

UNIVERSITY OF SOUTH BOHEMIA IN ČESKÉ BUDĚJOVICE  
FACULTY OF SCIENCE

# **Crystallographic studies of multimeric and ancestral haloalkane dehalogenases**

Ph.D. THESIS

Andrii Mazur, MSc

**Supervisor:** prof. Mgr. Ivana Kutá Smatanová, Ph.D.  
University of South Bohemia in České Budějovice, Faculty of Science

České Budějovice 2023



This thesis should be cited as:

Mazur, A. 2023: Crystallographic studies of multimeric and ancestral haloalkane dehalogenases. Ph.D. Thesis Series, No. 2. University of South Bohemia, Faculty of Science, České Budějovice, Czech Republic, 75 pp.

## **ANNOTATION**

Halogenated pollutants represent one of the greatest threats to nature. Scientists have been searching for a solution to degradation of such compounds and soil remediation for many years, and several suggestions have been completed until now. One of the ways is the biodegradation of halogenated compounds by bacterial enzymes, such as haloalkane dehalogenases, enzymes that cleave halogen bonds and degrade this kind of compounds. The wide substrate specificity of these enzymes pushes forward the research for new techniques to improve existing dehalogenases and revealing new features in unique properties of novel types.

In this work, X-ray crystallographic analysis was successfully used for structural-functional characterization of novel isolated and engineered haloalkane dehalogenases. The first part is devoted to the characterization of the novel haloalkane dehalogenase DpaA from *Paraglaciicola agarilytica* NO2. Unusual feature of this enzyme is its property to form multimeric construction, whereas other members of the same phylogenetic subfamily are present as monomers. We aim to expand the knowledge of the structure and formation of the multimeric haloalkane dehalogenases. The second part of the thesis deals with novel enzyme synthesized by use of the ancestral sequence reconstruction from dehalogenases with the

broadest range substrate specificity LinB and DmbA. Armed with a new technique the novel synthesized enzyme can enhance the effectiveness of the degradation of halogenated environmental pollutants.

**Declaration:**

I hereby declare that I am the author of this dissertation and that I have used only those sources and literature detailed in the list of references.

Ceske Budejovice,

.....

Andrii Mazur

**Financial support:**

The following funding is acknowledged: Grant Agency of the Czech Republic (GACR grant No. 17-24321S); DAAD mobility grant (grant No. DAAD-16-09); European Regional Development Fund-Project (ERDF grant No. CZ.02.1.01/0.0/0.0/15\_003/0000441) and Grant Agency of the University of South Bohemia (GAJU grant No. 17/2019/P)

## **Acknowledgements**

I would like to thank everybody, who supported me, during my PhD study and without them this work couldn't be possible.

Firstly, I would like to express my gratitude to my supervisor Ivana Kuta Smatanova, for her support and encouragement during all my study. I am very grateful for her patience, scientific and organizational advices, which was vital for doing the whole research.

My special thanks belong to Tatyana Prudnikova for her guidance and useful tips during my first steps into protein crystallography and Pavel Grinkevich for his help with data collections and analysis. Their help and support made a significant impact on my work.

My deepest gratitude belongs to my colleagues: Radka Chaloupkova, Jiri Damborsky, for providing me with protein samples and cooperation.

I would like to thank administration and technical staff of PrF for providing all necessary facilities to make my study and research easy and comfortable.

I want to thank all my friends from Nove Hradky and Ceske Budejovice: Saurabh Pandey, Vladyslav Bozhynov, Ann Platonova, Denis Birukov and many others, who made my life and work pleasant and productive.

Finally, I would like to express my grateful to my fiancé Oksana as well as to my parents for understanding and faith in me and helping me to make so many good decisions in my life.

*Dedicated to my family*

## List of papers and author's contribution

The thesis is based on the following papers:

- I. **Andrii Mazur**, Tatyana Prudnikova, Pavel Grinkevich, Jeroen R. Mesters, Daria Mrazova, Radka Chaloupkova, Jiri Damborsky, Michal Kutý, Petr Kolenko and Ivana Kuta Smatanova., (2021). The tetrameric structure of the novel haloalkane dehalogenase DpaA from *Paraglaciecola agarilytica* NO2. *Acta Cryst. D77*, 347–356. (IF = 5.699), DOI: <https://doi.org/10.1107/S2059798321000486>  
*Andrii Mazur performed the crystallization experiments, partial data collection, processed a part of the data, participated in writing of the manuscript. His contribution was 50 %.*
- II. **Andrii Mazur**, Pavel Grinkevich, Radka Chaloupkova, Petra Havlickova, Barbora Kascakova, Michal Kutý, Jiri Damborsky, Ivana Kuta Smatanova and Tatyana Prudnikova, (2021). Structural Analysis of the Ancestral Haloalkane Dehalogenase AncLinB-DmbA., *Int. J. Mol. Sci.*, 22(21), 11992. (IF = 6.208), DOI: <https://doi.org/10.3390/ijms222111992>  
*Andrii Mazur performed partial data collection, processed the data, participated in writing of the manuscript. His contribution was 50 %.*
- III. Tatyana Prudnikova, Barbora Kascakova, Jeroen R. Mesters, Pavel Grinkevich, Petra Havlickova, **Andrii Mazur**, Anastasiia Shaposhnikova, Radka Chaloupkova, Jiri Damborsky, Michal Kutý and Ivana Kuta Smatanova, (2019). Crystallization and Crystallographic Analysis of a *Bradyrhizobium Elkanii* USDA94 Haloalkane

Dehalogenase Variant with an Eliminated Halide-Binding Site., Crystals  
2019, 9, 375. (IF = 2.670)., DOI: <https://doi.org/10.3390/cryst9070375>  
*Andrii Mazur participated in part of the experimental work. His  
contribution was 10 %.*

## **Co-author agreement**

Ivana Kutá Smatanová, the supervisor of this Ph.D. thesis and co-author of all stated papers, fully acknowledges the stated contribution of Andrii Mazur to these manuscripts.

.....

Ivana Kutá Smatanová

# Contents

1. PROLOGUE AND AIMS OF RESEARCH .....	1
2. INTRODUCTION.....	3
2.1. Haloalkane dehalogenases .....	4
2.2. Structure of haloalkane dehalogenases .....	4
2.3. Multimeric nature of haloalkane dehalogenases.....	6
2.4. Catalytic mechanism of haloalkane dehalogenases .....	8
2.5. Haloalkane dehalogenases from marine organisms and their structural properties	10
2.5.1. <i>HLDs From Pollution-Degrading Microorganisms</i> .....	11
2.5.2. <i>HLDs From Symbiotic Microorganisms</i> .....	11
2.5.3. <i>HLDs From Genomic Databases and Metagenomic Libraries</i> .....	12
2.6. General properties of marine haloalkane dehalogenase.....	12
2.7. Quaternary Structure properties of marine haloalkane dehalogenases .....	13
2.8. Ancestral sequence reconstruction of haloalkane dehalogenases .....	14
3. MATERIALS AND METHODS.....	15
3.1. Protein crystallization .....	15
3.2. Methods of crystallization .....	17
3.3. Optimization .....	19
3.4. Crystal structure.....	20
3.5. X-ray diffraction .....	22
3.6. Solving structure and modeling .....	24
4. RESULTS AND DISCUSSION .....	29
4.1 The tetrameric structure of the novel haloalkane dehalogenase DpaA from Paraglaciecola agarilytica NO2 .....	29
4.2. Structural Analysis of the Ancestral Haloalkane Dehalogenase AncLinB-DmbA.	41
4.3. Crystallization and Crystallographic Analysis of a Bradyrhizobium Elkanii USDA94 Haloalkane Dehalogenase Variant with an Eliminated Halide-Binding Site.	56
5. CONCLUSIONS.....	69
6. REFERENCESS .....	71

## LIST OF ABBREVIATIONS

3D – Three-dimensional

Ala, A – Alanine

Arg, R – Arginine

Asn, N – Asparagine

Asp, D – Aspartic acid

ASR – Ancestral sequence reconstruction

Cys, C – Cysteine

DLS – Dynamic light scattering

FID – Free interface diffusion

Gln, Q – Glutamine

Glu, E – Glutamic acid

Gly, G – Glycine

His, H – Histidine

HLD – Haloalkane dehalogenase

Ile, I – Isoleucine

IR – Isomorphous replacement

Leu, L – Leucine

Lys, K – Lysine

MAD – Multi-wavelength anomalous dispersion

Met, M – Methionine

MIR – Multiple isomorphous replacement

MIRAS – Multiple isomorphous replacement with anomalous

MR – Molecular replacement

PDB – Protein data bank

PEG – Polyethylene glycol

Phe, F – Phenylalanine

Pro, P – Proline

SAD – Single-wavelength anomalous dispersion

SIR – Single isomorphous replacement

SIRAS – Single isomorphous replacement with anomalous

Trp, W – Tryptophan

RIP – Radiation damage-induced phasing

UV – Ultraviolet

Ser, S – Serine

Thr, T – Threonine

Trp, W – Tryptophan

Tyr, Y – Tyrosine

Val, V – Valine



# 1. PROLOGUE AND AIMS OF RESEARCH

---

With the rapid development of industry, the question of nature recovering becomes one of the most important topics nowadays. The remediation of soil after the use of industrial chemicals such as agricultural pesticides is one of the most urgent problems to be solved for society. The most common pesticide material is halogenated aliphatic compounds, which subsequently represents one of the largest groups of environmental pollutants. Several approaches of soil remediation have been proposed. One of them is the biodegradation of halogenated compounds by enzymatic reactions of bacterial origin. This reaction can be led by enzymes, which belong to a large group of  $\alpha/\beta$  hydrolases that are called haloalkane dehalogenases. They catalyze the cleavage of a carbon-halogen bond in halogenated compounds, resulting in the formation of a halide ion, a corresponding alcohol and a proton.

The structure of haloalkane dehalogenase as well as other proteins and macromolecules is a key point in the understanding their properties as well as biological mechanisms. There are several methods of structure determination have been developed up to date, one of them is X-ray crystallography. This is the most powerful method based on the determining the atomic structure of a protein crystal, in which the crystalline structure causes a beam of incident X-rays to diffract into many specific directions. Therefore, X-ray crystallography is used to determine the structure of novel types of haloalkane dehalogenases.

The aim of this research was to provide a structural analysis of the novel types of haloalkane dehalogenases to extend the knowledge of their nature and origin.

The specific research objectives were:

1. To determine the structure of DpaA protein, analyze its oligomeric state and compare DpaA with other selected proteins
2. To determine the structure of ancestral protein AncLinB-DmbA
3. To provide a comparative structural analysis of AncLinB-DmbA with its descendants LinB and DmbA
4. To provide a structural analysis of DbeA from *Bradyrhizobium Elkanii* USDA94 Haloalkane Dehalogenase Variant with an Eliminated Halide-Binding Site.

## 2. INTRODUCTION

---

The industrial revolution has enabled humanity to make a tremendous leap in development in the last few centuries. A wide variety of chemical compounds has been created for use in agriculture, industry, and on the battlefield. Unfortunately, many of them is very difficult to degrade, thus they pollute the biosphere and have enormous impact on human health. Therefore, many scientists are engaged in research on the degradation of these pollutants.

Halogenated organic compounds belong to one of the largest groups of environmental chemicals. The utility of these compounds can range from pesticides and soil fumigants (1,2-dibromoethane,  $\gamma$ -hexachlorocyclohexane, etc.) to solvents (1,2-dichloroethane, chloromethane, etc.) and industrial chemicals (vinyl chloride, 1,2-dichloroethane) [1-3]. The wide application of the above compounds and their high resistance to degradation make it difficult to abandon their use.

A number of methods have been developed to detect and reduce the toxicity of halogenated compounds, for example extraction, combustion, chemical degradation, sonochemical destruction, photochemical processes and bioremediation [4, 5]. Recent studies in microbiology have shown that some microorganisms have adapted to live in polluted areas and conditions [6, 7]. Further analyses revealed that these microorganisms contain enzymes capable of cleaving carbon-halogen bonds, and the products of the reactions are nontoxic compounds [6-9]. This makes microorganisms degrading xenobiotics very attractive for studying the evolution of enzymatic specificity and the construction of new metabolic pathways. Exploiting and engineering the catalytic potential of microbial enzymes lead to the development of new biological processes for the production of valuable chemicals.

## 2.1. Haloalkane dehalogenases

Haloalkane dehalogenases are enzymes (HLDs; EC 3.8.5.1) that catalyze hydrolytic cleavage of the carbon-halogen bond in various halogenated compounds into corresponding alcohols, halides, and protons. The HLDs have great potential applications in the detection and degradation of halogenated compounds such as 1,2-dichloroethane, 1,2,3-trichloropropane, 1,2-dibromoethane, etc. which are important in industry and the environment [9]. The HLDs can also be immobilized and engineered for higher activity, thermal stability and resistance to organic solvents. Other practical applications include detoxification, protein production, and cell imaging. The HLDs are also a good model for studying structure-function relationships between enzymes, making these enzymes an important component of cutting-edge research in biotechnology [8-10].

## 2.2. Structure of haloalkane dehalogenases

Structurally, haloalkane dehalogenase belongs to the  $\alpha/\beta$ -hydrolase superfamily and consists of 2 domains: the main (or core) domain and the cap domain (Fig 1). The core domain is conserved in different members of the  $\alpha/\beta$ -hydrolase superfamily and contains 8  $\beta$ -sheets surrounded by 6  $\alpha$ -helices and serves to maintain the arrangement of

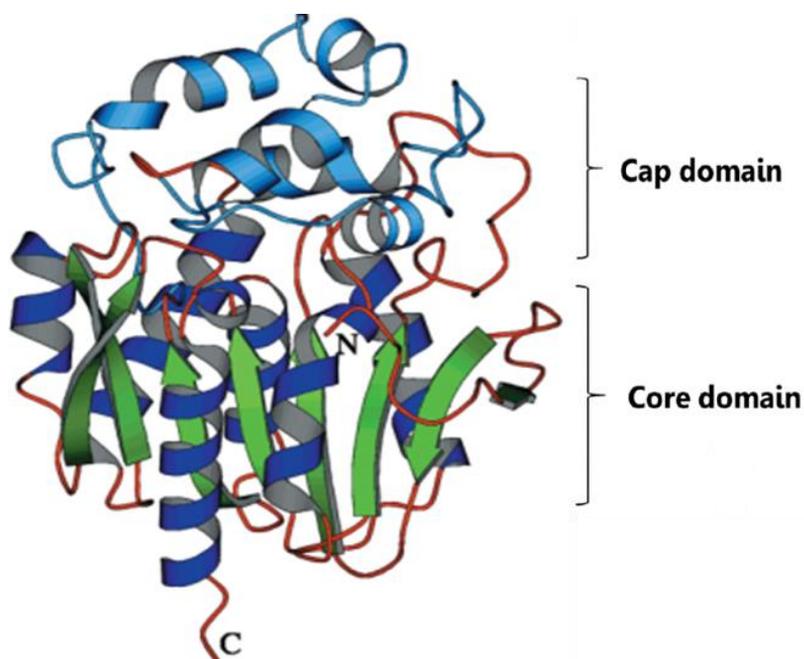


Figure 1. Overview of the tertiary structure of LinB.  $\alpha$ -Helices,  $\beta$ -sheets and loops forming the core domain are shown in dark blue, green and red, respectively; while the helices and loops from the cap domain are colored in light-blue [11].

the catalytic residues [11]. The helical cap domain is located at the top of the core domain and consists of multiple helices and loops. The cap domain is more flexible than the core domain and displays certain motions. It is known that this fact influences the substrate specificity of the enzymes [12].

The active site is located between two domains in a cavity composed of hydrophobic residues. It contains a catalytic triad of three amino acids (nucleophilic aspartate, basic histidine, catalytic aspartic acid) involved in the dehalogenated reaction. In addition, this cavity contains a binding site, which consists of two amino acids (tryptophan-tryptophan or tryptophan-asparagine) stabilizing halogen atom of the substrate during the dehalogenation reaction [13].

The phylogenetic analysis revealed that haloalkane dehalogenase is divided into three subfamilies: HLD-I, HLD-II and HLD-III [14]. They revealed some sequence and structural differences mostly in the cap domain. Also, great differences are presented in the catalytic pentad (Fig. 2).

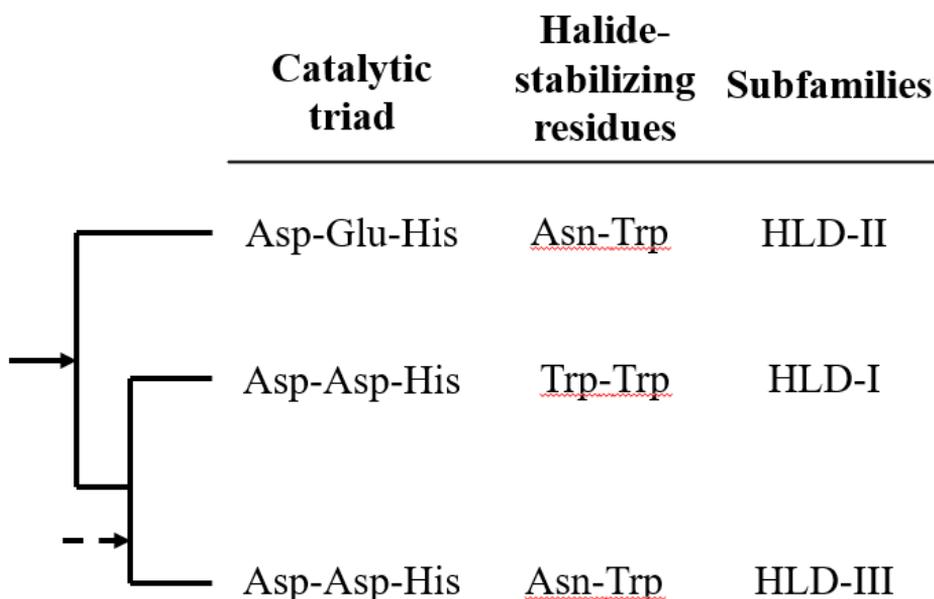


Figure 2. The catalytic pentad within the haloalkane dehalogenase family. The tree is rooted based on the results of outgroup analysis and its probable root is indicated by the solid arrow. An alternative root position is indicated by the dotted arrow. Adapted from [14].

### 2.3. Multimeric nature of haloalkane dehalogenases

Proteins in nature have tendencies to multimerization [15]. These molecules are formed into a symmetrical higher-order complex composed of subunits encoded by the same genetic locus. Depending on the numbers of subunits, such complexes are referred to as dimers, trimers, tetramers, etc. [16]. Such formation in nature gives proteins several advantages, which might be the factor of the evolution of this form

[16]. It is easier to fold multiple small proteins. Also the encounter rate of an enzyme and a small substrate is proportional to the effective radius of the enzyme [17]. A quite important advantage of multimeric formation is a more stable construction, which can decrease potential denaturation or other interactions due to a smaller surface area to volume [18].

In case of enzymes a good feature of oligomers might be the potential higher substrate specificity of catalytic sites due to their lower sensitivity to internal motions in higher-order structures rather than in monomers [16].

However, most of haloalkane dehalogenases are presented in nature as monomers. The quaternary structure among three subfamilies shows, that haloalkane dehalogenases mostly presented as monomers, with several dimers in HLD-II and HLD-III and some oligomeric states in HLD-III [14, 19]. The HLD-I subfamily dehalogenases were characterized only in monomeric state [20-22]. The dimer interaction has been discovered in haloalkane dehalogenases from organisms of various origin [23-27]. It may occur in dehalogenases from marine organisms, such as DmxA from *Marinobacter* sp. (PDB entry 5mxx; [23]), HanR from *Rhodobacteraceae* sp. (PDB entry 4brz; [24] ) or a DmmA from a metagenomic DNA of a marine microbial consortium (PDB entry 3u1t; [25]). Also, several dehalogenases with dimer interactions were identified in soil microorganisms: DccA from *Caulobacter crescentus* (PDB entry 5ers; [26]), DbeA from *Bradyrhizobium elkanii* (PDB entry 4k2a; [27]). All mentioned dimers do not share the same pattern between interacted domains. Some of them interact between cap domains, some between main domains.

On the other hand, several haloalkane dehalogenases show tendency to form a multimeric conformation. Haloalkane Dehalogenases DrbA and DmbC from HLD-III subfamily revealed a multimeric form on a size exclusion chromatography, but its structure has been not identified up to date [14, 19]. Quite interesting

properties have been found in DbjA from *Bradyrhizobium japonicum* USDA110. This enzyme shows high tolerance to pH changes, moreover its oligomerization is pH-dependent: monomer, dimer, tetramer and a high molecular weight cluster of the enzyme were distinguished in solution at different pH conditions [28]. The fact, that there is not so much data on multimeric haloalkane dehalogenases might be due to a difficulties of solving crystal structure of this oligomeric enzymes.

#### 2.4. Catalytic mechanism of haloalkane dehalogenases

The catalytic reaction of haloalkane dehalogenases is similar among each other and can be described by DhaA *R. rhodochrous* (Fig. 3) [29]. It is carried out in the active site by a catalytic triad (Asp-Glu-His) and two tryptophan residues (Asn-Trp) forming a halide and substrate binding site. This mechanism is provided in four steps: (i) substrate binding in the active site; (ii) nucleophilic attack of aspartate on the substrate, leading to the formation of a halide anion and an alkyl enzyme intermediate; (iii) nucleophilic addition of water activated by His and formation of an ester intermediate; (iv) release of the reaction products.

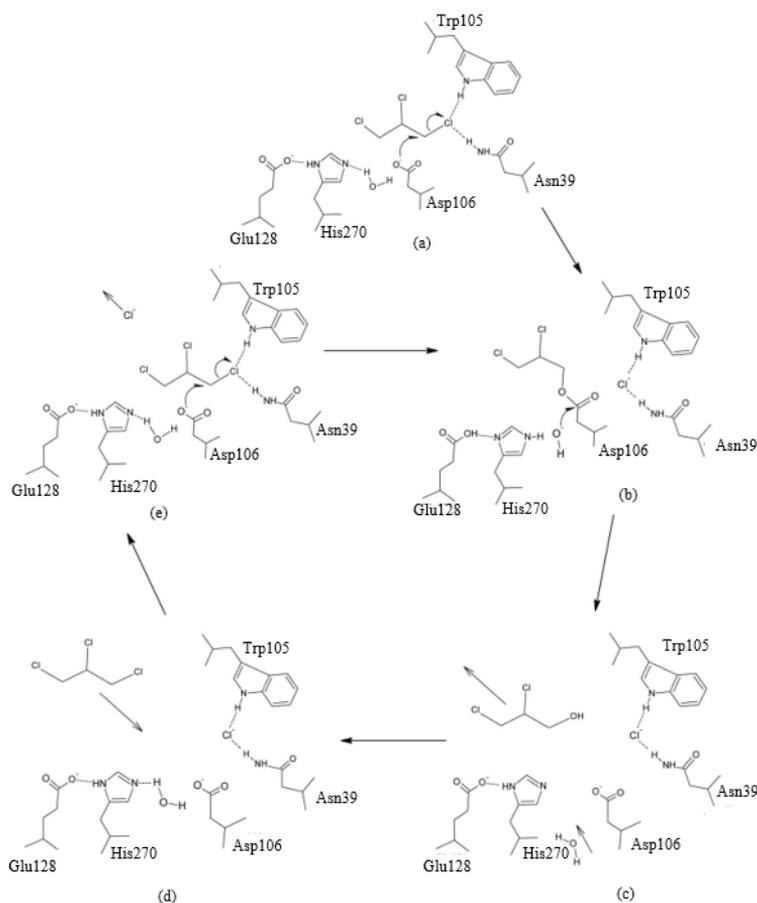


Figure 3 The catalytic mechanism of haloalkane dehalogenase DhaA from *R. rhodochrous* [29].

More specifically, the dehalogenation reaction is initiated when the substrate binds the active site of the enzyme through its halogen to the nitrogen atoms of two halide-stabilizing amino acid residues (Fig.3a). As a result, the halogen-bound  $sp^3$ -hybridised carbon atom of the substrate is placed toward the nucleophilic amino acid residue Asp. Next, one of the oxygen atoms of the Asp side chain attacks the halogen-bound carbon atom of the substrate (Fig.3b). This leads to the formation of a halide anion and an alkyl enzyme intermediate that remains associated with the side chains of two halide-stabilizing amino acid residues (Fig.3c). The second carboxylic oxygen remains free and its negative charge [30].

The third step of reaction consists of hydrolysis of the covalent alkyl enzyme intermediate by presenting water in the active site. The His base extracts a proton from the catalytic water molecule (Fig. 3d). The remaining hydroxyl anion attaches to the carboxyl carbon atom of Asp nucleophile side-chain, which form the so-called tetrahedral intermediate. To stabilize the positive charge of the His-imidazole ring, the third acid of the catalytic triad Glu bonds its side-chain oxygen to His. Then the tetrahedral intermediate breaks the bond with Asp, leading to the formation of the alcohol, while the hydroxyl anion being a source of oxygen and protons. The protonated His base could be an alternative source of protons. Finally, all three products alcohol, halide ion, and proton are released, highlighting the end point of the reaction [31, 32].

Described reaction mechanism is similar among all HLDs, however, the rate-limiting step in catalytic cycle might differ. In the case of DhIA reaction with 1,2-dichloroethane this is the halide release [33]. In another example: during the dehalogenation of 1,3-bromopropane by DhaA [32] the rate-limiting step is a release of alcohol and for dehalogenation of 1-chlorohexane and bromocyclohexane by LinB [34] it is hydrolysis of the alkyl-enzyme intermediate. The factors influencing the rate of the reaction are the catalytic residues composition, the active site cavity geometry, and number of access tunnels connecting the deeply buried active site cavity with the protein surface [31, 34, 35].

## 2.5. Haloalkane dehalogenases from marine organisms and their structural properties

One of the most unusual features of planet Earth is the amount of water – indeed, more than 70% of the planet surface is covered by ocean. The life occurred in the sea, and until these days, marine organisms represent the most diverse, but the least investigated biological group. Biochemical processes that take place in marine environment are of a particular interest. Apparently, the water hides not only living organisms, but also undiscovered biochemical reactions, some of which would be

potentially applied for the benefit of humankind. HLDs are no exception, as about 40% of haloalkane dehalogenases were first identified in the marine environment [20].

HLDs originating from marine organisms expose some interesting properties. Some marine HLDs show a high level of activity toward a broad spectrum of substrates, while the others are characterized by very high stability (including the highest stability of all wild-type enzymes). Selective conversion of racemic mixtures of valuable synthons  $\alpha$ -bromoalkanes and  $\beta$ -bromoesters was also reported [23, 36, 37]. Therefore, investigation of the catalytic properties of marine HLDs is a source of multiple discoveries in the field of biochemistry.

The HLDs initially discovered in marine organisms were further identified in a different types of living organisms beyond the ocean, such as in organisms degrading volcanic products or oil industry side products, as well as in some symbionts [25]. Thus, different approaches of HLDs investigation were employed, depending on their origin.

#### *2.5.1. HLDs From Pollution-Degrading Microorganisms*

Ocean pollution is one of the hallmarks of the anthropogenic footprint on our planet. Recently, several bacteria have been discovered that are able to utilize polluted compounds as a source of carbon and energy through a specialized enzymatic pathway. The adaptation of HLDs for biodegradation of halogenated compounds in the ocean provides a valuable direction for their investigation [38, 39].

#### *2.5.2. HLDs From Symbiotic Microorganisms*

Organohalogen producing organisms are often associated with symbiotic microorganisms. These microorganisms may use dehalogenase enzymes to degrade some of the toxic halogenated bioproducts of the host [40].

### 2.5.3. HLDs From Genomic Databases and Metagenomic Libraries

A fast development of sequencing technologies, leading to a wide number of sequences deposited in a GenBank allow us to study HLDs from marine organisms without their *in vitro* cultivation. Thus, metagenomic libraries help to overcome a challenge of marine microorganisms cultivation. Moreover, genomic databases and metagenomic libraries became a valuable tools for investigation, taking in account low similarity in amino-acid sequences of some HLDs [41].

Combination of proteomic and genomic analysis allowed identifying novel enzymes isolated from different marine habitats that catalyze a wide range of reactions in marine microenvironment [42-44].

## 2.6. General properties of marine haloalkane dehalogenase

Studies devoted to biochemical and biophysical characterization of HLD enzymes from marine organisms reported several properties common among HLDs. More importantly, several specific features in their enzyme activity and their physical and structural properties were revealed.

Several types of marine HLDs show a broad range of substrate specificities, similarly to other characterized HLDs [36, 45], however, their preference for brominated and iodinated substrates over the chlorinated compounds is uncommon among other HLDs [20].

The catalytic efficiency of marine HLDs seems to be similar to that of other characterized HLDs [46]. Nevertheless, the protein engineering allows reaching desired efficiency if the improvement of catalytic activity is required.

Interestingly, HLDs identified from marine organisms are exceptionally stable in various harsh environments, including high temperature and critical pH conditions. Further investigation of this unique feature in combination with the use

of protein engineering seems promising for a stable biocatalyst's generation [35, 46, 47].

## 2.7. Quaternary Structure properties of marine haloalkane dehalogenases

It is known that haloalkane dehalogenases from HLD-I subfamily are presented as monomers. Some dehalogenases from HLD-II subfamily can form dimers. Enzymes of HLD-III group represent an interesting exception, as they form high-order multimeric structures [19]. DbjA, on the other hand, can display various states (monomer, dimer, or a tetramer) depending on pH [28]. Marine dehalogenases, apparently, behave in a similar way. For example, DspA exists as a dimer in the absence of salt, but in the presence of 0.15 M NaCl displays predominantly monomeric form [37]. DmmA was reported to dynamically switch between a monomer and dimer states, although dimeric form is predominant [25, 36]. The DmxA also forms a dimer, connected by a disulfide bridge between surface-exposed C294 residues [23].

The DpaA, which was isolated from the psychrophilic and halophilic bacterium *Paraglaciecola agarilytica* NO2 and was found in marine sediment collected from the East Sea near Korea shows unique quaternary structure. This protein belongs to the HLD-I subfamily, however represents dimeric and tetrameric formation [20, 48]. It is unusual not only for HLD-I subfamily, but also for the HLDs from the marine organisms. The DpaA enzyme shows high enantioselectivity towards a racemic mixture of ethyl 2-bromopropionate and moderates enantioselectivity towards 2-bromopentane [21]. This feature shows a great potential for the further investigations.

## 2.8. Ancestral sequence reconstruction of haloalkane dehalogenases

Ancestral sequence reconstruction (ASR) is the method of determination of ancient protein sequences based on the employment of existing protein sequences. ASR combined with the experimental characterization of the corresponding proteins allows obtaining the most reliable results. Such analyses allow characterizing ancestral enzymes dating back to the millions of years, while deducing properties of the analogous organism. Furthermore, ASR enables the identification of amino acid residues that play a central role in the protein function, the aim often impossible to archive by only comparing extant proteins. Further applications include the longevity of mutations, the contribution of gene duplications to enzyme function, and the evolution of protein complexes [49, 50].

ASR method consists of three steps: (1) inference of an ancestral sequence based on a comparison of homologous amino acid sequences; (2) biochemical synthesis of a gene encoding the amino acid sequence of interest; and (3) gene expression *in vitro* in a host microorganism such as *Escherichia coli* [51].

There are several studies of ancestral reconstruction of HLDs [52-54]. They have revealed several advantages of this method, such as improved thermostability or broader substrate specificity with possible additional substrates involving in the reaction. Proteins reconstituted in this manner are often highly thermostable [51]. In addition, it is very likely that the thermostability of a natural protein can be increased by replacing putative ancestral amino acids in natural enzymes without compromising the catalytic activity [51]. Other studies were focused on the substrate specificity [55]. Because ancestral enzymes are exhibiting broad substrate specificity, but with a low catalytic efficiency, They can be a good target for a potential protein engineering [54]. Although mentioned studies have improved our understanding of the properties of HLDs, they did not provide any insights into ancestral proteins structure that is required for the next steps of their exploration and application, including protein engineering.

# 3. MATERIALS AND METHODS

---

## 3.1. Protein crystallization

Proteins consist of long polypeptide chains and are constructed of 20 different basic amino acids. These polypeptide chains form a secondary and tertiary structure of the protein. Despite the fact that the tertiary structure gives the protein molecule a globular shape, nevertheless the protein molecules actually have a rather unusual shape and are not ideal for crystal's formation. Therefore, protein crystals are fragile, soft and sensitive to any kind of external influences. However, it is not clear why some proteins are relatively easy to crystallize and others not. The principles of crystal growth have been the subject of intense research for many years [56]. As a result, the theoretical and practical aspects of crystallization of molecules such as salts or small organic compounds are now well known.

The mechanism of protein crystallization is much more difficult and can be explained as follows: When the concentration of the protein is higher than the limit of its solubility, the protein precipitates or (under certain conditions) goes into a supersaturated meta-stable phase and nucleate to a crystal. The protein must overcome an energy barrier to nucleate and form a crystal nucleus. The high-energy intermediate is the "critical nucleus" that seeds crystal growth [56]. To obtain the supersaturated phase, it is necessary to increase the effective protein concentration. For this purpose, the special mixture "precipitant solution" is used. It may be salt, polyethylene glycol or other compounds that remove water from the solution and allow protein to crystallize. The conditions required for crystallization also include any other factors, such as pH, temperature, type of precipitant, and concentration of protein and precipitant.

It is known that there are three phases of the crystallization process common to all macromolecules: nucleation, crystal growth and end of crystal growth (Fig. 4).

During nucleation, the molecules associate and form thermodynamically stable aggregates. These nuclei form directions for further bonding to other molecules and crystal growth [57].

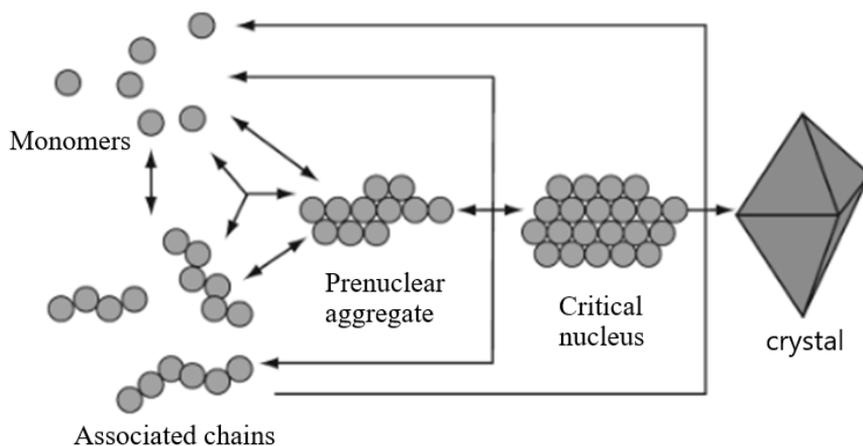


Figure 4. Reversible protein association from monomers to associated chains and aggregates forms into crystals [56].

The process of protein crystallization is described by phase diagram (fig. 5). This diagram illustrates the changing of crystallization parameters at different states of the macromolecule under specific conditions.

In precipitation zone, with a high concentration of macromolecules, precipitation of the amorphous precipitate of protein is observed. In the nucleation zone, where concentration is lower, the nucleation centers are formed and decrease the concentration of macromolecules and transfer to the metastable zone. In this case, only existing nuclei growth and make crystals. Then solution became undersaturated and no crystal grows, no nucleation occurs, and the protein passes into a dissolved state.

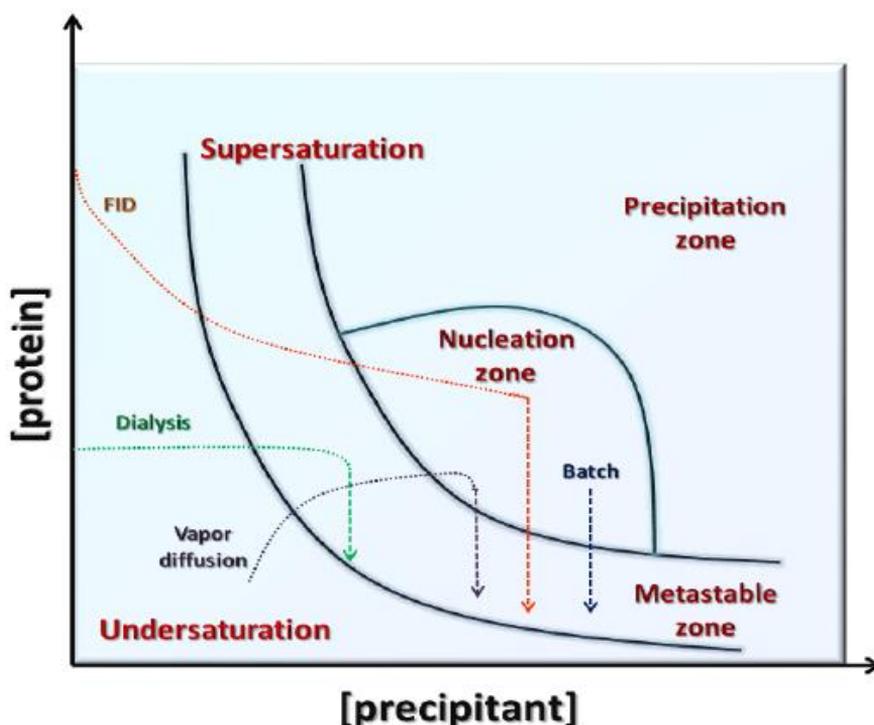


Figure 5. A simplified protein crystallization phase diagram. The different routes of reaching nucleation and metastable zones for the four main crystallization techniques are also shown [58].

### 3.2. Methods of crystallization

Nowadays, a number of techniques is used to crystallize macromolecules. The most common methods are vapor diffusion, microbatch, dialysis, and free interface diffusion (FID) [59-63]. Figure 5 shows how these techniques reach to the supersaturated state, nucleation, and growth of crystals. The most common techniques required for this action are shown in Figure 6.

The principle of the vapor diffusion method (Fig. 6) is based on the properties of water to evaporate and equilibrate the system. The droplet containing the protein and precipitant is placed on a cover slide (in the case of a hanging drop) or a base (in the case of a sitting drop) over the reservoir containing the precipitant solution.

The water evaporates from the drop and increases the concentration of the protein and precipitant, reaching the supersaturated state [59].

The microbatch method (Fig. 6) is a method, in which protein and precipitant are mixed and placed under the inertial oil to prevent evaporation. In this case, the supersaturated state is reached immediately after mixing of protein with the precipitant, and during crystallization, the crystallization concentration of the protein decreases and goes into metastable zone (Fig. 5) [59, 61].

In the microdialysis method, a semi-permeable membrane is used (Fig. 6). Small molecules such as salts, additives, and other precipitants can pass through the membrane with different porous size that is not permeable for biological macromolecules. The crystallization reagent diffuses out of or into the sample at a constant sample concentration, allowing the protein to reach the supersaturated state (Fig. 5). The advantage of this method is that the solution can be easily changed until the correct conditions are no longer found [59].

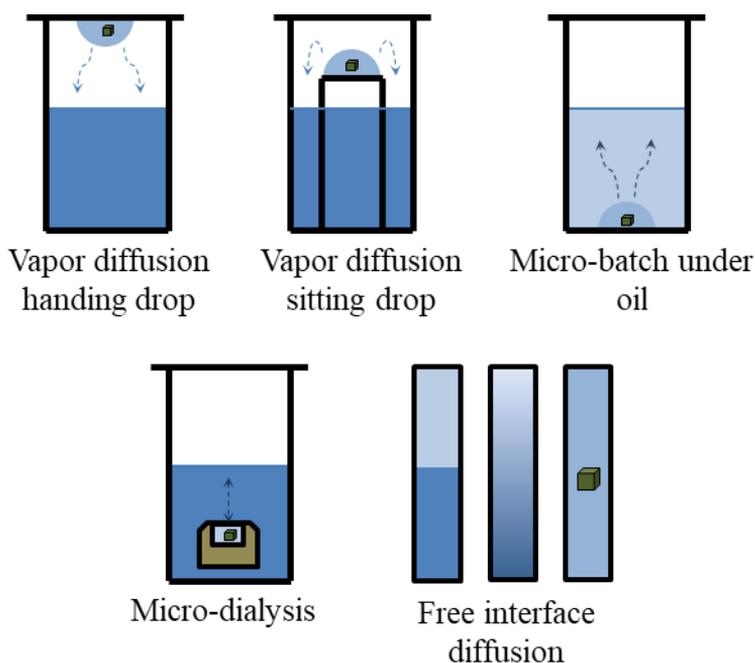


Figure 6. The schematic illustration of crystallization techniques: the most common hanging drop vapor diffusion and sitting drop vapor diffusion, microbatch under oil, microdialysis and counter diffusion (free interface diffusion). Adapted from [61].

The principle of the free interface diffusion method is based on the diffusion of the protein through the precipitant in a small glass capillary that creates a concentration gradient (Figs. 4, 5). The two solutions diffuse against each other creating the supersaturated gradient along the length of capillary. It helps to find the best ratio and concentration of precipitant and protein for growing of the largest crystals [62, 63].

### 3.3. Optimization

The first crystallization screening does not often lead to getting well-diffracted crystals. Many not good-shaped crystals can be observed under stereomicroscope, but they are not good enough for diffraction data measurement. Thus, different optimization strategies are used for improving crystals to prepare them in better quality.

The first step to grow better crystals should be variation parameters such as precipitant, pH, concentration of protein or precipitant and using different crystallization methods. Temperature dependence is also one of the sensitive conditions, which changes from one protein to another, so it is necessary to use different temperature regimes to grow better crystals [59].

The pH is also one of the most powerful and effective ways of inducing crystallization. Several proteins can be crystallized in fact in the absence of any precipitating agent simply by the very precise manipulation of pH [64].

The other strategy is using the additives. Additive is a chemical compound that promotes protein crystallization. Additives can be already present in the precipitant solution or extra-added to the mixture solution. Additives are classified into metal cations, small alcohols, detergents, reducing agents, etc. Metal cations

can improve crystallization by stimulation intermolecular interactions [65]. Detergents act as stabilizing agents and increase the solubility of the protein [65]. Reducing agents stabilize the protein by preventing cysteines oxidation and these are frequently added into protein during the protein preparation [66]. Moreover, physiological ligands, cofactors, substrates and inhibitors can function as stabilizing agents and improve crystal quality compared to apoprotein.

The third strategy for crystal optimization is using of seeding method. The seeding methods improve already existing crystals, by removing them from their nucleation solution (with high level of supersaturation), to the solution where supersaturation is lower. Crystals in this solution cannot nucleate anymore, but they still stay in metastable zone and move to the position on the phase diagram, where only crystal growth occurs. Therefore molecules of protein gather around seeded crystals, which increase their size and quality [67]. This method is divided into two types: micro- and macro- seeding. In macroseeding, already existing crystals are moved into a new solution. Microseeding is performed by transfer of crystal fragments crushed by vortex, seed beads, glass rods or other tools to the solution. A modification of microseeding is streak seeding, where seeds are added in solution by streaking with whisker or fiber through the new drop [68].

### 3.4. Crystal structure

A crystal structure consists of an arrangement of molecules, which are kept together by weak non-covalent interactions. A crystal is a solid, in which the atoms are arranged in a lattice constructed of identical block called “unit cell”. The unit cell is the parallelepiped shaped block that contains the “unit of pattern” that repeats throughout the whole crystal by translation along the directions of the parallelepiped edges (Fig. 7) [69]. It is described by three distances  $a$ ,  $b$ ,  $c$  and 3 inter-axial angles  $\alpha$ ,  $\beta$ ,  $\gamma$ .

Recall that the unit cell of a crystal is the smallest 3D geometric figure that can be stacked without rotation to form the lattice. The asymmetric unit is the smallest part of a crystal structure from that the complete structure is built using space group symmetry. The asymmetric unit may consist of only a part of a molecule, or can contain one or more molecules, if the molecules are not related by symmetry [70]. There are only rotation and translation or screw axes symmetry operations that are allowed in protein crystallography.

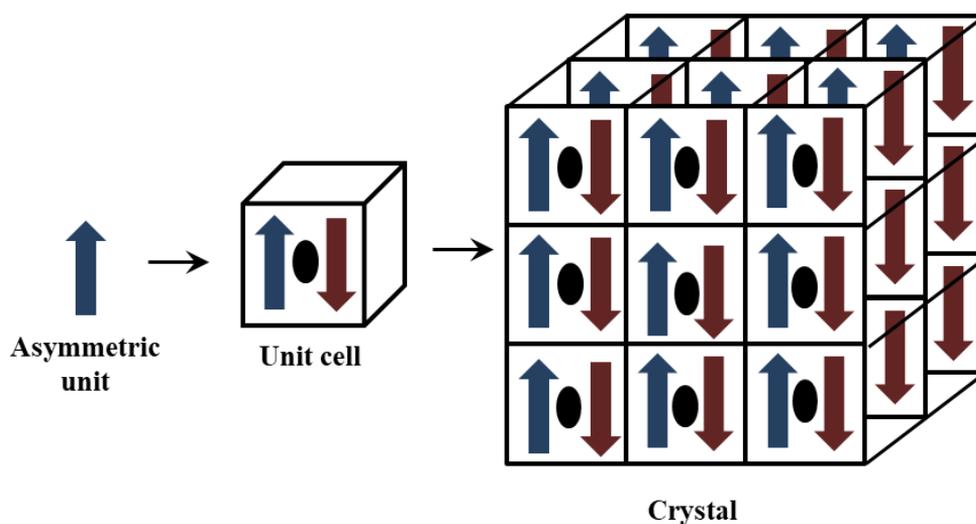


Figure 7. The arrangement of asymmetric unit into a unit cell and unit cell into whole crystal. The asymmetric unit (blue and red arrows) produces the copy by  $180^\circ$  rotating around the two-fold crystallographic symmetry axis (black ellipse) and forms into the unit cell, which is a brick of the whole crystal. Adapted from [70].

The internal symmetry of the unit cell is described by “space groups”. They are motifs, in which protein molecules are arranged in a 3D crystal and belong to 7 crystallographic systems (triclinic, monoclinic, orthorhombic, trigonal, tetragonal, hexagonal and cubic) [71]. There are 230 possible space groups, but biological macromolecules can crystallize only in 65 of them. The space group is found from

the diffraction pattern by looking for both symmetry in the intensity of the reflections (which give the rotational symmetry) and for missing reflections (which give the translational symmetries) [72, 73].

### 3.5. X-ray diffraction

X-rays are type of electromagnetic waves with wavelength range from  $10^{-3}$  to  $10^{-2}$  Å and high frequency, which make them easy to interact with the electrons of matter through oscillating electric field. The role of magnetic part of X-ray waves is lower, than electric and is usually neglected in measurements and calculations.

There are three mechanisms of X-ray interaction with electrons: elastic (Rayleigh) scattering, inelastic (Compton) scattering and absorption of X-rays by the electrons of an atom. Elastic or Rayleigh scattering is a pure scattering of photons without any deposits of energy to the electrons, but the direction of distributed changes. It makes the diffraction pattern for structure determination. Compton scattering has more energy than electron bonds and releases electron from the atom. After interaction photon and electron, the energy is loosed and moved on the reflected direction with other frequency and higher wavelength. This scattering increases the noise of the diffraction pattern. Absorption occurs, when the energy of photon is the same as the energy of electron in the atom. In this case photon gives it's all energy to electron, releasing it from electron shell. The shell can be filled by the electron from the higher level electron shell [70].

The reflected wave is found if the scattered wave and angle are known. A signal from the reflected wave is caught by X-ray detector if the waves have the same phase and interference occurs. In this condition, it is possible to calculate the structure of matter, which beam goes through, by using the Bragg's law equation and collect data from many angles.

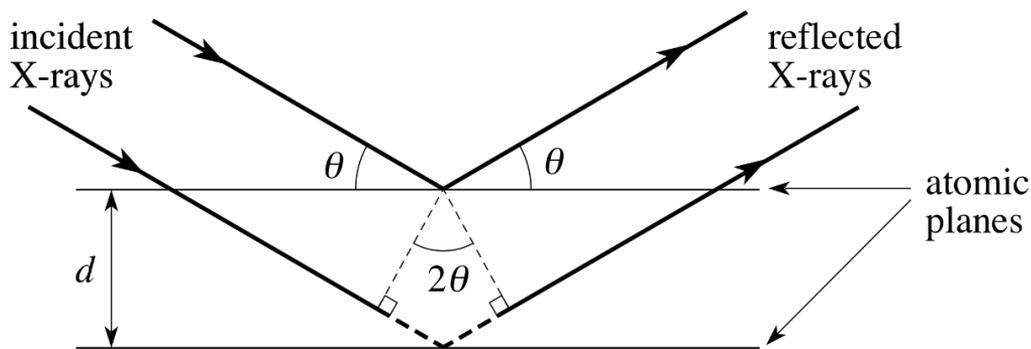


Figure 8. A schematic representation of the incoming X-ray photons scattering on electrons of atoms arranged in a crystal.

The equation for constructive interference is Bragg's law:

$$n\lambda = 2d \sin \theta$$

where  $\lambda$  is wavelength of the incoming X-ray,

$n$  is an integer,

$d$  is the distance between planes,

$\theta$  is the angle between incoming (scattered) wave and plane,

$2d\sin\theta$  is the path difference between waves [74].

The technique to obtain information from a crystal to diffraction pattern is called data collection. It is the last experimental step in crystallography [75]. To collect data from a crystal an intense source of X-rays (lab diffractometer or synchrotron source) is used. It starts from mounting a crystal to a cryoloop. Sometimes the mother liquor is needed to prevent total evaporation. A low temperature of experiment is used to minimize radiation damage and to prolong the crystal life, but it requires cryostream, attached to a goniometer or cooling a cryoloop before measurements. This needs to add a cryoprotector in mother liquor, to minimize damage from ice formation. After that, the cryoloop is mounted to a rotating goniometer, which is placed between X-ray source and detector (fig. 9). X-rays scatter by the electrons of every atom in the crystal in many directions, interfere with each other and produce the diffraction pattern of regularly arranged spots (reflections) on the detector. The intensities of measured spots contain the

information about the arrangement of molecules within a protein crystal in atomic detail [69].

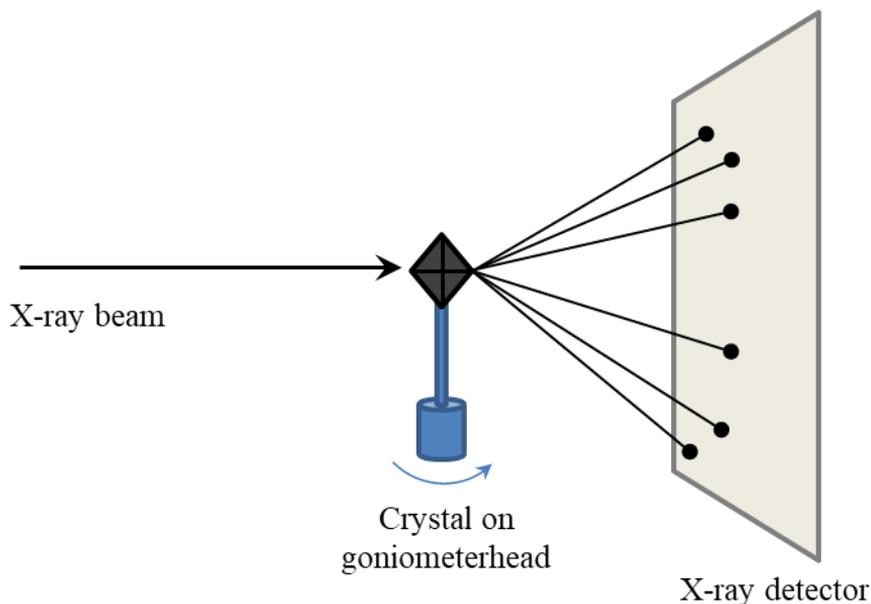


Figure 9. A schematic representation of X-ray diffraction experiment.

To obtain complete data it is necessary to measure as many reflections as possible. This is achieved by the rotation of the crystal around single axis with small angle during experiment. The result of the data collection is a set of diffraction images with recorded intensities for each single reflection and the position of reflections.

### 3.6. Solving structure and modeling

The main goal of the whole experiment is to calculate an electron density map (distribution of electrons in space). It allows obtaining the coordinates of atoms in the protein molecule. The data from the diffraction pattern are characterized by a structural factor. Structure factor ( $F_{hkl}$ ) is a complex number that represents the total scattering from all atoms of the unit cell and can be mathematically expressed

through Fourier transformation, from which the electron density for any point in the unit volume  $\rho(xyz)$  can be calculated [69]:

$$\rho(xyz) = \frac{1}{V} \sum_{hkl}^{\infty} (|F(hkl)|) \cdot e^{-2\pi i[hx+ky+lz-\phi(hkl)]}$$

where V - is the volume of the unit cell,

$hkl$  – are reflection Miller indices,

$|F_{hkl}|$  – The structure factor amplitudes,

$\Phi_{hkl}$  – The phase of the structure factor.

The structure factor amplitudes  $|F(hkl)|$  can be calculated as the square roots of the measured intensities ( $I_{hkl}$ ). The information about phases ( $\Phi_{hkl}$ ) is not obtained experimentally and has to be calculated. This is the main problem of X-ray crystallography called the phase problem. However, several methods exist, which can solve the phase problem and determine the structure of a protein. They are molecular replacement, isomorphous replacement, anomalous dispersion, isomorphous replacement with anomalous scattering, radiation damage-induced phasing.

Molecular replacement (MR) is the fastest and most common method for solving the structure of a protein. The phases are taken from an already known protein structure, which is homologous and has a similar fold and sequence to the measured structure. It can be achieved in two steps: rotation and translation. During rotation, the spatial orientation of the known and unknown molecules with relation to each other is determined. At the translation, the motion necessary to superimpose one molecule to another keeping one the orientation found during the rotation step calculations [74].

Isomorphous replacement (IR) is a method, where diffraction data is required not only from the native protein crystal, but also from the modified crystal of single

(SIR) or multiple (MIR) heavy atom derivatives. Soaking of protein crystals in heavy-atom solutions is used to create heavy-atom derivative crystals. Both types of crystals should have the same dimensions of the unit cells, so they are isomorphous. The differences in the structure factor amplitudes between native and derivative protein crystals can identify the positions of the heavy atoms. This information is used for finding phase. Phase determination with SIR leading to two possible solutions (phase ambiguity), while using of MIR unambiguous phase determination can be achieved. The disadvantages of this methods include a creation of the perfect isomorphous crystals [76].

Single and multiple anomalous dispersion (SAD and MAD) is a method that requires the presence of strong anomalously scattering atoms, included in the protein or soaked into a crystal. As X-ray wavelength approaching absorption edge, anomalous scattering occurs – Friedel pairs (Bragg's reflections, related by inversion through origin) are no longer equal in intensities. The experiment for a MAD method needs few data sets to collect at several wavelengths around the absorption edge of the element, where the anomalous scattering factors of the element significantly different from each other. In the SAD case, just one dataset is collected at the wavelength corresponding to the absorption peak of the element. The differences in structure factor amplitudes arising from anomalous scattering are used to solve the phase problem. Only one crystal is needed, but multiple data sets must be collected at three different wavelengths. Combination of IR and anomalous dispersion single and multiple isomorphous replacement with anomalous scattering (SIRAS and MIRAS) can be achieved [77, 78].

Radiation damage-induced phasing (RIP) is based on the radiation damage impact of X-ray or UV-induced on protein crystals. In this method, a single crystal is exposed of a high dose of X-ray or UV radiation after data collection. Afterward data are collected again. Radiation damage can make significant

differences in intensities between two datasets. These differences are used for determination of phases by a SIR-type method. RIP method works well for proteins, which contain disulfide bridges [48, 79].

When protein structure determination is finished, the structure refinement should be starting. Refinement is represented by iterative cycles of map calculation and model building to improve the agreement between the model and gathered data. The final step is deposition well of refined structure in the form of atomic coordinates list to the Protein Data Bank (PDB) to share it for scientific community [80].



# 4. RESULTS AND DISCUSSION

---

4.1 The tetrameric structure of the novel haloalkane dehalogenase DpaA from *Paraglaucicola agarilytica* NO2

This chapter is based on Paper I:

Andrii Mazur, Tatyana Prudnikova, Pavel Grinkevich, Jeroen R. Mesters, Daria Mrazova, Radka Chaloupkova, Jiri Damborsky, Michal Kutý, Petr Kolenko and Ivana Kuta Smatanova.,

*Acta Cryst.* (2021). D77, 347–356

## **Abstract:**

Haloalkane dehalogenases (EC 3.8.1.5) are microbial enzymes that catalyse the hydrolytic conversion of halogenated compounds, resulting in a halide ion, a proton and an alcohol. These enzymes are used in industrial biocatalysis, bioremediation and biosensing of environmental pollutants or for molecular tagging in cell biology. The novel haloalkane dehalogenase DpaA described here was isolated from the psychrophilic and halophilic bacterium *Paraglaucicola agarilytica* NO2, which was found in marine sediment collected from the East Sea near Korea. Gel-filtration experiments and size-exclusion chromatography provided information about the dimeric composition of the enzyme in solution. The DpaA enzyme was crystallized using the sitting-drop vapour-diffusion method, yielding rod-like crystals that diffracted X-rays to 2.0 Å resolution. Diffraction data analysis revealed a case of merohedral twinning, and subsequent structure modelling and refinement resulted in a tetrameric model of DpaA, highlighting an uncommon multimeric nature for a protein belonging to haloalkane dehalogenase subfamily I



## The tetrameric structure of the novel haloalkane dehalogenase DpaA from *Paraglaciecola agarilytica* NO2

**Andrii Mazur, Tatyana Prudnikova, Pavel Grinkevich, Jeroen R. Mesters, Daria Mrazova, Radka Chaloupkova, Jiri Damborsky, Michal Kutý, Petr Kolenko and Ivana Kuta Smatanova**

*Acta Cryst.* (2021). D77, 347–356



**IUCr Journals**

CRYSTALLOGRAPHY JOURNALS ONLINE

Copyright © International Union of Crystallography

Author(s) of this article may load this reprint on their own web site or institutional repository provided that this cover page is retained. Republication of this article or its storage in electronic databases other than as specified above is not permitted without prior permission in writing from the IUCr.

For further information see <https://journals.iucr.org/services/authorrights.html>

# The tetrameric structure of the novel haloalkane dehalogenase DpaA from *Paraglaciicola agarilytica* NO2

Andrii Mazur,<sup>a</sup> Tatyana Prudnikova,<sup>a</sup> Pavel Grinkevich,<sup>a</sup> Jeroen R. Mesters,<sup>b</sup> Daria Mrazova,<sup>a</sup> Radka Chaloupkova,<sup>c</sup> Jiri Damborsky,<sup>c,d</sup> Michal Kutý,<sup>a</sup> Petr Kolenko<sup>e</sup> and Ivana Kuta Smatanova<sup>a\*</sup>

Received 9 November 2020  
Accepted 13 January 2021

Edited by K. Djinovic Carugo, University of Vienna, Austria

**Keywords:** DpaA; crystallization; haloalkane dehalogenases; halogenated pollutants; tetrameric structure; oligomerization.

**PDB reference:** DpaA from *P. agarilytica* NO2, 7avr

**Supporting information:** this article has supporting information at journals.iucr.org/d

<sup>a</sup>Faculty of Science, University of South Bohemia in Ceske Budejovice, Branisovska 1760, 370 05 Ceske Budejovice, Czech Republic, <sup>b</sup>Institute of Biochemistry, Center for Structural and Cell Biology in Medicine, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany, <sup>c</sup>Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, Kamenice 5, 625 00 Brno, Czech Republic, <sup>d</sup>International Clinical Research Center, St Anne's University Hospital Brno, Pekarska 53, 656 91 Brno, Czech Republic, and <sup>e</sup>Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Břehová 7, 115 19 Prague, Czech Republic. \*Correspondence e-mail: ivanaks@seznam.cz

Haloalkane dehalogenases (EC 3.8.1.5) are microbial enzymes that catalyse the hydrolytic conversion of halogenated compounds, resulting in a halide ion, a proton and an alcohol. These enzymes are used in industrial biocatalysis, bioremediation and biosensing of environmental pollutants or for molecular tagging in cell biology. The novel haloalkane dehalogenase DpaA described here was isolated from the psychrophilic and halophilic bacterium *Paraglaciicola agarilytica* NO2, which was found in marine sediment collected from the East Sea near Korea. Gel-filtration experiments and size-exclusion chromatography provided information about the dimeric composition of the enzyme in solution. The DpaA enzyme was crystallized using the sitting-drop vapour-diffusion method, yielding rod-like crystals that diffracted X-rays to 2.0 Å resolution. Diffraction data analysis revealed a case of merohedral twinning, and subsequent structure modelling and refinement resulted in a tetrameric model of DpaA, highlighting an uncommon multimeric nature for a protein belonging to haloalkane dehalogenase subfamily I.

## 1. Introduction

Environmental pollution caused by human activities is one of the major global issues of our time. Halogenated organic compounds are a major group of chemicals that are used in agriculture and industry. Their applications vary from pesticides and soil fumigants (for example 1,2-dibromoethane and  $\gamma$ -hexachlorocyclohexane) to solvents (for example 1,2-dichloroethane and chloromethane) and industrial precursors (for example vinyl chloride and 1,2-dichloroethane) (Fishbein, 1979; Chaudhry & Chapalamadugu, 1991). The broad range of application of these compounds, as well as their high resistance to degradation, makes their decomposition difficult. Nevertheless, a number of technologies for the degradation of these compounds have already been developed (Swanson, 1999). Numerous microbiological surveys have shown that some microorganisms can survive in highly polluted areas (Janssen *et al.*, 2005; Fetzner, 1998). Their resistance is attributed to the presence of specific enzymes that are capable of degrading pollutants, which makes these microorganisms and their enzymes attractive targets for the development of new bioremediation technologies.

Haloalkane dehalogenases (HLDs; EC 3.8.1.5) represent a diverse group of enzymes that have attracted significant



© 2021 International Union of Crystallography

**Table 1**  
Production details for DpaA.

Source organism	<i>P. agarilytica</i> NO2
DNA source	Synthesized DNA (GI 495580204)
Restriction sites	NdeI/XhoI
Vector	pET-21b
Expression host	<i>E. coli</i> BL21(DE3)
Complete amino-acid sequence of the construct produced	MTIKALRTPPEERFSVLPAPPYQPNVYDDL GYESLRMAYIDEGDKDSEYTFLLHGE TWSYLRYKMPVFTDAGHRVVPDLFGF GRSDKPIEDSVYNFEPHRNSLIQLIEHL DLKNIVLVCQDWGGGLGLTTPMDMQDRF KKLIVMNTTISNGEPLAEAAVQWMAFNE TISELPVAGLVACDAGAAVNVMDALAYD APFPNKNYKVGKRFPPQMIPTNADDDAV KYGLRAIEFWSNEWSGESFMAIGMKDAV LGEAAMQLKTVIKGCPPEMKIEEAGHF VQEYGVVEAQALASFTMIHHHHHH

interest because of their ability to catalyse the hydrolysis of a wide range of halogenated aliphatic compounds. This reaction results in three products: (i) an alcohol, (ii) a halide ion and (iii) a proton (Fetzner & Lingens, 1994). HLDs belong to the  $\alpha/\beta$ -hydrolase superfamily, which represents one of the largest groups of structurally related enzymes with distinct catalytic functions (Carr & Ollis, 2009; Holmquist, 2000). Phylogenetic analysis divides HLDs into three subfamilies: HLD-I, HLD-II and HLD-III (Chovancová *et al.*, 2007; Vanacek *et al.*, 2018). Each family maintains a specific composition of the catalytic pentad, which is made up of a catalytic triad and two halide-stabilizing residues (Holmquist, 2000; Damborsky *et al.*, 2010). The catalytic pentad of the HLD-I subfamily consists of Asp-His-Asp/Trp-Trp, while HLD-II members predominantly contain Asp-His-Glu/Asn-Trp and HLD-III members contain Asp-His-Asp/Asn-Trp (Chovancová *et al.*, 2007). Because of their broad substrate specificity, catalytic mechanism and robustness, HLDs are used in industrial biocatalysis, bioremediation (Fetzner, 1998), biosensing of environmental pollutants (Bidmanova *et al.*, 2010; Campbell *et al.*, 2006; Los & Wood, 2007), cell imaging and protein analysis (Koudelakova *et al.*, 2011), and decontamination of warfare agents (Koudelakova *et al.*, 2011; Prokop *et al.*, 2006). Generally, haloalkane dehalogenases occur in a monomeric form; the exceptions are enzymes from the HLD-III group that form high-order multimeric structures and specific enzymes from the other groups that form dimers or dynamic oligomers (Kunka *et al.*, 2018). Oligomeric states of HLD-I enzymes are rare and therefore such specimens are of great interest. Moreover, the processing of X-ray diffraction data from oligomeric proteins can be challenging (Goodsell & Olson, 2000). Various problems such as pseudosymmetry, quasi-symmetry and twinning can affect the determination of the correct structure and lead to uncertainties in space-group assignment (Zwart *et al.*, 2008).

Here, we report the purification, crystallization and X-ray data analysis of DpaA, a novel haloalkane dehalogenase belonging to the HLD-I subfamily (Vanacek *et al.*, 2018) that was isolated from the psychrophilic and moderately halophilic bacterium *Paraglaciecola agarilytica* NO2. The bacterium was isolated from marine sediments collected in the East Sea near

**Table 2**  
Data-collection and refinement statistics for DpaA.

Data collection	
Space group	$P2_12_12_1$
<i>a</i> , <i>b</i> , <i>c</i> (Å)	117.39, 155.51, 155.61
$\alpha$ , $\beta$ , $\gamma$ (°)	90.0, 90.0, 90.0
Resolution range (Å)	47.44–2.00 (2.03–2.00)
Total No. of reflections	1289197 (59344)
No. of unique reflections	191831 (9393)
Completeness (%)	100 (100)
$\langle I/\sigma(I) \rangle$	7.90 (1.30)
$R_{\text{meas}}^{\ddagger}$	0.18 (1.61)
$CC_{1/2}$	99.80 (60.10)
Overall <i>B</i> factor from Wilson plot (Å <sup>2</sup> )	22
Refinement	
No. of reflections used for refinement	182376
$R_{\text{work}}^{\ddagger}/R_{\text{free}}^{\S}$ (%)	17.17/20.22
No. of non-H atoms	20815
No. of protein atoms	18810
No. of chloride ions	18
No. of water molecules	1987
Average <i>B</i> factor (Å <sup>2</sup> )	27.21
Ramachandran plot	
Most favoured (%)	96.04
Allowed (%)	3.62
Outliers (%)	0.34
R.m.s. deviations	
Bond lengths (Å)	0.02
Angles (°)	1.81
PDB code	7avr

$\ddagger R_{\text{meas}}$  is a redundancy-independent merging *R* factor.  $R_{\text{meas}} = \sum_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ , where  $\langle I(hkl) \rangle$  is the mean of the  $N(hkl)$  individual measurements  $I_i(hkl)$  of the intensity of reflection  $hkl$ .  $\S R_{\text{work}} = \sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum_{hkl} |F_{\text{obs}}|$ .  $\S R_{\text{free}}$  was monitored using 5% of the reflection data that were excluded from refinement.

Korea (Yong *et al.*, 2007). The genes of this extremophilic bacterium were identified and characterized by a combination of bioinformatics tools and automated laboratory screening techniques (Vanacek *et al.*, 2018). The DpaA enzyme shows high enantioselectivity towards a racemic mixture of ethyl 2-bromopropionate and moderate enantioselectivity towards 2-bromopentane (Vanacek *et al.*, 2018). This enzyme has also been demonstrated to form dimers in solution (Vanacek *et al.*, 2018), a feature that is not typical for a member of the HLD-I subfamily (Kunka *et al.*, 2018). Dimers and higher order oligomers probably lead to an unusual protein crystal packing that causes difficulties during X-ray data processing, space-group assignment and structure determination. The model of DpaA based on the crystallographic data is the first example of a tetrameric structure in haloalkane dehalogenase subfamily I. This finding contributes to the development of structure-solution procedures for oligomeric proteins that exhibit unusual crystal packing.

## 2. Materials and methods

### 2.1. Gene synthesis and cloning

The gene sequence encoding DpaA from *P. agarilytica* NO2 (GI 495580204) was optimized for expression in *Escherichia coli* and commercially synthesized by Bio Basic, Canada (Table 1). The recombinant gene was subcloned into the expression vector pET-21b conveying ampicillin resistance (Novagen, USA) employing NdeI/XhoI restriction sites. A tag

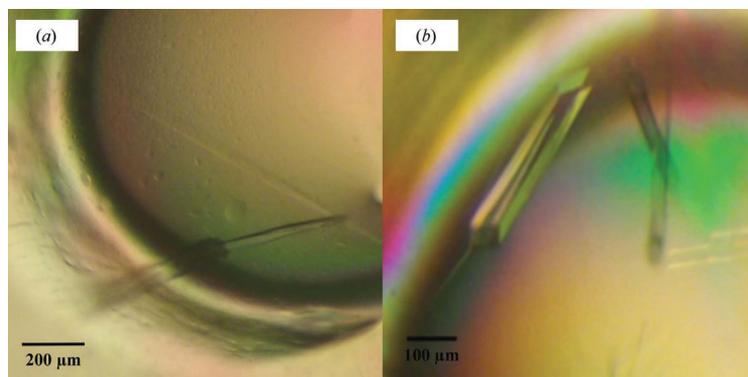
sequence of six histidine codons was attached downstream of the gene. Competent cells of *E. coli* strain DH5 $\alpha$  were transformed with the resulting constructs using the heat-shock method for plasmid propagation.

## 2.2. Protein production

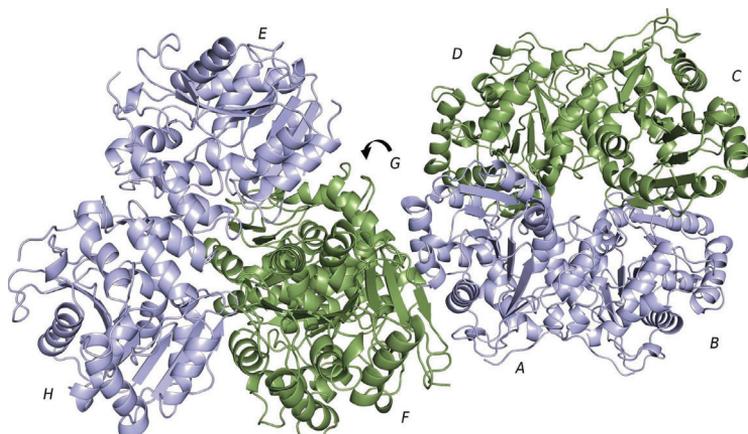
*E. coli* strain BL21(DE3) was used for routine high-level gene expression and protein production as outlined below. The competent cells were transformed with the pET-21b vector containing the *dpaA* gene using the standard heat-shock method, plated on agar plates with ampicillin (100  $\mu\text{g ml}^{-1}$ ) and grown overnight at 37°C. Single colonies were used to inoculate 10 ml lysogenic broth (LB) medium (Sigma–Aldrich, USA) with ampicillin (100  $\mu\text{g ml}^{-1}$ ) and the cells were grown overnight at 37°C. The overnight culture was used to inoculate 1000 ml LB medium. The cells were cultivated at 37°C with shaking at 110  $\text{rev min}^{-1}$  until an OD<sub>600</sub> of

0.4–0.6 was reached. Overexpression was induced by the addition of 0.5 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (Sigma–Aldrich, USA). The cells were then cultivated for 24 h at 20°C with shaking at 105  $\text{rev min}^{-1}$ . Biomass was harvested by centrifugation (3700g at 4°C for 10 min), washed with purification buffer A (20 mM K<sub>2</sub>HPO<sub>4</sub>/KH<sub>2</sub>PO<sub>4</sub> pH 7.5, 0.5 M NaCl, 10 mM imidazole) and resuspended in 10 ml buffer A per gram of wet biomass. The harvested biomass was frozen at –80°C. DNase I (New England Biolabs, USA) was added to a final concentration of 1.25  $\mu\text{g ml}^{-1}$  and the cell suspension was disrupted by a One-shot cell homogenizer (Constant Systems, UK) using one cycle at 1.5 kbar pressure. The cell lysate was centrifuged for 1 h at 21 000g and 4°C. The crude extract was decanted and the total protein concentration was determined by the Bradford method (Bradford, 1976).

The overexpressed histidine-tagged DpaA was purified from the crude extract using a single-step nickel-affinity chromatography approach. The cell-free extract was loaded onto a 5 ml nickel–nitrilotriacetic acid (Ni–NTA) Superflow column (Qiagen, Germany) in purification buffer A (20 mM potassium phosphate buffer pH 7.5, 0.5 M sodium chloride, 10 mM imidazole). Unbound and weakly bound proteins were washed with the same buffer until the baseline was reached. The target protein was eluted with elution buffer (purification buffer A containing 300 mM imidazole). The eluted protein was dialyzed overnight against 50 mM Tris–HCl buffer pH 7.5 at 4°C. Protein concentration was determined by the Bradford method and purified proteins were stored at 4°C.



**Figure 1**  
Crystals of DpaA from *P. agarilytica* NO2. (a) The DpaA crystals grown after initial screening; (b) the DpaA crystals used for data-collection experiments.



**Figure 2**  
Cartoon model of the asymmetric unit of DpaA: dimers AB and EH are shown in light blue; dimers DC and FG are shown in green.

## 2.3. Crystallization

Crystallization screening was performed on a Gryphon crystallization robot (Art Robbins Instruments, USA) and an Oryx4 crystallization robot (Douglas Instruments, UK) in MRC 2-well crystallization plates (Hampton Research, USA) and CombiClover Junior crystallization plates (Rigaku Reagents, USA), respectively, employing the sitting-drop vapour-diffusion method (Ducruix & Giegé, 2000). Initially, commercial crystallization screens were used: JCSG-*plus*, Structure Screen 1&2, Morpheus, the PGA screen, MIDAS (Molecular Dimensions, UK) and Index (Hampton Research, USA). Protein in 50 mM Tris–HCl buffer pH 7.5 was mixed with the precipitant solution in various ratios and

equilibrated against 50–70 µl precipitant for the MRC 2-well plates and 200 µl precipitant for the CombiClover Junior plates. Optimization of successful crystallization conditions was performed by varying the salt, PEG and protein concentrations and the temperature (McPherson & Cudney, 2014).

2.4. Data collection

Diffraction data were collected on beamline BL14.1 for macromolecular crystallography at the BESSY II electron-storage ring (Berlin-Adlershof, Germany) operated by the Helmholtz-Zentrum Berlin (Mueller *et al.*, 2015) and equipped with a PILATUS 6M detector (Dectris, Switzerland).

The crystals of DpaA were mounted in LithoLoops (Molecular Dimensions, UK) or nylon cryoloops (Hampton Research, USA) and were immediately flash-cooled in liquid nitrogen. As a cryoprotectant, 0.5 µl 50% (w/v) PEG 400 was

added to the droplet with crystals 5 min before harvesting. Diffraction experiments were performed at 100 K. Data-collection statistics are summarized in Table 2.

2.5. Structure solution and refinement

Data sets were indexed and integrated using the XDS software package (Kabsch, 2010) and were scaled using AIMLESS from the CCP4 suite (Winn *et al.*, 2011). The structure was solved using molecular replacement with MOLREP (Vagin & Teplyakov, 2010). Structure refinement was performed with REFMAC5 (Murshudov *et al.*, 2011) and Coot (Emsley *et al.*, 2010). A global overview of the interactions in assemblies in the asymmetric unit was obtained by making use of the PISA server (Krissinel & Henrick, 2007). All figures showing structural representations were prepared with PyMOL (DeLano, 2002). The surface model was gener-

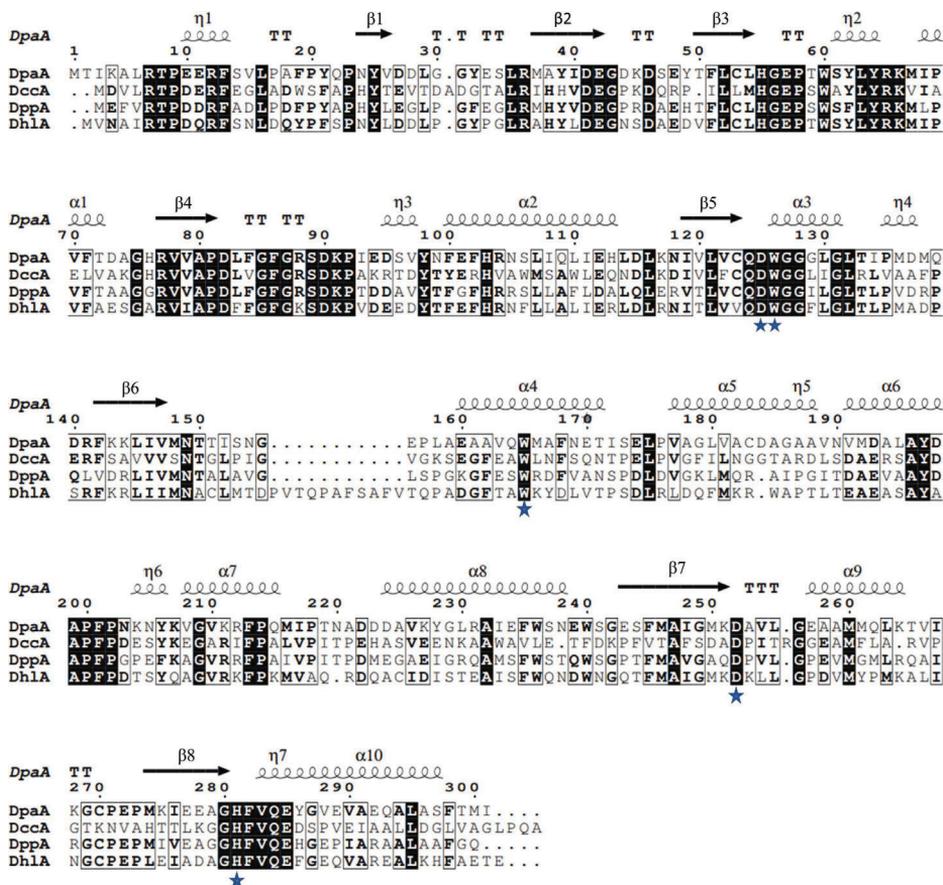


Figure 3 Multiple sequence alignment of the structurally described haloalkane dehalogenases DpaA from *Paraglaciocola agarilytica* NO2 (PDB entry 7avr), DccA from *Caulobacter crescentus* (PDB entry 5ers), DppA from *Plectocystis pacifica* SIR-1 (PDB entry 2xt0) and Dh1A from *Xanthobacter autotrophicus* GJ10 (PDB entry 2dhc). Catalytic residues are marked by blue stars. Identical residues are presented in white on a black background; similar residues are in bold. Secondary-structure elements of DpaA are shown above the sequences. The alignment was generated with Clustal Omega (Sievers *et al.*, 2011) and visualized using ESPript 3.0 (Robert & Gouet, 2014).

ated in *UCSF Chimera* 1.14 (Pettersen *et al.*, 2004), which shows solvent-excluded molecular surfaces composed of probe-contact, toroidal and re-entrant surfaces. Solvent-excluded molecular surfaces were created with the help of the *MSMS* package (Sanner *et al.*, 1996). In *UCSF Chimera*, amino-acid residues were assigned according to the *kdHydrophobicity* (Kyte & Doolittle, 1982), with values within the range  $-4.5$  (blue) to  $4.5$  (red). The more positive the value, the more hydrophobic the amino acids located in that region of the protein are. Hydrophobicity scales offer insight into the thermodynamics of the interactions that occur between the proteins. Data-refinement statistics are summarized in Table 2. The refined structure was deposited in the Protein Data Bank with PDB code 7avr.

### 3. Results and discussion

The recently isolated and characterized wild-type haloalkane dehalogenase DpaA (Vanacek *et al.*, 2018) from *P. agarilytica* NO2 was used in initial crystallization. Microcrystals and thin needle-shaped crystals (Fig. 1*a*) were found after seven days using a protein stock concentration of  $13.6 \text{ mg ml}^{-1}$  in a condition consisting of  $0.1 \text{ M}$  Na HEPES pH 7.5,  $2\%$  (w/v) PEG 400,  $2 \text{ M}$  ammonium sulfate at 293 K. To optimize the crystal size, various concentrations of protein and precipitant components and a lower temperature of 277 K were used. The best-diffracting DpaA crystal (Fig. 1*b*), with dimensions of  $350 \times 50 \times 30 \text{ }\mu\text{m}$ , grew within a week in a condition consisting of

$0.1 \text{ M}$  Na HEPES pH 7.5,  $4\%$  (w/v) PEG 400,  $2 \text{ M}$  ammonium sulfate using a protein concentration of  $8 \text{ mg ml}^{-1}$  and a 2:1 ratio of protein solution to precipitant.

Several diffraction data sets were collected with a maximum resolution ranging from 2.0 to 2.9 Å. The data sets exhibited space-group ambiguity and pseudotranslation, with values ranging from 27.8% to 47.6%. Data-collection statistics for the best-diffracting crystal are summarized in Table 2.

Data processing of the data set presented here was difficult because of the space-group ambiguity. Although *AIMLESS* suggested space groups with a tetragonal lattice, no successful molecular-replacement solution was found. Based on its sequence identity (52%), the haloalkane dehalogenase DppA from *Plesiocystis pacifica* SIR-I (PDB entry 2xt0; Hessler *et al.*, 2011) was used as a structural template. We attempted phasing in several space groups (for example  $P1$ ,  $P2_1$ ,  $P2_12_12$  and  $C222_1$ ). Phasing was successful in several of them, but the refinement proved to be stable in space group  $P2_12_12$  with nearly perfect merohedral twinning. A total of eight molecules of DpaA were placed into the electron-density map. After several cycles of manual and automated refinement (including amplitude twin refinement),  $R_{\text{work}}$  and  $R_{\text{free}}$  decreased to values of 17.2% and 20.2%, respectively.

The asymmetric unit of DpaA consists of two tetramers, chains *A*, *B*, *C* and *D* and chains *E*, *F*, *G* and *H*, which are shown in Fig. 2, with a dimer–dimer interaction in each. In the *ABCD* tetramer, two dimers, *AB* and *CD*, are held together by interactions between chains *B* and *C* and chains *D* and *A*.

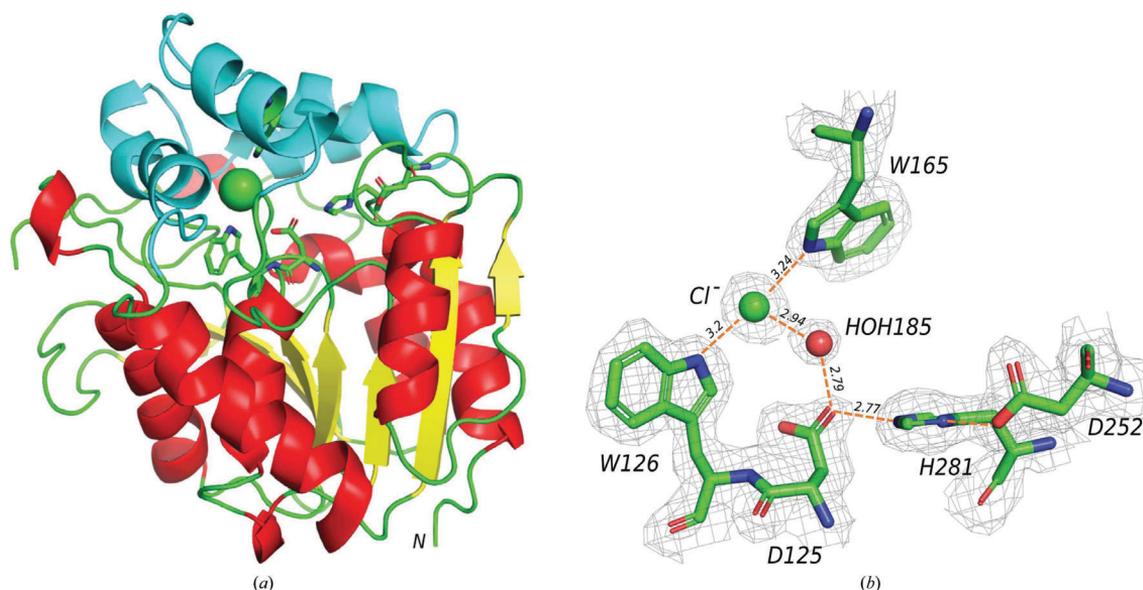


Figure 4

Cartoon model of the overall structure of DpaA (*a*) and the active site (*b*). (*a*) Overall structure.  $\alpha$ -Helices are shown in red,  $\beta$ -strands are shown in yellow and loops are in green for the core domain of DpaA. The cap domain is shown in cyan, the chloride ion in the active site is shown as a green sphere and the catalytic pentad is shown as green sticks. (*b*) Details of the active site. The catalytic pentad and the chloride ion and water molecule are shown as green and red spheres, respectively. Coordinating interactions are represented by orange dashed lines. The  $2F_o - F_c$  electron-density map for ions and interacting residues contoured at  $1.0\sigma$  is drawn as light grey mesh.

Analysis of the solvent-accessible surface area of the preliminary model of DpaA, using the *PISA* server (Krissinel & Henrick, 2007), suggested dimerization and tetramerization. Dimerization has previously been confirmed for this protein using size-exclusion chromatography (Vanacek *et al.*, 2018).

A multiple alignment with representatives of the HLD-I subfamily that show a dimeric nature and/or have a higher sequence identity (Carlucci *et al.*, 2016; Hesseler *et al.*, 2011; Silberstein *et al.*, 2007) was performed to compare DpaA with its relatives with previously determined crystal structures (Fig. 3). The alignment shows that DpaA has a highly conserved secondary structure in comparison with other

dehalogenases. The active-site residues are conserved among all compared proteins; the catalytic triad Asp125, Asp252 and His281 (numbering according to DpaA) and the two halide-binding residues (Trp126 and Trp165) are identical in all represented dehalogenases (Fig. 3). The main domain sequences are highly conserved among dehalogenases, in contrast to the amino-acid divergence of the cap-domain components. The amino acids of the main domain are responsible for the catalytic function, although the cap-domain residues are important for substrate recognition and specificity. Every HLD has its own wide or narrow range of substrate specificity, which is a result of the deviation of the

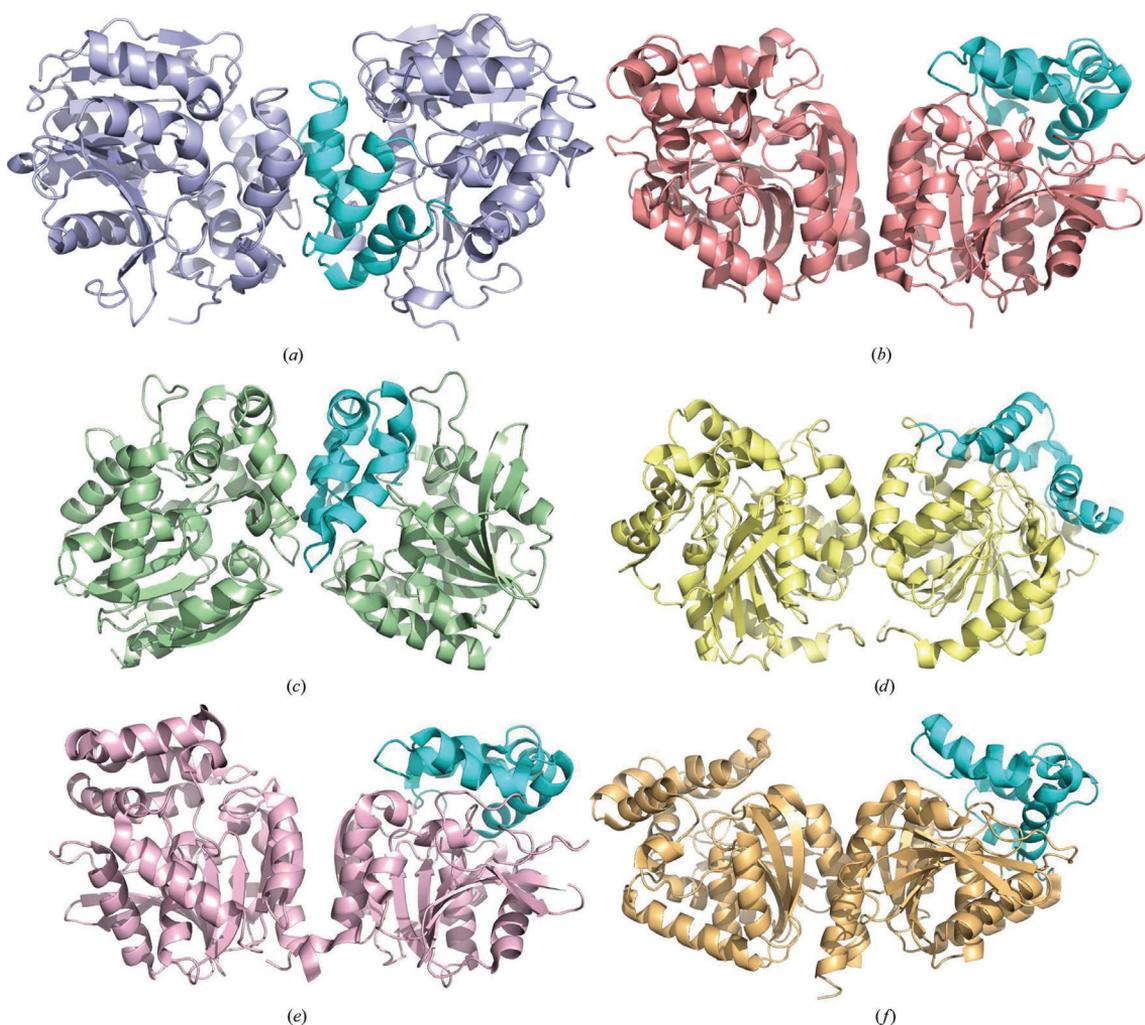


Figure 5

Cartoon models of the overall structures of various haloalkane dehalogenase dimers: (a) DpaA from *Paraglaciecola agarilytica* NO2 (PDB entry 7avr; this study), (b) DmmaA from *Moorea producta* (PDB entry 3u1t), (c) HanR from Rhodobacteraceae sp. (PDB entry 4brz), (d) DccA from *Caulobacter crescentus* (PDB entry 5ers), (e) DmxA from *Marinobacter* sp. (PDB entry 5mxp) and (f) DbeA from *Bradyrhizobium elkanii* (PDB entry 4k2a). The cap domain is shown in cyan.

amino-acid composition of the cap domain among the HLD family members.

### 3.1. Overall DpaA structure

DpaA has the typical organization of haloalkane dehalogenases from the HLD-I subfamily (Chovancová *et al.*, 2007), consisting of two domains (the main domain and the cap domain) with a deep cleft between them that harbours the active site (Fig. 4a). The main domain consists of an eight-stranded  $\beta$ -sheet with an antiparallel orientation of  $\beta$ 2. The  $\beta$ -strands are surrounded by six  $\alpha$ -helices: two on one side ( $\alpha$ 2 and  $\alpha$ 10) and four on the other side ( $\alpha$ 1,  $\alpha$ 3,  $\alpha$ 4 and  $\alpha$ 9). The cap domain, consisting of residues 156–221 inserted between the  $\beta$ 6 strand and the  $\alpha$ 8 helix of the main domain, is formed by four  $\alpha$ -helices ( $\alpha$ 4,  $\alpha$ 5,  $\alpha$ 6 and  $\alpha$ 7) and five connective loops (Fig. 4a).

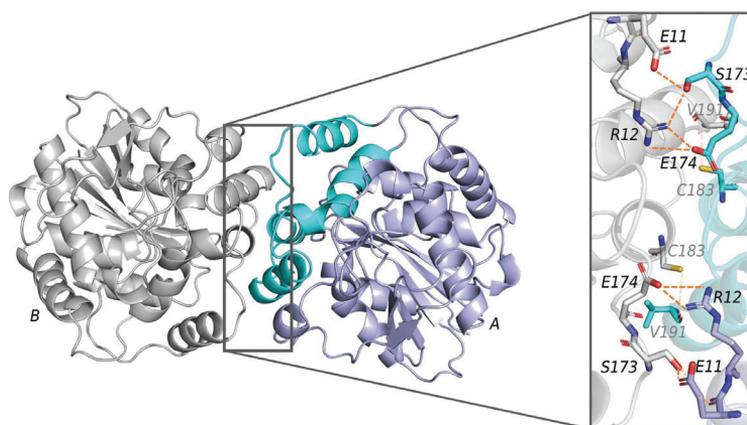
The active site of DpaA is buried in a hydrophobic cavity and consists of a catalytic pentad, which is a typical structural signature of the HLD-I subfamily (Chovancová *et al.*, 2007), namely two halide-stabilizing residues, Trp126 and Trp165, and a catalytic triad comprised of the nucleophile Asp125, the catalytic acid Asp252 and the catalytic base His281 (Fig. 4b). The nonprotein electron density in the active site was interpreted as water molecules and a chloride anion. The chloride anion occupies the halide-binding site and interacts with the N<sup>δ</sup>1 atoms of Trp126 and Trp165, the two halide-stabilizing residues, at distances of 3.20 and 3.24 Å, respectively. Further coordination of the chloride ion is provided by the N atom of Pro214 at a distance of 3.50 Å and the catalytic water molecule HOH185 at a distance of 2.94 Å (Fig. 4b). Asp125 O<sup>δ</sup>1 forms a hydrogen bond of 2.74 Å to the main-chain amino group of the halide-stabilizing Trp126 and connects to His281 and Asp252 of the catalytic triad (Fig. 4b). The whole catalytic pentad in DpaA has a similar conformation to the corre-

sponding catalytic residues in other members of HLD-I with known structures, such as DccA, DhIA and DppA.

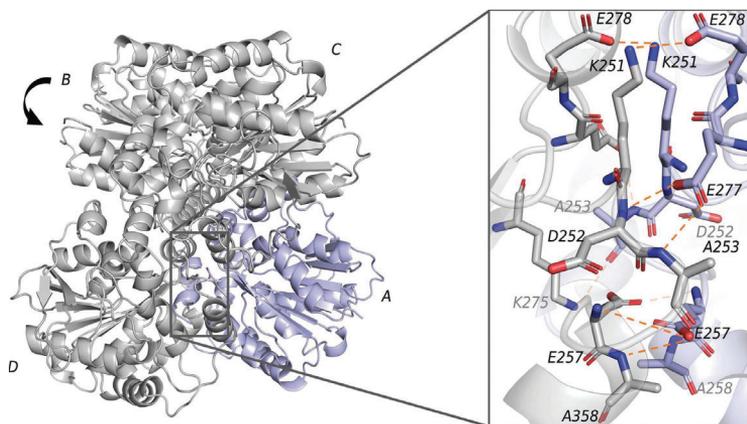
### 3.2. Oligomerization

Dimerization has previously been discovered in several HLD structures deposited in the Protein Data Bank: DmmA from *Moorea producta* (PDB entry 3ult; Gehret *et al.*, 2012), HanR from Rhodobacteraceae sp. (PDB entry 4brz; Novak *et al.*, 2014), DccA from *Caulobacter crescentus* (PDB entry 5ers; Carlucci *et al.*, 2016), DmxA from *Marinobacter* sp. (PDB entry 5mxp; Chrast *et al.*, 2019), DbeA from *Bradyrhizobium elkanii* (PDB entry 4k2a; Chaloupkova *et al.*, 2014) and DbjA from *B. japonicum* (PDB entry 3a2m; Prokop *et al.*, 2010) (Figs. 5b–5f). The DmmA dimers are formed by two C-terminal helices from the main domain (Fig. 5b). In the HanR structure, the dimers are formed by the contact of the external part of the cap domains, mainly by the residues in the loops connecting the  $\alpha$ 5 and  $\alpha$ 6 helices (Fig. 5c). In the DccA structure, the assembly is attained by an amino-acid interaction situated in the  $\alpha$ 2 helix on one side of the main domain (Fig. 5d). The dimerization in the DbjA and DbeA structures is identical and is achieved through residues in the C-terminal  $\alpha$ 11 helix (Fig. 5f). In the DmxA structure, the dimers are formed similarly to those of the DbeA protein but are also aided by a disulfide bridge between the C-terminal Cys294 residues (Fig. 5e).

The tertiary structure of DpaA represents a tetramer that consists of two homodimers. An analysis of intermolecular contacts in the modelled structure of DpaA suggested that the dimers are formed through noncovalent interactions mainly of the  $\alpha$ 5 (residues 173–183) and  $\alpha$ 6 (residues 191–198) helices. Further interaction is provided by the loop residues (11–18, 170–172 and 202–204; Figs. 5a and 6). Overall, 32 amino acids are involved in the formation of the protein-dimer interface. According to the protein interface analysis provided by

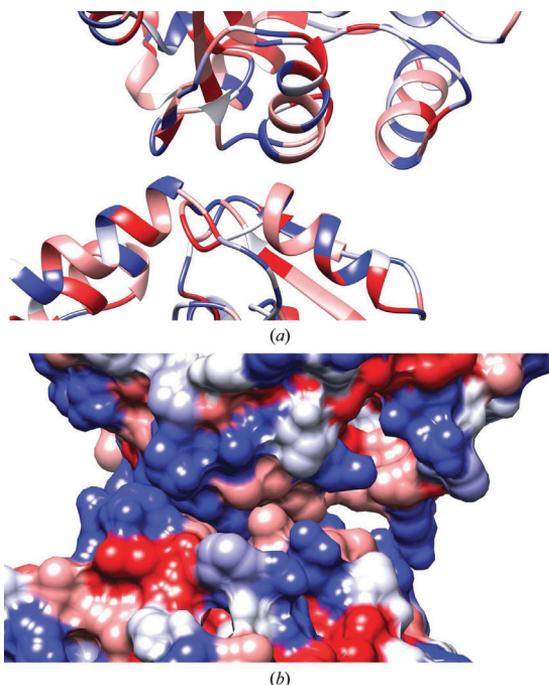


**Figure 6**  
Cartoon model of the dimer structure of DpaA and the interactions between chains A and B. Chain A is shown in light blue and chain B is in grey. The cap domain in chain A is shown in cyan. The amino acids involved in the interactions are shown as sticks. The interactions are shown as orange dashed lines.



**Figure 7**  
Cartoon model of the tetramer structure of DpaA and interactions between chains *A* and *D*. Chain *A* is shown in light blue; other chains are in grey. The interactions are shown as grey sticks with coloured elements for chain *D* and light blue sticks with coloured elements for chain *A*. The interactions are shown as orange dashed lines.

*PDBePISA* (Krissinel & Henrick, 2007), the buried solvent-accessible protein area of the dimer is 935.2 Å<sup>2</sup>, representing



**Figure 8**  
Cartoon representation of hydrophobic interactions between two dimers forming a tetramer, where the residues are coloured according to the *kdHydrophobicity* values from  $-4.5$  (shown in blue) to  $4.5$  (shown in red): (a) cartoon view of the interaction area and (b) surface representation of the contact area. Hydrophobicity values were calculated using *UCSF Chimera* (Pettersen *et al.*, 2004).

7.7% of the total solvent-accessible surface area of the dimer. Hydrogen bonds are formed from the N<sup>δ1</sup> atom of Arg12 in chain *B* to the O<sup>δ2</sup> atom of Glu174 and the O<sup>γ1</sup> atom of Ser173 in chain *A*, with distances of 2.77 and 3.10 Å, respectively. Further interaction is provided by the contact of O<sup>γ1</sup> of Ser173 in chain *B* and O<sup>ε1</sup> of Glu11 in chain *A*, with a distance of 2.62 Å. Another bond is represented by the interaction of the S atom of Cys183 in chain *A* with the N atom of Val191 in chain *D*, with a distance of 3.80 Å. Additional interaction is provided by the same residues of chain *D* and chain *A*. Chains *C* and *D* show the same set of interactions with similar distances.

The interactions which form a tetramer are shown in Fig. 7. This contact is weaker and is connected to the α9 helix and β8 sheet. The strongest hydrophilic interaction is provided by the O<sup>ε1</sup> and O<sup>ε2</sup> atoms of Glu277 from chain *A* and the N atoms of Asp252 and Ala253 from chain *D*, with distances of 3.00 and 3.10 Å, respectively. A hydrophilic interaction is provided by Glu278 from chain *A* and Lys251 from chain *D*, with a distance of 3.20 Å between the O<sup>ε2</sup> and N<sup>ε</sup> atoms, respectively. These interactions are similar for the same residues in chain *D* and have similar consistency in the interactions of chains *B* and *C*. Additional coordination is provided by hydrophobic interactions of one chain (residues 159–162, 250–255, 259–263, 269–271, 274 and 279) with the corresponding residues of the other chain (Fig. 8). Chains *B* and *C* show the same set of interactions with similar distances. To verify the crystallographic arrangement of monomeric units in the multimeric assembly, two tetramers were constructed by *PDBePISA* in space group *P2<sub>1</sub>2<sub>1</sub>2*, *ABCD* and *EFGH* tetramers, with values of Δ*G* for the interaction of  $-113.4$  and  $-165$  kcal mol<sup>-1</sup>, respectively. The *ABCD* tetramer is nearly identical to the *EFGH* tetramer in its organization. According to the protein interface analysis calculated by *PDBePISA*, the buried solvent-accessible area is

7147 Å<sup>2</sup>, representing 17% of the total solvent-accessible surface area of a tetramer.

#### 4. Conclusions

Crystallographic analysis of the DpaA protein revealed an unusual crystal packing in comparison to other members of the HLD-I subfamily. The model of DpaA reported in this paper represents the first tetrameric structure observed by X-ray crystallography in haloalkane dehalogenase subfamily I. Currently, all crystal structures of HLD-I subfamily members deposited in the PDB are monomeric. There are only a few examples of dimeric haloalkane dehalogenases from subfamilies HLD-II and HLD-III, such as DbjA (PDB entry 3a2m; Prokop *et al.*, 2010), DbeA (PDB entry 4k2a; Chaloupkova *et al.*, 2014), DmmA from the metagenomic DNA of a marine microbial consortium (PDB entry 3u1t; Gehret *et al.*, 2012), HanR from Rhodobacteraceae sp. (PDB entry 4brz; Novak *et al.*, 2014), DccA from *Caulobacter crescentus* (PDB entry 5ers; Carlucci *et al.*, 2016) and DmxA from *Marinobacter* sp. (PDB entry 5mxx; Chrast *et al.*, 2019). Crystal structures of enzymes from the HLD-III subfamily (DmbC and DrbA) that show oligomeric properties in size-exclusion chromatography (Jesenská *et al.*, 2009; Moore, 2018) have not been solved to date. The tendency of the DpaA protein to oligomerize is likely to cause problems during protein structure solution in a wide range of crystallization conditions and also across various crystal lattices.

#### Acknowledgements

The diffraction data were collected on BL14.1 at the BESSY II electron-storage ring operated by the Helmholtz-Zentrum Berlin. We would particularly like to acknowledge the help and support of Manfred S. Weiss during data collection. We also acknowledge Magda Vojtova for help with crystallization experiments.

#### Funding information

The following funding is acknowledged: Grant Agency of the Czech Republic (grant No. 17-24321S); DAAD mobility grant (grant No. DAAD-16-09); European Regional Development Fund-Project (grant Nos. CZ.02.1.01/0.0/0.0/15\_003/0000441, LM2015047 and LM2015055); GAJU (grant No. 017/2019/P).

#### References

- Bidmanova, S., Chaloupkova, R., Damborsky, J. & Prokop, Z. (2010). *Anal. Bioanal. Chem.* **398**, 1891–1898.
- Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
- Campbell, D. W., Müller, C. & Reardon, K. F. (2006). *Biotechnol. Lett.* **28**, 883–887.
- Carlucci, L., Zhou, E., Malashkevich, V. N., Almo, S. C. & Mundorff, E. C. (2016). *Protein Sci.* **25**, 877–886.
- Carr, P. D. & Ollis, D. L. (2009). *Protein Pept. Lett.* **16**, 1137–1148.
- Chaloupkova, R., Prudnikova, T., Rezacova, P., Prokop, Z., Koudelakova, T., Daniel, L., Brezovsky, J., Ikeda-Ohtsubo, W., Sato, Y., Kutý, M., Nagata, Y., Kuta Smatanova, I. & Damborsky, J. (2014). *Acta Cryst. D70*, 1884–1897.
- Chaudhry, R. G. & Chapalamadugu, S. (1991). *Microbiol. Rev.* **55**, 59–79.
- Chovancová, E., Kosinski, J., Bujnicki, J. M. & Damborský, J. (2007). *Proteins*, **67**, 305–316.
- Chrast, L., Tratsiak, K., Planas-Iglesias, J., Daniel, L., Prudnikova, T., Brezovsky, J., Bednar, D., Kuta Smatanova, I., Chaloupkova, R. & Damborsky, J. (2019). *Microorganisms*, **7**, 498.
- Damborsky, J., Chaloupkova, R., Pavlova, M., Chovancova, E. & Brezovsky, J. (2010). *Handbook of Hydrocarbon and Lipid Microbiology*, edited by K. N. Timmins, pp. 1081–1098. Berlin, Heidelberg: Springer.
- DeLano, W. L. (2002). *PyMOL*. <http://www.pymol.org>.
- Ducruix, A. & Giegé, R. (2000). *Crystallization of Nucleic Acids and Proteins. A Practical Approach*. Oxford University Press.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst. D66*, 486–501.
- Fetzner, S. (1998). *Appl. Microbiol. Biotechnol.* **50**, 633–657.
- Fetzner, S. & Lingens, F. (1994). *Microbiol. Rev.* **58**, 641–685.
- Fishbein, L. (1979). *Sci. Total Environ.* **11**, 111–161.
- Gehret, J. J., Gu, L., Geders, T. W., Brown, W. C., Gerwick, L., Gerwick, W. H., Sherman, D. H. & Smith, J. L. (2012). *Protein Sci.* **21**, 239–248.
- Goodsell, D. S. & Olson, A. J. (2000). *Annu. Rev. Biophys. Biomol. Struct.* **29**, 105–153.
- Hesseler, M., Bogdanović, X., Hidalgo, A., Berenguer, J., Palm, G. J., Hinrichs, W. & Bornscheuer, U. T. (2011). *Appl. Microbiol. Biotechnol.* **91**, 1049–1060.
- Holmquist, M. (2000). *Curr. Protein Pept. Sci.* **1**, 209–235.
- Janssen, D. B., Dinkla, I. J. T., Poelarends, G. J. & Terpstra, P. (2005). *Environ. Microbiol.* **7**, 1868–1882.
- Jesenská, A., Monincová, M., Koudeláková, T., Hasan, K., Chaloupková, R., Prokop, Z., Geerlof, A. & Damborský, J. (2009). *Appl. Environ. Microbiol.* **75**, 5157–5160.
- Kabsch, W. (2010). *Acta Cryst. D66*, 125–132.
- Koudelakova, T., Chovancova, E., Brezovsky, J., Monincova, M., Fortova, A., Jarkovsky, J. & Damborsky, J. (2011). *Biochem. J.* **435**, 345–354.
- Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
- Kunka, A., Damborsky, J. & Prokop, Z. (2018). *Methods Enzymol.* **605**, 203–251.
- Kyte, J. & Doolittle, R. F. (1982). *J. Mol. Biol.* **157**, 105–132.
- Los, G. V. & Wood, K. (2007). *Methods Mol. Biol.* **356**, 195–208.
- McPherson, A. & Cudney, B. (2014). *Acta Cryst. F70*, 1445–1467.
- Moore, B. (2018). *Marine Enzymes and Specialized Metabolism*. Cambridge: Academic Press.
- Mueller, U., Förster, R., Hellmig, M., Huschmann, F. U., Kastner, A., Malecki, P., Pühringer, S., Röwer, M., Sparta, K., Steffien, M., Ühlein, M., Wilk, P. & Weiss, M. S. (2015). *Eur. Phys. J. Plus*, **130**, 141–150.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst. D67*, 355–367.
- Novak, H. R., Sayer, C., Isupov, M. N., Gotz, D., Spragg, A. M. & Littlechild, J. A. (2014). *FEBS Lett.* **588**, 1616–1622.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.
- Prokop, Z., Opluštil, F., DeFrank, J. & Damborský, J. (2006). *Biotechnol. J.* **1**, 1370–1380.
- 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**, 6111–6115.
- Robert, X. & Gouet, P. (2014). *Nucleic Acids Res.* **42**, W320–W324.
- Sanner, M. F., Olson, A. J. & Spehner, J. C. (1996). *Biopolymers*, **38**, 305–320.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W.,

- Lopez, R., McWilliam, H., Rimmert, M., Söding, J., Thompson, J. D. & Higgins, D. G. (2011). *Mol. Syst. Biol.* **7**, 539.
- Silberstein, M., Damborsky, J. & Vajda, S. (2007). *Biochemistry*, **46**, 9239–9249.
- Swanson, P. (1999). *Curr. Opin. Biotechnol.* **10**, 365–369.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Vanacek, P., Sebestova, E., Babkova, P., Bidmanova, S., Daniel, L., Dvorak, P., Stepankova, V., Chaloupkova, R., Brezovsky, J., Prokop, Z. & Damborsky, J. (2018). *ACS Catal.* **8**, 2402–2412.
- Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. & Wilson, K. S. (2011). *Acta Cryst.* **D67**, 235–242.
- Yong, J.-J., Park, S.-J., Kim, H.-J. & Rhee, S.-K. (2007). *Int. J. Syst. Evol. Microbiol.* **57**, 951–953.
- Zwart, P. H., Grosse-Kunstleve, R. W., Lebedev, A. A., Murshudov, G. N. & Adams, P. D. (2008). *Acta Cryst.* **D64**, 99–107.

## 4.2. Structural Analysis of the Ancestral Haloalkane Dehalogenase AncLinB-DmbA

This chapter is based on Paper II:

Andrii Mazur, Pavel Grinkevich, Radka Chaloupkova, Petra Havlickova, Barbora Kascakova, Michal Kutý, Jiri Damborsky, Ivana Kuta Smatanova and Tatyana Prudnikova,

*Int. J. Mol. Sci.* **2021**, 22(21), 11992

### **Abstract:**

Haloalkane dehalogenases (EC 3.8.1.5) play an important role in hydrolytic degradation of halogenated compounds, resulting in a halide ion, a proton, and an alcohol. They are used in biocatalysis, bioremediation, and biosensing of environmental pollutants and also for molecular tagging in cell biology. The method of ancestral sequence reconstruction leads to prediction of sequences of ancestral enzymes allowing their experimental characterization. Based on the sequences of modern haloalkane dehalogenases from the subfamily II, the most common ancestor of thoroughly characterized enzymes LinB from *Sphingobium japonicum* UT26 and DmbA from *Mycobacterium bovis* 5033/66 was in silico predicted, recombinantly produced and structurally characterized. The ancestral enzyme AncLinB-DmbA was crystallized using the sitting-drop vapor-diffusion method, yielding rod-like crystals that diffracted X-rays to 1.5 Å resolution. Structural comparison of AncLinB-DmbA with their closely related descendants LinB and DmbA revealed some differences in overall structure and tunnel architecture. Newly prepared AncLinB-DmbA has the highest active site cavity volume and the biggest entrance radius on the main tunnel in comparison to descendant enzymes. Ancestral sequence reconstruction is a powerful technique to study molecular evolution and design robust proteins for enzyme technologies.



Article

# Structural Analysis of the Ancestral Haloalkane Dehalogenase AncLinB-DmbA

Andrii Mazur <sup>1</sup>, Pavel Grinkevich <sup>1</sup>, Radka Chaloupkova <sup>2,3</sup>, Petra Havlickova <sup>1</sup>, Barbora Kascakova <sup>1</sup>, Michal Kutý <sup>1</sup>, Jiri Damborsky <sup>2,4</sup>, Ivana Kuta Smatanova <sup>1,\*</sup> and Tatyana Prudnikova <sup>1,\*</sup>

- <sup>1</sup> Faculty of Science, University of South Bohemia in Ceske Budejovice, Branisovska 1760, 370 05 Ceske Budejovice, Czech Republic; mazura00@prf.jcu.cz (A.M.); pavel.grinkevich@gmail.com (P.G.); havlip04@prf.jcu.cz (P.H.); karafb00@prf.jcu.cz (B.K.); kutym@seznam.cz (M.K.)
- <sup>2</sup> Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, Kamenice 5, 625 00 Brno, Czech Republic; radka@chemi.muni.cz (R.C.); jiri@chemi.muni.cz (J.D.)
- <sup>3</sup> Enantis Ltd., Kamenice 771/34, 625 00 Brno, Czech Republic
- <sup>4</sup> International Clinical Research Center, St Anne's University Hospital Brno, Pekarska 53, 656 91 Brno, Czech Republic
- \* Correspondence: ivanaks@seznam.cz (I.K.S.); talianensis@gmail.com (T.P.)

**Abstract:** Haloalkane dehalogenases (EC 3.8.1.5) play an important role in hydrolytic degradation of halogenated compounds, resulting in a halide ion, a proton, and an alcohol. They are used in biocatalysis, bioremediation, and biosensing of environmental pollutants and also for molecular tagging in cell biology. The method of ancestral sequence reconstruction leads to prediction of sequences of ancestral enzymes allowing their experimental characterization. Based on the sequences of modern haloalkane dehalogenases from the subfamily II, the most common ancestor of thoroughly characterized enzymes LinB from *Sphingobium japonicum* UT26 and DmbA from *Mycobacterium bovis* 5033/66 was in silico predicted, recombinantly produced and structurally characterized. The ancestral enzyme AncLinB-DmbA was crystallized using the sitting-drop vapor-diffusion method, yielding rod-like crystals that diffracted X-rays to 1.5 Å resolution. Structural comparison of AncLinB-DmbA with their closely related descendants LinB and DmbA revealed some differences in overall structure and tunnel architecture. Newly prepared AncLinB-DmbA has the highest active site cavity volume and the biggest entrance radius on the main tunnel in comparison to descendant enzymes. Ancestral sequence reconstruction is a powerful technique to study molecular evolution and design robust proteins for enzyme technologies.



**Citation:** Mazur, A.; Grinkevich, P.; Chaloupkova, R.; Havlickova, P.; Kascakova, B.; Kutý, M.; Damborsky, J.; Kuta Smatanova, I.; Prudnikova, T. Structural Analysis of the Ancestral Haloalkane Dehalogenase AncLinB-DmbA. *Int. J. Mol. Sci.* **2021**, *22*, 11992. <https://doi.org/10.3390/ijms222111992>

Academic Editor: Torsten Herrmann

Received: 7 October 2021

Accepted: 2 November 2021

Published: 5 November 2021

**Keywords:** haloalkane dehalogenase; ancestral sequence reconstruction; structural analysis; halogenated pollutants

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Halogenated compounds are major components of herbicides, insecticides, fungicides, and other chemical agents that are widespread in industry and agriculture. Their environmentally safe disposal still remains an unsolved issue to date [1]. Haloalkane dehalogenases (HLDs) (EC 3.8.1.5) are hydrolytic enzymes that cleave carbon-halogen bonds in a broad range of halogenated aliphatic compounds, resulting in a corresponding alcohol, a halide ion, and a proton [2–4]. The interest in these enzymes is growing because of their utilization in biocatalysis, bioremediation, biosensing, and molecular imaging [5]. Structurally, HLDs belong to the  $\alpha/\beta$ -hydrolase fold superfamily [6,7]. The tertiary structure of HLDs is comprised of two domains: an  $\alpha/\beta$ -hydrolase core domain and a helical cap domain. The  $\alpha/\beta$ -hydrolase domain is composed of an eight-stranded mostly parallel  $\beta$ -sheet flanked by  $\alpha$ -helices and serves as a scaffold for the catalytic residues. The cap domain consists of a few helices inserted in the catalytic domain and is known to influence the substrate specificity of these enzymes [8]. The active site cavity is located between

the core main domain and the cap domain [9]. HLDs are divided into three subfamilies, HLD-I, HLD-II, and HLD-III, according to the composition of their catalytic residues and the anatomy of the cap domain. The catalytic residues in the HLD-I subfamily are D-H-D/W-W, while HLD-II members predominantly contain D-H-E/N-W and HLD-III members contain D-H-D/N-W [10,11].

LinB from *Sphingobium japonicum* UT26 [12,13] and DmbA from *Mycobacterium bovis* 5033/66 [14] are closely related enzymes sharing 68% of sequence identity [15] and both exhibiting a wide range of substrate specificity. The LinB breaks down many halogenated substrates including 1,2-dibromoethane, 1,3-dibromopropane, 1-bromo-3-chloropropane, 1-bromo-2-chloroethane, 4-bromobutanenitrile,  $\gamma$ -hexachlorocyclohexane and 1,2-dibromoethane [15,16]. It has been shown that its active site is one of the largest among other HLDs [17]. DmbA has a similar catalytic behavior as LinB but differs in access tunnels composition. It also catalyzes biodegradation of 1,2-dichloroethane, 2-iodobutane and 1,3-dichloropropene [15,18]. These properties make LinB and DmbA a good target for protein modification and bioengineering. A number of these enzyme variants have been successfully crystallized and their crystal structures have been solved [19–22].

Most of the enzymes have their active sites buried inside the protein core, rather than exposed to the solvent on the surface. These buried active sites are connected to the bulk solvent through tunnels, which act as exchange pathways for the substrate penetration and product release from the active site [23]. The structural and biochemical properties of these tunnels have a major impact on the catalysis and substrate specificity of an enzyme [24]. Therefore, the access tunnels became a frequent target of enzyme modifications [21]. Haloalkane dehalogenases contain several tunnels connecting the protein surface and the active site cavity. The tunnels identified in crystal structures of HLDs are usually referred to as the main tunnel and several slot tunnels; each of them can serve as a path for the substrate access to the active site, release of the reaction products from the active site, and exchange of solvent molecules. The main p1 tunnel is used for the halogenated substrate access to the active site as well as alcohol and halide products exchange. The p2 secondary slot tunnels are served for the alcohol release and water solvent exchange. There are several further variations of tunnel branches originating from p1 and p2 tunnels [21,23,25].

Ancestral sequence reconstruction represents a valuable tool for modification of the structure and biochemical properties of modern enzymes [26]. This technique allows the prediction of the sequence of a hypothetical common ancestor from an already known and related set of sequences of modern-day enzymes. Gene-encoding inferred ancestral sequence can then be synthesized, expressed in expression systems, and characterized. Here we report crystallization and X-ray structural analysis of the ancestral enzyme AncLinB-DmbA that combines structural features of LinB, DmbA and other closely related modern haloalkane dehalogenases.

## 2. Results and Discussion

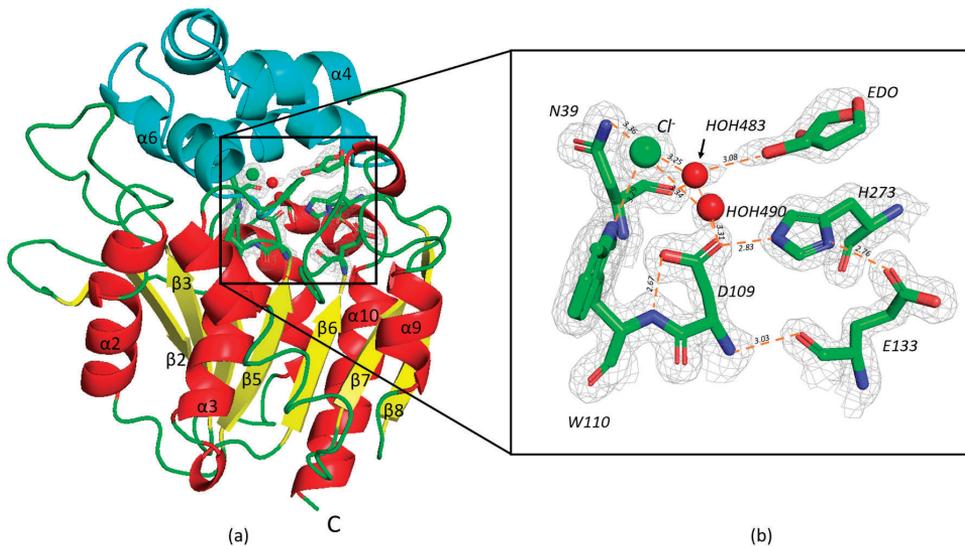
A freshly isolated and purified sample of AncLinB-DmbA in a concentration of 9.8 mg/mL in 50 mM Tris-HCl buffer pH 7.5 was used for the initial screening of suitable crystallization conditions. The initial crystallization screen resulted in the formation of several needle-shaped and 3D crystals. Three-dimensional (3D) crystals grew in the condition with precipitant composed of 30% PEG 4000, 0.1 M sodium citrate pH 5.6, 0.2 M ammonium acetate. Several crystals were prepared for diffraction analysis on the synchrotron. The diffraction dataset was collected at the resolution of 1.5 Å. The data collection statistics are presented in Table 1. The tetragonal  $P4_322$  space group was determined in POINTLESS [27] from CCP4 [28] program suit for further data processing. The model was constructed using molecular replacement and the structure of DmbA (PDB ID 2QVB) [14] sharing 83.2% sequence identity with AncLinB-DmbA was used as a template. After several cycles of restrained and anisotropic refinement as well as a manual rebuilding and refinement in COOT [29] the  $R_{work}$  and  $R_{free}$  of 15.9% and 17.1%, respectively, were achieved.

**Table 1.** The data collection and refinement statistics for AncLinB-DmbA.

Data Collection	
Space Group	P4 <sub>3</sub> 22
<i>a</i> , <i>b</i> , <i>c</i> (Å)	68.89, 68.89, 156.24
$\alpha$ , $\beta$ , $\gamma$ (°)	90.0, 90.0, 90.0
Resolution range (Å)	50–1.50 (1.59–1.50)
Total no. of reflections	641,156 (102,219)
No. of unique reflections	61,367 (9534)
Completeness (%)	99.6 (97.8)
$\langle I/\sigma(I) \rangle$	13.68 (1.65)
$R_{meas}$ *	102.9 (9.6)
CC <sub>1/2</sub>	0.99 (0.77)
Overall <i>B</i> factor from Wilson plot (Å <sup>2</sup> )	27.062
Refinement	
No. of reflections used for refinement	58,297
$R_{work}$ †/ $R_{free}$ ‡ (%)	15.94/17.10
No. of non-H atoms	2810
No. of protein atoms	2475
No. of chloride ions	5
No of ligands	1
No. of water molecules	319
Average <i>B</i> factor (Å <sup>2</sup> )	22.357
Ramachandran plot	
Most favored (%)	95.89
Allowed (%)	3.79
Outliers (%)	0
R.m.s. deviations	
Bonds (Å)	0.019
Angles (°)	1.958
PDB ID	7PW1

\*  $R_{meas}$  is a redundancy- independent merging R factor,  $R_{meas} = \sum_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ , where  $\langle I(hkl) \rangle$  is the mean of the  $N(hkl)$  individual measurements  $I_i(hkl)$  of the density of reflections  $hkl$ . †  $R_{work} = \sum_{hkl} ||F_{obs}| - |F_{calc}|| / \sum_{hkl} |F_{obs}|$ , ‡  $R_{free}$  was calculated using 5% of the reflection data that were excluded from refinement.

The structural organization of AncLinB-DmbA is similar to other HLDs-II subfamily members [9,10]; it is composed of two compact domains, a cap domain and a main domain. The cap domain is responsible for the substrate specificity, while the main domain carries out the catalytic function (Figure 1a). Eight  $\beta$ -strands, where  $\beta_2$  is antiparallel, form a twisted  $\beta$ -sheet surrounded by six  $\alpha$ -helices on both sides. The  $\alpha_2$ ,  $\alpha_3$ ,  $\alpha_8$ , and  $\alpha_9$  are located on the front side and the others on the back side. The cap domain consists of residues 149 to 211. It is formed by  $\alpha_4$ ,  $\alpha_5$ ,  $\alpha_6$ , and  $\alpha_7$  helices and connects to the main domain before  $\alpha_8$  (Figure 1a).



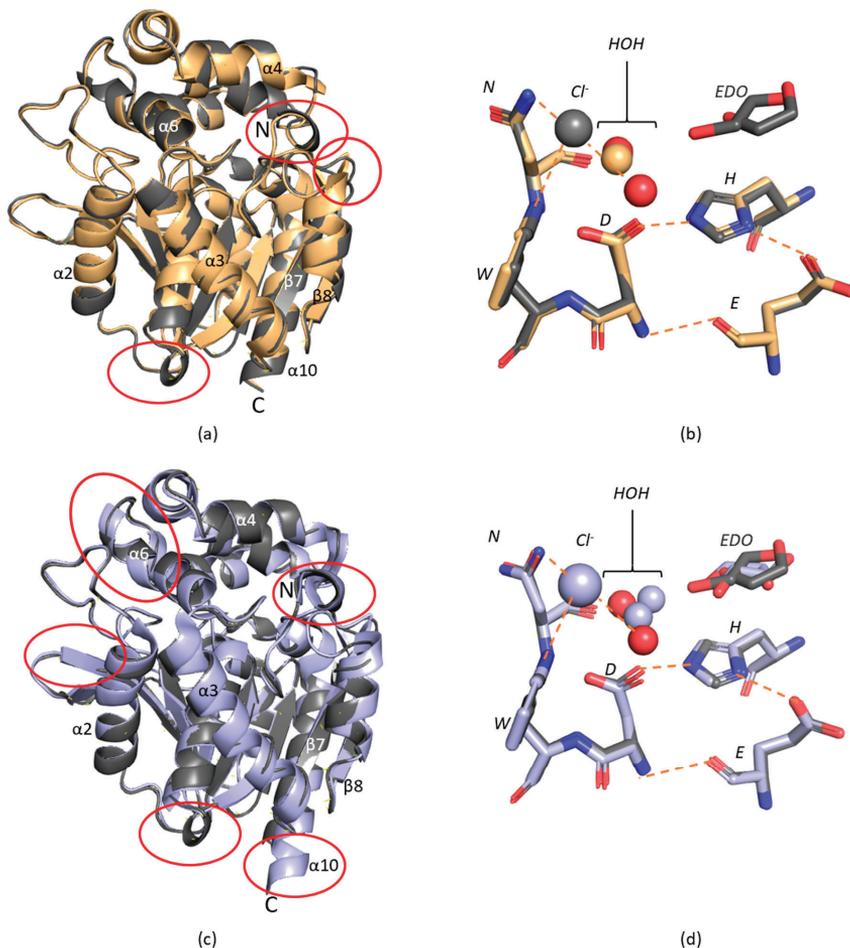
**Figure 1.** Cartoon model of the overall structure of AncLinB-DmbA (a) and the active site with bound ethylene glycol (EDO) (b). The overall structure of AncLinB-DmbA (a) is shown in cartoon representation.  $\alpha$ -helices are shown in red;  $\beta$ -strands are in yellow, and loops are in green for the core domain; the cap domain is shown in cyan. Details of the active site (b): the catalytic pentad and EDO are represented by sticks; the chloride ion and water molecules are shown as green and red spheres, respectively; coordinating interactions are represented by orange dashed lines; the  $2Fo-Fc$  electron density map for ions and interacting residues contoured at  $1.3\sigma$  is drawn as light grey mesh.

The active site of AncLinB-DmbA is similar to LinB, DmbA, and other HLD-II members [10]. It is buried in a hydrophobic cavity between two domains (Figure 1b). The catalytic triad contains nucleophile D109, catalytic acid E133, and catalytic base H273. W110 and N39 act as halide-stabilizing residues in AncLinB-DmbA (Figure 1b). All together three catalytic and two halide-binding amino acids represent the catalytic pentad. The electron density map in the vicinity of the active site was interpreted as a chloride anion coordinated by two water molecules (HOH483 and HOH490) and the ethylene glycol molecule (EDO) as a component of the precipitant in two alternative conformations (Figure 1b).

The chloride anion is coordinated by two halide stabilizing residues with distances between the  $N\epsilon 1$  atom in W110 and the  $N\delta 1$  atom in N39 of 3.29 and 3.37 Å, respectively. Further coordination of chloride ion is provided by N atom from P209 with 3.56 Å distance and two water molecules HOH483 and HOH490 with 3.25 and 3.34 Å distance, respectively (Figure 1b). The catalytic pentad is stabilized by the following hydrogen bonds: the N atom of the E109 interacts with the O atom of E133 with 3.03 Å distance, the D109 O $\delta 1$  atom is coordinated by the  $N\epsilon 1$  of H273 with a distance of 2.83 Å, and H273  $N\delta 1$  interacts with D109 O $\delta 1$  with O $\epsilon 1$  atom of E133 with 2.83 Å and 2.76 Å distances, respectively.

The 3D structure of AncLinB-DmbA was superposed with its closely related structures of descendant enzymes LinB and DmbA. As it is shown in Figure 2a the enzyme structure is highly conserved (RMSD of  $C\alpha$  atoms of AncLinB-DmbA and LinB is 0.778 Å and AncLinB-DmbA and DmbA is 0.895 Å), however, some small differences between the structure of ancestral and descendant enzymes were identified. These differences were uncovered not only in the cap domain but also in the main domain (Figure 2a). The  $\alpha 4$  in the cap domain of AncLinB-DmbA is shorter than the corresponding helices of LinB and DmbA and has a slightly different position; moreover,  $\alpha 6$  in the cap domain of DmbA is longer than the corresponding helices of AncLinB-DmbA and LinB (Figure 2a). Both these differences are presented in the positions, where the cap domain connects

with a core domain. Furthermore, some differences are presented in several loops: the loop near the N-terminus of AncLinB-DmbA forms an additional small helix and another additional helix is formed in AncLinB-DmbA between  $\alpha 3$  and  $\beta 6$ , which is not presented in DmbA and LinB structures. The core domain structure is generally more conserved among all compared enzymes, however, some structural differences in the arrangement of the secondary elements were also observed. The  $\beta 1$ ,  $\beta 2$ , and  $\beta 7$  are much shorter in AncLinB-DmbA than in DmbA and LinB. The  $\beta 1$  differs in AncLinB-DmbA not only by length but also by an angle that is almost orthogonal to the position of  $\beta 1$  in LinB and DmbA. The  $\beta 8$  sheet in AncLinB-DmbA is shorter than in the other two proteins and slightly shifted by 3 Å towards  $\alpha 10$  helix than the corresponding sheet in both modern enzymes. The  $\alpha 10$  helix is longer in DmbA by five amino acids compared to other enzymes (Figure 2a).



**Figure 2.** Structural comparison of AncLinB-DmbA (PDB ID 7PW1), LinB (PDB ID 1CV2) and DmbA (PDB ID 2QVB). (a,c) Superposition of the overall structure of AncLinB-DmbA with LinB (a) and of AncLinB-DmbA with DmbA (c).  $\alpha$  cartoon trace shows elements of the protein secondary structures. The AncLinB-DmbA is colored in grey; LinB is colored in light brown; DmbA is shown in light blue. The most significant structural differences are highlighted by red ellipses. (b,d) Superposition of AncLinB-DmbA with LinB (b) and AncLinB-DmbA with DmbA (d) active sites. Amino acids of AncLinB-DmbA are shown as grey sticks; amino acids of LinB are shown as light-yellow sticks; amino acids of DmbA are shown as light-blue sticks. The ions and other ligands are presented in the same color for related proteins.

The active site is located at the same position in all compared proteins (Figure 2b). The catalytic amino acids are conserved. The water molecules have similar distances to the corresponding coordinating residues in all tested enzymes. The chloride ion is presented in all structures at a similar position. Similarly, to AncLinB-DmbA, ethylene glycol in two conformations was found in the active site of DmbA as an artifact of used crystallization conditions. This leads to the similarities in bond distances of water molecules in both DmbA and AncLinB-DmbA.

The sequence alignment of AncLinB-DmbA (PDB ID 7PW1), LinB from *Sphingobium japonicum* UT26 (PDB ID 1CV2) [12]) and DmbA from *Mycobacterium bovis* 5033/66 (PDB ID 2QVB) [14]) was performed to compare AncLinB-DmbA with related dehalogenases in their access tunnel composition (Figure 3). The alignment shows a highly conserved secondary structure of AncLinB-DmbA to the modern dehalogenases. The active site residues are conserved among all compared proteins; the catalytic triad D109, E133, and H273 (numbers according to AncLinB-DmbA) and two halide-binding residues (N39 and W110) are identical in the tested dehalogenases (Figure 3), which is in agreement with the whole HLD-II subfamily [10,15,30]. However, comparison to the residues involved in access tunnel composition revealed some differences. The main differences in the composition of the access tunnels between the AncLinB-DmbA and the modern enzymes were found in positions 147, 135–139, and 250. The residue corresponding to position 147 in AncLinB-DmbA plays an important functional role, due to its location in the mouth of the p1 tunnel. AncLinB-DmbA contains glutamic acid in position 147, while DmbA contains alanine and LinB contains glutamine in the corresponding position. Another difference in tunnel composition was identified in position 250, where valine is found in AncLinB-DmbA, isoleucine in DmbA, and threonine in LinB. This position is also functionally important due to its location at the entrance of the p2 tunnel.

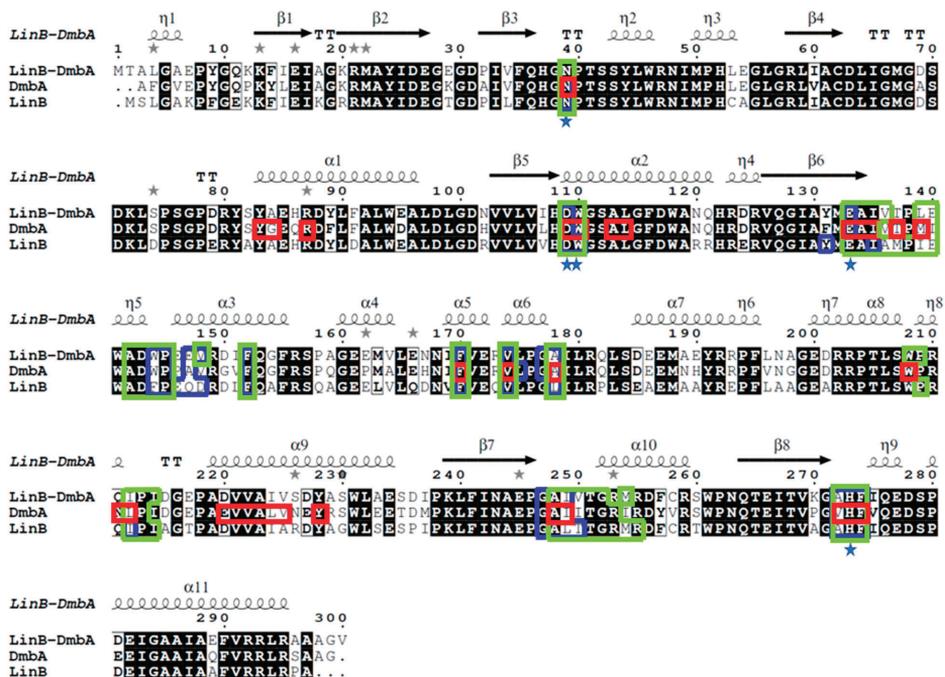
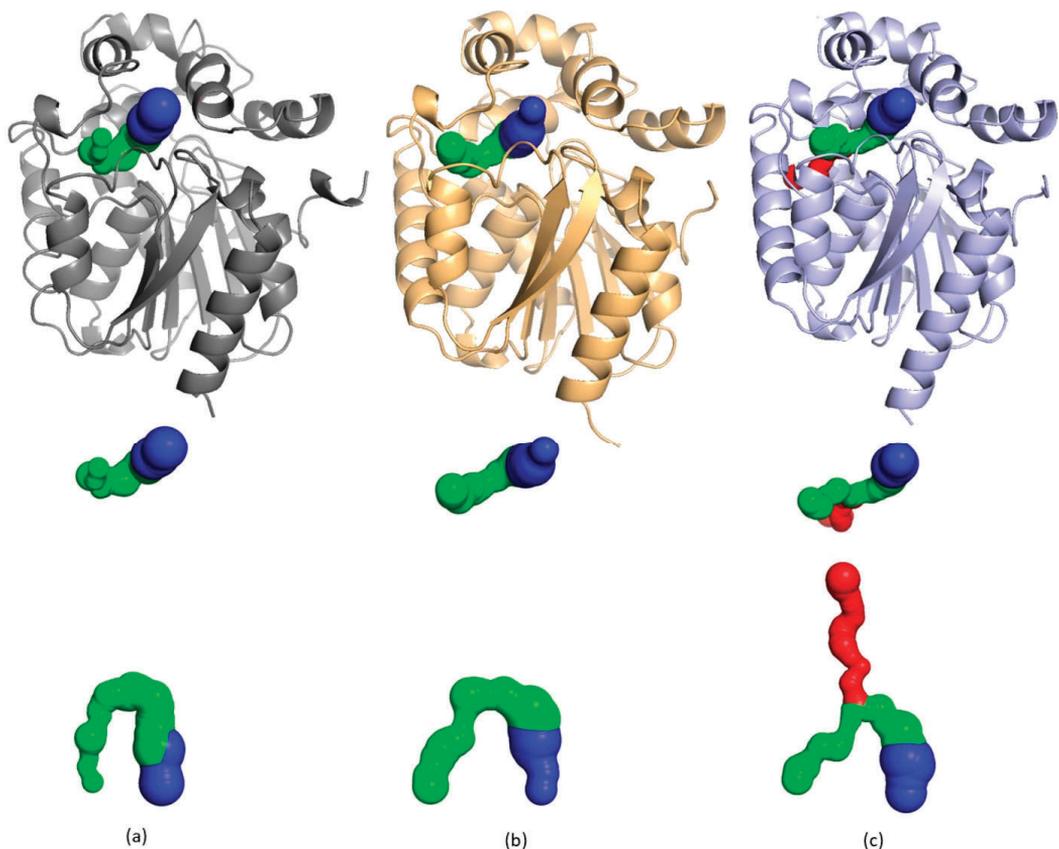


Figure 3. Sequence alignment of ancestor AncLinB-DmbA with modern-day counterparts LinB from *Sphingobium japonicum* UT26 (PDB ID 1CV2) and DmbA from *Mycobacterium bovis* 5033/66 (PDB ID 2QVB). Amino acid residues forming the p1 tunnel

are marked in blue box, the residues of the p2a tunnel are marked in green box and the residues of the p2b tunnel are marked in red box. Catalytic residues are marked by blue stars. Identical residues are highlighted by black background; similar residues are shown in bold typeface. The alignment was generated with Clustal Omega [31] and visualized using ESPript 3.0 [32].

The active site cavity is connected to the solvent on the protein surface by several tunnels. The physical properties of these tunnels play an important role in the substrate specificity of the enzyme [33]. The tunnel architecture was calculated using the program Caver WEB [34], and some differences in the tunnel and cavity constructions were found between the compared proteins (Figure 4). The AncLinB-DmbA and LinB contain one main tunnel p1 (highlighted in blue) and one slot tunnel p2a, while DmbA has one main tunnel p1 and two slot tunnels p2a and p2b (Figure 4). The information about the tunnel composition of the tested enzymes is listed in Table 2.



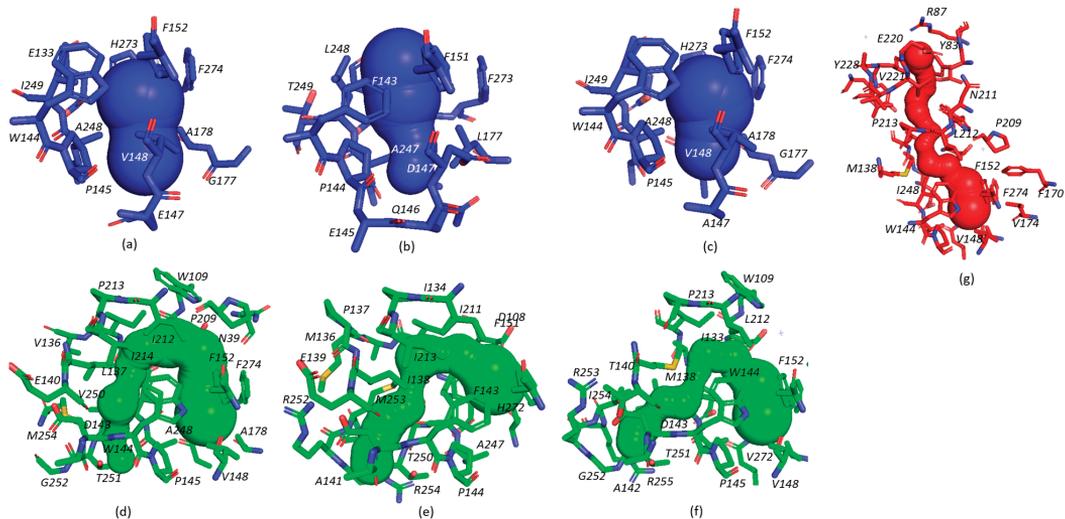
**Figure 4.** Access tunnels identified in crystal structure of AncLinB-DmbA (PDB ID 7PW1) (a), LinB (PDB ID 1CV2) (b) and DmbA (PDB ID 2QVB) (c). The crystal structures are presented in cartoon representation. Access tunnels are presented in two orientations: front (up) and top (down) view. The main (p1) tunnel is presented in blue, and slot p2a and p2b tunnels are presented in green and red, respectively.

**Table 2.** The properties and tunnel composition of selected haloalkane dehalogenases.

	AncLinB-DmbA	LinB	DmbA
PDB code	7PW1	1CV2	2QVB
Resolution (Å)	1.5	1.58	1.19
Sequence identity to AncLinB-DmbA (%)	-	81.7	83.2
R.m.s.d. for AncLinB-DmbA	-	0.778	0.995
Active site cavity volume (Å <sup>3</sup> )	460	406	375
p1 main tunnel characteristics			
Bottleneck radius (Å)	2.0	1.4	2.1
Number of residues	17	22	18
Length (Å)	3.9	6.9	3.9
Entrance radius (Å)	2.5	1.4	2.2
Distance to surface * (Å)	3.9	6.7	3.5
Curvature ‡	1	1	1.1
p2a slot tunnel characteristics			
Bottleneck radius (Å)	1	1	0.9
Number of residues	32	34	30
Length (Å)	22.4	17	19.9
Entrance radius (Å)	1.0	1.7	1.6
Distance to surface * (Å)	8.1	11.5	11.3
Curvature ‡	2.8	1.5	1.8
p2b slot tunnel characteristics			
Bottleneck radius (Å)	-	-	0.9
Number of residues	-	-	36
Length (Å)	-	-	24.3
Entrance radius (Å)	-	-	1.6
Distance to surface * (Å)	-	-	19.4
Curvature ‡	-	-	1.3

\* Distance to surface is a shortest length from the starting point (active site) and the surface in the direction of the tunnel. ‡ Curvature as a description of the shape of the tunnel as the ratio between the length of the tunnel and the shortest possible distance between the starting point and the tunnel ending point.

The p1 tunnel of AncLinB-DmbA is formed by 17 residues. The bottleneck consists of residues P145, E147, V148, V174, A178, A272, G177, and its radius is  $\sim 2$  Å (Figure 5a). The largest bottleneck radius of p1 is determined to be  $\sim 2.10$  Å in DmbA (Figure 5c). The smallest radius is in LinB ( $\sim 1.45$  Å on Figure 5b). The location of the bottleneck of the p2 tunnel in AncLinB-DmbA is close to the entrance to the surface when compared to its descendants. The p2 tunnel of AncLinB-DmbA consists of 32 residues, with the bottleneck formed by D143, W144, P145, A248, I249, T251, M254, V250 with a radius of  $\sim 1$  Å. LinB has the same bottleneck radius. The bottleneck radius of the p2a and p2b tunnels in DmbA is the smallest, both are at  $\sim 0.9$  Å.



**Figure 5.** The p1 and p2 access tunnels with corresponding amino acids of AncLinB-DmbA (PDB ID 7PW1) (a,d), LinB (PDB ID 1CV2) (b,e) and DmbA (PDB ID 2QVB) (c,f,g). The main p1 tunnel with its residues is colored blue, p2a slot tunnel with amino acids is colored green, p2b tunnel with its amino acids is colored red.

The p1 tunnel entrance is located between the  $\alpha 4$ -helix and the  $\beta 7$ - and  $\beta 8$ -sheets in all compared proteins. The p2 tunnel entrance is located near the  $\alpha 9$  in all proteins, however, the p2b tunnel entrance in DmbA is located between  $\alpha 2$  and  $\alpha 8$  on the opposite side of the molecule (Figure 5g). The radius of the entrance in AncLinB-DmbA is the largest among the descendants: it is  $\sim 2.5$  Å, compared to DmbA  $\sim 2.2$  Å and LinB  $\sim 1.4$  Å (Table 2). The entrance to the p1 tunnel additionally consists of residues P145, E147, A178, A272, G177, G247 (Figure 5a). The main difference in the p1 entrance among the tested enzymes is represented by the negatively charged residue E147 in AncLinB-DmbA, while to nonpolar Q146 in LinB and the nonpolar hydrophobic A145 in DmbA. Another major difference is the location of L177 in LinB (Figure 5b), while both DmbA and AncLinB-DmbA contain A178 (Figure 5a,c). The surface residue L177 in LinB is located in the mouth of the p1 tunnel where its side chain partially blocks the entrance of the substrates inside the enzyme active site. The presence of bottleneck residue L177 thus makes the bottleneck radius of LinB p1 tunnel the smallest when compared with AncLinB-DmbA and DmbA. The sidechain of D147 in LinB faces tunnel p1 and thus forms another bottleneck residue together with L177. It can be concluded that these two residues are responsible for the narrowest bottleneck in LinB, compared to other tested enzymes. Both D147 and L177 are important determinants of substrate specificity of LinB and therefore have been subjected to mutagenesis [35]. Other amino acids involved in the p1 tunnel formation are similar to each other and probably do not play a crucial role in enzyme functionality. The difference in entrance composition of p2 tunnel in AncLinB-DmbA is represented by bottleneck residue M254, where this residue is replaced by R253 and R252 in DmbA and LinB respectively. The other bottleneck residues of p2 tunnel in AncLinB-DmbA are L139 and with I249. In DmbA, the bulky M254 is replaced by I254 and thus the bottleneck is formed only by M139 and I249 residues. In LinB, M253 forms a bottleneck of the p2 tunnel together with I138 and L248 (Figure 5d,e). Another difference is W144 residue present in AncLinB-DmbA and DmbA in both tunnels p1 and p2a and F143 in LinB. At this position, the curvature of p2a in AncLinB-DmbA and DmbA is higher than in LinB (Figure 5d–f), which might indicate that W144 is responsible for the flexure. The other residues, participating in the p2a tunnel formation are the same or similar in all compared proteins.

The comparison of tunnels length revealed the longest p1 tunnel in LinB  $\sim 6.9$  Å. The shorter tunnels are in AncLinB-DmbA and in DmbA, both  $\sim 3.9$  Å. The length of p2 tunnels showed different situations: DmbA slot tunnels have length p2a  $\sim 19.9$  Å and p2b  $\sim 24.3$  Å; LinB p2a  $\sim 17$  Å (Table 2). The longest p2a tunnel length is in AncLinB-DmbA  $\sim 22.4$  Å. Due to a big curvature of AncLinB-DmbA, its p2a slot tunnel has the shortest distance to the surface among the others (Table 2). The overall tunnel properties of AncLinB-DmbA correlate more with DmbA than with LinB. The chemical properties of AncLinB-DmbA residues composing entrance to the p1 and p2a tunnels probably play a vital role in the substrate specificity of this protein.

The active site cavity is located between two domains in all dehalogenases and its volume is important in substrate specificity [36]. The volume of the active site cavity of AncLinB-DmbA is  $460$  Å<sup>3</sup>, which is larger compared to both LinB ( $406$  Å<sup>3</sup>) and DmbA ( $375$  Å<sup>3</sup>) (Table 2). Previous studies have demonstrated a correlation of active site cavity volume and access radius with substrate specificity in haloalkane dehalogenases [37,38]. The fact that AncLinB-DmbA has the biggest cavity volume leads to the conclusion that this enzyme exhibits a broader substrate specificity.

Finally, the structure of AncLinB-DmbA was solved and it shares a high degree of similarity with closely related modern-day haloalkane dehalogenases. Despite the highly conserved structure of AncLinB-DmbA and structural similarities with LinB and DmbA, the ancestral enzyme exhibits some differences in the overall structure. This is embodied in variations of  $\alpha$ -helices and  $\beta$ -sheets sizes and positions as well as the positioning of more flexible elements such as loops. Caver analysis revealed differences in several residues responsible for tunnels formation. Nevertheless, the core residues correspond to those in both LinB and DmbA. The differences in access tunnels of compared proteins revealed some advantages of AncLinB-DmbA from its descendants. The bottleneck radius in the p1 tunnel of AncLinB-DmbA is close to the wider bottleneck in DmbA. However, the ancestral protein revealed the widest entrance radius of the main tunnel among the other two proteins. The p2a bottleneck radius of AncLinB-DmbA has the same radius as LinB and from all compared proteins AncLinB-DmbA p2a bottleneck is located at the entrance of the tunnel. The biggest active site cavity was revealed in the AncLinB-DmbA. The largest active site cavity was found in AncLinB-DmbA. The knowledge of the composition of tunnel residues in AncLinB-DmbA may represent an advantage in further modification of its catalytic properties.

### 3. Materials and Methods

#### 3.1. Ancestral Sequence Reconstruction and Gene Synthesis

The ancestral sequence was reconstructed as previously described [39–41]. Protein sequences for the HLD-II subfamily were identified by database searching followed by clustering. The final nonredundant HLD-II dataset comprised several sequences and was used to infer the maximum likelihood phylogenetic tree of the HLD-II subfamily. The topology of the HLD-II tree agreed with the relevant parts of a previously published HLD family tree. The ancestral nodes along the evolutionary lineage to LinB and DmbA were predicted and the most common ancestor of the two enzymes was selected for laboratory resurrection. The selected node, the most probable ancestor sequence, was predicted by assigning to each position the ancestral state with the highest-weighted posterior probability. Positions with posterior probability less than 90% in the most likely ancestral state were considered ambiguous. Gene encoding inferred sequence of AncLinB-DmbA was synthesized artificially (GeneArt, Life technologies, Regensburg, Germany). The codon usage was automatically adapted to the codon bias of *Escherichia coli* genes by GeneArt's website service. For expression purposes, the AncLinB-DmbA gene was subcloned into the expression vector pET21b (Novagen, San Diego, CA, USA) between NdeI and BamHI restriction sites (Table 3).

**Table 3.** Production specifics for AncLinB-DmbA.

Source Organism	Artificial Gene
DNA source	-
Restriction sites	<i>NdeI/BamHI</i>
Vector	pET21b
Expression host	<i>E. coli</i>
Complete amino acid sequence of the construct produced	MTALGAEPYGQKKFIEIAGKRMAYIDEGEGDPIVF QHGNTSSYLWRNIMPHLEGLGRLIACDLIG MGDSKLSPSGPDYSYAEHRDYL FALWEALDL GDNVVLVIHDWGSALGFDWANQHRDRVQ GIAYMEAIVTPLEWADWPPEVRDIFQGFSP AGEEMVLENNIFVERVLP GAILRQLSDEEMAEY RRPFLNAGEDRRPTLSWPRQIPIDGEPADV VAIVSDYASWLAESDIPKLFINAEPGAIVTGRM RDFCRSWPNQTEITVKGAHFIQEDSPDEIGAA IAEFVRRRLAAAGV

### 3.2. Protein Expression and Purification

To overproduce AncLinB-DmbA in *E. coli*, expression of the corresponding gene (under control of the T7lac promoter) was induced by adding isopropyl  $\beta$ -D-thiogalactopyranoside (IPTG). *E. coli* strain BL21(DE3) cells containing recombinant plasmid pET21b: *ancLinB-DmbA* were grown in Luria broth medium containing ampicillin (100  $\mu$ g/mL) at 37 °C. When the cell culture reached an optical density of 0.6 at a wavelength of 600 nm, gene expression was induced by the addition of IPTG (final concentration 0.5 mM) and the cells were cultivated overnight at 20 °C. The cells were harvested, disrupted by sonication using a UP200S ultrasonic processor (Hielscher, Teltow, Germany), and centrifuged for 1 h at 4 °C and 21,000  $\times$  g. The supernatant was collected and further purified on a HiTrap IMAC HP 5 mL column charged with Ni<sup>2+</sup> ions (GE Healthcare, Uppsala, Sweden). The His-tagged enzyme was bound to the resin in an equilibrating buffer (20 mM potassium phosphate buffer, pH 7.5, containing 0.5 M sodium chloride and 10 mM imidazole). Unbound and nonspecifically bound proteins were eliminated by washing with a buffer containing 50 mM imidazole. The enzyme was eluted by a buffer containing 300 mM imidazole. The active fractions were pooled and dialyzed against 50 mM potassium phosphate buffer (pH 7.5) at 4 °C. The Bradford reagent (Sigma-Aldrich, St. Louis, MO, USA) was used to determine the enzyme concentration, with bovine serum albumin used as a standard. Enzyme purity was checked by sodium dodecyl polyacrylamide gel electrophoresis.

### 3.3. Crystallization

The crystallization process was performed manually in CombiClover crystallization plates (Emerald Biosystems, Bainbridge Island, WA, USA) by the sitting drop vapor diffusion method [42]. PegRx and Crystal screen (Hampton Research, Aliso Viejo, CA, USA) as well as NeXtal DWBlock Suites (Quigen, Crawley, UK) commercial crystallization screens were used. Finally, the crystals grew in the crystallization screen Crystal screen (Hampton Research, Aliso Viejo, CA, USA). One microliter of protein solution at a concentration of 9.8 mg/mL in a 50 mM Tris-HCl buffer pH 7.5 was mixed with a reservoir solution in the ratios 1:1 and 1:2 and equilibrated against 500–1000  $\mu$ L of the reservoir solution. The crystals grew with a size appropriate for synchrotron data collection without any further optimization.

### 3.4. Data Collection

The diffraction dataset was collected at the BESSY-II electron storage ring (Berlin-Adlershof, Germany) by the macromolecular crystallographic beamline 14.1, operated by the Helmholtz-Zentrum Berlin [43], equipped with a PILATUS detector (Dectris, Baden, Switzerland).

The crystals of AncLinB-DmbA were mounted in Litholoops (Molecular Dimensions Limited, Sheffield, UK) or nylon cryoloops (Hampton Research, Aliso Viejo, CA, USA) and then flash-cooled in liquid nitrogen. Diffraction experiments were performed at 100 K. All the data collection statistics are summarized in (Table 1).

### 3.5. Structure Solution and Refinement

The dataset was indexed and integrated using the XDS software package [44], and scaled using the program Scala from the CCP4 program package [28]. The structure was solved using the molecular replacement method by the MOLREP program [45]. The structure was refined by the REFMAC5 program [46] and manual building in COOT [29]. Figures with structural representations were prepared using PyMOL [47]. The data refinement statistics are summarized in (Table 1). The Caver Web v.1.0 program was used for tunnel detection and visualization [34].

## 4. Conclusions

In summary, the structural analysis of the reconstructed ancestral enzyme AncLinB-DmbA was performed, and the crystal structure was compared with closely related descendants LinB and DmbA. Despite high sequence similarities, all three proteins exhibited structural differences in the size and special arrangement of some secondary structure elements in both the main and the cap domains. Significant differences in the architecture of the access tunnels were also found. The tunnels of AncLinB-DmbA are more similar to those of DmbA in size and number of residues. The physical properties (length, curvature, radius of bottlenecks, and active site cavity volume) of tunnels in AncLinB-DmbA show that this protein has the best characteristics from its descendants. AncLinB-DmbA was found to have the largest volume of the active site cavity and the largest entrance radius of the p1 main tunnel among the compared proteins. In addition, the curvature of the slot p2a channel in AncLinB-DmbA is the highest in comparison to LinB and DmbA. In conclusion, the structural information on AncLinB-DmbA reported here improves the understanding of the enzyme properties and supports the method of ancestral sequence reconstruction as a valuable tool for enzyme modification.

**Author Contributions:** Conceptualization, I.K.S., J.D. and T.P.; methodology, R.C.; software, A.M., P.G. and M.K.; validation A.M., P.G., T.P., B.K., P.H. and M.K.; formal analysis, A.M., P.G., B.K., P.H. and T.P.; investigation and resources, I.K.S. and J.D.; data curation, A.M., P.G. and T.P.; writing—original draft preparation, A.M., P.G. and T.P.; writing—review and editing, A.M., P.G., T.P., B.K., P.H., I.K.S. and R.C.; visualization, A.M., P.G. and T.P.; supervision, I.K.S., J.D. and T.P.; project administration and funding acquisition, I.K.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by ERDF project, grant number CZ.02.1.01/0.0/0.0/15\_003/0000441, GAJU, grant number 17/2019/P and Czech Ministry of Education, grant number CZ.02.1.01/0.0/0.0/16\_026/0008451 and LM2018121.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the Corresponding author.

**Acknowledgments:** The diffraction data were collected at the beam line MX14.1 of the BESSY II electron storage ring operated by the Helmholtz-Zentrum Berlin. We would particularly like to acknowledge the help and support of Manfred S. Weiss during data collection. We would like to give our special thanks to Eva Šebestová for performing the sequence reconstruction calculations.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Gerba, C.P. Environmental Toxicology. In *Environmental and Pollution Science*, 3rd ed.; Brusseau, M.L., Pepper, I.L., Gerba, C.P., Eds.; Academic Press: New York, NY, USA, 2019; pp. 511–540.
2. Ollis, D.L.; Cheah, E.; Cygler, M.; Dijkstra, B.; Frolow, F.; Franken, S.M.; Harel, M.; Remington, S.J.; Silman, I.; Schrag, J.; et al. The alpha/beta hydrolase fold. *Protein Eng.* **1992**, *5*, 197–211. [[CrossRef](#)] [[PubMed](#)]
3. de Jong, R.M.; Tiesinga, J.J.; Rozeboom, H.J.; Kalk, K.H.; Tang, L.; Janssen, D.B.; Dijkstra, B.W. Structure and mechanism of a bacterial haloalcohol dehalogenase: A new variation of the short-chain dehydrogenase/reductase fold without an NAD(P)H binding site. *EMBO J.* **2003**, *22*, 4933–4944. [[CrossRef](#)] [[PubMed](#)]
4. Verschuere, K.H.; Seljée, F.; Rozeboom, H.J.; Kalk, K.H.; Dijkstra, B.W. Crystallographic analysis of the catalytic mechanism of haloalkane dehalogenase. *Nature* **1993**, *363*, 693–698. [[CrossRef](#)] [[PubMed](#)]
5. Koudelakova, T.; Bidmanova, S.; Dvorak, P.; Pavelka, A.; Chaloupkova, R.; Prokop, Z.; Damborsky, J. Haloalkane dehalogenases: Biotechnological applications. *Biotechnol. J.* **2013**, *8*, 32–45. [[CrossRef](#)]
6. Janssen, D.B.; Dinkla, I.J.; Poelarends, G.J.; Terpstra, P. Bacterial degradation of xenobiotic compounds: Evolution and distribution of novel enzyme activities. *Environ. Microbiol.* **2005**, *7*, 1868–1882. [[CrossRef](#)]
7. Fung, H.K.; Gadd, M.S.; Drury, T.A.; Cheung, S.; Guss, J.M.; Coleman, N.V.; Matthews, J.M. Biochemical and biophysical characterisation of haloalkane dehalogenases DmrA and DmrB in Mycobacterium strain J560 and their role in growth on haloalkanes. *Mol. Microbiol.* **2015**, *97*, 439–453. [[CrossRef](#)]
8. Holmquist, M. Alpha/Beta-hydrolase fold enzymes: Structures, functions and mechanisms. *Curr. Protein Pept. Sci.* **2000**, *1*, 209–235. [[CrossRef](#)]
9. Chaloupkova, R.; Prudnikova, T.; Rezacova, P.; Prokop, Z.; Koudelakova, T.; Daniel, L.; Brezovsky, J.; Ikeda-Ohtsubo, W.; Sato, Y.; Kutý, M.; et al. Structural and functional analysis of a novel haloalkane dehalogenase with two halide-binding sites. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2014**, *70*, 1884–1897. [[CrossRef](#)]
10. Chovancova, E.; Kosinski, J.; Bujnicki, J.M.; Damborsky, J. Phylogenetic analysis of haloalkane dehalogenases. *Proteins* **2007**, *67*, 305–316. [[CrossRef](#)]
11. Ang, T.F.; Maiangwa, J.; Salleh, A.B.; Normi, Y.M.; Leow, T.C. Dehalogenases: From Improved Performance to Potential Microbial Dehalogenation Applications. *Molecules* **2018**, *23*, 1100. [[CrossRef](#)]
12. Marek, J.; Vevodova, J.; Smatanova, I.K.; Nagata, Y.; Svensson, L.A.; Newman, J.; Takagi, M.; Damborsky, J. Crystal structure of the haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26. *Biochemistry* **2000**, *39*, 14082–14086. [[CrossRef](#)]
13. Nagata, Y.; Miyauchi, K.; Takagi, M. Complete analysis of genes and enzymes for gamma-hexachlorocyclohexane degradation in *Sphingomonas paucimobilis* UT26. *J. Ind. Microbiol. Biotechnol.* **1999**, *23*, 380–390. [[CrossRef](#)]
14. Mazumdar, P.A.; Hulecki, J.C.; Cherney, M.M.; Garen, C.R.; James, M.N. X-ray crystal structure of Mycobacterium tuberculosis haloalkane dehalogenase Rv2579. *Biochim. Biophys. Acta* **2008**, *1784*, 351–362. [[CrossRef](#)]
15. Koudelakova, T.; Chovancova, E.; Brezovsky, J.; Monincova, M.; Fortova, A.; Jarkovsky, J.; Damborsky, J. Substrate specificity of haloalkane dehalogenases. *Biochem. J.* **2011**, *435*, 345–354. [[CrossRef](#)]
16. Nagata, Y.; Miyauchi, K.; Damborsky, J.; Manova, K.; Ansorgova, A.; Takagi, M. Purification and characterization of a haloalkane dehalogenase of a new substrate class from a gamma-hexachlorocyclohexane-degrading bacterium, *Sphingomonas paucimobilis* UT26. *Appl. Environ. Microbiol.* **1997**, *63*, 3707–3710. [[CrossRef](#)] [[PubMed](#)]
17. Kmunicek, J.; Hynkova, K.; Jedlicka, T.; Nagata, Y.; Negri, A.; Gago, F.; Wade, R.C.; Damborsky, J. Quantitative analysis of substrate specificity of haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26. *Biochemistry* **2005**, *44*, 3390–3401. [[CrossRef](#)]
18. Jesenska, A.; Pavlova, M.; Strouhal, M.; Chaloupkova, R.; Tesinska, I.; Monincova, M.; Prokop, Z.; Bartos, M.; Pavlik, I.; Rychlik, I.; et al. Cloning, biochemical properties, and distribution of mycobacterial haloalkane dehalogenases. *Appl. Environ. Microbiol.* **2005**, *71*, 6736–6745. [[CrossRef](#)]
19. Degtjarik, O.; Chaloupkova, R.; Rezacova, P.; Kutý, M.; Damborsky, J.; Kuta Smatanova, I. Differences in crystallization of two LinB variants from *Sphingobium japonicum* UT26. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **2013**, *69*, 284–287. [[CrossRef](#)] [[PubMed](#)]
20. Okai, M.; Ohtsuka, J.; Imai, L.F.; Mase, T.; Moriuchi, R.; Tsuda, M.; Nagata, K.; Nagata, Y.; Tanokura, M. Crystal structure and site-directed mutagenesis analyses of haloalkane dehalogenase LinB from *Sphingobium* sp. strain MI1205. *J. Bacteriol.* **2013**, *195*, 2642–2651. [[CrossRef](#)]
21. Iermak, I.; Degtjarik, O.; Havlickova, P.; Kutý, M.; Chaloupkova, R.; Damborsky, J.; Prudnikova, T.; Kuta Smatanova, I. Description of Transport Tunnel in Haloalkane Dehalogenase Variant LinB D147C+L177C from *Sphingobium japonicum*. *Catalysts* **2021**, *11*, 5. [[CrossRef](#)]
22. Brezovsky, J.; Babkova, P.; Degtjarik, O.; Fortova, A.; Gora, A.; Iermak, I.; Rezacova, P.; Dvorak, P.; Smatanova, I.K.; Prokop, Z.; et al. Engineering a de Novo Transport Tunnel. *ACS Catal.* **2016**, *6*, 7597–7610. [[CrossRef](#)]
23. Kokkonen, P.; Slanska, M.; Dockalova, V.; Pinto, G.P.; Sanchez-Carnerero, E.M.; Damborsky, J.; Klan, P.; Prokop, Z.; Bednar, D. The impact of tunnel mutations on enzymatic catalysis depends on the tunnel-substrate complementarity and the rate-limiting step. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 805–813. [[CrossRef](#)] [[PubMed](#)]
24. Kokkonen, P.; Bednar, D.; Pinto, G.; Prokop, Z.; Damborsky, J. Engineering enzyme access tunnels. *Biotechnol. Adv.* **2019**, *37*, 107386. [[CrossRef](#)]

25. Li, A.; Shao, Z. Biochemical characterization of a haloalkane dehalogenase DadB from *Alcanivorax dieselolei* B-5. *PLoS ONE* **2014**, *9*, e89144. [[CrossRef](#)]
26. Harms, M.J.; Thornton, J.W. Analyzing protein structure and function using ancestral gene reconstruction. *Curr. Opin. Struct. Biol.* **2010**, *20*, 360–366. [[CrossRef](#)] [[PubMed](#)]
27. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *62*, 72–82. [[CrossRef](#)]
28. Winn, M.D.; Ballard, C.C.; Cowtan, K.D.; Dodson, E.J.; Emsley, P.; Evans, P.R.; Keegan, R.M.; Krissinel, E.B.; Leslie, A.G.; McCoy, A.; et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 235–242. [[CrossRef](#)]
29. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 486–501. [[CrossRef](#)] [[PubMed](#)]
30. Prokop, Z.; Monincová, M.; Chaloupková, R.; Klvana, M.; Nagata, Y.; Janssen, D.B.; Damborský, J. Catalytic mechanism of the haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26. *J. Biol. Chem.* **2003**, *278*, 45094–45100. [[CrossRef](#)]
31. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Soding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [[CrossRef](#)]
32. Robert, X.; Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **2014**, *42*, W320–W324. [[CrossRef](#)]
33. Brezovsky, J.; Kozlikova, B.; Damborsky, J. Computational Analysis of Protein Tunnels and Channels. *Methods Mol. Biol.* **2018**, *1685*, 25–42. [[CrossRef](#)] [[PubMed](#)]
34. Stourac, J.; Vavra, O.; Kokkonen, P.; Filipovic, J.; Pinto, G.; Brezovsky, J.; Damborsky, J.; Bednar, D. Caver Web 1.0: Identification of tunnels and channels in proteins and analysis of ligand transport. *Nucleic Acids Res.* **2019**, *47*, W414–W422. [[CrossRef](#)]
35. Damborsky, J.; Chaloupkova, R.; Pavlova, M.; Chovancova, E.; Brezovsky, J. Structure–Function Relationships and Engineering of Haloalkane Dehalogenases. In *Handbook of Hydrocarbon and Lipid Microbiology*; Timmis, K.N., Ed.; Springer: Berlin, Germany, 2010; pp. 1081–1098.
36. Liang, J.; Edelsbrunner, H.; Woodward, C. Anatomy of protein pockets and cavities: Measurement of binding site geometry and implications for ligand design. *Protein Sci.* **1998**, *7*, 1884–1897. [[CrossRef](#)]
37. Kunka, A.; Damborsky, J.; Prokop, Z. Haloalkane Dehalogenases From Marine Organisms. *Methods Enzymol.* **2018**, *605*, 203–251. [[CrossRef](#)]
38. Buryška, T.; Babkova, P.; Vavra, O.; Damborsky, J.; Prokop, Z. A Haloalkane Dehalogenase from a Marine Microbial Consortium Possessing Exceptionally Broad Substrate Specificity. *Appl. Environ. Microbiol.* **2018**, *84*. [[CrossRef](#)]
39. Babkova, P.; Sebestova, E.; Brezovsky, J.; Chaloupkova, R.; Damborsky, J. Ancestral Haloalkane Dehalogenases Show Robustness and Unique Substrate Specificity. *ChemBioChem* **2017**, *18*, 1448–1456. [[CrossRef](#)] [[PubMed](#)]
40. Chaloupkova, R.; Liskova, V.; Toul, M.; Markova, K.; Sebestova, E.; Hernychova, L.; Marek, M.; Pinto, G.P.; Pluskal, D.; Waterman, J.; et al. Light-Emitting Dehalogenases: Reconstruction of Multifunctional Biocatalysts. *ACS Catal.* **2019**, *9*, 4810–4823. [[CrossRef](#)]
41. Babkova, P.; Dunajova, Z.; Chaloupkova, R.; Damborsky, J.; Bednar, D.; Marek, M. Structures of hyperstable ancestral haloalkane dehalogenases show restricted conformational dynamics. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 1497–1508. [[CrossRef](#)]
42. McPherson, A.; Gavira, J.A. Introduction to protein crystallization. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **2014**, *70*, 2–20. [[CrossRef](#)]
43. Mueller, U.; Förster, R.; Hellmig, M.; Huschmann, F.U.; Kastner, A.; Malecki, P.; Pühringer, S.; Röwer, M.; Sparta, K.; Steffien, M.; et al. The macromolecular crystallography beamlines at BESSY II of the Helmholtz-Zentrum Berlin: Current status and perspectives. *Eur. Phys. J. Plus* **2015**, *130*, 141. [[CrossRef](#)]
44. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 125–132. [[CrossRef](#)] [[PubMed](#)]
45. Vagin, A.; Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 22–25. [[CrossRef](#)] [[PubMed](#)]
46. Murshudov, G.N.; Skubak, P.; Lebedev, A.A.; Pannu, N.S.; Steiner, R.A.; Nicholls, R.A.; Winn, M.D.; Long, F.; Vagin, A.A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 355–367. [[CrossRef](#)]
47. *The PyMOL Molecular Graphics System*; Version 2.4; Schrödinger, LLC: New York, NY, USA, 2020.

### 4.3. Crystallization and Crystallographic Analysis of a Bradyrhizobium Elkanii USDA94 Haloalkane Dehalogenase Variant with an Eliminated Halide-Binding Site

This chapter is based on Paper III:

Tatyana Prudnikova, Barbora Kascakova, Jeroen R. Mesters, Pavel Grinkevich, Petra Havlickova, Andrii Mazur, Anastasiia Shaposhnikova, Radka Chaloupkova, Jiri Damborsky, Michal Kutý and Ivana Kuta Smatanova,

*Crystals* **2019**, 9, 375

#### **Abstract:**

Haloalkane dehalogenases are a very important class of microbial enzymes for environmental detoxification of halogenated pollutants, for biocatalysis, biosensing and molecular tagging. The double mutant (Ile44Leu + Gln102His) of the haloalkane dehalogenase DbeA from *Bradyrhizobium elkanii* USDA94 (DbeA $\Delta$ Cl) was constructed to study the role of the second halide-binding site previously discovered in the wild-type structure. The variant is less active, less stable in the presence of chloride ions and exhibits significantly altered substrate specificity when compared with the DbeAwt. DbeA $\Delta$ Cl was crystallized using the sitting-drop vapour-diffusion procedure with further optimization by the random microseeding technique. The crystal structure of the DbeA $\Delta$ Cl has been determined and refined to the 1.4 Å resolution. The DbeA $\Delta$ Cl crystals belong to monoclinic space group *C121*. The DbeA $\Delta$ Cl molecular structure was characterized and compared with five known haloalkane dehalogenases selected from the Protein Data Bank.

Article

# Crystallization and Crystallographic Analysis of a *Bradyrhizobium Elkanii* USDA94 Haloalkane Dehalogenase Variant with an Eliminated Halide-Binding Site

Tatyana Prudnikova <sup>1,2,†</sup>, Barbora Kascakova <sup>1,†</sup>, Jeroen R. Mesters <sup>3</sup> , Pavel Grinkevich <sup>1</sup>, Petra Havlickova <sup>1</sup>, Andrii Mazur <sup>1,2</sup>, Anastasiia Shaposhnikova <sup>1,2</sup> , Radka Chaloupkova <sup>4</sup>, Jiri Damborsky <sup>4,5</sup> , Michal Kutý <sup>1,2</sup> and Ivana Kuta Smatanova <sup>1,2,\*</sup> 

<sup>1</sup> Faculty of Science, University of South Bohemia in Ceske Budejovice, Branisovska 1760, 37005 Ceske Budejovice, Czech Republic

<sup>2</sup> Center of Nanobiology and Structural Biology, Institute of Microbiology of the Czech Academy of Sciences, Zamek 136, 37333 Nove Hradky, Czech Republic

<sup>3</sup> Institute of Biochemistry, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany

<sup>4</sup> Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, Kamenice 5/A4, 62500 Brno, Czech Republic

<sup>5</sup> International Clinical Research Center, St. Anne's University Hospital Brno, Pekarska 53, 65691 Brno, Czech Republic

\* Correspondence: ivanaks@seznam.cz or talianensis@gmail.com

† These authors contributed equally to this work.

Received: 27 June 2019; Accepted: 19 July 2019; Published: 23 July 2019



**Abstract:** Haloalkane dehalogenases are a very important class of microbial enzymes for environmental detoxification of halogenated pollutants, for biocatalysis, biosensing and molecular tagging. The double mutant (Ile44Leu + Gln102His) of the haloalkane dehalogenase DbeA from *Bradyrhizobium elkanii* USDA94 (DbeA $\Delta$ Cl) was constructed to study the role of the second halide-binding site previously discovered in the wild-type structure. The variant is less active, less stable in the presence of chloride ions and exhibits significantly altered substrate specificity when compared with the DbeAwt. DbeA $\Delta$ Cl was crystallized using the sitting-drop vapour-diffusion procedure with further optimization by the random microseeding technique. The crystal structure of the DbeA $\Delta$ Cl has been determined and refined to the 1.4 Å resolution. The DbeA $\Delta$ Cl crystals belong to monoclinic space group C121. The DbeA $\Delta$ Cl molecular structure was characterized and compared with five known haloalkane dehalogenases selected from the Protein Data Bank.

**Keywords:** Haloalkane dehalogenase; halide-binding site; random microseeding

## 1. Introduction

Hazardous halogenated compounds are an important class of environmental pollutants. An obvious critical step in the potential biodegradation pathway is the dehalogenation process [1,2]. Haloalkane dehalogenases (HLDs) play an essential role in biodegradation of the halogenated pollutants. HLDs are predominantly bacterial enzymes that belong to the superfamily of  $\alpha/\beta$ -hydrolases and catalyze the hydrolytic conversion of a wide range of halogenated aliphatic compounds, and therefore play an important role in bioremediation [3] and industrial biocatalytic processes [4]. An aliphatic alcohol, a halide and a hydrogen cation are released during the enzymatic dehalogenation of haloalkanes by HLDs. The tertiary structures of HLDs are composed of a conserved  $\alpha/\beta$ -hydrolase core domain and an  $\alpha$ -helical cap domain [5]. The core domain is responsible for the catalytic reaction of the enzyme

and the cap domain is essential for substrate specificity and recognition [6]. A deep cleft is situated between these two domains, allowing the solvent to access the buried active site. The active site is composed of two halide-anion stabilizing residues and the catalytic triad consisting of a nucleophile, a base and an acid [7]. HLDs can be divided into three subfamilies, HLD-I, HLD-II and HLD-III, according to the composition of the catalytic residues and the anatomy of the cap domain [4].

A novel HLD DbeA from *B. elkanii* USDA94, a member of HLD-II subfamily [4], was structurally and biochemically characterized [7]. The structure of DbeA wild type was determined to 2.2 Å resolution and displays a typical topology of the  $\alpha/\beta$ -hydrolases (EC 3.8.1.5). The unique feature of the DbeA structure is the presence of two halide-binding sites, both fully occupied by chloride anions [7]. The first halide-binding site is located in the protein active site and is involved in substrate binding and stabilization of halogen ion produced during dehalogenation reaction. DbeA active site consists of five catalytic residues: two halide stabilizing residues (Trp104 and Asn38) and three amino acids essential for the catalytic activity of the enzyme [2,4]: the nucleophile Asp103, the catalytic base His271, and the catalytic acid Glu127. The second halide-binding site in DbeA is unique and has never been observed within HLD structures deposited in the PDB [8]. The second halide-binding site, which is buried in the protein core domain and located approximately 10 Å far from the first halide-binding site, is formed by five amino-acid residues: Ile44, Gln274, Gln102, Gly37 and Thr40 [7]. Superposition of the DbeA structure with other related HLD-II members revealed the presence of two unique amino acids in the second halide-binding site: Gln102 instead of a typical His and Ile44 as a substitution of an ordinary Leu, thereby sufficiently increasing the cavity volume to accommodate the second halide ion. The variant DbeA $\Delta$ Cl (Ile44Leu+Gln102His) was constructed and biochemically characterized to elucidate the role of the second halide-binding site in structure and function of DbeA [7].

Removal of the second halide binding site in DbeA significantly changed the substrate specificity of DbeA $\Delta$ Cl and reduced the catalytic activity by an order of magnitude towards most of the tested substrates [7]. Wild-type DbeA is more active, its melting temperature rises with an increasing concentration of chloride salts and the binding energy for chloride ions is higher when compared with the DbeA $\Delta$ Cl variant [7]. It was suggested that the chloride anion bound in a vicinity of second binding site may increase basicity of catalytic histidine and consequently accelerate the nucleophilic addition of water to the alkyl-enzyme intermediate [7]. In previous attempts, the crystallization of DbeA $\Delta$ Cl was unsuccessful. The obtained crystals were very unstable, sensitive to mechanical stress and poorly diffracted X-rays to about 10 Å resolution. It took several years to grow crystals with an improved diffraction quality. Here, we report the successful crystallization, structure determination and further characterization of DbeA $\Delta$ Cl variant.

## 2. Materials and Methods

### 2.1. Gene Synthesis, Cloning, Expression and Protein Purification

The recombinant gene dbeA $\Delta$ Cl-His6 (Ile44Leu + Gln102His) was synthesized artificially (Entelechon, Regensburg, Germany) according to the DbeA sequence [7] (Table 1). The restriction endonucleases NdeI and XhoI (Fermentas, Burlington, Canada) and T4 DNA ligase (Promega, Madison, USA) were applied to transfer the synthesized gene into the expression vector pET-21b (Novagen, Madison, USA). In order to overexpress DbeA $\Delta$ Cl in *E. coli* BL21(DE3) cells, the final genes were transcribed by T7 RNA polymerase, which is expressed by the isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG)-inducible lac UV5 promoter. Cells containing the plasmid were cultured in Luria broth medium at 310 K. When the culture reached an optical density of 0.6 at a wavelength of 600 nm, gene expression (at 293 K) was induced by the addition of 0.5 mM IPTG. The cells were subsequently harvested and disrupted by sonication using a Soniprep 150 (Sanyo Gallenkamp PLC, Loughborough, England). The supernatant was collected after centrifugation at 100,000 g for 1 h. The crude extract was further purified on a HiTrap Chelating HP 5 ml column charged with Ni<sup>2+</sup> ions (GE Healthcare, Uppsala, Sweden). The His-tagged enzyme was bound to the resin in the presence of 20 mM potassium phosphate buffer

pH 7.5, 0.5 M sodium chloride, 10 mM imidazole. Unbound and non-specifically bound proteins were washed out by a buffer containing 37.5 mM imidazole. The target enzyme was eluted with a buffer containing 300 mM imidazole. The active fractions were pooled and dialyzed overnight against 50 mM Tris-HCl pH 7.5. The DbeAΔCl enzyme was stored at 277 K in 50 mM Tris-HCl pH 7.5 buffer prior to analysis. The DbeAΔCl production information is summarized in Table 1.

**Table 1.** Production specifics for DbeAΔCl.

Source Organism	<i>Bradyrhizobium Elkanii</i> USDA94
DNA source	Artificially synthesized DNA
Transport vector	pMA
Expression vector	pET-21b
Expression host	<i>E. coli</i> BL21(DE3)
Complete amino acid sequence of DbeAΔCl	MTISADISLHHRAVLGSTMAYRETGRSDAPHVFLFHGNPTSSYL WRNIMPLVAPVGHCIAPDLIGYGQSGKPDISYRFFDQADY LDALIDELGIASAYLVAHDWGTALAFHLAARRPQLVRGLA FMEFIRPMRDWSDFHQHDAARETFRKFRTPGVGEAMILDN NAFVERVLPGSILRTLSEEEMAAYRAPFATRESRMPPTLML PRELPIAGEPADVTQALTAHAALAASTYPKLLFVGGSPGA LVSPAFAAEFAKTLKHC AVIQLGAGGHYLQEDHPEAIGRS VAGWIAGIEAASAQRHAALEHHHHHHH

## 2.2. Crystallization

The freshly isolated and purified DbeAΔCl protein was crystallized at a concentration of 30 mg.ml<sup>-1</sup> in 50 mM Tris-HCl buffer pH 7.5 by the sitting-drop vapour-diffusion procedure [9]. For initial screening several commercial precipitants kits were used: JBScreen Classic Kits № 1-10 and Wizard I-III (Jena Bioscience GmbH, Jena, Germany), Morpheus®HT-96, JCSG-plus™ HT-96, PACT premier™ HT-96 and Structure Screen 1 + 2 HT-96 kit (Molecular Dimensions Ltd (MDL), Suffolk, UK), Crystal Screen kit, PEGRx HT™ and PEG/Ion HT™ (Hampton Research (HR), Aliso Viejo, USA) and Axygen I-VIII crystallization kits (Axygen Biosciences, Union City, USA). The CombiClover 96 well plates (MDL, Suffolk, UK) for manual screening experiments as well as Swissi polystyrene MRC 2-drop plate (MDL, Suffolk, UK) were utilized for the initial screening on an Oryx3 robot (Douglas Instruments Ltd, Hungerford, UK) for the DbeAΔCl protein.

The hanging drop crystallization trials were carried out in Limbro 24 well plates (HR, Aliso Viejo, USA). The Douglas Instruments, USA Vapour Batch 96 well plates were used to perform the microbatch under oil crystallization [10]. Macroseeding experiments [11] were carried out by lowering the protein concentration to 20–25 mg.ml<sup>-1</sup>. The counter-diffusion crystallization was performed in single glass capillaries with inner diameters ranging from 0.1 to 0.4 mm (HR, Aliso Viejo, USA) using a three-layer configuration [12].

## 2.3. Data Collection, Processing and Structure Solution

X-ray diffraction data at 100 K were collected to the 1.4 Å resolution at the BESSY II electron storage ring on beamline MX 14.1 of the Helmholtz-Zentrum Berlin (Berlin-Adlershof, Germany; [13]). 2500 images were processed with the graphical user interface XDSAPP [14] for running XDS [15]. Phasing by molecular replacement was performed using MOLREP [16] and the structure of DbeA as the template (PDB code 4k2a; [7]). One molecule was found in the asymmetric unit of DbeAΔCl. Structure refinement and model building was performed using isotropic and anisotropic refinement protocols in REFMAC5 [17] and Coot [18] from the CCP4 package [19], respectively. The quality of the protein models was confirmed with MolProbity [20,21] and wwPDB [22] validation servers. The structure of the DbeAΔCl has been deposited to the Protein Data Bank under accession code 6s42.

Figures with the structure were prepared using the program PyMOL [23]. The complete information about the data collection, processing, and refinement statistics are provided in Table 2.

**Table 2.** Data collection and crystallographic statistics.

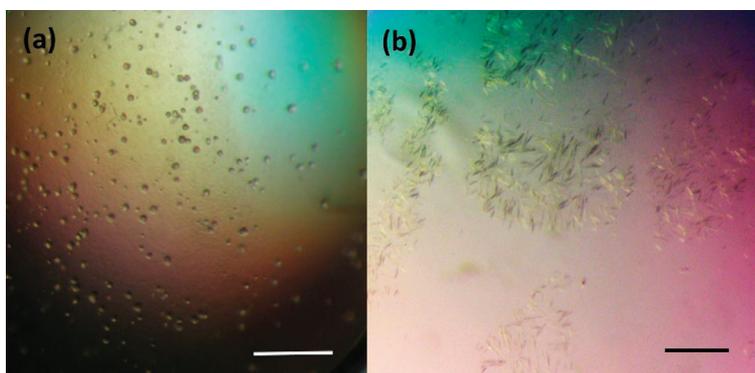
X-ray Diffraction Data Collection Statistics	
Space group	C121
Cell parameters (Å, °)	a = 128.95, b = 63.95, c = 46.05; $\alpha = \gamma = 90$ , $\beta = 106.27$
Wavelength (Å)	0.918
Resolution (Å)	1.39
Number of unique reflections	68,322
Redundancy	2.18 (2.20)
Completeness (%)	96.13 (92.17)
$R_{\text{merge}}^{\#}$	4.9 (26.3)
Average $I/\sigma(I)$	12.72 (3.19)
Wilson B (Å <sup>2</sup> )	20.8
Refinement Statistics	
Resolution range (Å)	41.96–1.4 (1.43–1.39)
No. of reflections in working set	64,907 (4,609)
R value (%) <sup>##</sup>	13.98
$R_{\text{free}}$ value (%) <sup>###</sup>	15.22
RMSD bond length (Å)	0.006
RMSD angle (°)	1.635
No. of atoms in AU	2,823
No. of protein atoms in AU	2,333
No. of water molecules in AU	470
No. of iodide ions in AU	8
No. of chloride ions in AU	3
Mean B value (Å <sup>2</sup> )	13.42
Ramachandran Plot Statistics	
- Residues in favoured regions (%)	97.2
- Residues in allowed regions (%)	100
PDB code	6s42

The data in parentheses refer to the highest-resolution shell. <sup>#</sup>  $R_{\text{merge}} = (|I_{\text{hkl}} - \langle I \rangle|) / I_{\text{hkl}}$ , where the average intensity  $\langle I \rangle$  is taken over all symmetry equivalent measurements and  $I_{\text{hkl}}$  is the measured intensity for any given reflection; <sup>##</sup> R-value =  $|F_o| - |F_c| / |F_o|$ , where  $F_o$  and  $F_c$  are the observed and calculated structure factors, respectively; <sup>###</sup>  $R_{\text{free}}$  is equivalent to R value but is calculated for 5% of the reflections chosen at random and omitted from the refinement process.

### 3. Results and Discussion

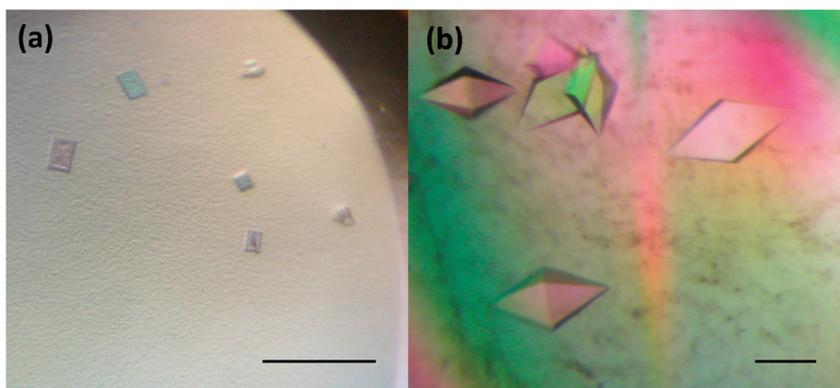
The crystallization procedure previously successfully used for growing of the DbeAwt protein [24] was applied to prepare crystals of freshly purified DbeAΔCl protein. Initially, only a light amorphous precipitation was observed. Further optimization of the crystallization conditions was carried out by variation of the protein and precipitant concentrations (PEG and salt) and drop protein-reservoir ratio composition. This optimization did not improve the results. Additional screening was conducted by applying further commercial crystallization kits: JBScreen Classic Kits № 1-10 and Wizard I-III (Jena Bioscience GmbH, Jena, Germany). Again, only very small and thin needle crystals or a heavy amorphous precipitation were observed. The next optimization step was based on moving the system closer to the metastable zone based on the phase transition diagram by decreasing the concentration of the crystallization drop components. The reduction of the DbeAΔCl concentration and variation of the precipitant concentrations yielded small microcrystals (Figure 1a) and a two-dimensional (2D) single needle crystals (Figure 1b) within a period of 1-3 weeks. The microcrystals were grown within three weeks from precipitant consisting of 28% (w/v) PEG 4000, 0.2 M Li<sub>2</sub>SO<sub>4</sub> in 0.1 M Tris pH 8.2 buffer and a 30 mg.ml<sup>-1</sup> protein concentration. The small needle crystals were observed after 10 days at a 10 to

30 mg.ml<sup>-1</sup> protein concentration in precipitant composed of 12– 17% (w/v) PEG 3350, 115– 125 mM MgCl<sub>2</sub> in 100 mM Tris-HCl 7.5 buffer.



**Figure 1.** Results of initial crystallization experiments of DbeAΔCl protein from *B. elkanii* USDA94: (a) microcrystals and (b) small 2D needle crystals. The scale bar represents 100 μm.

In order to improve the crystal quality, Morpheus®HT-96, JCSG-plus™ HT-96, PACT premier™ HT-96 and Structure Screen 1 + 2 HT-96 kit (Molecular Dimensions Ltd (MDL), Suffolk, UK), Crystal Screen kit, PEGRx HT™ and PEG / Ion HT™ (Hampton Research (HR), Aliso Viejo, USA) and Axygen I-VIII crystallization kits (Axygen Biosciences, Union City, USA) were applied. Finally, small 3D crystals (Figure 2a) with a dimension of about 15 × 5 × 35 μm were grown from a solution containing 26.57% (w/v) hexanediol at 295 K over 10 days. The crystals diffracted X-rays to a maximum resolution of 8–10 Å. The quality of these crystals did not allow to record good diffraction data and additional strategies were needed to improve the size and shape of obtained crystals.



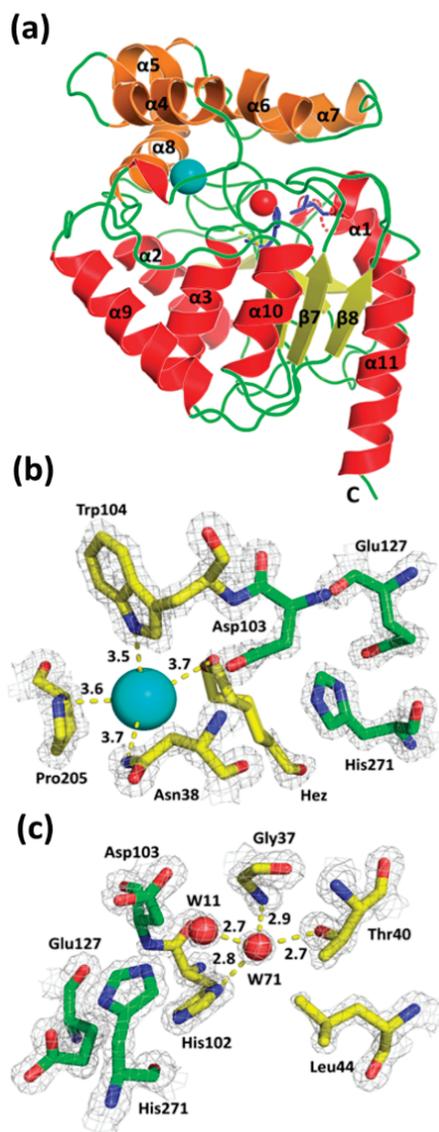
**Figure 2.** DbeAΔCl crystals used for diffraction analysis: (a) small 3D crystals grown at 26.5% (w/v) hexanediol and (b) big 3D crystals developed by random seeding experiments. The scale bar represents 100 μm.

Further optimization using the precipitant solution mentioned above was pursued by setting up the experiments at a lower temperature (193K), by application of the Additive Screen (HR) and by variation of the protein concentration from 10 to 40 mg.mL<sup>-1</sup>, however, without success. The same conditions were tested in a hanging drop vapour-diffusion, microbatch under oil, counter-diffusion and macroseeding procedure. All these experiments did not significantly improve the diffraction quality of the crystals. Next, the small crystals (Figure 2a) were used for random microseeding experiments [25] with application of two commercial crystallization kits: PACT premier™ HT-96 and Structure Screen 1 + 2 HT-96 kit (MDL), which finally resulted in the appearance of 3D crystals with an average dimensions 70 × 75 × 100 μm within a period of two weeks (Figure 2b). These crystals, later used for X-ray data collection experiments, were obtained by lowering the protein concentration to 20 mg.mL<sup>-1</sup> and applying the following precipitant: 0.2M sodium iodide and 20% (w/v) PEG 3350.

Crystals of DbeAΔCl diffracted X-rays to 1.4 Å resolution and belong to the monoclinic base-centered space group C121. The diffraction data allowed localizing 294 amino-acid residues fitting to the one molecule in the asymmetric unit. The overall shape of DbeAΔCl is like a block with 2,823 non-hydrogen atoms (470 water molecules, 8 iodide anions, 3 chloride anions and hexanediol molecule) and corresponds to the canonical architecture of HLDs of the α/β hydrolase fold superfamily (Figure 3a). The structural organization of DbeAΔCl displays two compact domains: an α/β hydrolase core domain and a helical cap domain with the active site located between them. The cap domain (residues 134–214) consists of five α-helices (α4, α5, α5', α6, α7 and α8) and six loops, together forming a lid, protecting the active site cavity. The core domain (residues 4–133 and 215–298) consists of a central twisted eight-stranded β-sheet with the β2 strand running antiparallel. The β-sheet region is flanked by six α-helices: two elements (α1 and α2) cover one side and the remaining four (α3, α9, α10 and α11) the other side [2,26] (Figure 3a). The protein displays a monomer as the biological unit according to the analysis of crystal contacts between molecules in the unit cell and crystal packing. The exploration of macromolecular interfaces by PDBePISA server [27] underpins the monomeric nature of the protein.

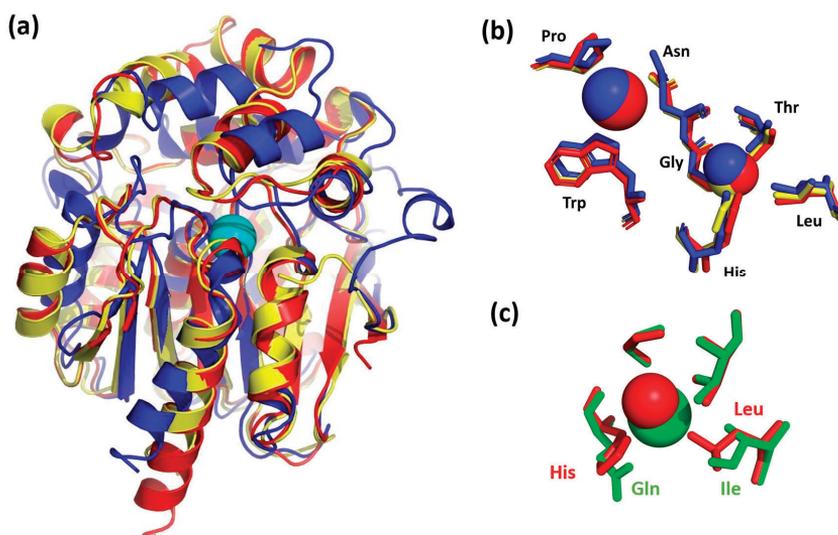
The DbeAΔCl active site displays a substrate-binding pocket typical for all haloalkane dehalogenases. The enzyme's active site cavity contains the catalytic triad consisting of Asp103, His 271 and Glu127. The nucleophile Asp103 is located at the turn between β-strand β5 and helix α3. The catalytic base His271 is positioned on the loop joining β8 and α11. The catalytic acid Glu127 is located behind β-strand β6. Inspecting the electron density map, one iodide anion and one molecule of hexanediol as components of precipitant cocktail were identified near the DbeAΔCl active site. The iodide ion is mainly stabilized by interactions with the N atoms of two halide-binding residues: Asn34 N<sup>δ2</sup> and Trp104 N<sup>ε1</sup> with distances of 3.72 Å and 3.47 Å, respectively. Further coordination is realized with the N atom of the pyrrolidine ring of Pro205 at a distance of 3.64 Å and an O atom of bound hexanediol at 3.68 Å distance (Figure 3b).

The substitutions Ile44Leu and Gln102His introduced into the structure of DbeA result in the elimination of the second halide-binding site (Figure 3c). The site is occupied by a water molecule in the mutant enzyme. Apparently, insufficient space is left for the positioning of the second halide anion as observed in the DbeAwt structure. The atoms of three residues: Thr40 O<sup>γ1</sup>, His102 N<sup>δ1</sup> and Gly37 N, with hydrogen bond distances of 2.74 Å, 2.79 Å, and 2.94 Å, respectively, coordinated water molecule W71 that was modelled instead of the halide anion (Figure 3c). Further coordination is realized by water molecule W11 at a distance of 2.73 Å. Water molecule W11 is situated between two halide-binding sites, 6.73 Å away from the iodide ion in the canonical active site of the protein. Water W11 is stabilized by interaction with the O<sup>δ2</sup> atom of catalytic nucleophile Asp103 at 2.71 Å distance.



**Figure 3.** The overall structure of DbeA $\Delta$ Cl (a), close-up view of the canonical DbeA $\Delta$ Cl active site (b) and second halide-binding site (c). The  $2F_o - F_c$  electron-density map contoured at  $2\sigma$  is shown in grey (a) C $\alpha$  ribbon trace shows elements of the protein secondary structure. The  $\alpha$ -helices are coloured red for the main domain and brown for the cap domain;  $\beta$ -strands are coloured yellow; loops are shown in green; iodide ion is presented as a cyan sphere, water molecule (W71) is shown as a red sphere; the two point substitutions Ile44Leu + Gln102His introduced into DbeA are highlighted as blue sticks. (b) The iodide ion in the active site is presented as a cyan sphere with coordination interaction distances in Å and highlighted by yellow dashed lines; hexanediol (Hez) (shown in two alternative conformations) and amino acids Asp103, Trp104, Pro205 coordinating the iodide ion are shown as sticks with carbon atoms coloured yellow. Carbon atoms of catalytic triad are highlighted in green. (c) The water molecules W11 and W71 are presented as red spheres; amino acids coordinating water molecule W71 are shown as sticks with carbon atoms coloured in yellow with and interactions with distances in Å shown by yellow dashed lines. Water molecule W71 is located in second-halide binding site.





**Figure 5.** Structural comparison of DbeA $\Delta$ Cl with homologous dehalogenases. (a) DbeA $\Delta$ Cl secondary structure superposition with DhaA and Dh1A. C $\alpha$  ribbon trace shows elements of the protein secondary structures. The DbeA $\Delta$ Cl is coloured in red; DhaA is coloured in blue; Dh1A is shown in yellow; iodide ion (cyan sphere) is placed in the canonical active site of DbeA $\Delta$ Cl. (b) Superposition of DbeA $\Delta$ Cl, DhaA and Dh1A active sites. Amino acids of DbeA $\Delta$ Cl are shown as red sticks; amino acids of DhaA are shown as blue sticks; amino acids of Dh1A are shown as yellow sticks. The iodide ion and water molecule W11 bound to DbeA $\Delta$ Cl are represented as red spheres; bromide ion and water are shown as blue spheres for the DhaA structure and the single water molecule in the Dh1A is shown as yellow sphere. (c) Superposition of second halide binding site residues of mutant and wild type DbeA. DbeA $\Delta$ Cl amino acids are shown as red sticks; amino acids of DbeAwt are shown as green sticks. Water W71 is shown as red sphere and Cl<sup>-</sup> ion is shown as green sphere.

In general, the secondary structure elements of the core domain are better conserved compared to the cap domain for the most dehalogenases. Significant differences in cap domains define substrate specificity and variability [1]. The considerable divergence among the superimposed structures was also observable at the relatively disordered N- and C-terminal parts of the proteins (Figure 5a). The position of the active site residues was well conserved among HLDs with some differences in Dh1A structure. DbeA $\Delta$ Cl, DbeA, LinB, DhaA and DmbA belong to the HLD-II subfamily with Asp-His-Glu catalytic triad and Asn-Trp halide binding residues [4] whereas Dh1A belongs to the HLD-I subfamily with an Asp-His-Asp + Trp-Trp catalytic pentad composition. The positions of the halide-binding residues Asn38 and Trp104 in the DbeA $\Delta$ Cl structure is conserved in comparison to the rest of HLD structures. The position of the catalytic triad (Asp103; His271 and Glu127) is well conserved among the dehalogenases from HLD-II subfamily. The side chain of the catalytic acid Asp in Dh1A is situated slightly deeper in the active site cavity in comparison to HLD-II structures.

Superposition of the iodide ion at the DbeA $\Delta$ Cl active site with halide ions in the vicinity of halide-stabilizing residues of the rest structures reveals some structural differences. Dh1A and DmbA contain halide ions inside the active site: Br<sup>-</sup> (Dh1A) and Cl<sup>-</sup> (DmbA) with shifts in 0.47 and 0.39 Å from the DbeA $\Delta$ Cl iodide ion position, respectively. However, there is only a water molecule in the LinB structure (0.27 Å away from iodide ion of DbeA $\Delta$ Cl) and an empty space in the DhaA structure (the closest water molecule is located 6.48 Å away from the iodide ion position).

Water molecule W71, which substitutes the halide ion at the second halide-binding site of the DbeA wild type structure, was found at the canonical place where only a water molecule is present in all the HLDs with a shift of 1.39 Å compared to the DhaA structure, 0.80 Å to the Dh1A, 0.55 Å to the

LinB and 0.43 Å to the DmbA, and 1.28 Å away compared to the second chloride anion in the DbeA structure. Superposition of halide binding sites of the DbeAΔCl structure with halide binding sites of the closest relative (DhaA) and the lowest identity HLD (DhlA) is shown in the Figure 5b.

#### 4. Conclusions

After many crystallization experiments and optimization cycles, we found that two point mutations deeply buried in the structure have a significant influence on the crystallization of the macromolecule. Finally, random microseeding experiments with a seed stock prepared from small 3D crystals obtained at a protein concentration of 20 mg.mL<sup>-1</sup> and 0.2 M sodium iodide plus 20% (w/v) PEG 3350 as the precipitant solutions produced crystals of the double mutant DbeAΔCl from *B. elkanii* USDA94. These crystals were of sufficient quality for X-ray diffraction experiments. The structure of DbeAΔCl was solved using molecular replacement and refined to 1.4 Å resolution. Overall, the structure is very similar to other HLDs structures of the α/β hydrolase fold superfamily (EC 3.8.1.5). The substitutions Ile44Leu and Gln102His resulted in a space reduction of the second halide-binding site and thus incapability of DbeAΔCl to bind a second halide anion as compared to the wild type structure. Instead of a chloride anion, a water molecule (W71) was found in the site of which the consequences are, DbeAΔCl is less active and less stable in the presence of chloride salts when compared with the DbeAwt enzyme [7]. The DbeAΔCl structure of *B. elkanii* was aligned and compared with five molecular structures of haloalkane dehalogenases selected from the PDB. The elements of the secondary structure and the catalytic pentad are well conserved. Superposition of the iodide ion at the DbeAΔCl active site with the other structures reveals some structural differences. The water molecule W71 located at the compromised second halide-binding site of DbeAΔCl was coordinated at the canonical place as compared to other HLDs.

**Author Contributions:** T.P., J.R.M. and I.K.S. designed the experiments and solved the structure. T.P., B.K., P.G., J.R.M. and M.K. analyzed the data. T.P. and B.K. wrote the manuscript. R.C., J.D., A.M., P.H. and A.S. provided technical support.

**Funding:** This work was supported by the Grant Agency of the Czech Republic 17-24321S; DAAD mobility grant DAAD-16-09; ERDF project CZ.02.1.01/0.0/0.0/15\_003/0000441; Ministry of Education, Youth and Sports of the Czech Republic (CZ.1.05/2.1.00/01.0024, CZ.1.05/2.1.00/01.0001, and LM2015055); GAJU 17/2019P.

**Acknowledgments:** The diffraction data were collected on the beam line MX14.1 at the BESSY II electron storage ring operated by the Helmholtz-Zentrum Berlin. We would particularly like to acknowledge the help and support of Manfred S. Weiss during data collection. Also, we would like to thank Stefan A. Kolek (Douglas Instruments Ltd, Hungerford, UK) for providing crucial information about the random seeding experiment during the 2014 FEBS-Instruct practical course PC14-005 at Nove Hrad, Czech Republic.

**Conflicts of Interest:** No interest conflict exists among the authors.

#### References

1. 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. Haloalkane dehalogenases: Structure of a Rhodococcus enzyme. *Biochemistry* **1999**, *38*, 16105–16114. [[CrossRef](#)] [[PubMed](#)]
2. Janssen, D.B.; Dinkla, I.J.T.; Poelarends, G.J.; Terpstra, P. Bacterial degradation of xenobiotic compounds: Evolution and distribution of novel enzyme activities. *Environ. Microbiol.* **2005**, *7*, 1868–1882. [[CrossRef](#)] [[PubMed](#)]
3. Prokop, Z.; Sato, Y.; Brezovsky, J.; Mozga, T.; Chaloupkova, R.; Koudelakova, T.; Jerabek, P.; Stepankova, V.; Natsume, R.; van Leeuwen, J.G.; et al. Enantioselectivity of haloalkane dehalogenases and its modulation by surface loop engineering. *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 6111–6115. [[CrossRef](#)] [[PubMed](#)]
4. Chovancova, E.; Kosinski, J.; Bujnicki, J.M.; Damborsky, J. Phylogenetic analysis of haloalkane dehalogenases. *Proteins* **2007**, *62*, 305–306. [[CrossRef](#)] [[PubMed](#)]
5. Holmquist, M. Alpha/beta-hydrolase fold enzymes: Structures, functions and mechanisms. *Curr. Protein Pept. Sci.* **2000**, *1*, 209–235. [[CrossRef](#)] [[PubMed](#)]

6. Prokop, Z.; Oplustil, F.; DeFrank, J.; Damborsky, J. Enzymes fight chemical weapons. *Biotechnol. J.* **2006**, *1*, 1370–1380. [[CrossRef](#)]
7. Chaloupkova, R.; Prudnikova, T.; Rezacova, P.; Prokop, Z.; Koudelakova, T.; Daniel, L.; Brezovsky, J.; Ikeda-Ohtsubo, W.; Sato, Y.; Kutý, M.; et al. Structural and functional analysis of a novel haloalkane dehalogenase with two halide-binding sites. *Acta Cryst.* **2014**, *70*, 1884–1897. [[CrossRef](#)]
8. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [[CrossRef](#)]
9. Ducruix, A.; Giegé, R. *Crystallization of Nucleic Acids and Proteins*; Oxford University Press: Oxford, UK, 1999. [[CrossRef](#)]
10. Chayen, N.E.J. Crystallization with oils: A new dimension in macromolecular crystal growth. *J. Cryst. Growth* **1999**, *196*, 434–441. [[CrossRef](#)]
11. Bergfors, T.M. *Protein Crystallization: Techniques, Strategies and Tips*; International University Line: La Jolla, CA, USA, 1999. [[CrossRef](#)]
12. Gavira, J.A.; Jesus, W.; Camara-Artigas, A.; Lopez-Garriga, J.; Garcia-Ruiz, J.M. Crystallization and diffraction patterns of the oxy and cyano forms of the *Lucina pectinata* haemoglobins complex. *Acta Cryst.* **2006**, *62*, 196–199. [[CrossRef](#)]
13. Gerlach, M.; Mueller, U.; Weiss, M.S. The MX beamlines BL14.1-3 at BESSY II. *JLSRF* **2016**, *2*, 1–6. [[CrossRef](#)]
14. Sparta, K.M.; Krug, M.; Heinemann, U.; Mueller, U.; Weiss, M.S. XDSAPP2.0. *J. Appl. Cryst.* **2016**, *49*, 1085–1092. [[CrossRef](#)]
15. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **1993**, *26*, 795–800. [[CrossRef](#)]
16. Vagin, A.; Teplyakov, A. MOLREP: An automated program for Molecular Replacement. *J. Appl. Cryst.* **1997**, *30*, 1022–1025. [[CrossRef](#)]
17. Murshudov, G.N.; Skubak, P.; Lebedev, A.A.; Pannu, N.S.; Steiner, R.A.; Nicholls, R.A.; Winn, M.D.; Long, F.; Vagin, A.A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Cryst.* **2011**, *67*, 355–367. [[CrossRef](#)]
18. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and development of Coot. *Acta Cryst.* **2010**, *66*, 486–501. [[CrossRef](#)]
19. Winn, M.D.; Ballard, C.C.; Cowtan, K.D.; Dodson, E.J.; Emsley, P.; Evans, P.R.; Keegan, R.M.; Krissinel, E.B.; Leslie, A.G.W.; McCoy, A.; et al. Overview of the CCP4 suite and current developments. *Acta Cryst.* **2011**, *67*, 235–242. [[CrossRef](#)]
20. 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. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Cryst.* **2010**, *66*, 12–21. [[CrossRef](#)]
21. Hintze, B.J.; Lewis, S.M.; Richardson, J.S.; Richardson, D.C. Molprobity's ultimate rotamer-library distributions for model validation. *Proteins* **2016**, *84*, 1177–1189. [[CrossRef](#)]
22. Gore, S.; Velankar, S.; Kleywegt, G.J. Implementing an X-ray validation pipeline for the Protein Data Bank. *Acta Cryst.* **2012**, *68*, 478–483. [[CrossRef](#)]
23. Schrodinger, L.L.C. The PyMOL Molecular Graphics System, Version 2.0. 2019. Available online: <https://pymol.org/2/> (accessed on 23 July 2019).
24. Prudnikova, T.; Mozga, T.; Rezacova, P.; Chaloupkova, R.; Sato, Y.; Nagata, Y.; Brynda, J.; Kutý, M.; Damborsky, J.; Kuta-Smatanova, I. Crystallization and Preliminary X-ray Analysis of a Novel Haloalkane Dehalogenase DbeA from *Bradyrhizobium elkanii* USDA94. *Acta Cryst.* **2009**, *65*, 353–356. [[CrossRef](#)]
25. Shaw Stewart, P.D.; Kolek, S.A.; Briggs, R.A.; Chayen, N.E.; Baldock, P.F.M. Random Microseeding: A Theoretical and Practical Exploration of Seed Stability and Seeding Techniques for Successful Protein Crystallization. *Cryst. Growth Des.* **2011**, *11*, 3432–3441. [[CrossRef](#)]
26. Ollis, D.L.; Cheah, E.; Cygler, M.; Dijkstra, B.; Frolow, F.; Franken, S.M.; Harel, M.; Remington, S.J.; Silman, I.; Schrag, J.; et al. The alpha/beta hydrolase fold. *Protein Eng.* **1992**, *5*, 197–211. [[CrossRef](#)] [[PubMed](#)]
27. Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **2007**, *372*, 774–797. [[CrossRef](#)] [[PubMed](#)]
28. Pikkemaat, M.G.; Ridder, I.S.; Rozeboom, H.J.; Kalk, K.H.; Dijkstra, B.W.; Janssen, D.B. Crystallographic and kinetic evidence of a collision complex formed during halide import in haloalkane dehalogenase. *Biochemistry* **1999**, *38*, 12052–12061. [[CrossRef](#)] [[PubMed](#)]

29. Marek, J.; Vevodova, J.; Smatanova, I.K.; Nagata, Y.; Svensson, L.A.; Newman, J.; Takagi, M.; Damborsky, J. Crystal structure of the haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26. *Biochemistry* **2000**, *39*, 14082–14086. [[CrossRef](#)]
30. Mazumdar, P.A.; Hulecki, J.C.; Cherney, M.M.; Garen, C.R.; James, M.N.G. X-ray crystal structure of *Mycobacterium tuberculosis* haloalkane dehalogenase Rv2579. *Biochim. Biophys. Acta* **2008**, *1784*, 351–362. [[CrossRef](#)]
31. Koudelakova, T.; Bidmanova, S.; Dvorak, P.; Pavelka, A.; Chaloupkova, R.; Prokop, Z.; Damborsky, J. Haloalkane dehalogenases: Biotechnological applications. *Biotechnol. J.* **2013**, *8*, 32–45. [[CrossRef](#)]
32. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

## 5. CONCLUSIONS

---

The overall goal of the present PhD study was to determine the structure of the novel types of haloalkane dehalogenases, which represent a promising tool for degradation of halogenated compounds. Better understanding of their biology and structural geometry is necessary for further improvement of their substrate specificity and catalytic activity.

The structural characterization of the novel marine dehalogenase from *Paraglaciecola agarilytica* NO2 bacterium as well as newly synthesized by ancestral sequence reconstruction dehalogenase AncLinB-DmbA were performed.

Crystallographic analysis of the DpaA protein was done. An unusual for HLD-I subfamily crystal packing of the DpaA protein was revealed. Specifically, using X-ray crystallography DpaA tetramers formation was observed, phenomenon, that has not been described before for haloalkane dehalogenase subfamily I. Usually, the HLD-I subfamily members display monomeric arrangement in crystal structures, while some members of subfamily HLD-II were found in dimers formation, and some members of HLD-III subfamily – in oligomeric form (however, their structure is not solved yet). The multimeric haloalkane dehalogenases provide unique characteristics and expand the range of properties for such enzymes. However, the tendency of the DpaA protein to oligomerize challenges protein structure solution in a wide range of crystallization conditions and across various crystal lattices.

The structure of AncLinB-DmbA was studied in the second part of the thesis. Structural analysis revealed a high degree of similarity between ancestral enzyme AncLinB-DmbA and closely related modern haloalkane dehalogenases LinB from *Sphingobium japonicum* UT26 and DmbA from *Mycobacterium bovis* 5033/66. Despite the highly conserved structure, AncLinB-DmbA exhibits some

differences in the overall construction, such as variations in the sizes and positions of  $\alpha$ -helices and  $\beta$ -sheets and the presence of more flexible elements such as loops. Nevertheless, the core residues of ancestor enzyme correspond to those in LinB and DmbA. In addition, caver analysis revealed differences in several residues responsible for the formation of tunnels, that might indicate some advantages of AncLinB-DmbA compared with its descendants. The radius of the bottleneck in the p1 tunnel of AncLinB-DmbA is similar to the wider bottleneck in DmbA, while the ancestral protein displays the widest main tunnel entrance radius compared to the other two proteins. In addition, AncLinB-DmbA displays the largest active site cavity. Knowledge of the composition of tunnel residues in AncLinB-DmbA is required for further modification of its catalytic properties.

## 6. REFERENCES

---

1. Fishbein, L., *Potential halogenated industrial carcinogenic and mutagenic chemicals: IV. Halogenated aryl derivatives*. Science of The Total Environment, 1979. **11**(3): p. 259-278.
2. Häggblom, M. and I. Bossert, *Halogenated Organic Compounds - A Global Perspective*. 2004. p. 3-29.
3. Commandeur, L.C. and J.R. Parsons, *Degradation of halogenated aromatic compounds*. Biodegradation, 1990. **1**(2-3): p. 207-20.
4. Hitchman, M.L., et al., *Disposal methods for chlorinated aromatic waste*. Chemical Society Reviews, 1995. **24**(6): p. 423-430.
5. Dwivedi, A. and U.C. Pande, *PHOTOCHEMICAL DEGRADATION OF HALOGENATED COMPOUNDS: A REVIEW*. Scientific Reviews and Chemical Communications, 2012. **2**.
6. Falta, R.W., *The Potential for Ground Water Contamination by the Gasoline Lead Scavengers Ethylene Dibromide and 1,2-Dichloroethane*. Groundwater Monitoring & Remediation, 2004. **24**(3): p. 76-87.
7. Janssen, D.B., et al., *Bacterial degradation of xenobiotic compounds: evolution and distribution of novel enzyme activities*. Environmental Microbiology, 2005. **7**(12): p. 1868-1882.
8. Janssen, D.B., et al., *Degradation of halogenated aliphatic compounds by Xanthobacter autotrophicus GJ10*. Appl Environ Microbiol, 1985. **49**(3): p. 673-7.
9. Damborský, J., et al., *Structure–Function Relationships and Engineering of Haloalkane Dehalogenases*. 2010. p. 1081-1098.
10. Prokop, Z., et al., *Enantioselectivity of Haloalkane Dehalogenases and its Modulation by Surface Loop Engineering*. Angewandte Chemie International Edition, 2010. **49**(35): p. 6111-6115.
11. Marek, J., et al., *Crystal structure of the haloalkane dehalogenase from Sphingomonas paucimobilis UT26*. Biochemistry, 2000. **39**(46): p. 14082-6.
12. Ollis, D.L., et al., *The alpha/beta hydrolase fold*. Protein Eng, 1992. **5**(3): p. 197-211.
13. Koudelakova, T., et al., *Substrate specificity of haloalkane dehalogenases*. Biochem J, 2011. **435**(2): p. 345-54.
14. Chovancová, E., et al., *Phylogenetic analysis of haloalkane dehalogenases*. Proteins, 2007. **67**(2): p. 305-16.
15. Hashimoto, K., et al., *Caught in self-interaction: evolutionary and functional mechanisms of protein homooligomerization*. Phys Biol, 2011. **8**(3): p. 035007.
16. Lynch, M., *Evolutionary diversification of the multimeric states of proteins*. Proceedings of the National Academy of Sciences, 2013. **110**(30): p. E2821-E2828.
17. Janin, J., *Physical biology of the cell, Second Edition*. Crystallography Reviews, 2013. **19**.

18. Bershtein, S., W. Mu, and E.I. Shakhnovich, *Soluble oligomerization provides a beneficial fitness effect on destabilizing mutations*. Proceedings of the National Academy of Sciences, 2012. **109**(13): p. 4857-4862.
19. Jesenská, A., et al., *Biochemical characterization of haloalkane dehalogenases DrbA and DmbC, representatives of a novel subfamily*. Applied and Environmental Microbiology, 2009. **75**(15): p. 5157-5160.
20. Kunka, A., J. Damborsky, and Z. Prokop, *Haloalkane dehalogenases from marine organisms*. Methods in enzymology, 2018. **605**: p. 203-251.
21. Vanacek, P., et al., *Exploration of enzyme diversity by integrating bioinformatics with expression analysis and biochemical characterization*. Acs Catalysis, 2018. **8**(3): p. 2402-2412.
22. Silberstein, M., J. Damborsky, and S. Vajda, *Exploring the binding sites of the haloalkane dehalogenase DhIA from Xanthobacter autotrophicus GJ10*. Biochemistry, 2007. **46**(32): p. 9239-49.
23. Chrast, L., et al., *Deciphering the Structural Basis of High Thermostability of Dehalogenase from Psychrophilic Bacterium Marinobacter sp. ELB17*. Microorganisms, 2019. **7**(11).
24. Novak, H.R., et al., *Biochemical and structural characterisation of a haloalkane dehalogenase from a marine Rhodobacteraceae*. FEBS Lett, 2014. **588**(9): p. 1616-22.
25. Gehret, J., et al., *Structure and activity of Dmma, a marine haloalkane dehalogenase*. Protein science : a publication of the Protein Society, 2012. **21**: p. 239-48.
26. Carlucci, L., et al., *Biochemical characterization of two haloalkane dehalogenases: DccA from Caulobacter crescentus and DsaA from Saccharomonospora azurea*. Protein Science, 2016. **25**(4): p. 877-886.
27. Chaloupkova, R., et al., *Structural and functional analysis of a novel haloalkane dehalogenase with two halide-binding sites*. Acta Crystallographica Section D, 2014. **70**(7): p. 1884-1897.
28. Chaloupkova, R., et al., *Stereoselectivity and conformational stability of haloalkane dehalogenase DbjA from Bradyrhizobium japonicum USDA110: the effect of pH and temperature*. The FEBS Journal, 2011. **278**(15): p. 2728-2738.
29. Lahoda, M., et al., *Crystallographic analysis of 1,2,3-trichloropropane biodegradation by the haloalkane dehalogenase DhaA31*. Acta Crystallographica Section D, 2014. **70**(2): p. 209-217.
30. Damborský, J. and J. Koča, *Analysis of the reaction mechanism and substrate specificity of haloalkane dehalogenases by sequential and structural comparisons*. Protein Engineering, 1999. **12**(11): p. 989-998.
31. Janssen, D.B., *Evolving haloalkane dehalogenases*. Current opinion in chemical biology, 2004. **8**(2): p. 150-159.
32. Bosma, T., et al., *Steady-state and pre-steady-state kinetic analysis of halopropane conversion by a Rhodococcus haloalkane dehalogenase*. Biochemistry, 2003. **42**(26): p. 8047-8053.

33. Schanstra, J.P. and D.B. Janssen, *Kinetics of halide release of haloalkane dehalogenase: evidence for a slow conformational change*. *Biochemistry*, 1996. **35**(18): p. 5624-5632.
34. Prokop, Z., et al., *Catalytic mechanism of the haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26*. *Journal of Biological Chemistry*, 2003. **278**(46): p. 45094-45100.
35. Stepankova, V., et al., *Expansion of Access Tunnels and Active-Site Cavities Influence Activity of Haloalkane Dehalogenases in Organic Cosolvents*. *ChemBioChem*, 2013. **14**(7): p. 890-897.
36. Buryska, T., et al., *A haloalkane dehalogenase from a marine microbial consortium possessing exceptionally broad substrate specificity*. *Applied and environmental microbiology*, 2018. **84**(2): p. e01684-17.
37. Fortova, A., et al., *DspA from *Strongylocentrotus purpuratus*: the first biochemically characterized haloalkane dehalogenase of non-microbial origin*. *Biochimie*, 2013. **95**(11): p. 2091-2096.
38. Kasai, Y., et al., *Predominant growth of *Alcanivorax* strains in oil-contaminated and nutrient-supplemented sea water*. *Environmental Microbiology*, 2002. **4**(3): p. 141-147.
39. Abraham, W.-R., H. Meyer, and M. Yakimov, *Novel glycine containing glucolipids from the alkane using bacterium *Alcanivorax borkumensis**. *Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism*, 1998. **1393**(1): p. 57-62.
40. Huang, J., et al., *Phylogenetic diversity and characterization of 2-haloacid degrading bacteria from the marine sponge *Hymeniacidon perlevis**. *World Journal of Microbiology and Biotechnology*, 2011. **27**(8): p. 1787-1794.
41. Ferrer, M., et al., *Metagenomics for mining new genetic resources of microbial communities*. *Microbial Physiology*, 2009. **16**(1-2): p. 109-123.
42. Hårdeman, F. and S. Sjöling, *Metagenomic approach for the isolation of a novel low-temperature-active lipase from uncultured bacteria of marine sediment*. *FEMS microbiology ecology*, 2007. **59**(2): p. 524-534.
43. Hu, Y., et al., *Novel lipolytic genes from the microbial metagenomic library of the South China Sea marine sediment*. *FEMS microbiology ecology*, 2010. **72**(2): p. 228-237.
44. Kennedy, J., J.R. Marchesi, and A.D. Dobson, *Marine metagenomics: strategies for the discovery of novel enzymes with biotechnological applications from marine environments*. *Microbial cell factories*, 2008. **7**(1): p. 1-8.
45. Li, A. and Z. Shao, *Biochemical characterization of a haloalkane dehalogenase *DadB* from *Alcanivorax dieselolei* B-5*. *PloS one*, 2014. **9**(2): p. e89144.
46. Kokkonen, P., et al., *Structure-function relationships and engineering of haloalkane dehalogenases*. *Aerobic utilization of hydrocarbons, oils and lipids*, 2017: p. 1-21.
47. Stepankova, V., J. Damborsky, and R. Chaloupkova, *Organic co-solvents affect activity, stability and enantioselectivity of haloalkane dehalogenases*. *Biotechnology Journal*, 2013. **8**(6): p. 719-729.

48. Ravelli, R.B., et al., *Phasing in the presence of radiation damage*. Journal of synchrotron radiation, 2005. **12**(3): p. 276-284.
49. Gumulya, Y. and E.M. Gillam, *Exploring the past and the future of protein evolution with ancestral sequence reconstruction: the 'retro' approach to protein engineering*. Biochem J, 2017. **474**(1): p. 1-19.
50. Merkl, R. and R. Sterner, *Ancestral protein reconstruction: techniques and applications*. Biological Chemistry, 2016. **397**(1): p. 1-21.
51. Furukawa, R., et al., *Ancestral sequence reconstruction produces thermally stable enzymes with mesophilic enzyme-like catalytic properties*. Scientific reports, 2020. **10**(1): p. 1-13.
52. Babkova, P., et al., *Structures of hyperstable ancestral haloalkane dehalogenases show restricted conformational dynamics*. Computational and structural biotechnology journal, 2020. **18**: p. 1497-1508.
53. Babkova, P., et al., *Ancestral Haloalkane Dehalogenases Show Robustness and Unique Substrate Specificity*. ChemBioChem, 2017. **18**(14): p. 1448-1456.
54. Chaloupkova, R., et al., *Light-Emitting Dehalogenases: Reconstruction of Multifunctional Biocatalysts*. ACS Catalysis, 2019. **9**(6): p. 4810-4823.
55. Huang, R., et al., *Enzyme functional evolution through improved catalysis of ancestrally nonpreferred substrates*. Proceedings of the National Academy of Sciences, 2012. **109**(8): p. 2966-2971.
56. Weber, P.C., *Physical principles of protein crystallization*. Adv Protein Chem, 1991. **41**: p. 1-36.
57. Drenth, J. and C. Haas, *Protein crystals and their stability*. Journal of Crystal Growth, 1992. **122**(1): p. 107-109.
58. Russo Krauss, I., et al., *An overview of biological macromolecule crystallization*. Int J Mol Sci, 2013. **14**(6): p. 11643-91.
59. Bergfors, T., *Protein Crystallization: Techniques, Strategies, and Tips*. 1999: International University Line.
60. McPherson, A., *Introduction to protein crystallization*. Methods, 2004. **34**(3): p. 254-265.
61. Rupp, B. and J. Wang, *Predictive models for protein crystallization*. Methods, 2004. **34**(3): p. 390-407.
62. García-Ruiz, J.M., *Counterdiffusion methods for macromolecular crystallization*, in *Methods in Enzymology*. 2003, Elsevier. p. 130-154.
63. Otálora, F., et al., *Counterdiffusion methods applied to protein crystallization*. Progress in biophysics and molecular biology, 2009. **101**(1-3): p. 26-37.
64. McPherson, A., *Crystallization of biological macromolecules*. 1999: Cold Spring Harbor Laboratory Press.
65. McPherson, A., et al., *The effects of neutral detergents on the crystallization of soluble proteins*. Journal of Crystal Growth, 1986. **76**(3): p. 547-553.
66. McPHERSON, A., *Review Current approaches to macromolecular crystallization*. EJB reviews 1990, 1990: p. 49-71.
67. Bergfors, T., *Seeds to crystals*. Journal of structural biology, 2003. **142**(1): p. 66-76.

68. Stura, E.A. and I.A. Wilson, *Applications of the streak seeding technique in protein crystallization*. Journal of Crystal Growth, 1991. **110**(1): p. 270-282.
69. Rhodes, G., *Crystallography made crystal clear: a guide for users of macromolecular models*. 2010: Elsevier.
70. Rupp, B., *Biomolecular crystallography: principles, practice, and application to structural biology*. 2009: Garland Science.
71. Lockwood, E.H. and R.H. Macmillan, *Geometric symmetry*. 1978: CUP Archive.
72. McRee, D.E., *Practical protein crystallography*. 1999: Elsevier.
73. Aroyo, M.I., *International Tables for Crystallography*. 2013: Wiley Online Library.
74. Drenth, J., *Principles of protein X-ray crystallography*. 2007: Springer Science & Business Media.
75. Waseda, Y., E. Matsubara, and K. Shinoda, *X-ray diffraction crystallography: introduction, examples and solved problems*. 2011: Springer Science & Business Media.
76. Faust, A., et al., *A tutorial for learning and teaching macromolecular crystallography*. Journal of Applied Crystallography, 2008. **41**(6): p. 1161-1172.
77. Taylor, G., *The phase problem*. Acta Crystallographica Section D: Biological Crystallography, 2003. **59**(11): p. 1881-1890.
78. Faust, A., et al., *Update on the tutorial for learning and teaching macromolecular crystallography*. Journal of Applied Crystallography, 2010. **43**(5): p. 1230-1237.
79. Nanao, M.H., G.M. Sheldrick, and R.B. Ravelli, *Improving radiation-damage substructures for RIP*. Acta Crystallographica Section D: Biological Crystallography, 2005. **61**(9): p. 1227-1237.
80. Berman, H.M., *The protein data bank: a historical perspective*. Acta Crystallographica Section A, 2008. **64**(1): p. 88-95.

© for non-published parts Andrii

Mazur Mazura00@prf.jcu.cz

Mazurandrey92@gmail.com

## **Crystallographic studies of multimeric and ancestral haloalkane dehalogenases**

Ph.D. Thesis Series, 2023, No. 2

All rights reserved

For non-commercial use only

Printed in the Czech Republic by

Typodesign Edition of 10 copies

University of South Bohemia in České Budějovice

Faculty of Science

Branišovská 1760

CZ-37005 České Budějovice, Czech

Republic Phone: +420 387 776 201

[www.prf.jcu.cz](http://www.prf.jcu.cz), e-mail: [sekret-fpr@prf.jcu.cz](mailto:sekret-fpr@prf.jcu.cz)