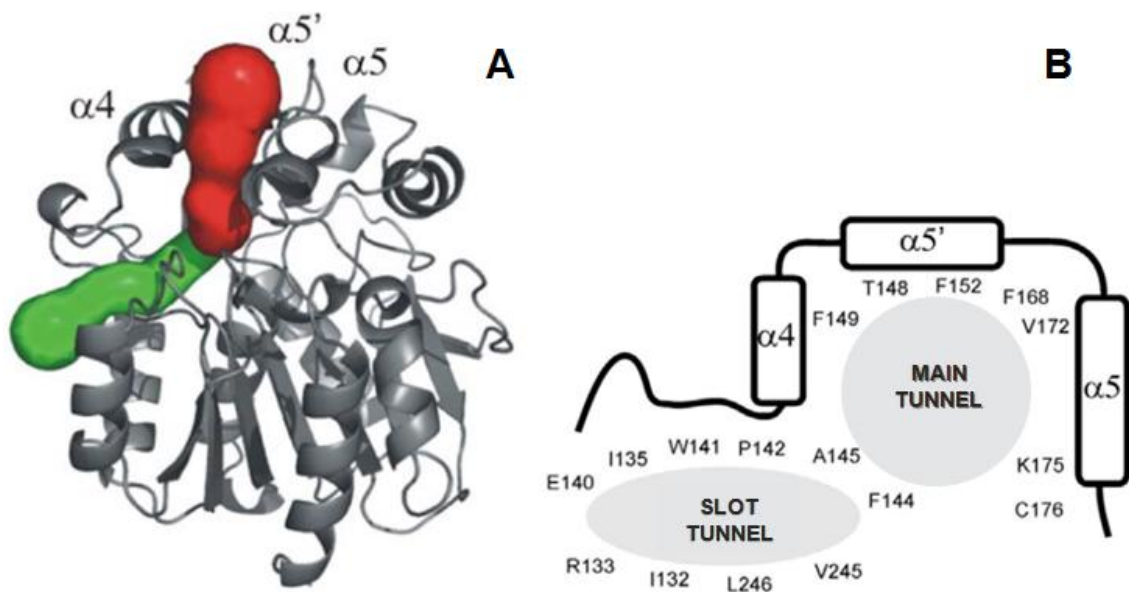
## 2.5. DhaA representative of haloalkane dehalogenases

The DhaA haloalkane dehalogenase from *Rhodococcus rhodochrous* NCINB13064 is involved in initial dehalogenation of several haloalkanes. It is a key step in the aerobic mineralization of many halogenated compounds marked as environmental pollutants (Janssen *et al.*, 2005). The native substrate for DhaA protein is *1-chlorobutane* (Poelarends *et al.*, 2000). Particular interest lies in studying the modification of the DhaA to improve the conversion of various chemicals substrates, such as *1,2,3-trichloropropane* (Gray *et al.*, 2001; Bosma *et al.*, 2002; Pavlova *et al.*, 2009). The good targets for DhaA mutagenesis are the tunnel residues located outside of active site as their replacement does not lead to loss their functionality by disruption of the active site architecture (Pavlova *et al.*, 2009; Stsiapanava *et al.*, 2010).

DhaA has globular structure composed of two domains. Central catalytic domain with an  $\alpha/\beta$  hydrolase fold consists of eight-stranded  $\beta$ -sheet and six  $\alpha$ -helices ( $\beta 1-\beta 2-\beta 3-\alpha 1-\beta 4-\alpha 2-\beta 5-\alpha 3-\beta 6$ ,  $\alpha 8-\beta 7-\alpha 9-\beta 8-\alpha 10$ ). Second domain is cap domain composed of five  $\alpha$ -helices ( $\alpha 4-\alpha 5'-\alpha 5-\alpha 6-\alpha 7$ ). The active site cavity is located between two protein domains and connected with the protein surface by two major access tunnels: the main tunnel and the slot tunnel used by ligands for their exchange between the active site and surrounding solvent. A catalytic pentad of DhaA includes a nucleophile (Asp106), a base (His272), a catalytic acid (Glu130) and two halide-stabilizing residues (Asn41 and Trp107) (Fig.7).



**Figure 7.** Cartoon representation of the DhaA31 structure. The  $\alpha/\beta$ -hydrolase core domain is shown in green; the helical cap domain is shown in blue. Catalytic pentad residues are shown in magenta stick representation.



**Figure 8.** DhaA tunnels. (A) The main tunnel is colored in red. Slot is highlighted in green. (B) Schematic representation of access paths for DhaA identified by protein crystallography and molecular dynamic simulations (adapted from Petrek *et al.*, 2006).

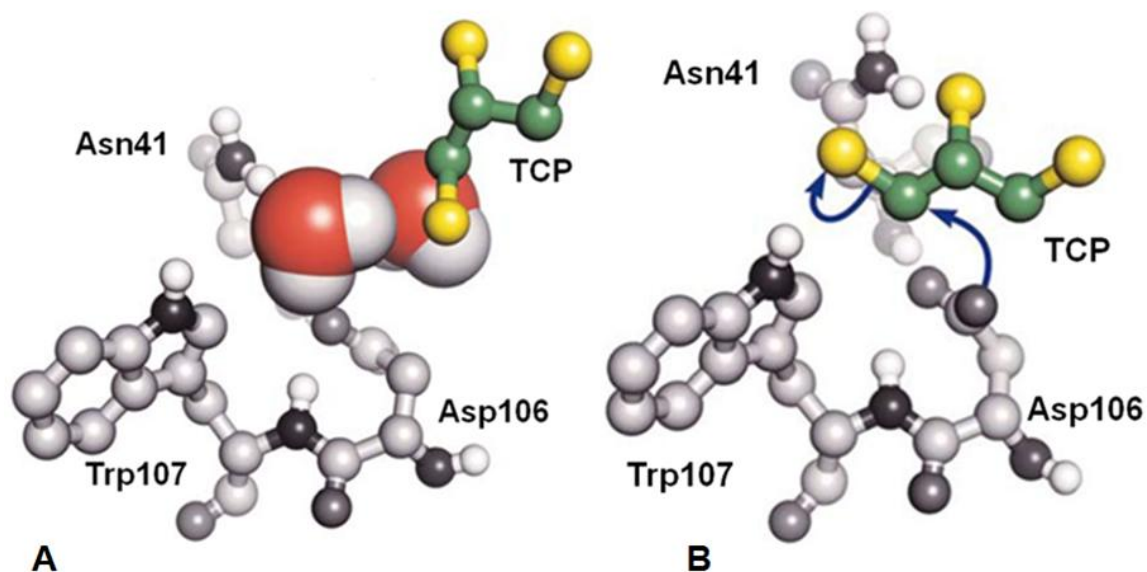
The main tunnel is located between  $\alpha 4$  and  $\alpha 5$  helices of the cap domain and is formed by nonpolar residues (such as Ala145, Phe144, Phe149, Phe168, Ala172, and Cys176), and polar residues as Thr148 and Lys175. The slot tunnel is located between  $\alpha 4$  helix of the cap domain and loops connecting helix  $\alpha 4$  with strand  $\beta 3$  of the core domain and strand  $\beta 3$  with helix  $\alpha 8$ . This tunnel is surrounded by nonpolar amino acids Ile132, Ile135, Trp141, Pro142, Leu246 and Val245, and polar residues Arg133 and Glu140 (Fig.8) (Newman *et al.*, 1999; Otyepka & Damborsky, 2002; Petrek *et al.*, 2006; Klvana *et al.*, 2009). The DhaA belongs to the *HLD-II* subfamily according with latest phylogemetic analysis of haloalkane dehalogenases superfamily (Chovancova *et al.*, 2007).

## 2.6. DhaA protein variants

### 2.6.1. DhaA31 mutant

The haloalkane dehalogenase DhaA from *Rhodococcus rhodochrous* NCIMB 13064 (Kulakova *et al.*, 1997) can slowly convert the industrial pollutant and suspected human carcinogen *1,2,3-trichloropropane* (TCP) to *2,3-dichloropropane-1-ol* (DCL) under laboratory conditions. The efficiency of this reaction is too low for biotechnological processes. Moreover, due to the low conversion rate, cells are longer exposed to toxic effects of TCP, which can inhibit their growth (Schindler *et al.*, 1999; Bosma *et al.*, 1999; Bosma *et al.*, 2003). Several variants of DhaA were constructed to increase the efficiency of this reaction (Bosma *et al.*, 2002; Gray *et al.*, 2001, Pavlova *et al.*, 2009, Klvana *et al.*, 2009, Stsiapanava *et al.*, 2010). Finally, focused directed evolution was used to construct the most successful variant DhaA31 with up to 32-fold higher catalytic activity and 26-fold higher catalytic efficiency, than parent wild type enzyme. The positions for mutagenesis were selected by computer modeling and randomized experimentally. The DhaA31 mutant protein contains five mutations. The C176Y mutation was found to modify the mouth of the main tunnel connecting the buried active site with the surrounding solvent. The I135F and L246I mutations located in the slot tunnel and the V245F and Y273F mutations located in the main tunnel were found to constrict access of tunnels. Computer analysis revealed that the active site of the wild type DhaA is easily accessible to water molecules in the active site compete with TCP for interaction with the nucleophile, Asp106. Thus, the narrowed tunnels of DhaA31 shield the active site from the bulk solvent, and the activated complex is more stable than in the parent DhaA protein (Pavlova *et al.*, 2009) (Fig.9).

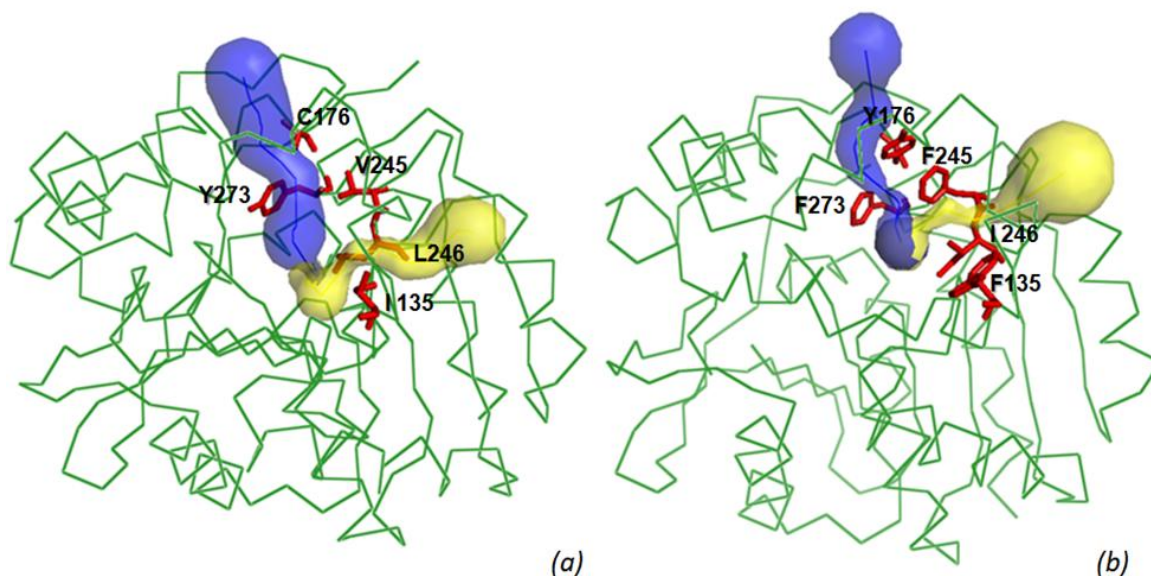




**Figure 9.** Comparison of the wild type *DhaA* and the *DhaA31* mutant. (A) nonreactive configuration of TCP competing with waters for the nucleophile in the wild type *DhaA*. (B) near-attack configuration of TCP in the novel haloalkane dehalogenase mutant *DhaA31* (adapted from Pavlova *et al.*, 2009).

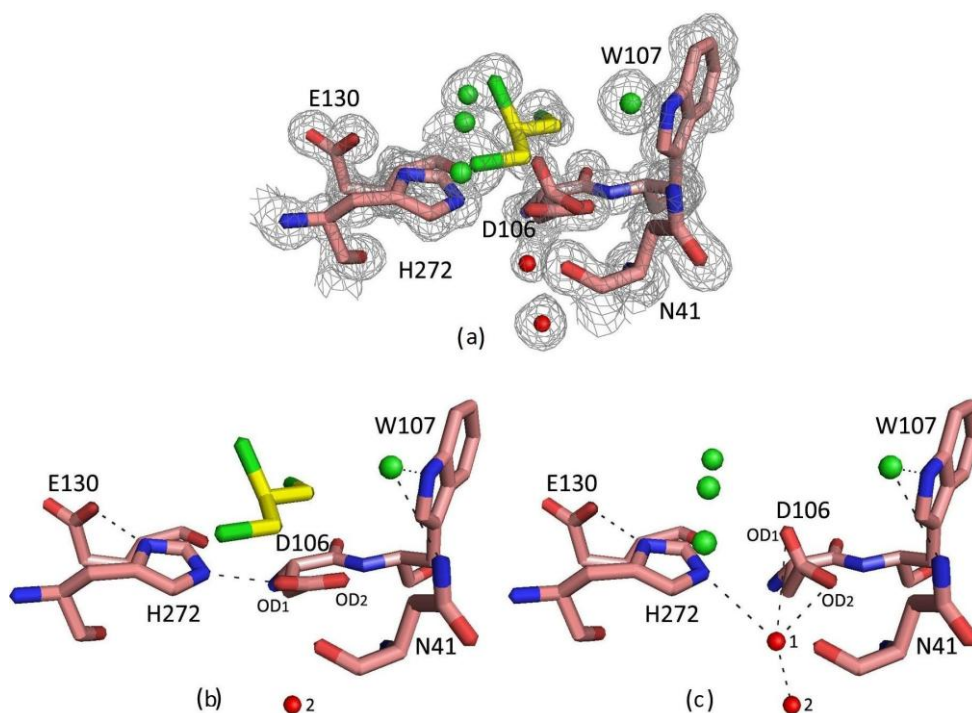
Moreover, the study of these two *DhaA* access tunnels shows that the dehalogenation product DCL exchanged more frequently through the main tunnel but it is also possible through the slot tunnel. The TCP substrate's entry to or exit from the active site released through the slot tunnel. DCL release through the main tunnel of *DhaA31* mutant is reduced and completely prevented through the slot tunnel because of I135F, C176Y and V245F bulky substitutions (Fig.10).

Kinetic analyses confirmed that the mutations improved carbon-halogen bond cleavage. The mutant *DhaA31* has  $k_{\text{cat}}$  value equal  $1.26 \text{ s}^{-1}$  and  $K_{\text{m}}$  value equal 1.2 mM, while the wild type has  $k_{\text{cat}}$  value equal  $0.04 \text{ s}^{-1}$  and  $K_{\text{m}}$  value equal 1.0 mM. Thus, optimization tunnel anatomy demonstrated changes in physicochemical properties in *DhaA31* mutant by enhance activity with TCP by decreasing accessibility of the active site for water molecules, thus promoting formation of an activated complex (Pavlova *et al.*, 2009).



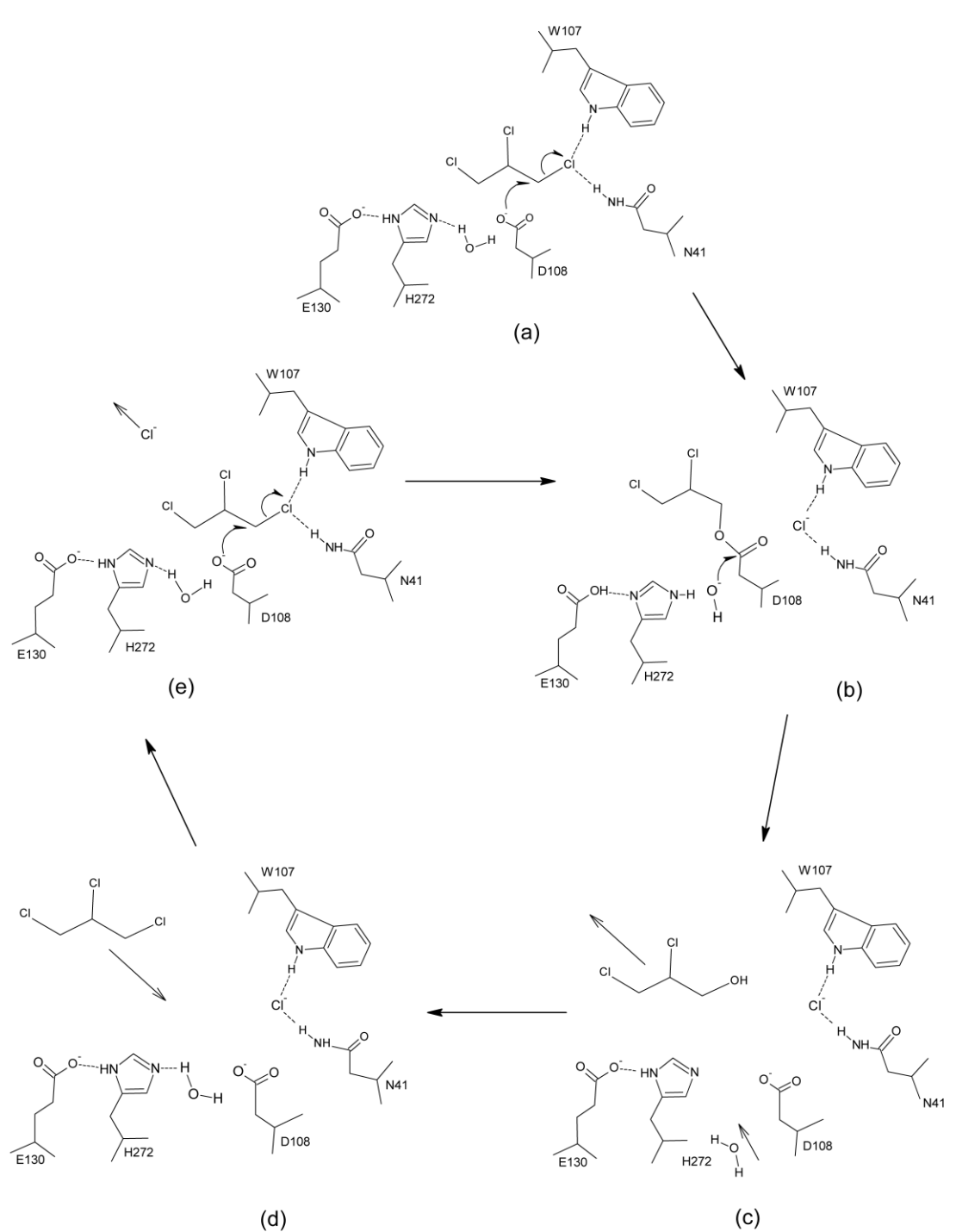
**Figure 10.** Ribbon representations of the structures of wild-type DhaA (a) and DhaA31 (b). Main tunnel shown in blue. Slot tunnel shown in yellow. The residues selected for mutagenesis (a) and the mutated residues (b) are shown as red sticks (adapted from Lahoda *et al.*, 2012).

In turn, crystal structure of DhaA31-substrate complex contains a combination of two different states inside active-site cavity. The first one is characterized by post-attack conformation of Asp106 (alternative A, occupancy of 0.6) and pre-near-attack orientation of substrate molecule (Fig.11b and 12d). The catalytic Asp106 has already undergone a reaction cycle as evidenced by the absence of the water molecule (H<sub>2</sub>O-1) which was used for extracted a proton during nucleophilic attack (Fig.12b). At this state His272 side chain protonated due to low pH and no acting as a base (Verschueren *et al.* 1993c). The O<sup>δ1</sup> atom of Asp106 interacts with N<sup>ε2</sup> atom of His 272, 2.53Å (Fig.12c). TCP (occupancy of 0.6) has just entered into active-site cavity does not bind because of chloride product still bound with nitrogen atoms of Asn41 and Trp107 (canonical halide binding site). Substrate molecule must compete with the chloride ion for the halide-binding site before formation of substrate-binding in the active site (the first reaction step) (Fig.12d-e).



**Figure 11.** *DhaA31* active site in the complex with TCP (a). TCP carbon and chloride atoms are coloured yellow and green, respectively. Green spheres represent chloride anions and red spheres represent oxygen atoms of water molecules. The alternative conformation A of the catalytic residue Asp106 in the post-attack configuration and TCP (b). The alternative conformation B of the catalytic residue Asp106 with halide product living the active site (c).

The second state is characterized by active conformation of Asp106 (alternative B, occupancy of 0.4) and presence of chloride anions in active-site cavity (Fig.11c). The distance between O<sup>δ1</sup> atom of Asp106 and N<sup>ε2</sup> atom of His 272 changes to 3.14 Å, at the same time, the N<sup>δ1</sup> atom of His272 hydrogen-bonds with the side chain of Glu130 (2.75 Å). Thus, the catalytic residue Asp106 at conformation B can performs the second reaction step: transition state for nucleophilic attack. The hydrolytic water molecule (H<sub>2</sub>O-1, occupancy of 0.4) was presented in this state. With that, the chloride ions were presented in non-canonical halide-binding sites show the release pathway for the second product from the previous reaction cycle. While the chloride ion bond in canonical halide-binding site is the second product formed during the new reaction cycle. This chloride ion will be replaced by substrate molecule when the next reaction cycle will start (Fig.12e).



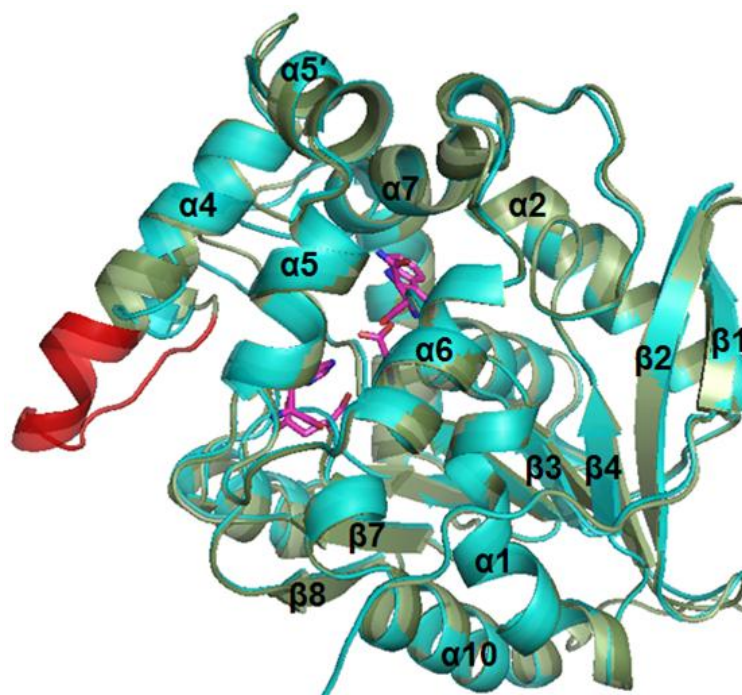
**Figure 12.** Proposed reaction mechanism of DhaA from *Rhodococcus rhodochrous* (based on Verschueren et al., 1993). The reaction schema: Substrate binding and nucleophilic attack of Asp106 (a); Alkyl-enzyme intermediate followed by nucleophilic attack of water (b); Halogen remains bound to catalytic residues Asp106 and His272, alcohol product leaves the active site and new hydrolytic water enters the active site (c); New substrate molecule (pre- or near-attack orientation) enters the active site (d) and replaces the chloride ion in the canonical halide-binding site (e).

### 2.6.2. DhaA12 mutant

DhaA12 protein variant was constructed during “enantioselectivity project” leading to characterization and improving quality of selected haloalkane dehalogenases. Enantioselectivity transformation is very important aspect in pharmacology as the working drugs substance is often one of two enantiomers, while the other one is responsible for adverse effects. The haloalkane dehalogenase DhaA shows a weak enantioselectivity with 1,2- and 1,3-dihaloalkanes, and non-enantioselectivity with  $\beta$ -bromoalkanes while another haloalkane dehalogenase DbjA shows high enantioselectivity with  $\beta$ -bromoalkanes. The active-site cavity of DbjA is the largest of all *HLDs* with known structure. The large active sites are responsible for the broad substrate specificity of enzymes (Koudelakova *et al.*, 2011). The cavity of DhaA is smaller and therefore cannot accommodate larger substrates. It was further identified that DhaA and DbjA are the structurally related to each other – 59% sequence similarity (Sykora *et al.*, 2012).

The DhaA12 is experimental mutant protein combining the nature of two parent’s proteins DhaA and DbjA. DhaA12 mutant contains several amino acid residues typical for haloalkane dehalogenase DbjA that were introduced into DhaA by an insertion of loop  $_{139}\text{HHTEVAEEQDH}_{149}$  (Fig.13) and eight additional mutations W141F, P142A, F144A, G171R, A172V, K175G, C176G, V245A.

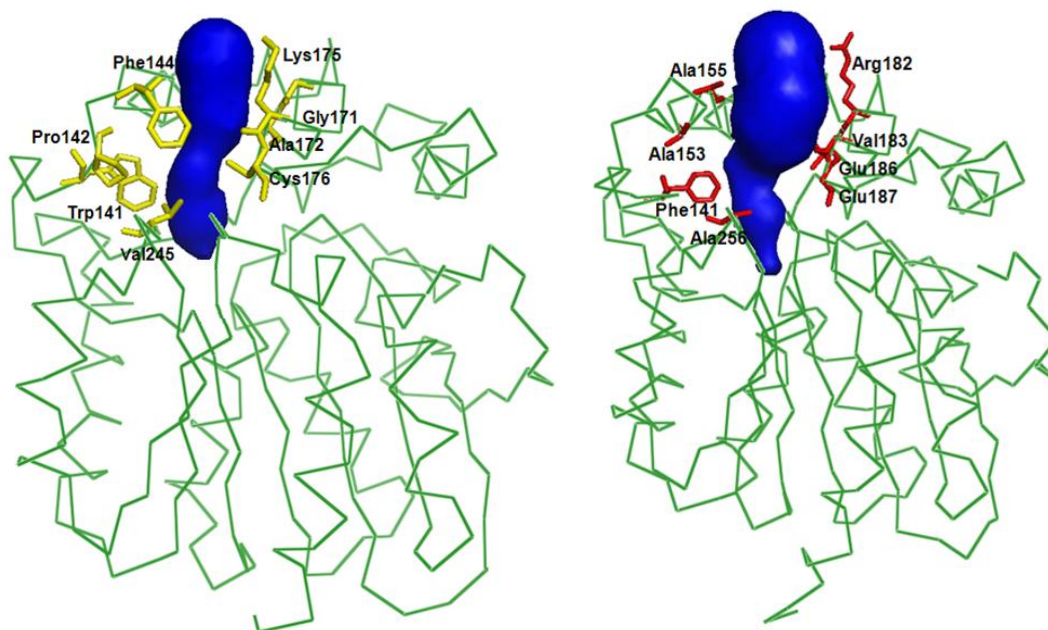
Enantioselectivity of DhaA and its cumulative mutants were tested towards 2-bromobutane, 2-bromopentane, 2-bromohexane, 2-bromoheptane and ethyl 2-bromopropionate using GC analysis. In this experiment the level of protein activity increased with the exception of the 2-bromobutane substrate and the enantioselectivity towards 2-bromopentane was not changed (Chrobakova *et al.*, 2006).



**Figure 13.** *Cartoon representation of the DhaA12 and DhaA structures alignment. The DhaA12 is shown in green. The DhaA wild type is shown in blue and insertion loop is shown in red. Catalytic pentad residues are shown in magenta stick representation.*

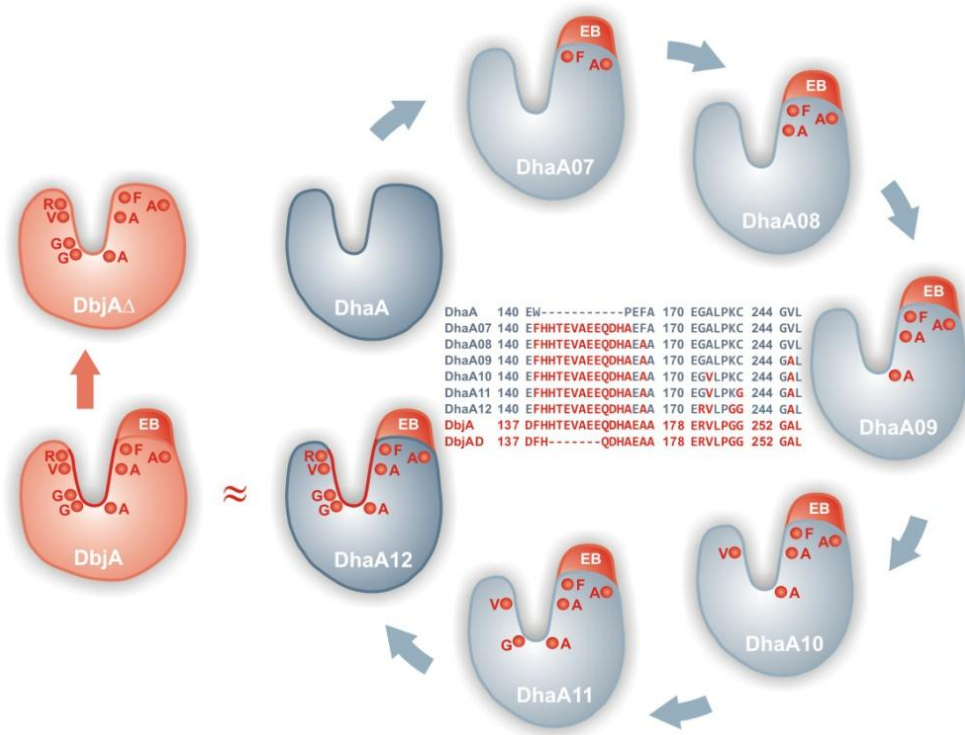
DhaA12 possesses the identical active site and the access tunnel with DbjA. The main tunnel of DhaA12 structure was mutational changed using eight point substitutions. Loop region characterizing of great flexibility (Emmer J., 2007) and forms an additional protein surface. The combined effect of the loops and mutations is possible included in the process of formation of the active-site pocket. The substitutions are changed the anatomy of the main tunnel to make it wider, expecting that the exchange of substrates in DhaA12 will increase and large substrates may have good access to the active site (Fig.14). However, enantioselectivity of DhaA12 with  $\beta$ -bromoalkanes does not reach the level of DbjA, suggesting that other factors comprising broader protein context in addition to the geometrical aspects influence the enzyme enantioselectivity. Loop region inserted to DhaA12 does not possess such

dynamics and hydration as in parent DbjA, presumably due to different protein context (Sykora *et al.*, 2013).



**Figure 14.** *Ribbon representations of the structures of wild-type DhaA (a) and DhaA12 (b) with the main tunnel shown in blue. The residues selected for mutagenesis are shown as yellow sticks. The mutated residues are shown as red sticks.*

The comparison of DhaA12 and DbjA confirms successful transplantation of the active site and the access tunnel of DbjA into DhaA in a geometrical sense (Fig.15). The active site and the access tunnel of DhaA12 match DbjA. So, engineering of DhaA12 provides a critical point of our understanding of protein structure-function relationships and future rational protein designs.



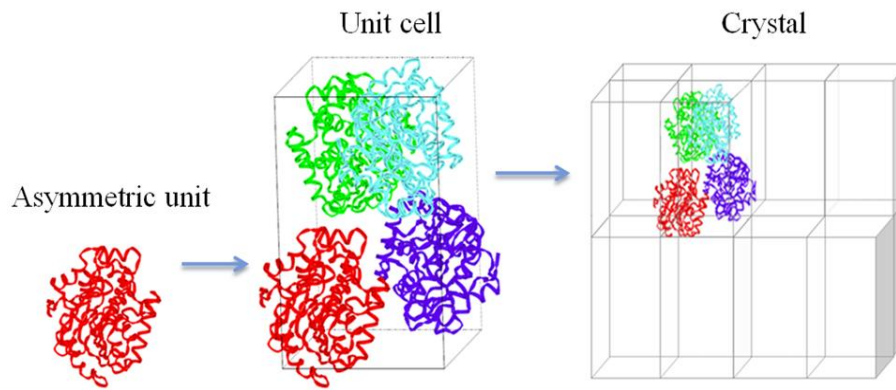
**Figure 15.** Transplantation of the active site and the access tunnel from DbjA to DhaA. The active site and the access tunnel of DbjA were transplanted to DhaA by cumulative introduction of mutations through DhaA07–12. EB fragment (loop insertion) and eight substitutions shows in red (adapted from Sykora et al., 2013).



### 3. Experimental methods

In the modern science the study of three-dimensional structure of proteins has been fundamental interest. It is important to understand precisely, which structural features of protein are responsible for its biological activity. The most common experimental method allowing this analysis is *single-crystal X-ray crystallography*. The main object in an *X-ray* diffraction experiments is *macromolecular crystal*. All of the macromolecules are polymers and they are precisely ordered forming the three-dimensional arrays of molecules inside of crystal. The crystal consists of numerous building blocks called the *asymmetric units*. An asymmetric unit is the smallest unit of crystal that contains all structural information and that can reproduce the *unit cell* by symmetry operations. The unit cell is a small box containing one or more asymmetric units that contains all structural and symmetry information (Fig.16). The most common symmetry operations for crystals of biological macromolecules are rotations, translations and screw axes (combinations of rotation and translation).

In the case the full content of unit cell is exact, we can imagine the whole crystal as an efficiently packed array of many unit cells stacked beside and on top of each other, more or less like identical boxes (Rhodes, 2006). Using the crystallographic software gives us a further opportunity to interpretate this data and to analyze the structure of the protein.



**Figure 16.** *Creation of a crystal from a fundamental asymmetric unit, illustrated by the DhaA12 mutant protein.*

### 3.1. Macromolecular crystallization

#### 3.1.1. Basic of method

Before the crystallography research is starting, obtaining of protein crystal is needful and limited condition. Crystallization process is induced in supersaturated solution that is far from equilibrium. The system of this experimental solution reaches equilibrium through the formation of a solid state presented by crystal (Fig.17). If no adequate amount of energy ever becomes available the system will remain in a metastable zone, where the nucleation state and subsequent growth of crystal cannot be observed. If precipitation occurs rapidly it may produce many small crystals (nuclei), or spherulites (an amorphous solid). The common crystallization strategy includes slow precipitation state that can produce crystal.

An ideal crystallization strategy would be to start with conditions corresponding to the labile zone of the phase diagram (Fig.18), and then, when nuclei form, shift it into the metastable zone, where growth, but not additional nucleation, can occur (Mc Pherson, 1999; Rhodes, 2006).

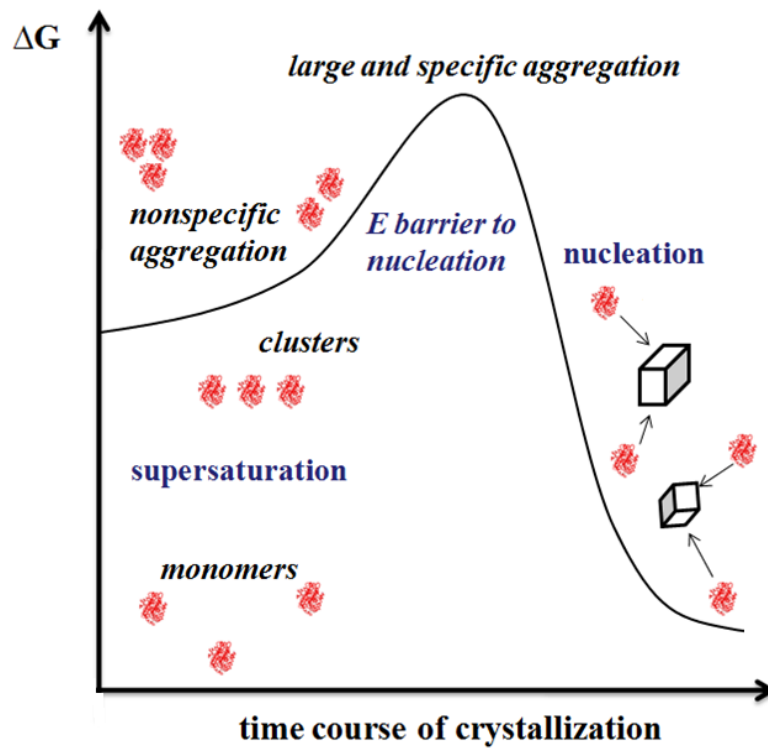


Figure 17. Schematic representation of growth and nucleation levels.

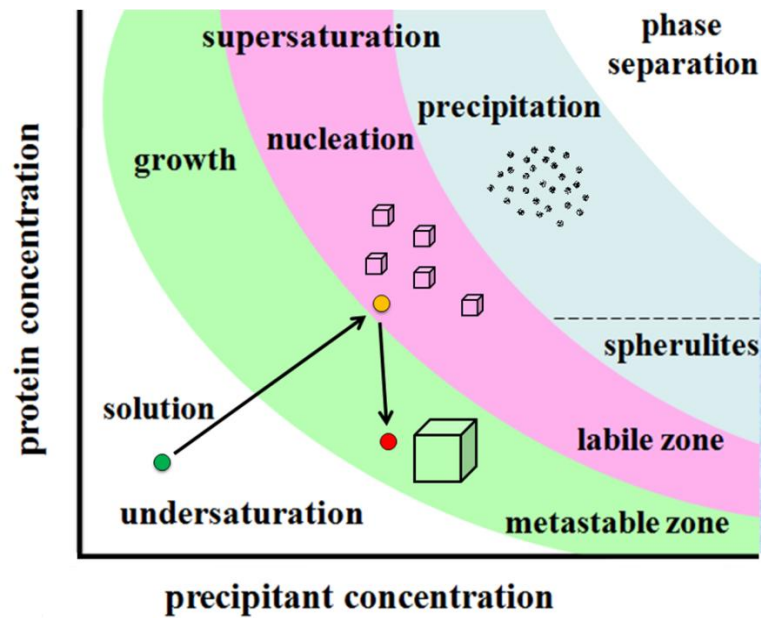


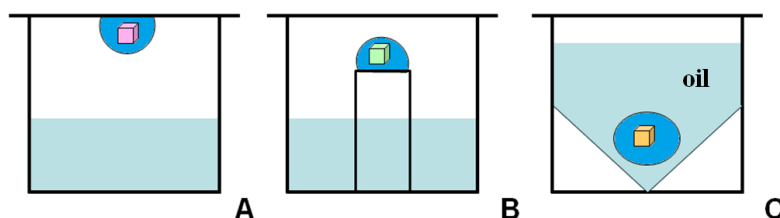
Figure 18. Phase diagram for ideal crystallization strategy.

### 3.1.2. Crystallization techniques

Crystallization techniques are subdivided into three main classes: conventional, advanced and alternative. The most *common* techniques are based on vapor diffusion. These are sitting and hanging drop, batch and microbatch under oil (Fig.19).

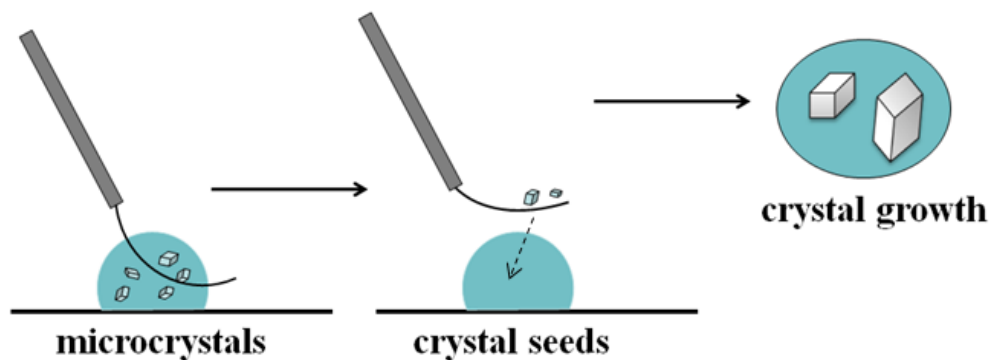
Sitting and hanging drop techniques are easy to perform using a small amount of protein sample (1-10  $\mu\text{l}$ ) and screening or optimization solution inside in the plate's reservoir (100-800  $\mu\text{l}$ ). In sitting drop technique, one places a small droplet of the protein combined with an equal or in different proportions of reservoir solution on a plate's platform and in hanging drop technique on a siliconized glass cover slide. Then crystallization plates covered by tape for sitting drop method and a siliconized glass cover slide inverted over the reservoir for hanging drop method. Over time the reservoir will pull water from the droplet in a vapor phase such that equilibrium will exist between the drop and the reservoir solution.

Microbatch under oil is a technique where a small drop of the sample is mixed with the crystallization solution and placed under a layer of paraffin or silicon oil. Using batch methodology the protein sample is simply mixed with precipitant into Petri dish or similar items.



**Figure 19.** Schematic representation of conventional crystallization techniques: (A) hanging drop, (B) sitting drop, (C) microbatch under oil.

This type of experiments can be also performed in capillaries. Crystallization in capillaries belongs to the *advanced* techniques which are used for improvement of crystal growth and optimization of crystallization conditions. Sometimes many small crystals appeared without growth during nucleation stage. In this case we can use *alternative* seeding techniques (Fig.20). In seeding the experimental principle is the same as in previous principles except that each droplet is seeded by a few small crystals. Crystal seeds start to grow increasing the size and quality of crystals (Bergfors, 1999; McPherson, 1999).



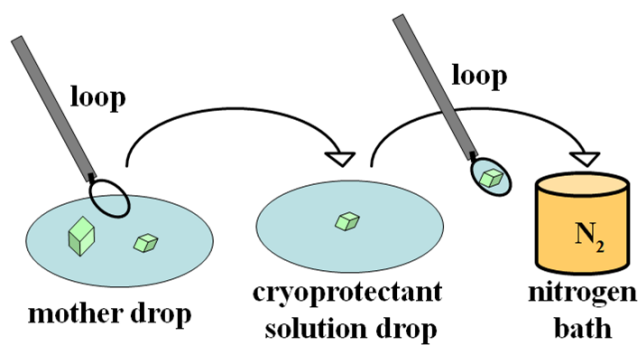
**Figure 20.** Schematic representation of streak seeding technique.

### **3.1.3. Preparation of crystals to diffraction measurements**

The crystal quality plays an important role in the results of diffraction measurements. The first question is the nature of the crystal. Different tests are used for examination of the crystals: dye binding test, dehydration test, crush test, run a gel and X-ray diffraction. The IZIT dye (Hampton Research, CA, USA) binding test is simple, just place 1  $\mu$ l of colorant in the experimental drop and wait for an hour. IZIT will fill the solvent channels in protein crystals during the test time. The protein crystals become blue in clear drop. In dehydration test the reservoir with the sample drop is simply opened. The solvent channels inside of protein crystals lose the water molecules and crystals begin to break down whereas salt crystals do not. Crush test as physical manipulation involves the destruction of crystals. Protein crystal will crush easily when touched with a needle. The most convincing proof of the crystal nature is X-ray diffraction.

The second question is – Why the protein crystal without any defects does not diffract? Many parameters influence the formation of macromolecular crystals. The most important parameters are purity of the protein, the solubility of the protein in suitable solvent and the presence of ligands. It is difficult to predict optimal conditions for obtaining suitable diffraction quality protein crystals. The specific values of a variable and components must be carefully identified for each protein. The protein in complex with ligand may be more likely to crystallize compared with the apo protein. The complex is more stable than the free native protein because the conformational change inside the protein makes it more adaptable to crystallization. Sometimes crystals are extremely fragile to X-rays. The third step in preparation of crystals to diffraction measurements is saving of the crystals. Cryoprotection of sensitive crystal includes adding cryoprotecting agent such as glycerol, sugars, ethylene glycol and polyethylene glycol into

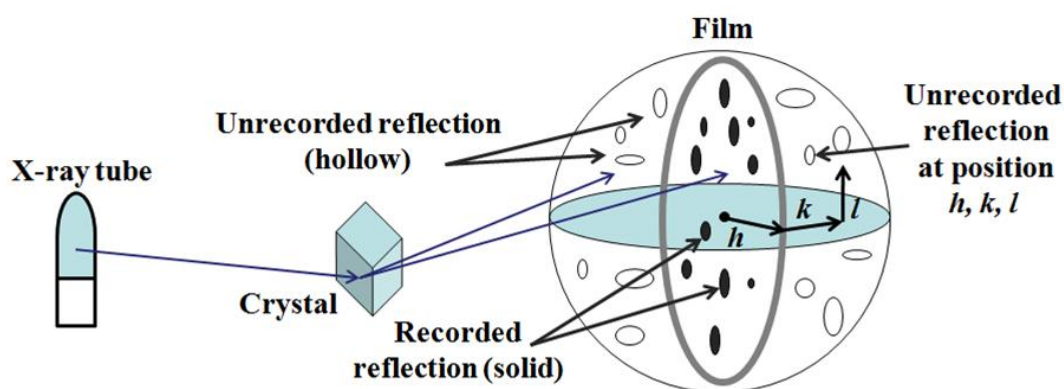
the crystal. Crystal is transferred from initial drop into final solution with some of cryoprotecting agents (Fig.21). Finally, the diffraction experiments and data collection can be achieved when all above-mentioned steps are completed.



**Figure 21.** Schematic representation the major steps involved in cryocooling macromolecular crystals (adapted from <http://www.mitegen.com>).

### 3.2. Crystallographic data collection and the Phase problem

The data collection starts from the installation of the crystal between an X-ray source and X-ray detector. A collimated beam of monochromatic X-rays is directed through the crystal and rays are scattered in all directions by the electrons of every atom in the object and a different diffraction pattern is recorded. Each diffraction pattern is recorded on a film. A hypothetical three-dimensional diffraction pattern with the reflections would be produced by all possible orientations of a crystal in the X-ray beam (Fig.22). Only one plane of the three-dimensional diffraction pattern (solid spots) is superimposed on the film. In order to measure the directions and intensities of all additional reflections (hollow spots) the crystallographer must collect diffraction patterns from all unique orientations of the crystal with respect to the X-ray beam. The main result of crystallographic data collection is a list of intensities for each point in the three-dimensional reciprocal lattice.



**Figure 22.** Crystallographic data collection, showing reflection measured at one particular crystal orientation (solid, on the film) and those that could be measured at other orientation (hollow, within the sphere but not on the film). Each reflection is located by its three-dimensional coordinates  $h$ ,  $k$  and  $l$  (adapted from Rhodes 2006).



Each reflection can be assigned three coordinates or indices in the three-dimensional reciprocal space. In the reciprocal space of the diffraction pattern usually used  $h, k, l$  coordinates. The central reflection (the round solid spot at the center of the film in Fig. 22) is taken as the origin in reciprocal space and assigned the coordinates  $(h, k, l) = (0, 0, 0)$ . The other reflections are assigned whole-number coordinates counted from this origin. The position of a reflection is related to the angle by which the diffracted beam diverges from the source beam. Another coordinates system is the unit cell (Fig.16) of the crystal (real space) where an atom's position is described by its coordinates  $x, y, z$ .

A vertex of the unit cell, or any other convenient position, is taken as the origin with coordinates  $(x, y, z) = (0, 0, 0)$ . Coordinates in real space designate real spatial position within the unit cell. Real-space coordinates are usually given in angstroms ( $\text{\AA}$ ) or nanometers (nm). As the dimensions of reciprocal space are the inverse of dimensions in the real space of the crystal, distance in reciprocal space are expressed in the reciprocal angstroms  $\text{\AA}^{-1}$ .

Each diffracted ray is a complicated wave, the sum of diffractive contributions from all atoms in the unit cell. For a unit cell containing  $n$  atoms the structural factor  $F_{hkl}$  is the sum of all the atomic  $f_{hkl}$  values for individual atoms:

$$F_{hkl} = \sum_{j=1}^n f_j e^{2\pi i(hx_j + ky_j + lz_j)} \quad [1]$$

Thus the parameters can be measured and analyze in the X-ray diffraction pattern are the position  $hkl$  and the intensity  $I_{hkl}$  of each reflection and the electron density of a volume element centered at  $(x, y, z)$ . For this reason the

Fourier transformation is used. Each reflection is characterized by its structural factors. It has amplitude and a phase. The measurable intensity  $I_{hkl}$  of one reflection gives the amplitude  $|F(hkl)|$  but the phase  $\alpha$  of each reflection is not recorded on any kind of detector.

So, the electron density is transformed by Fourier transformation of the structural factors, as follows:

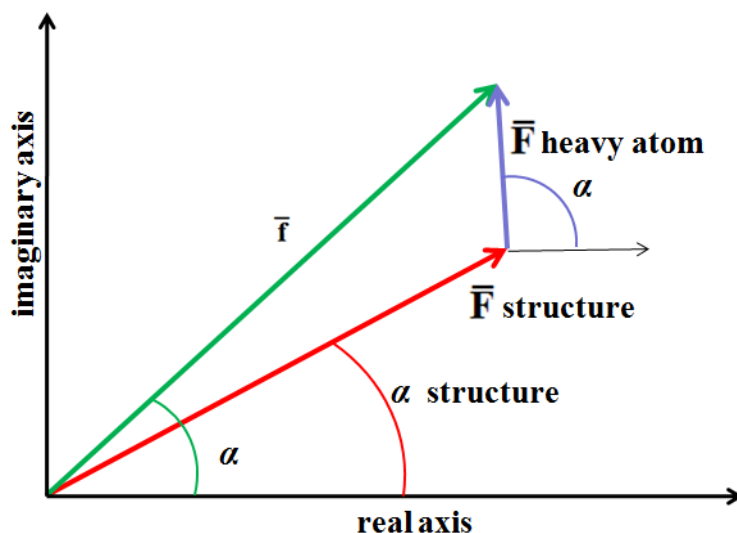
$$\rho(x, y, z) = \frac{1}{V} \sum_{h,k,l} |F(hkl)| e^{-2\pi i(hx+ky+lz)} \quad [2]$$

The phase is unknown and it is only additional information the crystallographer needs in order to compute  $\rho(x,y,z)$  and thus obtain an image of the unit cell contents. This problem is known as the “phase problem” in crystallography. There are a number of methods for overcoming this problem. These include the *isomorphous replacement*, *anomalous dispersion* and *molecular replacement* (Wilmanns & Weiss, 2005; Rhodes, 2006).

### 3.2.1. Isomorphous replacement

Isomorphous replacement is also called heavy-atom method. In this method diffraction data sets of two crystals are used. The first one is crystal of the native macromolecule. The second one is the derivative crystal of the protein soaked in solutions of heavy ions in order to see perturbation against the signal of hundreds or thousands of smaller atoms. This crystal must be isomorphous with native crystals, i.e. the heavy atom must not disturb crystal packing or the conformation of the protein. The substructure can be determined from differences between these two diffraction data sets. The changes must be clearly detectable and large enough to measure accurately. The structural factor amplitude and the phase of the substructure can be calculated. Then the phase of the reflection of the macromolecule crystal can

be calculated using the heavy-atom phase. The structure factor  $\bar{F}_{hkl}$  for any  $hkl$  can be represented as a vector composed of a scattering contributions  $\bar{f}(\mathbf{j})_{hkl}$  for every atom  $\mathbf{j}$  in the unit cell where the individual  $\bar{f}(\mathbf{j})_{hkl}$  are also vectors (Fig. 23).



**Figure 23.** A structure factor  $\bar{f}$  for the heavy-atom derivative is the sum of contributions from the native structure  $\bar{F}$  and  $\bar{F}$  of the heavy atom. The structure factor of the protein whose phase we are seeking is red. The structure factor of the heavy atom is blue. The structural factor of the heavy atom derivative of the protein is green (adapted from McPherson, 2003).

Once the position of the heavy atom is known the phase  $\alpha_{hkl}$  of the heavy atom can be computed using only the heavy atom contribution. Then this can be used as an approximation to the phase angle for the entire structure  $\alpha_{hkl}$  in a Fourier synthesis  $\rho(x,y,z)$ . This synthesis of the electron density will be directly measured (from the diffraction pattern)  $|\mathbf{F}(hkl)|$  and the phase  $\alpha_{hkl}$  calculated from the heavy atom. In succeeding Fourier maps as the phase  $\alpha_{hkl}$  of the heavy atom more nearly equals the correct phase  $\alpha_{hkl}$ . They are included to produce better phase approximations unit. So, the coordinates of

all the atoms in the unit cell are known and a model of the structure can be created (McPherson, 2003).

### **3.2.2. Anomalous dispersion**

In this method a diffraction experiment can be carried out at two (or more) different X-ray energies. An element exhibits anomalous scattering when the X-ray wavelength is near the element's absorption edge. Absorption edges for the light atoms in the unit cell are not near the wavelength of X-ray used in crystallography: carbon, nitrogen and oxygen do not contribute to anomalous scattering. However, absorption edges of heavy atoms are in this range and X-rays of varying wavelength are available. X-ray data can be collected under conditions that maximize anomalous scattering by the heavy atom. So, under anomalous scattering can be used atoms of protein containing selenomethionine instead of the normal sulfur-containing methionine, atoms that occur naturally in the macromolecules and heavy atoms introduced into the crystal by soaking or co-crystallization.

The absorption edge is usually done before data collection acquisition by a fluorescence scan. Several (multiwavelength anomalous dispersion) or only one wavelength (single anomalous dispersion) at the absorption peak experiment data are collected to use the difference in Friedel's pair to establish phase. Then the establishing phases can be used to find the native phases (McPherson, 2003; Rhodes, 2006).

### **3.2.3. Molecular replacement**

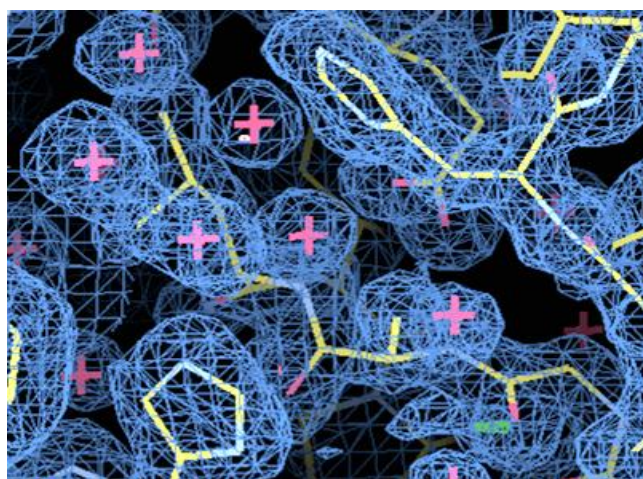
Sometimes crystallographer can use the phases from structural factors of a known protein as initial estimates of phase for a new protein. The known protein is referred as a phasing model. The known structure is placed in the unit cell of the new protein structure and the structure factor  $\bar{F}$  is calculated.

The phase  $\alpha_{hkl}$  of the structure factor amplitude  $|\mathbf{F}(hkl)|$  will be approximately corrected and can be used as a starting point for further refinement. This method is called molecular replacement.

The first we need to superimpose the structure of the model on the structure of the new protein in its unit cell and determine the correct orientation and translation of the search model. Then the structural factors can be calculated of the properly positioned model and use the phases of these computed structure factors as initial estimates of the required phases (McPherson, 2003; Rhodes, 2006).

#### 3.2.4. Obtaining the molecular model and structure refinement

Obtaining a detailed molecular model of the unit cell contents entails calculating  $\rho(x,y,z)$  using amplitudes  $|\mathbf{F}(hkl)|$  computed from measured intensities  $I_{hkl}$  in the native data set and phases  $\alpha_{hkl}$  computed from heavy-atom data, anomalous dispersion, or molecular replacement method. The structural factors are computed in this model using Eq. [1]. As the model is constructed, the viewer sees the current atom locations in the form of three-dimensional coordinates (Fig.24).



**Figure 24.** *The electron-density map of the DhaA12 mutant protein*

Once the model has been obtained the mathematical optimization procedure called refinement can be started. The progress of refinement is typically monitored by looking at the *R-factor*:

$$R = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|} \quad [3]$$

The *R-factor* parameter is the correspondence between the expected structure-factor amplitudes from the model in a given trial location and the actual amplitudes derived from the native data on the studied protein. A more revealing criterion of model quality and of improvements during refinement is the so-called *free R-factor*.  $R_{free}$  is computed with a small set of randomly chosen intensities (the “test set”), which are set aside from the beginning and not used during refinement. At any stage in refinement  $R_{free}$  measures how well the current atomic model predicts a subset of the measured intensities that were not included in the refinement. Whereas *R* measures how well the current model predicts the entire data set that produced model (Rhodes, 2006).

## References

Bergfors, T. M. (1999). *Protein crystallization: Techniques, strategies and tips*.

International University Line, La Jolla, USA.

Bidmanova, S., Chaloupkova, R., Damborsky, J., Prokop, Z. (2010). *Anal. Bioanal. Chem.* In press. DOI: 10.1007/s00216-010-4083-z.

Bosma, T., Damborsky, J., Stucki, G. & Janssen, D. B. (2002). *Appl. Env. Microbiol.* **68**, 3582-3587.

Bosma, T., Kruizinga, E., de Bruin, E.J., Poelarends, G.J. & Janssen, D.B. (1999). *Appl. Environ. Microbiol.* **65**, 4575–4581.

Bosma, T., Pikkemaat, M., G., Kingma, J., Dijk, J. & Janssen, D., B. (2003). *Biochemistry* **42**, 8047–8053.

Campbell, D. W., Muller, C. & Reardon, K. F. (2006). *Biotechnol. Lett.* **28**, 883–887.

Chovancova, E., Kosinski, J., Bujnicki, J. M. & Damborsky, J. (2007). *Proteins* **67**, 305–316.

Chrobakova, T., Prokop, Z., Grodecka, L. (2006). Prog. Rep.: Protein Engineering Group.

Damborsky, J., Koca, J. (1999). *Protein Eng. Des. Sel.* **12**, 989–998.

Emmer J. (2007). Master's thesis. Structural studies of the haloalkane dehalogenase mutant (DhaA12) from *Rhodococcus rhodochrous*.

Gray, K.A. Richardson, T. H., Kretz, K., Short, J. M., Bartnek, F., Knowles, R., Kan, L., Swanson, P. E., Robertson, D. E. (2001). *Adv. Synth. Catal.* **343**, 607–617.

Janssen, D. B. (2004). *Cur. Opin. Chem. Biol.* **8**, 150–159.

- Janssen, D. B., Dinkla, I. J. T., Poelarends, G. J., Terpstra, P. (2005). *Environ. Microbiol* **7**, 1868–1882.
- Janssen, D. B., Pries, F., Van der Ploeg, J. R. (1994). *Annu. Rev. Microbiol.* **48**, 163–191.
- Janssen, D. B., Schanstra, J. P. (1994). *Cur. Opin. Biotechnol.* **5**, 253–259.
- Klvana, M., Pavlova, M., Koudelakova, T., Chaloupkova, R., Dvorak, P., Prokop, Z., Stsiapanava, A., Kutý, M., Kuta-Smatanova, I., Dohnalek, J., Kulhanek, P., Wade, R.C., Damborsky, J. (2009). *J. Mol. Biol.* **392**, 1339–1356.
- Koudelakova, T., Chovancova, E., Brezovsky, J., Monincova, M., Fortova, A., Jarkovsky, J., Damborsky, J. (2011). *Biochemistry* **435**, 345–354.
- Kulakova, A. N., Larkin, M. J., Kulakov, L. A. (1997). *Microbiology* **143**, 109–115.
- Lahoda, M., Chaloupkova, R., Stsiapanava, A., Damborsky, J. and Kuta-Smatanova, I. (2011). *Acta Cryst.*, **F67**, 397-400
- Mena-Benitez, G.L., Gandia-Herrero, F., Graham, S., Larson, R.T., McQueen-Mason, S.J., French, C.E., Rylott, E.L., Bruce, C.N. (2008) *Plant Physiology*, Vol. 147, 1192–1198.
- McPherson, A. (1999). *Crystallization of biological macromolecules*. Cold Spring Harbor Laboratory Press, New York, 586 pp.
- McPherson, A. (2002). *Introduction to macromolecular crystallography*, John Wiley & Sons, Inc., 248 pp.
- Naested, H., Fennema, M., Hao, L., Andersen, M., Janssen, D.B., Mundy, J. (1999) *Plant J.* **18**, 571-576.
- Nagata, Y., Miyauchi, K., Damborsky, J., Manova, K., Ansorgova, A., Takagi, M. (1997). *Appl Environ Microbiol*, **63**, 3707-3710.
- Nardini, M., Dijsktra, B. W. (1999). *Curr. Opin. Struct. Biol.* **9**, 732–737.



- Newman, J., Peat, T. S., Richard, R., Kan, L., Swanson, P. E., Affholter, J.A., Holmes, I. H., Schindler, J. F., Unkefer C. J., Terwilliger, T. C. (1999). *Biochemistry* **38**, 16105–16114.
- Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., Schrag, J., Sussman, J. L., Verschueren, K. H.G., Goldman, A. (1992). *Protein Eng.* **5**, 197–211.
- Otyepka, M., Banas, P., Magistrato, A., Carloni, P., Damborsky, J. (2007). *Proteins* **70**, 707–717.
- Otyepka, M., Damborsky, J. (2002). *Protein Sci.* **11**, 1206–1217.
- Pavlova, M., Klvana, M., Prokop, Z., Chaloupkova, R., Banas, P., Otyepka, M., Wade, R. C., Tsuda, M., Nagata, Y., Damborsky J. (2009). *Nature Chem. Biol.* **5**, 727–733.
- Petrek, M., Otyepka, M., Banas, P., Kosinova, P., Koca, J. & Damborsky, J. (2006). *BMC Bioinformatics* **7**, 316pp.
- Poelarends, G. J., Van Hylckama Vlieg, J. E. T., Marchesi, J. R., Freitas Dos Santos, L. M., Janssen, D.B. (1999). *J. Bacteriol.* **181**, 2050-2058.
- Poelarends, G. J., Zandstra, M., Bosma, T., Kulakov, L.A., Larkin, M.J., Marchesi, J.R., Weightman, A.J., Janssen, D. B. (2000). *J. Bacteriol.* **182**, 2725-2731.
- Pries, F., Van der Ploeg, J. R., Dolfing, J., Janssen, D. B. (1994). *FEMS Microbiol. Rev.* **15**, 277–295.
- Prokop, Z., Damborsky, J., Oplustil, F., Jesenska, A., Nagata, Y. (2005). Patent WO 2006/128390 A1.
- Prokop, Z., Oplustil, F., DeFrank, J., Damborsky, J. (2006). *Biotech. J.* **1**, 1370–1380.
- Prokop, Z., Sato, Y., Brezovsky, J., Mozga, T., Chaloupkova, R., Koudelakova, T., Jerabek, P., Stepankova, V., Natsume, R., Leeuwen, J. G. E., Janssen, D. B., Florian, J., Nagata, Y., Senda, T., Damborsky, J. (2010). *Angew. Chem. Int. Ed.* **49**, 6111–6115.

Rhodes, G. (2006). *Crystallography made crystal clear*. 3rd ed., Academic Press., 352 pp.

Sato, Y., Monincova, M., Chaloupkova, R., Prokop, Z., Ohtsubo, Y., Minamisawa, K., Tsuda, M., Damborsky, J. (2005). *Applied and Environ. Microbiology*. **71**, 4372-4379.

Schindler, J.F., Naranjo, P.A., Honaberger, D.A., Chang, C.-H., Brainard, J.R., Vanderberd, L.A., Unkefer, C.J. (1999). *Biochemistry*. **38**, 5772-5778.

Stsiapanava, A., Dohnalek, J., Gavira, J., A., Kutý, M., Koudelakova, T., Damborsky, J., Smatanova, I., K. (2010). *Acta Cryst. D***66**, 962–969

Sykora, J., Brezovsky, J., Koudelakova, T., Lahoda, M., Chaloupkova, R., Chernavets, T., Fortova, A., Stepankova, V., Prokop, Z., Kuta-Smatanova, I., Hof, M., Damborsky, J. (2013) (in the press)

Stucki, G., Thuer, M. (1995). *Environ. Sci. Technol.* **29**, 2339–2345.

Swanson, P. E. (1999). *Cur. Opin. Biotechnol.* **10**, 365–369.

Verschueren, K.H., Kingma, J., Rozeboom, H.J., Kalk, K.H., Janssen, D.B., Dijkstra, B.W. (1993a). *J. Mol. Biol.* **232**, 856-72.

Verschueren, K.H., Kingma, J., Rozeboom, H.J., Kalk, K.H., Janssen, D.B., Dijkstra, B.W. (1993b). *Biochemistry*.**32**, 9031-9037.

Verschueren, K. H. G., Seljee, F., Rozeboom, H. J., Kalk, K. H. & Dijkstra, W. (1993c). *Nature* **363**, 693–698.

Wilmanns, M. & Weiss, M. S. (2005). *Molecular Crystallography. Encyclopedia of Condensed Matter Physics*, Academic Press, 453–458.

<http://www.pdb.org>

<http://www.mitegen.com>

## 4. Summary of papers

### Paper I

**Crystallization and crystallographic analysis of the *Rhodococcus rhodochrous* NCIMB 13064 DhaA mutant DhaA31 and its complex with 1,2,3-trichloropropane**

Reproduced with permission from

Lahoda, M., Chaloupkova, R., Stsiapanava, A., Damborsky, J. and Kuta Smananova, I., *Acta Crys.*, **F67**, 397-400 (2011)

Maryna Lahoda,<sup>a</sup> Radka  
Chaloupkova,<sup>b</sup> Alena  
Stsiapanava,<sup>a</sup> Jiri Damborsky<sup>b</sup>  
and Ivana Kuta Smatanova<sup>a,c\*</sup>

<sup>a</sup>Institute of Physical Biology, University of  
South Bohemia, Ceske Budejovice, Zamek  
136,

373 33 Nove Hradky, Czech Republic,

<sup>b</sup>Loschmidt Laboratories, Department of  
Experimental Biology and Research Centre for  
Toxic Compounds in the Environment,  
Faculty

of Science, Masaryk University, Kamenice  
5/A4, 625 00 Brno, Czech Republic, and

<sup>c</sup>Institute of Systems Biology and Ecology,  
Academy of Science of the Czech Republic,  
Zamek 136, 373 33 Nove Hradky, Czech  
Republic

Correspondence e-mail: ivanaks@seznam.cz

Received 24 November 2010

Accepted 10 January 2011

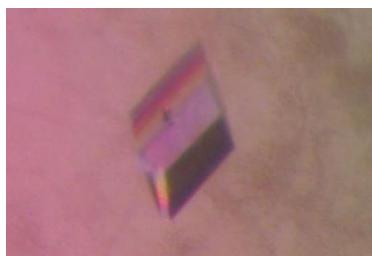
## Crystallization and crystallographic analysis of the *Rhodococcus* rhodochrous NCIMB 13064 DhaA mutant DhaA31 and its complex with 1,2,3-trichloropropane

Haloalkane dehalogenases hydrolyze carbon–halogen bonds in a wide range of halogenated aliphatic compounds. The potential use of haloalkane dehalogenases in bioremediation applications has stimulated intensive investigation of these enzymes and their engineering. The mutant DhaA31 was constructed to degrade the anthropogenic compound 1,2,3-trichloropropane (TCP) using a new strategy. This strategy enhances activity towards TCP by decreasing the accessibility of the active site to water molecules, thereby promoting formation of the activated complex. The structure of DhaA31 will help in understanding the structure–function relationships involved in the improved dehalogenation of TCP. The mutant protein DhaA31 was crystallized by the sitting-drop vapour diffusion technique and crystals of DhaA31 in complex with TCP were obtained using soaking experiments. Both crystals belonged to the triclinic space group P1. Diffraction data were collected to high resolution: to 1.31 Å for DhaA31 and to 1.26 Å for DhaA31 complexed with TCP.

### 1. Introduction

The haloalkane dehalogenase DhaA was isolated from the Gram-positive bacterium *Rhodococcus rhodochrous* NCIMB 13064 (Kulakova *et al.*, 1997). The enzyme catalyzes the hydrolytic dehalogenation of various halogenated aliphatic hydrocarbons. The structures of several DhaA mutant variants have been published (Newman *et al.*, 1999; Klvana *et al.*, 2009; Stsiapanava *et al.*, 2010). The main  $\alpha/\beta$ -hydrolase domain is conserved in the members of the  $\alpha/\beta$ -hydrolase superfamily and serves as a scaffold for the catalytic residues. The second helical cap domain is structurally more variable and is known to influence the substrate specificity of these enzymes (Chovancova *et al.*, 2007). The active-site cavity is located between two protein domains and is connected to the protein surface by two major access tunnels: the main tunnel and the slot tunnel used by the ligands for exchange between the active site and the surrounding solvent. A catalytic pentad of DhaA which performs hydrolysis has been identified. It includes a nucleophile (Asp106), a base (His272), a catalytic acid (Glu130) and two halide-stabilizing residues (Asn41 and Trp107). The dehalogenation reaction proceeds in the active site via nucleophilic attack of a nucleophile on a C atom carrying a halogen on the substrates. This leads to cleavage of the carbon–halogen bond, displacement of the halide and the formation of a covalent alkyl-enzyme intermediate. The alkyl-enzyme intermediate is subsequently hydrolyzed by a water molecule activated by a catalytic base (Pavlova *et al.*, 2009).

Some haloalkane dehalogenases are involved in the biodegradation of important environmental pollutants. The enzyme DhaA is of particular interest because it enables a very slow conversion of the toxic artificial compound and suspected human carcinogen 1,2,3-trichloropropane (TCP) to 2,3-dichloropropane-1-ol under laboratory conditions (Schindler *et al.*, 1999; Bosma *et al.*, 1999). To increase the efficiency of this reaction, focused directed evolution was used to construct the mutant DhaA31, which has up to 32-fold higher catalytic activity and 26-fold higher catalytic efficiency than the parent wild-type enzyme (Pavlova *et al.*, 2009).



## crystallization communications

The mutant DhaA31 has  $k_{\text{cat}}$  value equal  $1.26 \text{ s}^{-1}$  and  $K_{\text{m}}$  value equal  $1.2 \text{ mM}$ , while the wild type has  $k_{\text{cat}}$  value equal  $0.04 \text{ s}^{-1}$  and  $K_{\text{m}}$  value equal  $1.0 \text{ mM}$ . The positions for mutagenesis were selected by computer modelling and were randomized experimentally. DhaA31 has large aromatic residues at two of the three randomized positions and at two positions modified by site-directed mutagenesis. These changes enhance activity towards

TCP by decreasing the accessibility of the active site to water molecules, thus promoting the formation of an activated complex. Kinetic analyses confirmed that the mutations improved carbon–halogen bond cleavage and shifted the rate-limiting step to the release of products (Pavlova *et al.*, 2009). Comparison of the wild type DhaA and the mutant DhaA31 structures both in the free form and in complex with TCP could provide valuable structural information on the molecular basis of the enhanced catalysis.

## 2. Materials and methods

### 2.1. Construction of the mutant, protein expression and purification

The recombinant gene *dhaA31* carrying I135F, C176Y, V245F, L246I and Y273F mutations was obtained by focused directed evolution. The methodology has been described previously by Pavlova *et al.* (2009). The *dhaA31* gene was carried by a pAQN vector downstream of the *tac* promoter (P *tac*), which was under the control of *lacI<sup>f</sup>*. *Escherichia coli* BL21 cells with the plasmid were cultured in 4 l Luria–Bertani medium at 310 K. The induction of enzyme expression at 303 K was initiated by the addition of isopropyl  $\beta$ -D-1-thiogalactopyranoside to a final concentration of 0.5 mM. The cells were harvested and disrupted by sonication using a Soniprep 150 (Sanyo Gallenkamp PLC, Loughborough, England). The supernatant was used after centrifugation at 21 000g for 1 h. The enzyme was purified by immobilized metal-affinity chromatography as described previously (Pavlova *et al.*, 2009). The enzyme was bound to the resin in equilibration buffer (20 mM potassium phosphate buffer pH 7.5, 0.5 M sodium chloride and 10 mM imidazole). Unbound and nonspecifically bound proteins were washed out with buffer containing 37.5 mM imidazole. The target enzyme was eluted using buffer containing 300 mM imidazole. The active fractions were pooled and dialyzed overnight against 100 mM Tris–HCl buffer pH 7.5 without additives. The DhaA31 enzyme was stored in the same buffer. The entire purification process and storage took place at 277 K.

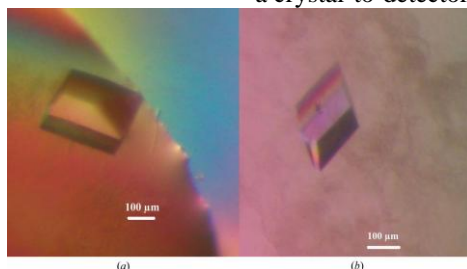
The active fractions were pooled and dialyzed overnight against 100 mM Tris–HCl buffer pH 7.5 without additives. Protein concentrations were determined by the Bradford method (Bradford, 1976). SDS–PAGE was run on 15% polyacrylamide gels to check the purity of the protein.

### 2.2. Crystallization

Freshly isolated and purified DhaA31 enzyme in 100 mM Tris–HCl buffer pH 7.5 was used in crystallization experiments. Initial crystallization experiments were carried out by the sitting-drop vapour-diffusion method (Ducruix & Giege, 1999) using Combi-Clover crystallization plates (EBS plates; Emerald BioSystems, Bainbridge Island, USA) at room temperature and subsequently at the lower temperature of 277 K. Protein microcrystals were obtained using a couple of precipitants from the commercial crystallization kit Crystal Screen (Hampton Research, Aliso Viejo, California, USA) and PEG-based conditions from the JBScreen Classic kit (Jena Bioscience GmbH, Jena, Germany). Three-dimensional crystals of the mutant protein DhaA31 were obtained at 277 K during a four month optimization procedure. The droplets, which consisted of a 2–4 ml total volume with precipitant: protein ratios of 1:1 (drop volume 2 ml), 1:2 (3 ml) and 1:3 (4 ml), were equilibrated over 800 ml reservoir solution. DhaA31 microcrystals were grown in 2 ml droplets consisting of equal volumes of protein solution at a concentration of  $7 \text{ mg ml}^{-1}$  in 50 mM Tris–HCl buffer pH 7.5 and reservoir solution consisting of 29% (w/v) PEG 4000 and 100 mM MES sodium salt pH 6.4.

### 2.3. Data collection and processing

Diffraction data for DhaA31 and for DhaA31 complexed with TCP were collected on beamline X12 at EMBL Hamburg Outstation (Germany) at a fixed wavelength of  $1.033 \text{ \AA}$  using a MAR CCD 225 mm detector. Prior to data collection, several crystals of DhaA31 were soaked with TCP (Fig. 1b). In this soaking experiment, 50  $\mu\text{l}$  TCP was added to 800  $\mu\text{l}$  reservoir solution and left for 6 h at room temperature. Crystals were mounted in nylon loops (Hampton Research, Aliso Viejo, USA; Teng, 1990) directly from the test drop and were flash-frozen in a stream of nitrogen gas at 100 K without additional cryoprotection. The diffraction data for DhaA31 were collected in two steps. 120 low-resolution images were recorded with an oscillation angle of  $3^\circ$  and a crystal-to-detector distance of 300 mm.



**Figure 1** Crystals of the mutant haloalkane dehalogenase DhaA from *R. rhodochrous* NCIMB 13064: (a) DhaA31 and (b) DhaA31 complexed with TCP.

**Table 1**

Data-collection statistics for crystals of DhaA31 and DhaA31 with TCP. Values in parentheses are for the highest resolution shell.

	DhaA31	DhaA31 with TCP
Beamline	DESY X12	DESY X12
Detector	MAR CCD 225 mm	MAR CCD 225 mm
Wavelength (Å)	1.033	1.033
Resolution range (Å)	50–1.31 (1.35–1.31)	50–1.26 (1.27–1.26)
Space group	<i>P</i> 1	<i>P</i> 1
Unit-cell parameters (Å, °)	$a = 42.55, b = 44.37, c = 46.41$ $\alpha = 115.3, \beta = 98.5, \gamma = 109.5$	$a = 42.49, b = 44.39, c = 46.53$ $\alpha = 115.3, \beta = 97.7, \gamma = 109.5$
Measured reflections	580117	537263
Unique reflections	297181	162936
Redundancy	4.3	2.4
Completeness (%)	94.6 (90.3)	92.4 (82.7)
$R_{\text{merge}}^{\dagger}$ (%)	4.4 (8.1)	3.3 (6.1)
$I/\sigma(I)$	35.6 (18.3)	41.1(17.6)

$R_{\text{merge}}^{\dagger} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ , where  $I_i(hkl)$  is the intensity of the  $i$ th measurement of reflection  $hkl$  and  $\langle I(hkl) \rangle$  is the average intensity of the reflection

A high-resolution set of 360 images (50–1.31 Å) was collected with an oscillation angle of 1° and a crystal-to-detector distance of 100 mm. Similarly, diffraction data for DhaA31 in complex with TCP were collected in two steps. A set of 400 high-resolution images (50–1.26 Å) was recorded with a 0.5° oscillation angle and a crystal-to-detector distance of 100 mm and this was followed by the collection of 130 low-resolution image frames with a 1.5° oscillation angle and a crystal-to-detector distance of 300 mm. Data for both protein crystals were integrated and scaled using the HKL-2000 program package (Otwinowski & Minor, 1997).

#### 2.4. Preliminary structure solution

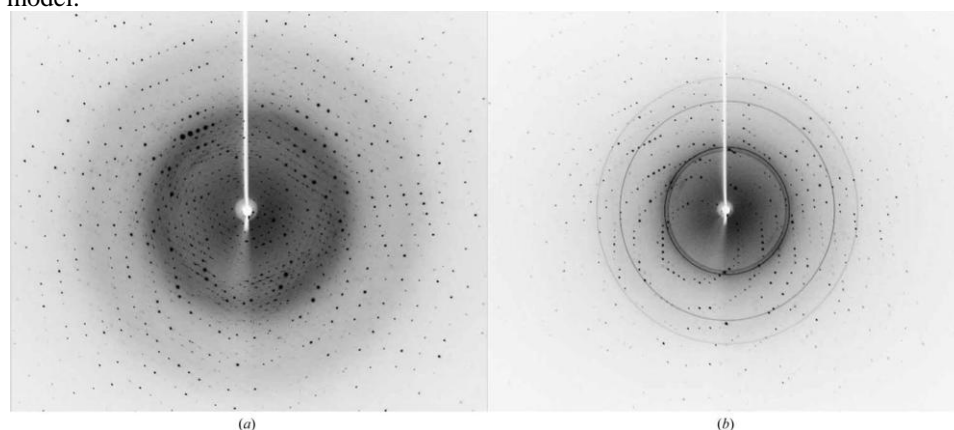
The structures of DhaA31 and of its complex with TCP were solved by the molecular-replacement method in MOLREP (Vagin & Teplyakov, 2010) using the structure of haloalkane dehalogenase DhaA from *R. rhodochrous* (PDB code 3fbw; Stsiapanava *et al.*, 2010) as the search model.

Rigid-body refinement was performed using the REFMAC5 program from the CCP4 program suite (Murshudov *et al.*, 1997).  $R$  and  $R_{\text{free}}$  were 0.234 and 0.257, respectively, for DhaA31 and 0.251 and 0.279, respectively, for DhaA31 complexed with TCP.

### 3. Results and discussion

The initial crystallization conditions for mutant haloalkane dehalogenase DhaA31 were identified using solution C3 of Crystal Screen (Hampton Research, Aliso Viejo, USA), which consisted of 0.1 *M* Na HEPES pH 7.5, 2% (v/v) PEG 400, 2.0 *M* ammonium sulfate, and solution D4 of the JBScreen Classic kit (Jena Bioscience GmbH, Jena, Germany), composed of 30% (w/v) PEG 4000, 100 *mM* sodium acetate pH 4.6, 200 *mM* ammonium sulfate. Using these conditions, DhaA31 protein microcrystals were grown in one week. During the optimization procedure, the initial crystallization conditions were slightly modified and resulted in the use of a final precipitant solution consisting of 29% (w/v) PEG 4000 and 100 *mM* MES sodium salt pH 6.4. Three-dimensional crystals of DhaA31 were grown at 277 K and reached final dimensions of 0.28 x 0.19 x 0.13 mm (Fig. 1a) in three weeks.

Three-week-old single crystals of DhaA31 were used for soaking experiments followed by diffraction data collection at a synchrotron radiation source. Single crystals of DhaA31 and of DhaA31 complexed with TCP were used for X-ray diffraction measurements. Data sets for DhaA31 and DhaA31 complexed with TCP were collected to maximum resolutions of 1.31 Å (Fig. 2a) and 1.26 Å (Fig. 2b), respectively. The crystals belonged to the triclinic space group *P*1, with unit-cell parameters  $a = 42.55, b = 44.37, c = 46.41$  Å,  $\alpha = 115.3, \beta = 98.5, \gamma = 109.5^\circ$  for DhaA31 and  $a = 42.49, b = 44.39, c = 46.53$  Å,  $\alpha = 115.3, \beta = 97.7, \gamma = 109.5^\circ$  for DhaA31 complexed with TCP. The asymmetric units of the crystals of DhaA31 and DhaA31 complexed with TCP contained one monomer, with solvent contents of 40.0 and 40.6%, respectively. The calculated Matthews coefficient (VM; Matthews, 1968) for one monomer of DhaA31 in the asymmetric unit was 2.05 Å<sup>3</sup> Da<sup>-1</sup> and that for DhaA31 in complex with TCP was 2.07 Å<sup>3</sup>.



**Figure 2** Diffraction images of crystals of the mutant haloalkane dehalogenase DhaA from *R. rhodochrous* NCIMB 13064: (a) DhaA31, (b) DhaA31 complexed with TCP. The crystals diffracted to resolutions of 1.31 and 1.26 Å for DhaA31 and DhaA31 complexed with TCP, respectively.

Data-collection statistics for DhaA31 and its complex with TCP are presented in Table 1. The structures are currently being refined and their interpretation will follow.

This work was supported by the Ministry of Education of the Czech Republic (LC06010, MSM6007665808, MSM0021622412 and CZ.1.05/2.1.00/01.0001) and by the Academy of Sciences of the Czech Republic (AV0Z60870520 and IAA401630901). We thank Paul Tucker for assistance with data collection and the EMBL for access to the X12 beamline at the DORIS storage ring of DESY in Hamburg. The research work of ML was supported by the University of South Bohemia, grant GAJU 170/2010/P.

## References

- Bosma, T., Kruizinga, E., De Bruin, E. J., Poelarends, G. J. & Janssen, D. B. (1999). *Appl. Environ. Microbiol.* **65**, 4575–4581.
- Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
- Chovancova, E., Kosinski, J., Bujnicki, J. M. & Damborsky, J. (2007). *Proteins*, **67**, 305–316.
- Ducruix, A. & Giege, R. (1999). *Protein Crystallography: A Practical Approach*, 2nd ed. Oxford: IRL Press
- Klvana, M., Pavlova, M., Koudelakova, T., Chaloupkova, R., Dvorak, P., Prokop, Z., Stsiapanava, A., Kutý, M., Kuta-Smatanova, I., Dohnalek, J., Kulhanek, P., Wade, R. C. & Damborsky, J. (2009). *J. Mol. Biol.* **392**, 1339–1356.
- Kulakova, A. N., Larkin, M. J. & Kulakov, L. A. (1997). *Microbiology*, **143**, 109–115.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst. D53*, 240–255.
- Newman, J., Peat, T. S., Richard, R., Kan, L., Swanson, P. E., Affholter, J. A., Holmes, I. H., Schindler, J. F., Unkefer, C. J. & Terwilliger, T. C. (1999). *Biochemistry*, **38**, 16105–16114.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pavlova, M., Klvana, M., Prokop, Z., Chaloupkova, R., Banas, P., Otyepka, M., Wade, R. C., Tsuda, M., Nagata, Y. & Damborsky, J. (2009). *Nature Chem. Biol.* **5**, 727–733.
- Schindler, J. F., Naranjo, P. A., Honabberger, D. A., Chang, C.-H., Brainard, J. R., Vanderberg, L. A. & Unkefer, C. J. (1999). *Biochemistry*, **38**, 5772–5778.
- Stsiapanava, A., Dohnalek, J., Gavira, J. A., Kutý, M., Koudelakova, T., Damborsky, J. & Kuta-Smatanova, I. (2010). *Acta Cryst. D66*, 962–969.
- Teng, T.-Y. (1990). *J. Appl. Cryst.* **23**, 387–391.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst. D66*, 22–25.

### **Crystallographic analysis of 1,2,3-trichloropropane biodegradation by haloalokane dehalogenase DhaA31.**

Reproduced with permission from

Lahoda, M., Mesters, J., Stsiapanava, A., Chaloupkova, R., Damborsky, J. and Kuta Smatanova, I., (2013) *manuscript submitted to Acta Cryst D*



# Crystallographic analysis of 1,2,3-trichloropropane biodegradation by haloalkane dehalogenase DhaA31

Maryna Lahoda<sup>a,b</sup>, Jeroen R. Mesters<sup>c</sup>, Alena Stsiapanava<sup>b</sup>, Radka Chaloupkova<sup>d</sup>, Michal Kutý<sup>a,b,e</sup>, Jiri Damborsky<sup>d</sup> and Ivana Kuta Smatanova<sup>a,b,e\*</sup>

<sup>a</sup>University of South Bohemia in Ceske Budejovice, Institute of Complex Systems FFPW and CENAKVA, Zamek 136, 373 33 Nove Hradky, Czech Republic

<sup>b</sup>University of South Bohemia in Ceske Budejovice, Faculty of Science, Branisovska 31, 37005 Ceske Budejovice, Czech Republic

<sup>c</sup>University of Lübeck, Institute of Biochemistry, Center for Structural and Cell Biology in Medicine, Ratzeburger Allee 160, 23538 Lübeck, Germany

<sup>d</sup>Masaryk University Brno, Loschmidt Laboratories, Department of Experimental Biology and Research Centre for Toxic Compound in the Environment, Faculty of Science, Kamenice 5/A13, 62500 Brno, Czech Republic

<sup>e</sup>Academy of Science of the Czech Republic, Institute of Nanobiology and Structural Biology GCRC, Zamek 136, 373 33 Nove Hradky, Czech Republic

\*E-mail: [ivanaks@seznam.cz](mailto:ivanaks@seznam.cz)

**Synopsis** The *Rhodococcus rhodochrous* haloalkane dehalogenase variant DhaA31 showing improved activity on 1,2,3-trichloropropane, carries five substitutions in its structure. Crystallization of DhaA31 with 1,2,3-trichloropropane at room temperature and low pH facilitated the arrest of a protein-substrate complex. Comparison of the DhaA31 substrate-free structure with the DhaA31-substrate complex reveals two alternative conformations for the nucleophilic residue Asp106 and a possible pathway for halide release, while comparison of DhaA31 with the wild-type DhaA reveals reduced size and solvent accessibility of the active site.

## Abstract

Haloalkane dehalogenases catalyze the hydrolytic cleavage of carbon-halogen bonds, which is a key step in aerobic mineralization of many environmental pollutants. One important pollutant is the toxic and anthropogenic compound 1,2,3-trichloropropane (TCP). Rational design was combined with saturation mutagenesis to obtain the haloalkane dehalogenase variant DhaA31, which displays an increased catalytic activity towards TCP. Here we report the 1.31 Å crystal structure of substrate-free DhaA31, the 1.26 Å structure of DhaA31 in a complex with TCP, and the 1.85 Å wild-type DhaA structure. Crystals of the enzyme-substrate complex were successfully obtained by adding volatile TCP to the reservoir at crystallization at room temperature and pH 6.4. Comparison of the substrate-free structure with the DhaA31 enzyme-substrate complex reveals that the nucleophilic Asp106 changes its conformation from an inactive to an active state during the catalytic cycle. The positions of three chloride ions found inside the enzyme's active site indicate a possible pathway for halide release from the active site through the main tunnel. Comparison of the mutant DhaA31 with wild-type DhaA revealed that introduced substitutions reduced volume and solvent accessibility of the active site pocket.

**Keywords:** *Rhodococcus rhodochrous*; chlorinated compound; alkyl-enzyme intermediate; biodegradation; vapour diffusion crystallization; X-ray diffraction

## 1. Introduction

Halogenated compounds can serve as a carbon and energy source for certain aerobic microorganisms. Such organisms have specifically developed a variety of enzyme systems to degrade these compounds. On the other hand, many synthetically prepared halogenated chemicals, which represent important environmental pollutants, such as 1,2-dichloroethane, 1,2-dichloropropane, or 1,2,3-trichloropropane, are not easily biodegraded (Janssen *et al.*, 2005). Haloalkane dehalogenases have the potential to convert these halogenated hydrocarbons by hydrolytic cleavage of their carbon-halogen bonds to the corresponding alcohols (Pries *et al.*, 1994; Bosma *et al.*, 2002). The first haloalkane dehalogenase reported to have detectable activity towards 1,2,3-trichloropropane (TCP) was DhaA, isolated from the Gram-positive bacterium *Rhodococcus rhodochrous* NCIMB 13064 (Kulakova *et al.*, 1997). Wild-type DhaA slowly converts ( $k_{\text{cat}} = 0.08 \text{ s}^{-1}$ ) the toxic and potentially carcinogenic TCP to 2,3-dichloropropane-1-ol (DCL) under laboratory conditions (Schindler *et al.*, 1999; Bosma *et al.*, 1999; Pavlova *et al.*, 2009). The same reaction mechanism, but significantly slower conversion of TCP to DCL ( $k_{\text{cat}} = 0.005 \text{ s}^{-1}$ ) has been reported for the related haloalkane dehalogenase LinB (Monincova *et al.*, 2007). The haloalkane dehalogenases consist of two domains: the main  $\alpha/\beta$ -hydrolase domain and an  $\alpha$ -helical cap domain. The active site of haloalkane dehalogenases is located between the two domains in a hydrophobic cavity, connected with the protein surface by two major access tunnels, designated the main tunnel and slot tunnel, and three additional transient tunnels have been identified by molecular modeling (Klvana *et al.*, 2009). Ligands can freely move between the buried active site and the surrounding solute through these tunnels. The dehalogenation reaction catalyzed by haloalkane dehalogenases takes place in the buried active site. The halogenated substrate binds to form the substrate-enzyme complex and the nucleophilic residue Asp106 attacks a  $sp^3$ -hybridized carbon atom of the bound substrate, resulting in the formation of a covalent alkyl-enzyme intermediate by cleavage of the carbon-halogen bond and release of the halide ion. The alkyl-enzyme intermediate is subsequently hydrolyzed by a water molecule, activated by the catalytic base His272 (Schindler *et al.*, 1999; Otyepka *et al.*, 2007; Pavlova *et al.*, 2009).

Several variants of DhaA have been constructed and characterized with the goal of improving the enzymatic kinetic properties for TCP conversion (Bosma *et al.*, 2002; Gray *et al.*, 2001; Pavlova *et al.*, 2009). The most successful variant, DhaA31, has exhibited 32-fold higher catalytic activity and 26-fold higher catalytic efficiency for TCP at an optimal pH value of 8.6 and 37 °C than the wild-type enzyme (Pavlova *et al.*, 2009). This variant contains large aromatic residues in four out of five positions targeted by mutagenesis. The C176Y mutation, located in the cap domain, was suggested to modify the mouth of the main tunnel that connects the buried active site with the surrounding solvent. V245F is located in the main tunnel, and I135F and L246I are located in the slot tunnel and the active site, respectively (Bosma *et al.*, 2002; Pavlova *et al.*, 2009).

In this study, the variant DhaA31 was crystallized in a substrate-free form and in the presence of TCP at pH 6.4 and differing temperatures. Analogous crystallographic analysis has been previously carried out with haloalkane dehalogenase Dh1A from *Xanthobacter autotrophicus* GJ10 (Verschueren *et al.*, 1993c) where different stages of the enzyme's reaction pathway were trapped by varying pH and temperature during soaking experiments with 1,2-dichloroethane. At pH 5.0 and 4 °C, the enzyme-substrate complex was captured. Increasing the temperature resulted in the accumulation of a covalently bound intermediate of the active site, and increasing the pH to 6.2 allowed for the release of the alcohol and chloride ion. At pH 8.2, an optimal condition for

catalysis, significant movement in the position of catalytic residues was observed. The catalytic histidine was mostly neutral, which results in a loss of the salt-bridge between the catalytic base and the nucleophile, as observed in the pH 6.2 structure. This left the negatively charged aspartate side-chain ready to perform a nucleophilic attack on the carbon atom of the substrate molecule (Verschuere *et al.*, 1993a,c).

Here, we analyzed and compared the crystal structures of DhaA31 in its free form with wild-type DhaA and the variant DhaA31 crystallized in the presence of TCP. This comparison revealed two alternative conformations for the nucleophilic residue Asp106 in the DhaA31-TCP complex, and enabled the identification of seven chlorine atoms/chloride ions locations within the active site. The position and spatial arrangement of three of these atoms/ions can be explained by modeling with a TCP molecule. Variant DhaA31 also carries five amino acid substitutions that reduce the volume of the active site cavity and accessibility for water molecules, which may provide an explanation for the higher catalytic efficiency with TCP.

## 1. Materials and methods

### 1.1 Protein crystallization

The standard sitting-drop vapour-diffusion technique (Ducruix & Giege, 1999) was performed in CombiClover crystallization plates (EBS plate, Emerald BioStructures, WA, USA) at 277 K. Three-dimensional crystals of DhaA31 were grown from 2  $\mu$ l droplets consisting of equal volumes of protein solution, 7 mg/ml in 50 mM Tris-HCl buffer, pH 7.5, and reservoir solution containing 29% (w/v) PEG 4000 and 100 mM MES-NaOH buffer, pH 6.4, equilibrated against the reservoir solution. Crystals grew to final dimensions of 0.28 x 0.19 x 0.13 mm<sup>3</sup> in three weeks. For complex formation, crystals were placed over an 800  $\mu$ l reservoir solution with 50  $\mu$ l of TCP and left for 6 hours at room temperature to ensure slow penetration of substrate into the crystals by TCP vapor diffusion, thereby avoiding crystal destruction, which takes place through direct soaking (Lahoda *et al.*, 2011). Crystals of wild-type DhaA were obtained in 5 days by mixing 1  $\mu$ l protein solution at a concentration of 11 mg/ml in 50  $\mu$ l Tris-HCl buffer, pH 7.5, with an equal volume of precipitant solution composed of 39% (w/v) PEG 4000, 8% (v/v) TCP and 100 mM sodium acetate. Drops were equilibrated over a 700  $\mu$ l reservoir solution (Stsiapanava *et al.*, 2011). Prior to data collection, crystals of wild-type DhaA were soaked in a solution containing 25% (v/v) glycerol, 20% PEG 4000, 4% TCP and 50 mM sodium acetate, and were flash frozen in liquid nitrogen.

### 1.2 Data collection and structure solution

Diffraction data for substrate-free DhaA31 and DhaA31 in complex with TCP were collected at EMBL Hamburg outstation DESY X12 beamline with a fixed wavelength of 1.033 Å, and equipped with a MAR CCD 225 mm detector. Crystals were mounted directly into a nylon loop (Hampton Research, Aliso Viejo, USA) (Teng, 1990) and flash-frozen in a cold nitrogen stream at 100 K without additional cryoprotection. DhaA31 data sets were collected in two steps, 120 low-resolution images (50–2.89 Å) were recorded with an oscillation angle of 3° and a crystal-to-detector distance of 300 mm. A high-resolution data set of 360 images (50–1.31 Å) was collected with an oscillation angle of 1° and a crystal-to-detector distance of 100 mm. For DhaA31 in complex with TCP, 400 high-resolution images (50–1.26 Å) were collected with a 0.5° oscillation angle and a crystal-to-detector distance of 100 mm. This was followed by a collection of 130 image frames with a 1.5° oscillation angle and a crystal-to-detector distance of 300 mm. Data for both crystals were integrated and scaled using the *HKL2000* program package (Otwinowski & Minor, 1997).

Diffraction data for wild-type DhaA were collected at the beamline 14.2 at the BESSY II electron-storage ring (Mueller *et al.*, 2012) operated by the Joint Berlin MX-Laboratory (Berlin-Adlershof, Germany). Experiments were done at a single wavelength of 1.9 Å using a Rayonix 225 mm CCD detector. A single data set (resolution, 22–1.85 Å) of 360 images was collected with an oscillation angle of 1° per image and a crystal-to-detector distance of 60 mm. The data were indexed and integrated using the *XDS* program (Kabsch 1993, 2010*a,b*), and were scaled with *Scala* from the *CCP4* program suite (Collaborative Computational Project, Number 4, 1994). The data-collection statistics are summarized in Table 1.

### 1.3 Structure refinement

The structures of DhaA31, its complex with TCP and wild-type DhaA were solved by the molecular replacement method using *MOLREP* (Vagin & Teplyakov, 1997) and the structure of haloalkane dehalogenase DhaA from *R. rhodochrous* (PDB ID code 3FBW) (Stsiapanava *et al.*, 2010) as the search model. Both structures of DhaA31 were refined employing the *SHELXL* program (Sheldrick 2008). The *Coot* package was used for manual building (Emsley *et al.*, 2010). After structure refinement, the  $|F_o| - |F_c|$  Fourier difference map of the active site was interpreted. The difference density in the active site of the substrate-free DhaA31 structure was interpreted as a chloride ion by anomalous difference maxima and the typical distances of Cl<sup>-</sup> – HN contacts. The structure of DhaA31 in complex with substrate displays electron density for 1,2,3-trichloropropane molecule and four possible positions for chloride-ion binding. Anisotropic atomic displacement parameters were refined for all protein and ligand atoms and ions. For water O atoms, the isotropic atomic displacement parameters were kept. The combined isotropic/anisotropic ADP refinement, as compared to standard isotropic refinement (respectively), resulted in a lower R and R<sub>free</sub> value (R = 0.234 and R<sub>free</sub> = 0.257 to R = 0.128 and R<sub>free</sub> = 0.160 for DhaA31; R = 0.251 and R<sub>free</sub> = 0.279 to R = 0.135 and R<sub>free</sub> = 0.164 for DhaA31 in complex with TCP).

Restrained refinement for wild-type DhaA was performed in *Refmac5* (Murshudov *et al.*, 1997) and model building was done in *Coot*. Atomic displacement parameters of all protein and water O atoms, and of the chloride ion were refined isotropically. Structure-refinement statistics are presented in Table 1.

### 1.4 Model validation and deposition

Structural models with respect to experimental data were assessed using the program *SFCHECK* (Vaguine *et al.*, 1999). Geometric parameters were validated by *Coot* (Emsley *et al.*, 2010) and *MolProbity* (Chen *et al.*, 2010). 90.0% of residues in the substrate-free DhaA31 structure and 89.6% of residues in the DhaA31-TCP complex are located in the “core” region of the Ramachandran plot. For both structures, two residues are located in the “generous” region. These Ramachandran outliers with clear electron densities are the catalytic residue Asp106 and variant residue Phe245, located in the active site and the main tunnel, respectively. In the case of wild-type DhaA, 96.7% of all residues were in the most favored regions of the Ramachandran plot and 100.0% of all residues were in allowed regions. The structures and original diffraction intensities were deposited in the Protein Data Bank. The substrate-free DhaA31 structure, the DhaA31-TCP complex structure and wild-type DhaA were assigned PDB ID codes 3RK4, 4FWB and 4HZG, respectively. The volumes of the active site cavities were calculated using the 3V server with the following settings (Voss & Gerstein, 2010): Grid resolution of 0.5 Å, outer probe radius of 3 Å and inner probe radius of 1.4 Å. The Van-der-Waals volume of the TCP molecule was calculated by the Geometry Plugin of Marvin 5.11.1 (ChemAxon, Budapest).

## 2. Results and discussion

## 2.1 Overall structure of the protein variant DhaA31

The structure of wild-type haloalkane dehalogenase DhaA from *Rhodococcus rhodochrous* NCINB13064 has previously been described by Newman *et al.* (1999). For both structures of protein variant DhaA31, electron density was clearly visible for amino-acid residues 4 to 295. DhaA31 has a globular structure and is composed of two domains analogous to the wild-type enzyme (Fig. 1). The main domain with the  $\alpha/\beta$  hydrolase fold consists of an eight-stranded  $\beta$ -sheet surrounded by six  $\alpha$ -helices (topology:  $\beta 1-\beta 2-\beta 3-\alpha 1-\beta 4-\alpha 2-\beta 5-\alpha 3-\beta 6$ ,  $\alpha 8-\beta 7-\alpha 9-\beta 8-\alpha 10$ ). The second domain, called the cap domain, is composed of five  $\alpha$ -helices ( $\alpha 4-\alpha 5'-\alpha 5-\alpha 6-\alpha 7$ ). The active site is located between the two domains. The catalytic residues, forming the catalytic pentad, are located in this cavity: Asn41, Asp106, Trp107, Glu130 and His272 (Chovancova *et al.*, 2007; Stsiapanava *et al.*, 2010).

### 1.1 Analysis of the active site

The active-site cavity of the largely refined substrate-free DhaA31 structure contained a Fo - Fc difference electron-density peak (greater than 10  $\sigma$  above the mean) that was finally interpreted as a chloride ion. Coordinated to Asn41 N <sup>$\delta 2$</sup>  (at 3.33 Å distance) and Trp107 N <sup>$\epsilon 1$</sup>  (3.32 Å), the chloride ion occupies the canonical halide-binding site (Fig. 2a) and most probably originates from the buffer solution. In addition, two water molecules were modeled close to the side chain of Asp106. The catalytic residue Asp106 displays a C <sup>$\alpha$</sup>  - C <sup>$\beta$</sup>  - C <sup>$\gamma$</sup>  bond angle of - 114.65°. The N <sup>$\delta 1$</sup>  atom of His272 forms a hydrogen-bond with the side chain of Glu130 (2.73 Å), and the N <sup>$\epsilon 2$</sup>  atom of His272 forms a hydrogen-bond with the O <sup>$\delta 1$</sup>  atom (2.87 Å) of the side chain of Asp106. The O <sup>$\delta 2$</sup>  atom of Asp106 accepts hydrogen bonds from the main chain amides of Asn41 (2.80 Å) and Trp107 (2.83 Å). This constellation is in good agreement with the hydrogen-bonding network observed in the DhlA structure at low pH values (Verschueren *et al.*, 1993a).

For crystals soaked with TCP, the largely refined DhaA31 structure exhibited in the active site four Fo - Fc difference electron-density peaks greater than 10  $\sigma$  above the mean. In an effort to interpret the difference electron-density, chloride ions, TCP substrate, alkyl-enzyme and product were modeled independently. We found that the three chlorine atoms of 1,2,3-trichloropropane superimposed correctly with three out of the four Fo - Fc difference electron-density peaks. The 2,3-dichloropropane-1-ol product molecule does not superimpose well because of the shorter distance between the C and O of an alcohol product (about 1.3 Å) compared with C and Cl distances in the substrate molecule (about 1.8 Å). The largest Fo - Fc difference electron-density peak (greater than 30  $\sigma$  above the mean) superimposed with the canonical halide binding site. Nonetheless, we cannot fully exclude the concomitant existence of a near attack and/or alkyl-enzyme state. Unfortunately, modeling of these states had to be abandoned because of low occupancy. Finally, one chloride ion (N <sup>$\delta 2$</sup>  Asn41 - Cl<sup>-</sup>, 3.34 Å and N <sup>$\epsilon 1$</sup>  Trp107 - Cl<sup>-</sup>, 3.27 Å; occupancy 1.0) in the canonical halide binding site and one TCP molecule (occupancy 0.6) was modeled into the difference electron-density peaks. In addition, a water molecule (occupancy 0.4) and two conformations of the catalytic Asp106 residue were modeled with a C <sup>$\alpha$</sup>  - C <sup>$\beta$</sup>  - C <sup>$\gamma$</sup>  bond angle of 126.52° (alternative A, N <sup>$\epsilon 2$</sup>  His 272 - O <sup>$\delta 1$</sup>  Asp106, 2.49Å) and 106.93° (alternative B, N <sup>$\epsilon 2$</sup>  His 272 - O <sup>$\delta 1$</sup>  Asp106, 3.11Å). The occupancies of alternatives A and B are 0.6 and 0.4, respectively (Fig. 2b). With the aforementioned interpretation of the difference electron density map, three small parts of the initial peaks remained unsolved. Because of their very close proximity to the Cl-2 and Cl-3 atoms of the TCP substrate, these peaks were interpreted finally as low occupancy (0.2 to 0.4) chlorine atoms and we hypothesize that these may represent alternative conformations of low occupancy substrate/product molecules, or, that these chlorine atoms reveal a possible route for halide release. Two distinct conformations of the nucleophile have been previously observed in the crystal structure of the haloalkane dehalogenase LinB solved at 0.95 Å resolution (Oakley *et al.*, 2004).

The active-site cavity of the largely refined wild-type DhaA structure crystallized in the presence of TCP contained merely one clear Fo - Fc difference electron-density peak (greater than 10  $\sigma$  above the mean), which was modeled as a chloride ion occupying the canonical halide-binding site. Additionally, three small Fo - Fc difference electron-density peaks (less than 4.5  $\sigma$  above the mean) were located close to Cys176 but distal from Asp106. Modeling of a low occupancy TCP molecule into the Fo - Fc difference electron-density peaks was finally abandoned because the chlorine atoms superposed incorrectly with the three Fo - Fc difference electron-density peaks.

#### Implications for the reaction mechanism

The crystal structure of the DhaA31-TCP complex displays a combination of two different states of the active-site cavity. One state is characterized by a post-attack conformation of Asp106 (alternative A, occupancy of 0.6). Here, the catalytic Asp106 must already have undergone a reaction cycle as evidenced by the absence of the catalytic water molecule (H<sub>2</sub>O-1) used in the cleavage of the intermediate state (alkyl-enzyme), releasing the product 2,3-dichloropropane-1-ol (Fig. 2c). At this point, the His272 side chain most probably remains partly protonated due to the pH of the crystallization drop (Verschuere *et al.*, 1993c) and, as a consequence, the O <sup>$\delta$ 1</sup> atom of Asp106 interacts with the N <sup>$\epsilon$ 2</sup> atom of His272, at a 2.53 Å distance, forming an ion pair (Fig. 4d). Apparently, the TCP substrate (occupancy of 0.6) in the active site did not replace the chloride product coordinated to the nitrogen atoms of Asn41 and Trp107 (canonical halide-binding site), and the substrate presumably resides in a pre- or near-attack orientation (Fig. 4e). During initial modeling of the TCP substrate, alkyl-enzyme and 2,3-dichloropropane-1-ol product, we noticed that the three chlorine atoms of 1,2,3-trichloropropane also superimposed correctly with the Fo - Fc difference electron-density peak at the halide binding site and two out of the three Fo - Fc difference electron-density peaks used in the modeling of the TCP molecule in the pre- or near-attack orientation. The modeling of this TCP orientation was abandoned however because of a negative Fo - Fc difference electron-density peak at one of its carbon atoms. Nonetheless, from the model it became clear that the relocation of one particular chlorine atom is sufficient to switch from the pre- or near-attack orientation to the proper orientation necessary for catalysis to proceed.

The second state is characterized by a conformation of the Asp106 side chain (alternative B, occupancy of 0.4) ready to perform a nucleophilic attack on the carbon atom of a bound substrate molecule (Fig. 4b). Compared to the first state, the distance between the O <sup>$\delta$ 1</sup> atom of Asp106 and the N <sup>$\epsilon$ 2</sup> atom of His272 changes from 2.53 Å to 3.14 Å. Here, the catalytic water molecule (H<sub>2</sub>O-1; occupancy of 0.4) is located 2.76 Å away from the C <sup>$\gamma$</sup>  atom of the Asp106, 3.17 Å from the N <sup>$\epsilon$ 2</sup> atom of His272 and 2.59 Å from the carbonyl O atom of Asn41. These observations are consistent with the interactions described by Verschuere *et al.* 1993c. The chloride ion bound in the canonical halide-binding site is a product formed during the reaction that will be replaced by a substrate molecule to start the next reaction cycle (Fig. 4f). Furthermore, three chloride ions with partial occupancies were modeled in non-canonical halide-binding positions, revealing a possible route for halide release in the variant DhaA31. These non-canonical halide-binding sites, not observed in the wild-type DhaA, are close to the V245F, C176Y and Y273F mutations located at the mouth of the main tunnel. Actually, these substitutions play a crucial role in the reshaping of the DhaA hydrophobic pocket. Consequently, the volume of the active site of DhaA31 (127 Å<sup>3</sup>) is clearly reduced in comparison to wild-type DhaA (187 Å<sup>3</sup>). When taken into account that a TCP molecule has a calculated volume of 105 Å<sup>3</sup>, the variant DhaA31 has an active site complementary for a TCP molecule, leaving a little space for water solvent. The bottlenecks of both the main and side access tunnels are also significantly altered in DhaA31. Reduced accessibility of the active site for water molecules, promoting formation of the activated complex, has been previously proposed as a key mechanism for the improved activity of DhaA31 with TCP (Pavlova *et al.*, 2009).

## Conclusions

Crystals of DhaA31 were obtained in two different forms: (i) one substrate-free form and (ii) one form in complex with 1,2,3-trichloropropane. The crystal structures provide new information for certain stages of the enzyme's catalytic cycle. Most interestingly, the catalytic residue Asp106 displays two alternative conformations. Conformation *A* represents a post-attack conformation for Asp106 with the catalytic His272 protonated, creating a salt-bridge with the nucleophile at low pH (Verschuere *et al.*, 1993*a,c*). The alternative conformation *B* is observed as a result of the loss of the salt-bridge between the latter two catalytic residues. The water molecule (H<sub>2</sub>O-1) located close to the base His272 (distance 3.17 Å) and the C' atom of the nucleophile Asp106 (distance 2.76 Å for conformation *B*), most probably serves as the catalytic water for cleavage of the alkyl-enzyme bond. The absence of a catalytic water near Asp106 in the post-attack conformation indicates that hydrolysis of a covalent alkyl-enzyme bond intermediate took place (Fig. 2*c* and 3*b*). After hydrolysis of the covalently bound intermediate, this catalytic water should be replenished by another water molecule (for example H<sub>2</sub>O-2) (Fig. 2*c,d* and 4*d*).

Comparative analysis of the wild-type DhaA structure and the DhaA31-substrate complex structure revealed differences in the architecture of the hydrophobic cavity and two access tunnels. The introduced substitutions, located mainly distal from the catalytic residue Asp106, lead to a volume reduction of the hydrophobic pocket and decreased bottlenecks of the access tunnels. In conjunction with the pH of the crystallization drop and the gentle vapor diffusion of TCP, an enzyme-substrate complex could be obtained. Upon entering the active site of DhaA31, TCP has less available space and binds in a defined manner proximal to Asp106. The defined binding results from specific interactions between the substrate's chlorine atoms on one hand, and Phe149, Phe168, Phe245, Phe273, and Trp141 on the other (Fig. 3*b*).

In our opinion, the observed pre- or near-attack orientation of TCP in the complex in the absence of water molecules is a crucial first step for the enhanced catalysis by DhaA31. The hydrophobic pocket of wild-type DhaA is larger and significantly more accessible to water solvent, which may explain the low turnover. For the reaction to proceed in DhaA31, relocation of one particular TCP chlorine atom is sufficient to switch from the observed pre- or near-attack state to the proper conformation necessary for catalysis *i.e.*, an obligatory exchange through competition of the chloride ion bound to the canonical halide-binding site between Asn41 and Trp107 and a chlorine atom of the substrate is required. The small Fo - Fc difference electron-density peaks distal from Asp106 as observed in the different structures, may possibly underscore the role of the main tunnel in substrate and product exchange. This would be in good agreement with a previous study of the product release pathways, where insertion of bulky residues into two access tunnels of the wild-type DhaA, decreased the accessibility of the individual tunnels for the halide ions and alcohol products (Klvanová *et al.*, 2009).

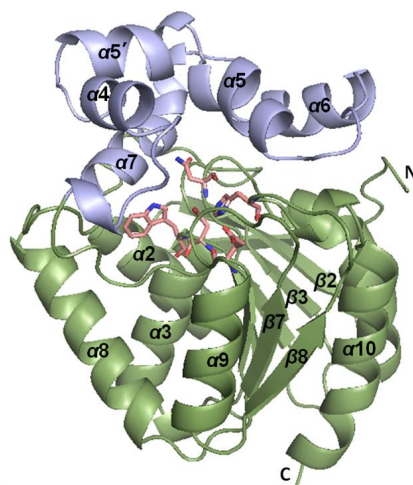
In summary, crystallization of DhaA31 with 1,2,3-trichloropropane at low pH facilitated the arrest of an enzyme-substrate complex, revealing two alternative conformations for the nucleophilic residue Asp106. The low pH contributed to the retention of the chloride product at different positions inside the new hydrophobic pocket, illustrating a possible trajectory for chloride ion release through the main tunnel. Comparative analysis of three crystal structures provided information on the possible structural basis of enhanced conversion of TCP by DhaA31, which is due to reduced volume and redesign of the active site pocket.

**Table 1** Data-collection and refinement statistics

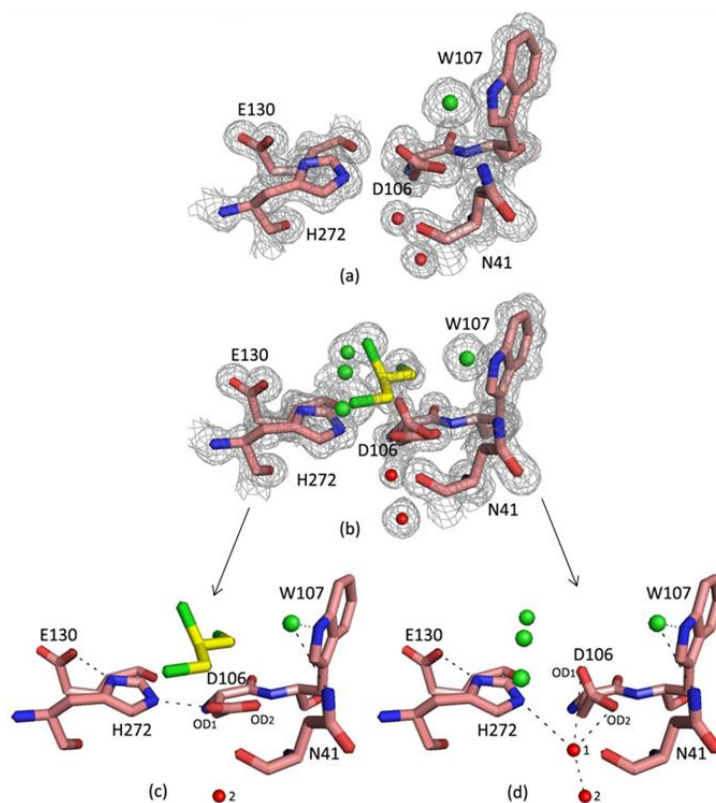
	DhaA31	DhaA31 with TCP	DhaA
<b>Data collection</b>			
Resolution range (Å)	50–1.31 (1.35–1.31)	50–1.26 (1.27–1.26)	22–1.85 (1.94–1.85)
Space group	<i>P1</i>	<i>P1</i>	<i>P1</i>
Unit-cell parameters (Å, °)	<i>a</i> = 42.55, <i>b</i> = 44.37, <i>c</i> = 46.41, $\alpha$ = 115.3, $\beta$ = 98.5, $\gamma$ = 109.5	<i>a</i> = 42.49, <i>b</i> = 44.39, <i>c</i> = 46.53, $\alpha$ = 115.3, $\beta$ = 97.7, $\gamma$ = 109.5	<i>a</i> = 44.77, <i>b</i> = 44.47, <i>c</i> = 46.51, $\alpha$ = 115.3, $\beta$ = 97.8, $\gamma$ = 109.2
Unique reflections (observed)	73,847	74,995	20,934
Redundancy	4.3	2.4	3.7
Completeness (%)	94.6 (90.3)	92.4 (82.7)	87.8 (81.5)
$R_{\text{merge}}^{\dagger}$ (%)	4.4 (8.1)	3.3 (6.1)	15.8 (56.8)
$I/\sigma(I)$	35.6 (18.3)	41.1 (17.6)	5.3 (2.0)
<b>Refinement</b>			
No. of reflections used for refinement	61,990	69,189	17,017
Max. resolution (Å)	1.31	1.26	1.95
$R_{\text{work}}^{\ddagger}/R_{\text{free}}^{\S}$ (%)	12.8/16.0	13.5/16.4	21.6/27.3
No. of non-H atoms	2,855	2,824	2,645
No. of protein atoms	2,469	2,461	2,412
No. of ligand atoms	1	10	1
No. of chloride ions	1	4	1
No. of water molecules	385	353	232
<b>R.m.s. deviations</b>			
Bond lengths (Å)	0.012	0.012	0.014
Bond angles (°)	0.028	0.029	1.665
<b>PDB code</b>	3RK4	4FWB	4HZG

$^{\dagger}R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$ , where  $I_i(hkl)$  is the intensity of the  $i$ th measurement of reflection  $hkl$  and  $\langle I(hkl) \rangle$  is the average intensity of the reflection.  $^{\ddagger}R_{\text{work}} = \frac{\sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum_{hkl} |F_{\text{obs}}|}$ .  $^{\S}R_{\text{free}}$  was monitored using 5% of the reflection data that were excluded from refinement.

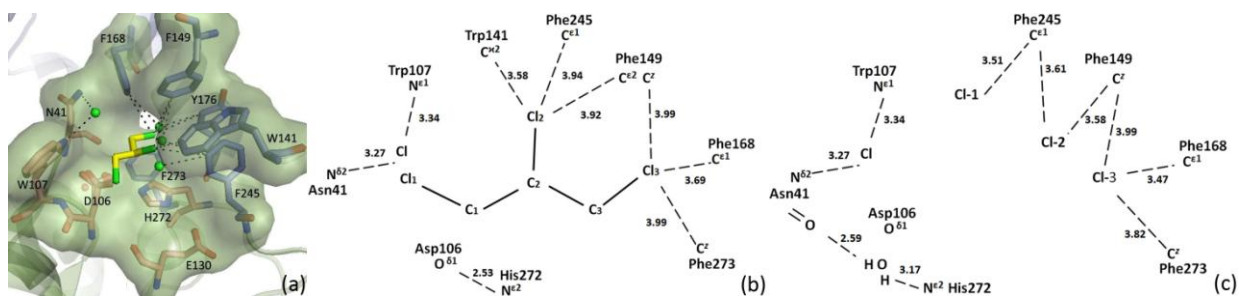




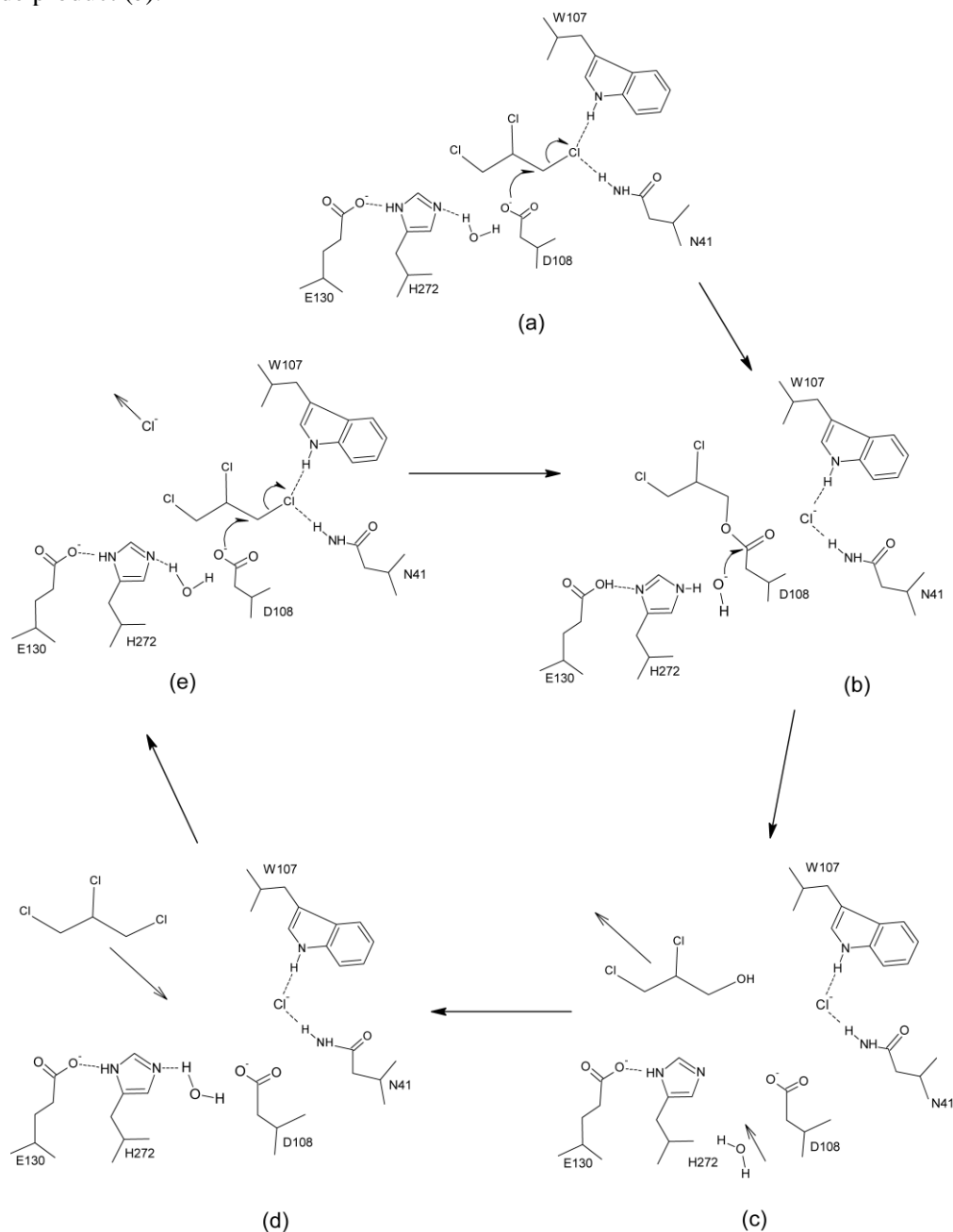
**Figure 1** Cartoon representation of the DhaA31 structure. The  $\alpha/\beta$ -hydrolase core domain is shown in green; the helical cap domain is shown in blue. Catalytic pentad residues are shown in magenta stick representation. Figures were prepared using *PyMOL* (DeLano, 2002).



**Figure 2** The active site of DhaA31. Active-site residues of substrate-free DhaA31 with a fixed position for Asp106 (a) and DhaA31 in the complex with TCP (b–d) are shown in stick representation. The  $2|F_o| - |F_c|$  electron density map, contoured at  $1\sigma$  above the mean, around the active site is shown as a grey mesh (a,b). TCP carbon and chlorine atoms are colored yellow and green, respectively. Green spheres represent chloride anions and red spheres represent oxygen atoms of water molecules. Alternative conformation A of the catalytic residue Asp106 in the post-attack configuration with the substrate (c) and alternative conformation B of the catalytic residue Asp106 with halide product leaving the active site (d).



**Figure 3** The DhaA31 active-site cavity with TCP (a). Schematic representation of the DhaA31 hydrophobic cavity: The interaction of hydrophobic residues with substrate TCP (b) and the chloride product (c).



**Figure 4** Proposed reaction mechanism of DhaA from *Rhodococcus rhodochrous* (based on Verschueren et al., 1993). The reaction schema: Substrate binding and nucleophilic attack of

Asp106 (a); Alkyl-enzyme intermediate followed by nucleophilic attack of water (b); Halogen remains bound to catalytic residues Asp106 and His272, alcohol product leaves the active site and new hydrolytic water enters the active site (c); New substrate molecule (pre- or near-attack orientation) enters the active site (d) and replaces the chloride ion in the canonical halide-binding site (e).

**Acknowledgements** We thank the staff at EMBL (DESY, Hamburg) and BESSY, (Berlin) for assistance with data collection as well as Rolf Hilgenfeld (University of Lübeck) for enabling and arranging a short-term laboratory visit. This work has been supported by the Grant Agency of the Czech Republic project P207/12/0775. Support of the Ministry of Education of the Czech Republic and the Academy of Sciences of the Czech Republic is also appreciated.

## References

- Bosma, T., Damborsky, J., Stucki, G., Janssen, B.D. (2002). *Appl. Environ. Microbiol.* 68, 3582–3587.
- Bosma, T., Kruizinga, E., de Bruin, E.J., Poelarends, G.J. & Janssen, D.B. (1999). *Appl. Environ. Microbiol.* 65, 4575–4581.
- Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C. (2010). *Acta Cryst.* D66, 12–21.
- Chovancova, E., Kosinski, J., Bujnicki, J.M., Damborsky, J. (2007). *Proteins* 67, 305–316.
- Collaborative Computation Project, Number 4 (1994). *Acta Cryst.* D50, 760–763.
- DeLano, W.L. (2002). The *PyMOL Molecular Viewer*. [Http://www.pymol.org](http://www.pymol.org).
- Crystallization of Nucleic Acids and Proteins: A Practical Approach (Second Edition)*, eds. Ducruix, A. and Giege, R. (1999). Oxford, IRL Press.
- Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K. (2010). *Acta Cryst.* D66, 486–501.
- Gray, K. A., Richardson, T.H., Kretz, K., Short, J.M., Bartnek, F., Knowles, R., Kan, L., Swanson, P.E., Robertson, D.E. (2001). *Adv. Synth. Catal.* 343, 607–616.
- Janssen, D.B., Dinkla, I.J.T., Poelarends, G.J., Terpstra, P. (2005). *Appl. Environ. Microbiol.* 7, 1868–1882.
- Kabsch, W. (1993) *J. Appl. Cryst.* 26, 795–800.
- Kabsch, W. (2010a). *Acta Cryst.* D66, 125–132.
- Kabsch, W. (2010b). *Acta Cryst.* D66, 133–139.
- Klvana, M., Pavlova, M., Koudelakova, T., Chaloupkova, R., Dvorak, P., Prokop, Z., Stsiapanava, A., Kutý, M., Kuta-Smatanova, I., Dohnalek, J., Kulhanek, P., Wade, C.R., Damborsky, J. (2009). *J. Mol. Biol.* 392, 1339–1356.
- Kulakova, A.N., Larkin, M.J., Kulakov, L.A. (1997). *Appl. Environ. Microbiol.* 143, 109–115.
- Lahoda, M., Chaloupkova, R., Stsiapanava, A., Damborsky, J., Kuta Smatanova, I. (2011). *Acta Cryst.* F67, 397–400.
- Monincova, M., Prokop, Z., Vevodova, J., Nagata, Y., Damborsky, J. (2007). *Appl. Environ. Microbiol.* 73, 2005–2008.
- Mueller, U., Darowski, N., Fuchs, M.R., Förster, R., Hellmig, M., Paithankar, K.S., Pühringer, S., Steffien, M., Zocher, G., Weiss, M.S. (2012). *J. Synchrotron Radiat.* 19, 442–449.
- Murshudov, G.N., Vagin, A.A., Dodson, E.J. (1997). *Acta Cryst.* D53, 240–255.
- Newman, J., Peat, T.S., Richard, R., Kan, L., Swanson, P.E., Affholter, J.A., Holmes, I.H., Schindler, J.F., Unkefer, C.J., Terwilliger, T.C. (1999). *Biochemistry* 38, 16105–16114.
- Otwinowski, Z., Minor, W. (1997). *Methods Enzymol.* 276, 307–326.
- Oakley, A., Klvana, M., Otyepka, M., Nagata, Y., Wilce, M.C.J., Damborsky, J. (2004). *Biochemistry* 43, 870–878.
- Otyepka, M., Banas, P., Magistrato, A., Carloni, P., Damborsky J. (2007). *Proteins* 70, 707–717.
- Pavlova, M., Klvana, M., Prokop, Z., Chaloupkova, R., Banas, P., Otyepka, M., Wade, R. C., Tsuda, M., Nagata, Y., Damborsky, J. (2009). *Nature Chem. Biol.* 5, 727–733.
- Petrek, M., Otyepka, M., Banas, P., Kosinova, P., Koca, J., Damborsky, J. (2006). *BMC Bioinformatics* 7, 316.
- Pries, F., Van der Wijngaard, A. J., Bos, R., Pentenga, M., Janssen, D. B. (1994). *J. Biol. Chem.* 269, 17490–17494.
- Pries, F., Van der Ploeg, J.R., Dolfing, J., Janssen, D.B. (1994). *FEMS Microbiol. Rev.* 15, 277–295.
- Schindler, J.F., Naranjo, P.A., Honaberger, D.A., Chang, C.-H., Brainard, J.R., Vanderberd, L.A., Unkefer, C.J. (1999). *Biochemistry* 38, 5772–5778.
- Sheldrick, G.M. (2008). *Acta Cryst.* A64, 112–122.
- Stsiapanava, A., Chaloupkova, R., Jesenska, A., Brynda, J., Weiss, M.S., Damborsky, J., Kuta Smatanova, I. (2011). *Acta Cryst.* F67, 253–257.

- Stsiapanava, A., Dohnalek, J., Gavira, J.A., Kutý, M., Koudelakova, T., Damborsky, J., Kuta Smatanova, I. (2010). *Acta Cryst.* D66, 962–969.
- Teng, T.-Y. (1990). *J. Appl. Cryst.* 23, 387–391.
- Vagin, A., Teplyakov, A. (1997). *J. Appl. Cryst.* 30, 1022–1025.
- Vaguine, A.A., Richelle, J., Wodak, S.J. (1999). *Acta Cryst.* D55, 191–205.
- Verschuere, K.H., Kingma, J., Rozeboom, H.J., Kalk, K.H., Janssen, D.B., Dijkstra, B.W. (1993a). *J. Mol. Biol.* 232, 856–72.
- Verschuere, K.H., Kingma, J., Rozeboom, H.J., Kalk, K.H., Janssen, D.B., Dijkstra, B.W. (1993b). *Biochemistry* 32, 9031–9037.
- Verschuere, K.H., Seljée, F., Rozeboom, H.J., Kalk, K.H., Dijkstra, B.W. (1993c). *Nature* 363, 693–698.
- Voss, N. R. and Gerstein M. (2010). *Nucleic Acids Res.* 38 (Web Server issue): W555–W562.

**Crystal structure of haloalkane dehalogenase mutant DhaA12 from Gram-positive bacterium *Rhodococcus rhodochrous* NCIMB13064.**

Reproduced with permission from

Lahoda, M., Koudelakova, T., Stsiapanava, A., Mesters, J., Damborsky, J. and Kuta Smatanova, I., (2013) *prepared for submission to Acta Cryst F*

# Crystal structure of haloalkane dehalogenase mutant variant DhaA12 from *Rhodococcus rhodochrous* NCIMB13064

Maryna Lahoda<sup>a,b</sup>, Tana Koudelakova<sup>c</sup>, Alena Stsiapanava<sup>b</sup>, Jiri Damborsky<sup>c</sup>  
and Ivana Kuta Smatanova<sup>a,b,d</sup>

<sup>a</sup>University of South Bohemia in Ceske Budejovice, Institute of Complex Systems FFPW and CENAKVA, Zamek 136, 37333 Nove Hradky, Czech Republic

<sup>b</sup>University of South Bohemia in Ceske Budejovice, Faculty of Science, Branisovska 31, 37005 Ceske Budejovice, Czech Republic

<sup>c</sup>Masaryk University, Faculty of Science and Loschmidt Laboratories, Department of Experimental Biology and Research Centre for Toxic Compound in the Environment, Kamenice 5/A13, 62500 Brno, Czech Republic

<sup>d</sup>Institute of Nanobiology and Structural Biology GCRC, Academy of Science of the Czech Republic, Zamek 136, 37333 Nove Hradky, Czech Republic

\*E-mail: ivanaks@seznam.cz

**Synopsis** The structure of *Rhodococcus rhodochrous* haloalkane dehalogenase variant DhaA12 was solved and revealed successful insertion of additional loop fragment.

## Abstract

The haloalkane dehalogenase DhaA, a microbial  $\alpha/\beta$ -hydrolase, shows a weak enantioselectivity with 1,2- and 1,3-dihaloalkanes, and non-enantioselectivity with  $\beta$ -bromoalkanes, while similar haloalkane dehalogenase DbjA shows high enantioselectivity with  $\beta$ -bromoalkanes. It was further identified that DhaA and DbjA are the structurally related to each other. This finding fact resulted toward a possibility in construction of mutant protein that enantioselective ability approaching DbjA level. The DhaA12 was prepared as experimental mutant protein combining the nature of two parent's proteins DhaA and DbjA. The diffraction data of DhaA12 crystal were collected to the resolution of 1.78 Å and the crystal structure of enzyme was solved belonging to the orthorhombic space group  $P2_12_12_1$ . Structure analysis of mutant protein is important for future protein design and understanding of enantioselectivity properties of haloalkane dehalogenase's family.

**Keywords:** *Rhodococcus rhodochrous*; enantioselectivity; mutant protein; crystallization; vapor diffusion method; X-ray diffraction

## 3. Introduction

Haloalkane dehalogenases (HLDs) structurally belong to the  $\alpha/\beta$ -hydrolase superfamily (Ollis *et al.*, 1992; Nardini & Dijkstra, 1999), in which different representatives have broad substrate specificities. HLDs are able to convert halogenated alkenes, ethers, alcohols, ketones and cyclic dienes. The haloalkane dehalogenase DhaA refers to the HLD-II subfamily. This subfamily merges multiple members of experimentally characterized haloalkane dehalogenases possessed identical catalytic pentad: a nucleophile – Asp, another catalytic acid – Glu with two halide-stabilizing residues Asn and Trp (Chovancova *et al.*, 2007). Moreover three members of this

subfamily as LinB from *Sphingobium japonicum*, DhaA from *Rhodococcus sp.* and DbjA from *Bradyrhizobium japonicum*, show excellent enantioselectivity for  $\alpha$ -bromoesters, and DbjA also shows high enantioselectivity with two  $\beta$ -bromoalkanes (Prokop *et al.*, 2010). This finding is very interesting because a study of these proteins is one of the main steps for further understanding and designing of selective enzymes, which may play an important role in enantioselectivity transformation.

To separate enantiomers is very problematic process as enantiomers have identical physical and chemical properties. For this reason, enantioselectivity is an important aspect in pharmacology. The mutant variant of DhaA12 was constructed with the aim to characterize and improve quality of selected haloalkane dehalogenase. DhaA12 mutant variant contains several amino acid residues typical for haloalkane dehalogenase DbjA that were introduced into DhaA enzyme by an insertion of loop <sub>139</sub>HHTEVAEEQDH<sub>149</sub> and eight additional mutations in both the active site and the access tunnel: W141F, P142A, F144A, G171R, A172V, K175G, C176G, V245A. Enantioselectivity of DhaA and its cumulative mutants were tested towards to 2-bromobutane, 2-bromopentane, 2-bromohexane, 2-bromoheptane and ethyl 2-bromopropionate using GC analysis. In this experiment the level of protein activity increased with the exception of the 2-bromobutane substrate but the enantioselectivity towards 2-bromopentane was not changed (Chrobakova *et al.*, 2006).

The engineering of DhaA12 mutant variant was to enhance the enantioselectivity of DhaA enzyme. The crystal structure of DhaA12 was obtained (PDB ID code 2V9Z; Emmer *et al.*, 2007), but this structure did not clarify to understand the changes in the structural features due to the low resolution of 3.00 Å. On the base of this fact the structure of DhaA12 measured to the high resolution of 1.78 Å is necessary for structure analysis and understanding the structural determinants of enantioselectivity.

## 4. Materials and methods

### 4.1 Construction of mutant, protein expression and purification

The recombinant gene *dhaA12* with loop insertion <sub>139</sub>HHTEVAEEQDH<sub>149</sub> and mutations W141F+P142A+F144A+G171R+A172V+K175G+C176G+V245A were constructed and recloned into expression vector pET21b. Recombinant plasmids were transformed to *Escherichia coli* BL21 (DE3). DhaA12 protein was purified to homogeneity using immobilized metal affinity chromatography. Eluted protein was dialyzed against 100 mM Tris-HCl buffer pH 7.5 without additives. Protein concentrations were determined by the Bradford method. SDS-PAGE was run using 15% polyacrylamide gels.

### 4.2 Protein crystallization

The purified DhaA12 mutant variant was crystallized using sitting-drop vapor-diffusion technique (Ducruix & Giege, 1999). Crystallization experiments have been performed in CombiClover Crystallization Plates (EBS plate, Emerald BioStructures, WA, USA). Needle-shaped crystals of mutant variant DhaA12 have been obtained using commercial crystallization screening kits such as kits of Hampton Research and Jena BioScience. The optimization procedure lasted 3 months and was performed at 277 K. The first single crystals appeared after 5 days using a 1:1 ratio composed of 8 mg/ml of protein in 50 mM Tris-HCl buffer at pH 7.5 and reservoir solution. The reservoir solution consisted of 20% PEG 4000 and 100 mM MES sodium salt at pH 6.1. Crystal's growth lasted for 4 weeks following by crystals dissolving. In after 3 days new crystals appeared in the same drops and grew in 2 weeks to the dimension of 0.32 x 0.12 x 0.07 mm.

### 4.3 Data collection and processing

Diffraction data set for DhaA12 mutant variant was collected from single crystal on the beamline 14.1 at the BESSY-II electron storage ring (Mueller *et al.*, 2012) operated by the Joint Berlin MX-Laboratory (Berlin-Adlershof, Germany). Crystal was mounted directly into nylon loop (Teng, 1990) and flash-frozen in a cold nitrogen stream at 100 K without additional cryoprotection. 100 image frames were recorded with an oscillation angle of 1° and a crystal-to-detector distance of 170 mm. The XDS program package (Kabsch, 1993; Kabsch, 2010*a,b*) was used and data were scaled with *Scala* from the *CCP4* program suite (Collaborative Computational Project, Number 4, 1994). Data-collection statistics for DhaA12 are summarized in Table 1.

### 4.4 Structure refinement

The structure of DhaA12 mutant variant was solved by the molecular replacement using *MOLREP* (Vagin & Teplyakov, 1997). The known structure of *Rhodococcus rhodochrous* haloalkane dehalogenase DhaA mutant was used as a template (PDB ID code 3FBW; Stsiapanava *et al.*, 2010). Rigid-body refinement was performed using *REFMAC5* (Murshudov *et al.*, 1997). Further DhaA12 structure was refined using *SHELXL* program (Sheldrick *et al.*, 2008). *Coot* (Emsley *et al.*, 2010) program was used for manual building of enzyme's parts. The isotropic ADP refinement was applied to whole protein, waters and for chloride ion, resulting in lowering of R and  $R_{\text{free}}$  characteristics from R = 0.251 and  $R_{\text{free}}$  = 0.284 to R = 0.176 and  $R_{\text{free}}$  = 0.207. The data-collection and refinement statistics are present in Table 1.

### 4.5 Model validation and deposition

The structural models with respect to the experimental data were assessed using the program *SFCHECK* (Vaguine *et al.*, 1999). Geometric parameters were validated using *Coot* (Emsley *et al.*, 2010) and *MolProbity* (Chen *et al.*, 2010). The residues of DhaA12 most favorable regions are 90.2%, two residues (catalytic residue Asp106, Leu282) are located in the generous allowed region and one (Asp167) is located in the disallowed region. All of these Ramachandran outliers have clear electron density. The structure and original diffraction intensities were deposited in the Protein Data Bank under PDB ID code 3SKO.

### 4.6 Tunnel calculation and visualization

The tunnels in DhaA wild type and DhaA12 mutant variant were calculated by *CAVER PyMOL* plugin (Petrek *et al.*, 1999) and visualized by *PyMOL* system (DeLano, 2002). The residue Asp106 from active site was used as starting point for calculation pathways.

## 5. Result and discussion

### 5.1 Overall structure of DhaA12 mutant

The overall structure of DhaA12 mutant variant is similar to the wild type DhaA haloalkane dehalogenase from *R. rhodochrous* as previously described Newman *et al.* (1999). The DhaA12 electron density was well identified for residues Nos. 4 – 311. The mutant variant of DhaA12 protein is composed from the main  $\alpha/\beta$ -hydrolase domain and the helical cap domain (Fig. 1). The main  $\alpha/\beta$ -hydrolase domain consists of eight-stranded  $\beta$ -sheet and six  $\alpha$ -helices, in which it is separated by the cap domain insertion between  $\beta$ -strand  $\beta_6$  and  $\alpha$ -helix  $\alpha_8$  ( $\beta_1$ - $\beta_2$ - $\beta_3$ - $\alpha_1$ - $\beta_4$ - $\alpha_2$ - $\beta_5$ - $\alpha_3$ - $\beta_6$ ,  $\alpha_8$ - $\beta_7$ - $\alpha_9$ - $\beta_8$ - $\alpha_{10}$ ; Chovancova *et al.*, 2007; Stsiapanava *et al.*, 2010). Insertion of loop  $_{139}\text{HHTEVAEEQDH}_{149}$  from DbjA protein is arranged in 141\_142 position of DhaA protein. This loop region extends the helical cap domain of DhaA12. The cap domain includes five  $\alpha$ -helices ( $\alpha_4$ - $\alpha_5'$ - $\alpha_5$ - $\alpha_6$ - $\alpha_7$ ) and the loop located in  $\alpha_4$ -helix. The active-site cavity is located



between the main and cap domains. The catalytic residue presents as a pentad typical for haloalkane dehalogenases: Asn41, Asp106, Trp107, Glu130, His283 (Stsiapanava *et al.*, 2010).

## 5.2 Analysis of active site

Active site of DhaA12 protein structure contains five water molecules and a chloride ion (Fig.2). The location of this chloride ion in the active site is consistent with the three-dimensional structure of three haloalkane dehalogenase's mutants DhaA04, DhaA14, DhaA15, which have been solved earlier (Stsiapanava *et al.*, 2010). The chloride ion position coordinates by hydrogen bonds with Asn41 N<sup>δ2</sup> and Trp107 N<sup>ε1</sup> of the canonical halide-binding site (Verschueren *et al.*, 1993). The distances between the chloride ion and the N atoms of Asn41 N<sup>δ2</sup> and Trp107 N<sup>ε1</sup> are 3.44 and 3.21 Å, respectively. The water molecules interact with Asp106 O<sup>δ1</sup>, His142 N<sup>δ2</sup> and the hydroxyl group of Tyr284 occupying a free space within the active site.

## 5.3 Analysis of access tunnels

All known haloalkane dehalogenases contain a variety of tunnels through which the active site of a single protein interacts with their environment. DhaA12 mutant variant as other haloalkane dehalogenases has several tunnels (Klvana *et al.*, 2009). The main tunnel of DhaA12 enzyme was mutational changed using eight point substitutions. The W141F, P142A, F144A, V245A are located in  $\alpha$ -helix  $\alpha4$  and formed non-polar residues. Other four mutations: non-polar A172V, K175G, C176G and polar G171R, are located in  $\alpha$ -helix  $\alpha5$ . In this case, the main tunnel is positioned between these two  $\alpha$ -helices  $\alpha4$  and  $\alpha5$  and it is wider than in DhaA wild type (Fig.3). It provides easier penetration of water molecules and consequently larger substrates, such as 2-bromohexane or 2-bromoheptane, into the active site. Moreover,  $\alpha$ -helix  $\alpha4$  includes insertion of loop <sub>139</sub>HHTEVAEEQDH<sub>149</sub> that forms an additional protein surface (Fig.4). Loop region characterizes great flexibility (Emmer 2007) and carries a number of alternatives: Glu140 and Glu149. The residue His152 is represented by two conformations and the C <sup>$\alpha$</sup>  – C <sup>$\beta$</sup>  – C <sup>$\gamma$</sup>  bond angle changes from 117.96° (alternative A) to 118.86° (alternative B). The occupancies of A and B alternatives are the same - 0.5 for both. In turn, the residue Glu149 interacts with His152 and the C <sup>$\beta$</sup>  – C <sup>$\gamma$</sup>  – C <sup>$\delta$</sup>  bond angle changes from 110.42° (alternative B) to 116.33° (alternative A). It is suggesting that the motions of these residues are involved in the partial lateral closure of the main tunnel.

## 6. Conclusions

The active site of *Rhodococcus rhodochrous* haloalkane dehalogenase DhaA12 mutant variant contains clear identified five water molecules and chloride ion that comes from protein buffer. The positions of the chloride ion and catalytic pentad in DhaA12 structure are conserved in all earlier described mutant haloalkane dehalogenase's structures (3FBW, 3G9X, 3FWH, Stsiapanava *et al.*, 2010; 3RK4, 4FWB, Lahoda *et al.*, 2011). Structural alignment of the DhaA wild-type and mutant variant DhaA12 shows the absolute homology between selected haloalkane dehalogenase's structures excluding only new loop region (Fig.4). The loop insertion <sub>139</sub>HHTEVAEEQDH<sub>149</sub> is localizing on the protein surface and carrying alternatives conformations are flexible region of the structure. The combined effect of the loop and alternative conformations of Glu149 and His152 amino acids is possible including in the process of active site pocket formation.

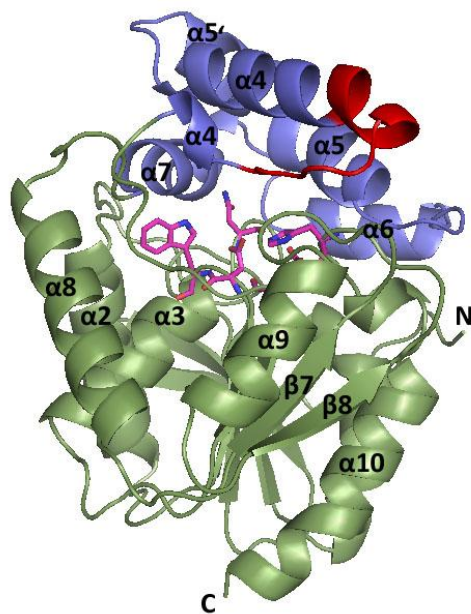
Comparative analysis of DhaA and DhaA12 structures demonstrates that the substitutions of W141F, P142A, F144A, G171R, A172V, K175G, C176G and V245A are changing the anatomy of the main tunnel to make it wider, expecting increased exchange of substrates in DhaA12. The

results of this work will be used for subsequent computational design and modeling of novel enzymes with enhanced enantioselectivity.

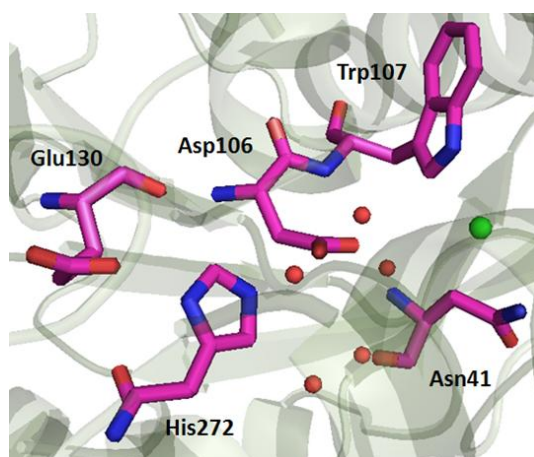
**Table 1** Data-collection and refinement statistics

	DhaA12
<b>Data collection</b>	
Beamline	14.1/BESSY
Detector	MAR Mosaic CCD 225 mm
Wavelength (Å)	0.918
Resolution range (Å)	31.39 – 1.78 (1.88 – 1.78)
Space group	$P2_12_12_1$
Unit-cell parameters (Å, °)	$a= 51.58, b= 68.69, c= 84.37$ $\alpha= \beta= \gamma= 90$
Unique reflections (observed)	29,443
Redundancy	4.1
Completeness (%)	100.0 (99.9)
$R_{\text{merge}}^\dagger$ (%)	8.4 (37)
$I/\sigma(I)$	14.4 (4.2)
<b>Refinement</b>	
No. of reflections used for refinement	34,658
Maximum resolution (Å)	1.78
$R_{\text{work}}^\ddagger/R_{\text{free}}^\S$ (%)	17.6/20.7
No. of non-H atoms	2,813
No. of protein atoms	2,537
No. of chloride ion	1
No. of water molecules	275
<b>R.m.s. deviations</b>	
Bond lengths (Å)	0.007
Bond angles (°)	0.024
<b>PDB code</b>	
	3SKO

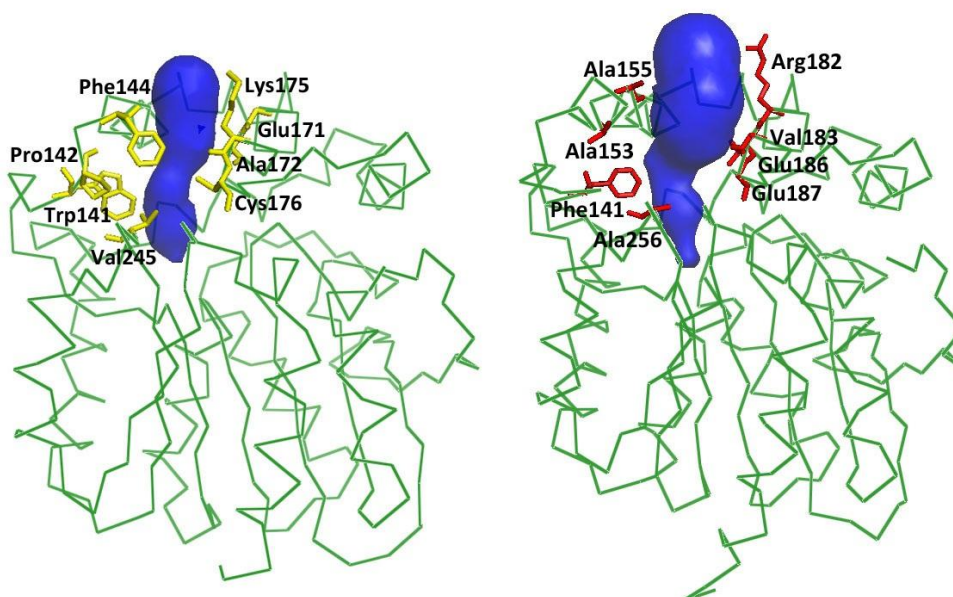
$^\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ , where  $I_i(hkl)$  is the intensity of the  $i$ th measurement of reflection  $hkl$  and  $\langle I(hkl) \rangle$  is the average intensity of the reflection.  $^\ddagger R_{\text{work}} = \sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum_{hkl} |F_{\text{obs}}|$ .  $^\S R_{\text{free}}$  was monitored using 5% of the reflection data that were excluded from refinement.



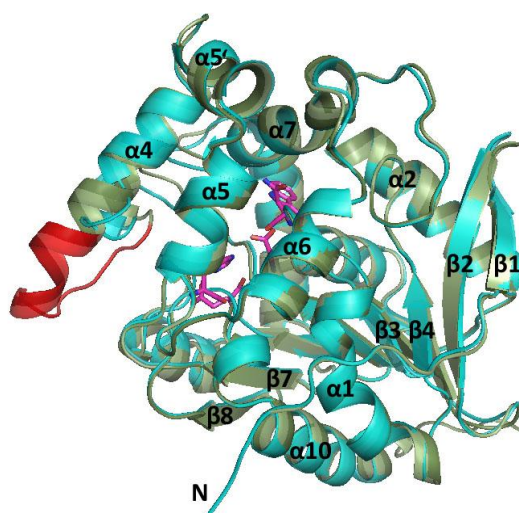
**Figure 1** Cartoon representation of the DhaA12 enzyme structure. The  $\alpha/\beta$ -hydrolase core domain is shown in green, the helical cap domain is shown in blue and insertion loop is shown in red. Catalytic pentad residues are shown in magenta stick representation. Figure was prepared using *PyMOL* (DeLano, 2002).



**Figure 2** The DhaA12 active site in stick representation. The catalytic pentad is shown in magenta, the green sphere represents chloride anions and red spheres represent the O atoms of water molecules.



**Figure 3** Ribbon representations of the structures of wild-type DhaA (left) and mutant variant DhaA12 (right) with the main tunnel shown in blue. The residues selected for mutagenesis are shown as yellow sticks. The mutated residues are shown as red sticks. The tunnels were calculated using *CAVER* (Petrek et al., 2006).



**Figure 4** Cartoon representation of the DhaA12 and DhaA structures alignment. The DhaA12 is shown in green. The DhaA wild type is shown in blue and insertion loop is shown in red. Catalytic pentad residues are shown in magenta stick representation.

**Acknowledgements** We thank the staff at BESSY II (Berlin) for assistance with data collection as well as Rolf Hilgenfeld (University of Lübeck) for enabling and arranging a short-term laboratory visit. This work has been supported by the Grant Agency of the Czech Republic project P207/12/0775. Support of the Ministry of Education of the Czech Republic and the Academy of Sciences of the Czech Republic is also appreciated. We would like to acknowledge support through the EC under ELISA grand (agreement number: 226716) for access to BESSY II.

## References

- Chovancova, E., Kosinski, J., Bujnicki, J.M., Damborsky, J. (2007). *Proteins* **67**, 305–316.
- Chrobakova, T., Prokop, Z., Grodecka, L. (2006). *Prog. Rep.: Protein Engineering Group*. Collaborative Computational project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- DeLano, W. L. (2002). The *PyMOL Molecular Viewer*. [Http://www.pymol.org](http://www.pymol.org).
- Ducruix, A., Giege, R. (1999). *A Practical Approach*. 2nd Ed. Oxford: IRL Press.
- Emmer J. (2007). Master's thesis. Structural studies of the haloalkane dehalogenase mutant (DhaA12) from *Rhodococcus rhodochrous*.
- Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
- Kabsch, W. (2010a). *Acta Cryst.* **D66**, 125–132.
- Kabsch, W. (2010b). *Acta Cryst.* **D66**, 133–139.
- Klvana, M., Pavlova, M., Koudelakova, T., Chaloupkova, R., Dvorak, P., Prokop, Z., Stsiapanava, A., Kutý, M., Kuta-Smatanova, I., Dohnalek, J., Kulhanek, P., Wade, C.R., Damborsky, J. (2009). *J. Mol. Biol.* **392**, 1339–1356.
- Mueller, U., Darowski, N., Fuchs, M.R., Förster, R., Hellmig, M., Paithankar, K.S., Pühringer, S., Steffien, M., Zocher, G., Weiss, M.S. (2012). *J. Synchrotron Radiat.* **19**, 442–449.
- Murshudov, G. N., Vagin, A. A., Dodson, E. J. (1997). *Acta Cryst.* **D53**, 240–255.
- Nardini, M., Dijkstra, B.W. (1999). *Curr. Opin. Struct. Biol.* **9**, 732–737.
- Newman, J., Peat, T. S., Richard, R., Kan, L., Swanson, P. E., Affholter, J. A., Holmes, I. H., Schindler, J.F., Unkefer, C. J., Terwilliger, T. C. (1999). *Biochemistry* **38**, 16105–16114.
- Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M., Harel, M., Remington, S. J., Petrek, M., Otyepka, M., Banas, P., Kosinova, P., Koca, J., Damborsky, J. (2006). *BMC Bioinformatics* **7**, 316.
- Prokop, Z., Sato, Y., Brezovsky, J., Mozga, T., Chaloupkova, R., Koudelakova, T., Jerabek, P., Stepankova, V., Natsume, R., van Leeuwen, J. G., Janssen, D. B., Florian, J., Nagata, Y., Senda, T., Damborsky, J. (2010). *Angew. Chem. Int. Ed.* **49**, 1–6.
- Sheldrick, G. M. (2008). *Acta Cryst.* **A64**, 112–122.
- Silman, I., Schrag, J., Sussman, J. L., Verschueren, K. H., Goldman, A. (1992). *Protein Eng.* **5**, 197–211.
- Stsiapanava, A., Dohnalek, J., Gavira, J.A., Kutý, M., Koudelakova, T., Damborsky, J., Kuta Smatanova, I., (2010). *Acta Cryst.* **D66**, 962–969.
- Teng, T.-Y. (1990). *J. Appl. Cryst.* **23**, 387–391.
- Vagin, A., Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
- Verschueren, K.H., Seljée, F., Rozeboom, H.J., Kalk, K.H., Dijkstra, B.W. (1993). *Nature* **363**, 693–698.

### **Dynamics and Hydration: Two Parameters to be Challenged by Rational Design of Enzyme Catalysts**

Reproduced with permission from

- I. Sykora, J., Brezovsky, J., Koudelakova, T., Lahoda, M., Chaloupkova, R., Chernavets, T., Fortova, A., Stepankova, V., Prokop, Z., Kuta Smatanova, I., Hof, M. and Damborsky, J., (2013) *prepared for submission to Nature Chem.*