UNIVERSITY OF SOUTH BOHEMIA IN ČESKÉ BUDĚJOVICE

FACULTY OF SCIENCE

Crystallographic study of biotechnologically attractive haloalkane dehalogenases DpcA and DmxA

Ph.D. THESIS

Katsiaryna Doleželová (Tratsiak), MSc

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

České Budějovice 2019

This thesis should be cited as:
Tratsiak K., 2019: Crystalographic study of biotechnologically attractive haloalkane dehalogenases DpcA and DmxA. Ph.D. Thesis. Series, No. 7. University of South Bohemia, Faculty of Science, České Budějovice, Czech Republic, 165 pp.

Annotation

Since 1991, when the first haloalkane dehalogenase (HLD) from *Xanthobacter autotrophicus* GJ10 was described, nearly 20 HLDs have been characterized biochemically and fourteen HLDs were analysed structurally. These enzymes belong to α/β-hydrolases and, owing to their ability to bring about the hydrolytic conversion of toxic halogenated compounds, are a source of broad biotechnological applications. DpcA from *Psychrobacter cryohalolentis* K5 and DmxA from *Marinobacter* sp. ELB17 are among them. Although their overall structures are quite similar to other HLD members, they possess several unique properties. Information on their 3D structure may significantly contribute to the understanding of protein function. DpcA and DmxA structures will allow us to gain insights into structural determinants of specificity and stability. This thesis aims to elucidate the tertiary and quartery structures of these enzymes using of X-ray crystallography.

Declaration [in Czech]

Prohlašuji, že svoji disertační práci jsem vypracovala samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své disertační práce, a to v úpravě vzniklé vypuštěním vyznačených částí archivovaných Přírodovědeckou fakultou elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

České Budějovice, 11.03.2019 .....

Katsiaryna Doleželová (Tratsiak)

This thesis originated from the Faculty of Science, University of South Bohemia, supporting doctoral studies in the Biophysics study programme.

Financial support

This research was supported by the GACR grants P207/12/0775 and 17-24321S.

List of publications and author's contribution

The thesis is based on the following publications (listed chronologically):

1. **K. Tratsiak,** O. Degtjarik, I. Drienovska, L. Chrast, P. Rezacova, M. Kuty, R. Chaloupkova, J. Damborsky and I. Kuta Smatanova. Crystallographic analysis of new psychrophilic haloalkane dehalogenases: DpcA from *Psychrobacter cryohalolentis* K5 and DmxA from *Marinobacter sp*. ELB17. *Acta Cryst.* (2013). F69, 683-688. (IF $_{2013}$ = 0.568)

   Katsiaryna Tratsiak conceived the experiments (jointly with IKS), performed protein crystallization and soaking, analyzed the X-ray diffraction data (jointly with PR, OD), solved (jointly with TP), refined and validated the protein structure (jointly with PR), prepared the manuscript and figures (jointly with other co-authors).

2. **K. Tratsiak,** T. Prudnikova, I. Drienovska, J. Damborsky, J. Brynda, P.Pachl, M. Kuty, R. Chaloupkova, P. Rezacova and I. Kuta Smatanova. Crystal structure of cold-adapted haloalkane dehalogenase DpcA from *Psychrobacter cryohalolentis* K5. Accepted for publication in *Acta Cryst. F* (2019). (IF $_{2017}$ = 0.989)

   Katsiaryna Tratsiak devised the experiments (jointly with IKS), crystallized protein, solved (jointly with PR, TP), refined and validated the protein structure (jointly with TP), performed structural analysis (jointly with PR, TP, RCH), deposited the structure in the PDB database, prepared the manuscript and figures (jointly with other co-authors).

3. L.Chrast, **K. Tratsiak,** L. Daniel, T. Prudnikova, J. Brezovsky, I. Kuta Smatanova, R. Chaloupkova, J. Damborsky. Deciphering of paradoxically high thermostability of dehalogenase from psychrophilic bacterium.

   Katsiaryna Tratsiak devised the experiments (jointly with IKS, RCH), crystallized protein, solved (jointly with TP), refined and validated the protein structure (jointly with TP), performed structural analysis (jointly with PR, TP, RCH), deposited the structure in the PDB database, prepared the manuscript and figures (jointly with other co-authors).

ABSTRACT

This Ph.D. thesis is devoted to the crystallographic study of selected haloalkane dehalogenases. It provides structural characterization of two extremozymes. Both are isolated from psychrophilic organisms. DpcA from *Psychrobacter cryohalolentis* K5 (PDB ID 6F9O) is a cold-adapted enzyme and demonstrates the highest activity at 25 ºC, while retaining more than 25% of its activity at 5 ºC. DmxA from *Marinobacer* sp. ELB17 (PDB ID 5MXP) exhibits the highest thermal stability (Tm = 65.9 ± 0.1 °C) of all biochemically characterized members of the haloalkane dehalogenase (HLD) family. The paradigm of thermophilic enzymes from psychrophiles has not yet been conclusively explained and the understanding of cold-adaptation is not clear either, although numerous comparative sequence and structure analyses have been carried out.

DpcA and DmxA are applied to catalyse the conversion of highly toxic halogenated compounds, which spread in the biosphere as a result of anthropogenic activities and appear to be a real and potential hazard not only environmentally, but are also of health concern, having a limited biodegradability. Selected enzymes are also advantageous for various biotechnological applications.

The objective of this study was to describe the structural features of DpcA and DmxA enzymes. To describe the macromolecular structures, we determined the influence of physical and chemical parameters on the crystallization of DpcA and DmxA, and crystallized them in an optimized condition. We obtained the diffraction data of crystals at atomic resolution (DpcA crystals diffracted to the resolution of 1.05Å and DmxA of 1.45 Å) and solved the protein structures using coordinate of HLD homologues. That enhances the available structural information about HLD.

The herein presented thesis provides the structural characterizations of DpcA. We recognized two alternative conformations for the side chain of the catalytic nucleophile D123. We carried out structural comparison with other HLD members and identified substantial differences in the architecture of the active site and access tunnels of DpcA. This may indicate that the main tunnel can be explored by substrate and alcohol product molecules, while the other is used for halide ions or water molecules. Moreover, comparative analysis revealed

major differences in the region of the α4 helix of the cap domain, which is one of the key determinants of the tunnels' properties, that possesses tunnels of the shortest length among other members.

The DpcA structural information reported here establishes a base for understanding its enzymatic properties and can guide modification of this cold-adapted enzyme for enzyme application for various biotechnological applications.

The herein presented thesis provides the structural characterizations of DmxA. We analysed the DmxA macromolecular structure. It possesses a unique composition of catalytic machinery (halide-stabilizing residue, Q40, instead of the conventional N typical of other members of the subfamily). We present that the enzyme contains narrow access tunnels for solvent, forms a dimeric structure via a covalent bond and it is the first example of oligomerization via disulfide bridge formation reported in the HLD enzyme family thus far, however it has no effect on enzyme stability or functionality. We revealed the important residues, affective the size of access tunnels, which was identified as the crucial feature by determining the high stability of DmxA. In this study, several different structural features of DmxA were analysed in detail in order to explain its extraordinary stability.

Furthermore, this thesis focuses on the structure-function relationship of the enzymes. Numerous attempts to obtain the enzyme structures in complex with the substrates let us to identify halogen atoms in the vicinity of the active site.

Dedicated to

My Family

ACKNOWLEDGEMENTS

...Ab imo pectore (sincerely from the bottom of my heart). Upřímně z celého srdce.
Из самой глубины души, от всего сердца, с полной откровенность.
Шчыра ад усяго сэрца.

I would like to express my thanks to all of the teaching staff at the University of South Bohemia for sharing their knowledge. My gratitude for the excellent study programme and organization of the project goes out to the Dean of the Faculty of Scienceat the University of South Bohemia, prof. RNDr. František Vácha, Ph.D; Vice-Dean for Science, doc. RNDr. Milan Předota, Ph.D.; Deputy Head of the Institute of Physics, prof. RNDr. Tomáš Polívka, Ph.D.; faculty secretary, Hana Sýkorová; prof. RNDr. Rüdiger Ettrich, Ph.D.; prof. RNDr. Dalibor Štys, CSc.; Ing. Dagmar Kaftanová, from the study programme office; and Bc. Petra Korcová from the department of student affairs.

This work would not be possible without my supervisor doc. Mgr. Ivana Kutá Smatanová, Ph.D., and especially I would like to express my thanks for this opportunity to be involved in the entire process, for the ideas, advices and the organization of the team, and the perfect management of all of the project's requirements and for connecting people and their ideas. I would also like to thank her for her open-minded support at any time and offering inspiration and opportunities to realize the project's aims. I thank Mgr. Michal Kutý, Ph.D., for his support from the chemical point of view.

I give my acknowledgement to the group of prof. Mgr. Jiří Damborský, Dr., for their great collaboration, and doc. Mgr. Radka Chaloupková, Ph.D., for project support, and giving me a deeper insight into the project – and for their energy and scientific contribution during our discussions and the opportunity to come to their laboratory, and Lukas Chrast and Ivana Drienovska – for the biochemical characterization of protein.

I would like to show my appreciation to RNDr. Pavlína Řezáčová, Ph.D., for her valued collaboration on my project, for her personal caring and critical view and advices, and for enabling me to acquire more useful experience working on other projects within the wonderful team of the Research Group of Structural Biology at the Institute of Organic Chemistry and Biochemistry ASCR, v.v.i. I would also like to pay my respect to the entire Structural Biology group of IOCB, especially doc. RNDr. Jiří Brynda, Ph.D., and Petr Pachl, Ph.D., for their open minds and answers to my questions, Vaclav Veverka, Ph.D. and Pavel Srb, Ph.D., for their help with the NMR experiment, and Aleš Hnízda, Ph.D., for his work on the project.

I would like to acknowledge dr. Juan Antonio Gavira and all of the team from the Institute of Laboratorio de Estudios Cristalográficos, IACT, Spanish National Research Council (CSIC) for the opportunity to have a scientific stay in a pleasant atmosphere, and for his positive and intelligent ideas. I thank dr. Manfred S. Weiss and dr. Sandra Pühringer for their assistance with data collection on MX14.2 beam line operated by the Helmholtz-Zentrum Berlin (HZB) at the BESSY II electron storage ring (Berlin-Adlershof, Germany), dr. Andrey Lebedev for the for the willingness to support on the software issues.

I want to thank all my friends from Nove Hrady and Ceske Budejovice. Many thanks to Tanya Prudnikova who first helped me with delving into the crystallization trials and further advice and collaboration on projects. Oksana Degtjarik for her critical point of view and advice. My thanks go out as well to Hanna Huryna, Karina Romanova, Valentina Kuznetsova, Katya Shamaeva, Gurkan Keskan, Alina Kevorkova, Pavel Grinkevich, Alex Pautsino, Tatsiana Chernavetc, and Julia Volkova, for their friendly support and cherished memories during the study.

Finally, with all my heart I am deeply thankful to my entire family in Belarus and in the Czech Republic, for their emotional and financial support during my study. Especially, I want to thank my mother for believing in me, filling me with energy and helping me by providing great advice. My whole heart and gratitude belong to my beloved husband Petr who always supported me in good and bad moments, filled me with positive energy, believed in me and made my life happier.

I would also like to thank any who helped but who were not mentioned here for being a part of my Ph.D. study life.

CONTENTS

LIST OF ABBREVIATIONS

3D – tree-dimensional
Ala, A – Alanine
Arg, R – Arginine
Asn, N – Asparagine
Asp, D – Aspartic acid
CLECs – cross-linked enzyme crystals
Cryo-EM- cryo -electron microscopy
Cys, C – Cysteine
etc. – an abbreviation of et cetera (“and more”)
Ex.- example
Gln, Q – Glutamine
Glu, E – Glutamic acid
Gly, G – Glycine
His, H – Histidine
Ile, I – Isoleucine
Leu, L – Leucine
Lys, K– Lysine
Met, M – Methionine
NMR - Nuclear magnetic resonance
PDB – Protein data bank
Phe, F – Phenylalanine
Pro, P – Proline
SAXS – Small-angle X-ray scattering
Ser, S – Serine
TMOS – tetramethyl orthosilicate
Thr, T – Threonine
Trp, W – Tryptophan
TLS – translation/ libration/ screw group
Tyr, Y – Tyrosine
Val, V – Valine

1. Prologue and aims of thesis

In the drama of life on a molecular scale, proteins are where the action is.
(Lesk, 2001)

Why do we study proteins? Why are they so important for human existence? Every living cell and all biological processes depend on proteins (Gunton *et al.*, 2007).

Our life without proteins, abundant organic molecules, which play a crucial role in all aspects of the cell structure and function, would not exist. The "ancient alphabet" – is the letter code of just 20 amino acids that are common building blocks of all proteins provide the possibility to write genius "poems" in the form of billions of protein sequences (Berg JM, 2002).

According to research that was done under the leadership of Grant Brown (Ho *et al.*, 2018), there are 42 billion proteins in a single eukaryotic cell (*Cerevisiae*) or there are 2–4 million proteins per cubic micron in bacteria, yeast, and mammalian cells, as estimated by Milo (Milo, 2013). Nevertheless, we do not know all organism's protein structures, functions and interactions in the systems – it is exciting and highly important to have such knowledge of the compositions surrounding us. Such information will be useful for prolonging life and health benefits, to contribute to the knowledge of the universe and to add input to the comfort of life and advantageous technologies.

X-ray crystallography is one of the key tools in discovering protein structures.

Crystallography is a method used to study crystals, their properties and protein structures. Having a long history, this field of science is dynamic and rapidly changing. Thanks to crystallography, the understanding of many biological processes at a molecular level has already led to a fundamental change in the treatment of many chronic diseases. Computer software development has revolutionized in such a way that crystal structures can be represented and compared beautifully, comparable to the best examples of abstract art (Tilley, 2006).

In order to contribute to a large research project on haloalkane dehalogenases, we had two enzymes under study, DpcA from *Psychrobacter cryohalolentis* K5 and DmxA from *Marinobacter* sp ELB17, which are useful for different biological applications.

The **main goals** of this thesis were to understand the relation between protein structure and function as fundamental knowledge of biological mechanisms through:

1) crystallization and to acquiration of high-resolution X-ray data of DpcA and DmxA crystals;
2) characterization the structure of DpcA from *Psychrobacter cryohalolentis* K5 and uncovering the structural features of its cold-adaptation;
3) characterization the structure of DmxA from *Marinobacter* sp ELB17 and revealing its high thermostability.

The information obtained from their molecular structures will hopefully contribute to engineering of new enzymes with improved catalytical characteristics and a broader environmental activity range.

2. Introduction

2.1. Enzymes and their function

Life depends on the organized sets of chemical reactions, proceeding in a way to maintain life. The key role in this process is allocated to enzymes. Enzymes are mostly proteins, (excluding some catalytic RNA molecules, ribozymes, playing an important role in gene expression) (Talini et al., 2011). Enzymes are highly specific and accelerate reactions significantly. This occurs by decreasing the activation energy of a specific reaction. Enzymes are not consumed during the reaction. Binding to one or more ligands, called substrates, they convert them into one or more chemically modified products and thus they act as biocatalysts. It is the catalysis of organized sets of chemical reactions by enzymes that creates and maintains the cell, making life possible (Alberts, Johnson, Lewis, Raff, *et al.*, 2002, Copeland, 2000).

Enzymes are the most important components of the metabolism of living organisms. The glucose oxidation in the organisms, for example, is the reaction of an energy exchange, where the enzymes take part at each stage of reactions (depolymerization of polysaccharides, a glucose phosphorylation, substrate-level and oxidative phosphorylation, the Krebs cycle, an electron-transport chain). On this reaction processes we can notice that the existence of the developed organism without enzymes would not be possible, as the organism needs to receive quickly a great deal of energy in the form of molecules of ATP, and without the participation of the enzymes many reactions would take a long time (Berg *et al.*, 2015, Lodish & Lodish, 1999, Alberts, Johnson, Lewis, Walter, *et al.*, 2002). The carbonic anhydrase enzyme is among the fastest known. It catalyzes the reaction of the transfer of CO_2 from the tissue into the blood (Berg JM, 2002).

The following properties of enzymes can be mentioned: 1) all of them represent globular proteins; 2) information on them, as well as on other proteins, is coded in molecules of DNA; 3) a large number of the enzymes work extremely effectively, it means very small amount of enzyme causes transformation of a large amount of substrate; 4) enzymes accelerate reactions by decreasing $\Delta G^{\ddagger}$, the free energy of activation in order to overcome the reaction energy barrier by facilitating the transition state formation (Fig. 1); 5) their presence does not influence neither the nature, nor properties of the final product (products) of reaction; 6) enzymes are highly specific, i.e. one enzyme catalyzes usually only one reaction; 7) the reaction catalyzed by enzyme is reversible, the same enzyme

catalyzes both forward, and reverse reaction; 8) the activity of enzyme changes depending on pH, temperature, pressure and also concentrations of enzyme and substrate; 9) there are enzymes for which energy is necessary for work (Soper *et al.*, 1990, Berg JM, 2002, Cooper, 2000).

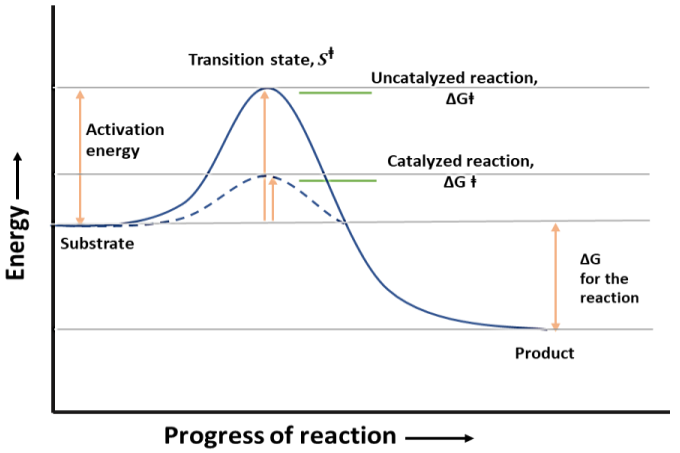

Figure 1. Energy diagram for uncatalyzed and catalyzed reactions. The conversion of the reactants (substrate) to a product. The acceleration of the reaction by the enzymes by decreasing ΔG‡, the free energy of activation. Adapted from Berg JM (Berg JM, 2002).

The two main models of enzyme-substrate binding are described:

1. Fischer's version (1890) – a theory of "key lock" (Lichtenthaler, 1995).

Fischer explained the high specificity of enzymes with a special form of a molecule ("lock"), precisely corresponding form of a molecule of substrate (“key”). The formed enzyme-substrate complex is leading to the reaction products formation, resulting the molecules of the product to be changed in shape and they release from the active site, which can accept a new substrate.

2. Koshland version (1959) is an induced fit theory (a new interpretation of the “key lock” theory) (Koshland Jr, 1953).

From the obtained data, Koshland concluded that the enzyme – substrate interaction is more dynamical. Their binding causes some changes in the structure of the enzyme active site. The amino-acids of the active site change the conformation making the active site of the enzyme take a certain form which gives the chance to the enzyme in the most effective way to perform the function (Soper *et al.*, 1990).

2.2. Enzyme activity

As the enzymes enhance rates of reactions, the kinetic description of their activity is needed to understand their function. A general scheme of the enzyme reaction model is proposed by two scientists in 20th century, Leonor Michaelis and Maud Leonora Menten, known as Michaelis-Menten Kinetics (Berg JM, 2002). The model serves the explanation how an enzyme can cause kinetic rate enhancement of a reaction and also explains how reaction rates depend on the concentration of enzyme and substrate, is the following:

When the molecules of an enzyme and a substrate appear in a system, they are attracted to each other under the influence of electrostatic forces (Fried & Boxer, 2017), followed by the process of the orientation of a molecule of substrate in the active site of an enzyme and the further process of molecular recognition and the formation of the enzyme-substrate complex. Then the enzyme changes the conformation under the influence of a molecule of substrate and external condition. At the same time conditions for substrate conversion are created. Eventually, the product is formed. The product releases from the active site, and the enzyme returns to the initial conformation to be ready to provide the next round (Lodish *et al.*, 2007, B L Blaney & Ewing, 1976).

The Michaelis-Menten model (1) is the one of the simplest and best-known approaches to enzyme kinetics, the derivation was provided by George Briggs and J.B.S. Haldane in 1925 (Briggs & Haldane, 1925):

$$E + S \underset{k'_r}{\overset{k_f}{\rightleftharpoons}} [ES] \xrightarrow{k_{cat}} E+P \qquad (1)$$

Where enzyme – E reacts with the substrate, S, *via* enzyme-substrate complex, ES, which then reacts irreversibly to generate a product, P. k_f - is the rate constant of association of enzyme-substrate binding, bimolecular; *k'r*- is the rate constant of dissociating of the ES complex to regenerate free enzyme and substrate, unimolecular; and k_{cat} is the turnover rate.

The Michaelis-Menten equation (*Eq.*1) for this system is:

Equation 1

$$v = \frac{V_{\max}[S]}{K_M + [S]}$$

A plot of the Michaelis-Menten equation's (Michaelis & Menten, 2007) is represented in Fig. 2. A plot of the reaction velocity (V_0) as a function of concentration [S] of a substrate for an enzyme that obeys Michaelis-Menten kinetics. It shows that the maximal velocity (V_{max}) is approached asymptotically. K_M - is the substrate concentration at which the reaction velocity is 1/2 of the V_{max}.

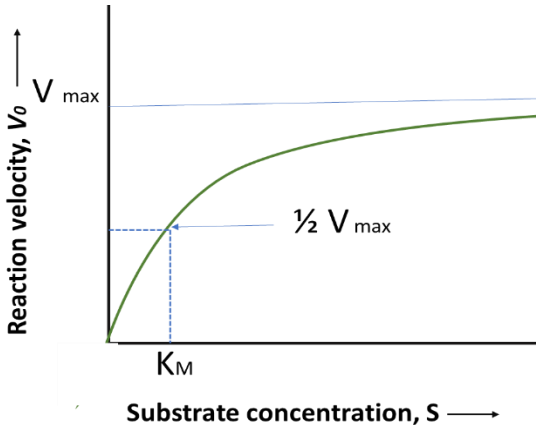

Figure 2. Michaelis-Menten kinetics (Michaelis 1913: 334, Michaelis & Menten, 2007). Adapted from Berg (Berg JM, 2002).

Currently there are about 150.357 protein strutures (on February, 19th, 2019) in wwPDB (worldwide Protein Data Bank) (Berman *et al.*, 2003). Many of them already have a known structure that came from X-ray crystallography. In 1926, James Sumner, was the first who published a paper on urease crystallization – the first crystallization of an enzyme (Sumner, 1926). Sumner carried out detailed analysis that allowed him first to show that enzymes can be crystallized and that the crystals were composed of a protein, that their dissolution in the solvent led to enzymatic activity, which was important for the development of enzymology, as a science (Simoni *et al.*, 2002). In 1946, Sumner received the Nobel Prize in Chemistry "for his discovery that enzymes can be crystallized." The prize was also awarded to J. H. Northrup and Wendell Stanley "for their preparation of enzymes and virus proteins in pure form" (Foundation, 1999).

Later, in 1957, the first 3D structure of a protein, myoglobin, was deduced by Sir John Cowdery Kendrew from X-ray diffraction. Since that time, numerous protein crystal structures were solved using this method.

The enzymes are also available to scientists, as an indispensable tool for research purposes, participating in molecular biological reactions. Selecting the enzyme with the most appropriate specific activity at the optimal reaction conditions needed for a specific purpose utilized in a specific protocol will certainly improve the quality of research (Rittié & Perbal, 2008).

2.3. Haloalkane dehalogenases

2.3.1. The spreading of toxic halogenated compound

Xenobiotics are constantly accumulating in the environment. They spread in the biosphere as a result of anthropogenic activities and appear to be a real and potential hazard for humankind.

These are various pesticides, solvents and monomers of organic synthesis, dyes and others (Bhatt *et al.*, 2007). Most often substrates are chlorderivatives that are industrially important organic compounds and are produced in large amount, belong to an important class of environmental pollutants found in ground and surface water, in food and in atmosphere (Wang *et al.*, 2013). Haloalkanes are widely distributed by industry: they are used as pesticides, fumigants (Liu *et al.*, 2013), solvents, in petroleum industry, vinyl chloride production (Tomasi *et al.*, 1984), refrigerants (Davidson & William, 1958) and as antiknock substances (Fetzner, 1998), and they are not only of environmental, but also of health concern, having a limited biodegradability (Belkin, 1992, Slater *et al.*, 1995).

Such substances are known to have kidney toxicity (Bláha *et al.*, 1998), hepatotoxic (Ugazio *et al.*, 1973, Weber *et al.*, 2003), cancerogenic and/or mutagenic effects (Tomasi *et al.*, 1984, Sasaki *et al.*, 1998, Roldan-Arjona *et al.*, 1991), causing the irreversible alkylation of DNA bases (Hemminki, 1983, Boysen *et al.*, 2009, Sobol *et al.*, 2007), which leads to the progression of degenerative diseases or cancer under their exposure (Akers *et al.*, 1999, Nakai *et al.*, 2016).

However, some haloalkanes are also produced naturally (Verschueren, Franken, *et al.*, 1993, Gribble, 2003). Chloromethane (CH_3Cl), is the most copious halocarbon in the atmosphere, which emission is connected with the biomass burning and other natural sources (Yokouchi *et al.*, 2000). Numerous halocarbons have volcanic origin (Jordan *et al.*, 2000). Haloalkanes were also discovered in minerals, rocks and shales (Isidorov *et al.*, 1993). It was found, that

chloromethane, bromomethane and iodomethane are produced by several species of marine algae (Scarratt & Moore, 1996, Tait & Moore, 1995). Several studies identified phytoplankton groups, that produce another halogenated compounds (Moore, 1996, Liu *et al.*, 2013).

The availability of halogens atoms causes their high toxicity and resistance to degradation in the environment. Numerous bacteria have attracted researchers with the real prospects of use in biotechnological applications for many years. Such bacteria have the unique ability to carry out the biotransformation of various natural and synthesized connections and also to act as destructors of halogenated aromatic xenobiotics. The ability to carry out dehalogenation is mentioned in the literature for many species of bacteria: *Pseudomonas, Bacillus, Corynebacterium, Nocardia, Rhodococcus, Streptomyces, Arthrobacter,* etc (van Pée & Unversucht, 2003, Fetzner, 1998). Bacteria have different mechanisms of dehalogenation: reductive dehalogenation, oxygenolytic dehalogenation, "thiolytic" dehalogenation, intramolecular, substitution, dehydrohalogenation and hydrolytic dehalogenation (Fetzner, 1998). A major factor of the detoxication of xenobiotics in the soil and water is the activity of bacteria, and particularly those, which possess the enzymes carrying out a dehalogenation reaction – dehalogenases, which play an important role in the biodegradation of chlorinated pollutants (Copley, 1998).

However, extremophilic microorganisms are the highest objects of interests because of their ability to survive in high or low temperatures, high salinity and either high alkalinity or high acidity and high pressure (Gomes & Steiner, 2004, Niehaus *et al.*, 1999, Margesin & Schinner, 2001). They are able to produce the enzymes that are extremely resistant to surrounding conditions, comparable to those prevailing in different industrial processes. These enzymes can be exploited as a model to design and construct proteins with new properties that are attractive for industrial applications (Niehaus *et al.*, 1999, van den Burg, 2003). Haloalkane dehalogenases are among of them.

2.3.2. Haloalkane dehalogenases

Haloalkane dehalogenases (HLDs; EC 3.8.1.5) are enzymes with catalytic activity for the conversion of highly toxic brominated, chlorinated, iodinated aliphatic compounds yielding a corresponding alcohol, a halide ion, and a proton

(Janssen, 2004, Janssen *et al.*, 1988). HLDs are able to convert halogenated alkanes cycloalkanes, alkenes, alcohols, ketones, ethers and cyclic dienes (Chovancová *et al.*, 2007, Janssen *et al.*, 1988, Poelarends *et al.*, 1998).

The catalyzed reaction is significant for the application of these enzymes in bioremediation (Stucki & Thueer, 1995, Prokop *et al.*, 2006), for constructing biosensors (Campbell *et al.*, 2006) and due to their enantioselectivity HLDs are also involved in the synthesis of optically pure compounds (Pieters *et al.*, 2001, Westerbeek *et al.*, 2011), protein analysis and intracellular imaging (Ohana *et al.*, 2009, Los & Wood, 2007, Murrey *et al.*, 2015).

HLDs within other enzymes: esterases, peptidases, lipases or epoxide hydrolases belong to the α/β-hydrolase protein fold family, which was identified in 1992 (Ollis *et al.*, 1992, Holmquist, 2000, Nardini & Dijkstra, 1999).

To date, a number of putative HLDs were identified by phylogenetic analysis, slightly fewer than 20 of them were biochemically characterized (Supplementary materials, Table S1), (Chovancová *et al.*, 2007) some proteins were crystallized and the 3D structure determined. Some HLDs and their crystallization conditions and collection data are presented in Table 1. The enzymes crystallization conditions vary widely in the composition of compounds and the crystal outcome.

Table 1. List of crystallized haloalkane dehalogenases

HLD crystal shape	**Organism from which was isolated**	**Crystal-lization method**	**Crystallization conditions of the diffracted crystal**	**Resolution of the collected X-ray diffraction data, space group, number of molecules in the asymmetric unit**	**HLD-subfa-mily, Reference**
DhlA (2had) n.a	*Xanthobacter autotrophicus GJ10*	hanging-drop vapour-diffusion	bis-Tris-H2SO4, 64%, ammonium sulphate pH 6.2.	2.4 Å (rotating anode X-ray generator) P21212, one monomeric molecule Matthews: 2.03, density percent sol: 39.39%	I (Franken *et al.*, 1991)

DppA (2xto) 0.04 x 0.11x 0.2 mm, needle-shaped	*Plesiocystis pacifica* SIR-I	hanging-drop vapour-diffusion	0.1 M Na MES pH 6.5, 1.8 M ammonium sulfate, 5% PEG 400, concentration - 8 mg ml^{-1}	1.95 Å (in-house X-ray generator), orthorhombic space group $P2_12_12$, one molecule in the asymmetric unit, cryoprotection: soaking 30s in 5%(*v/v*) PEG 400 in reservoir solution Matthews: 2.6, density percent sol: 53%	I (Bogdanovic *et al.*, 2010)
DccA (5esr) 0.1x 0.1x 0.2 mm	*Caulobacter crescentus*	sitting-drop vapor diffusion	Temp: 298 K, reservoir solution contained 0.005M Cobalt chloride, 0.005M Magnesium chloride, 0.005M Cadmium chloride, 0.005M Nickel chloride, 0.1M HEPES:- NaOH, pH 7.5, 12% PEG 3350; Protein concentration - 10 mg ml^{-1}	1.4 Å, beamline LRL-CAT, $C222_1$ 20% glycerol as cryo-protectant. Matthews: 2.2, density percent sol: 44.12%	I (Carlucci *et al.*, 2016)
DmxA (5mxp) 0.16 x 0.1 x 0.15 mm rhombo-hedral	*Marinobacter sp.* ELB17	sitting-drop vapor diffusion	Temp: 293 K, 0.2M ammonium acetate, 30%(w/v) PEG 4000 and 0.1M sodium citrate tribasic dihydrate, pH 5.6, additionally added sarcosine	1.45 Å, ESRF beamline ID29, space group $P2_12_12_1$, Matthews: 1.89, density percent sol: 34.85%	II (Tratsiak *et al.*, 2019, accepted for pub-lishing)
DmrA (4mj3) n.a. rhombo-hedra	*Moorea producta*	vapor diffusion	Temp: 298 K, 0.1 M bis-Tris propane, pH 6.5, 0.2 M potassium thiocyanate, 16% PEG 3350, reservoir solution: 2.2 M sodium malonate pH 7.0, 5% glycerol, Protein at 3 mg ml^{-1}	2.2 Å, data were collected at the GM/CA-CAT beamline 23ID-D, Australian synchrotron beamline MX2, Matthews: 2.23, density percent sol: 44.74%	II (Gehret *et al.*, 2012a)
DatA (3wib) 0.20x 0.07x 0.05 mm	*Agrobacterium tumefaciens* C58	sitting-drop vapour-diffusion	At 293 and 277 K, reservoir solution: 0.1 M CHES pH 8.6, 1.0 M potassium sodium tartrate, 0.2 M lithium sulfate, 0.01 M barium chloride.	1.70 Å, the Photon Factory, beamline: NE3A, primitive tetragonal space group *P*422, Matthews: 2.45, density percent sol: 49.72% two molecules in the asymmetric unit.	II (Mase *et al.*, 2012)

DbeA (4k2a) 0.05x 0.05x 0.30 mm needle-shaped	*Bradyrhizobium elkani* USDA94	sitting-drop vapour-diffusion	Temp: 277 K, 0.1 M Tris–HCl, pH 7.5, 20%(w/v) PEG 3350 or 4000 and 0.15 M calcium acetate, prot at a concentration of 4–6 mg ml $^{-1}$ in 0.1M Tris–HCl (pH 7.5).	2.2 Å (BESSY BEAMLINE 14.1), primitive orthorhombic space group $P2_12_12_1$, Matthews: 2.36, density percent sol: 46% one molecules in asymmetric unit	II (Prudnikova *et al.*, 2009)
DbjA (3a2m) 0.4 x 0.12 x0.1 mm Rectan-gular shape	*Bradyrhizobium japonicum* USDA110	hanging-drop vapour-diffusion method, using the microseeding technique	Temp: 293 K, reservoir solution: 17–19.5% (w/v) PEG 4000, 0.2 M calcium acetate and 0.1 M Tris–HCl, pH 7.7–8.0.	1.75 Å, Rigaku rotating-anode X-ray generator (FR-D), $P2_12_12$, Matthews: 2.4, density percent sol: 40 % orthorhombic space group, four molecules in one asymmetric unit	II (Sato *et al.*, 2007)

2.3.3. The overall HLD structure and the reaction mechanism

The overall HLD structure is quite similar among all enzymes and the globular protein is composed of a core (conserved main domain) and the cap domain. The main domain plays an important role in structural stability, is comprised of 8 stranded mostly parallel β-sheets, β2 is oriented antiparallel. β-sheets are surrounded by 6 α-helices (two are from the one side, four – from the other). The cap domain is inserted between β-strand β6 and α-helix α8 and is formed of several α-helices, connected by six loop insertions. The cap domain (involving residues 184–211 in DhlA, residues 155-220 in DpcA, residues 135-216 in DmxA) is less conserved in evolution and it plays a decisive role in the determination of substrate specificity (Pries *et al.*, 1994, Newman *et al.*, 1999, Pikkemaat *et al.*, 2002, Otyepka & Damborský, 2002) (Fig. 3).

Both domains contribute to the formation of a hydrophobic pocket in which the active site is buried, accessible via two well-defined tunnels (Marek et al. 2000, Kmunicek et al. 2005, Petrek et al. 2006). An active site is buried inside a hydrophobic core and is represented by five residues.

Based on the phylogenetic analysis, composition of the active site there are HLD subdivisions on three subfamilies: Asp-His-Asp+Trp-Trp in HLD-I, Asp-His-Glu+Asn-Trp in HLD-II, and Asp-His-Asp+Asn-Trp in HLD-III (Chovancova et al., 2007).

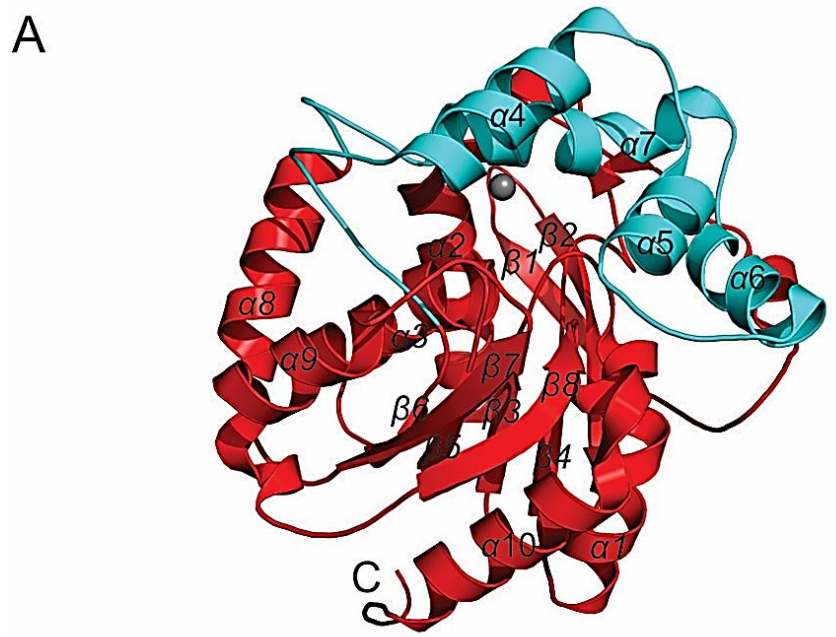

Figure 3. The overall tertiary structure of haloalkane dehalogenases (represented as the structure of DpcA). The main domain is represented in red, the cap domain in blue. α-helices and β-sheets are labelled. Adapted from Tratsiak *et al.*, 2019, (accepted for publication).

Firstly, the catalytic mechanism of haloalkane dehalogenase was described by Koen and Verschueren from the study of the DhlA enzyme from *Xanthobacter autotrophicus* GJ10 in complex with 1,2-dichloroethane. It was discovered by different methods (Pikkemaat *et al.*, 2002, Verschueren, Seljee, *et al.*, 1993) and revealed that the reaction is realized *via* S_N2 substitution mechanism, which involves the catalytic triad Asp-His-Asp/Glu and two residues stabilizing halide (Trp-Trp or Trp-Asn).

The four-step reaction mechanism is initiated by the substrate binding to the active site. Two tryptophanes with the nitrogen atom of their rings are positioned towards the cavity to stabilize by hydrogen-bonding a halide ion of the substrate. During this process, the carbon-halogen bond is cleaved upon a nucleophilic attack of an Oδ1 atom of nucleophilic aspartate (Asp 124/123/105) on the substrate's C1 atom. An alkyl-enzyme intermediate is formed (Fig. 4) and is hydrolyzed by a water molecule that is activated by the base histidine (His 289/280/273). The removal of a proton by Asp and His from the molecule of water is in order to facilitate the hydrolysis of the alkyl-enzyme intermediate.

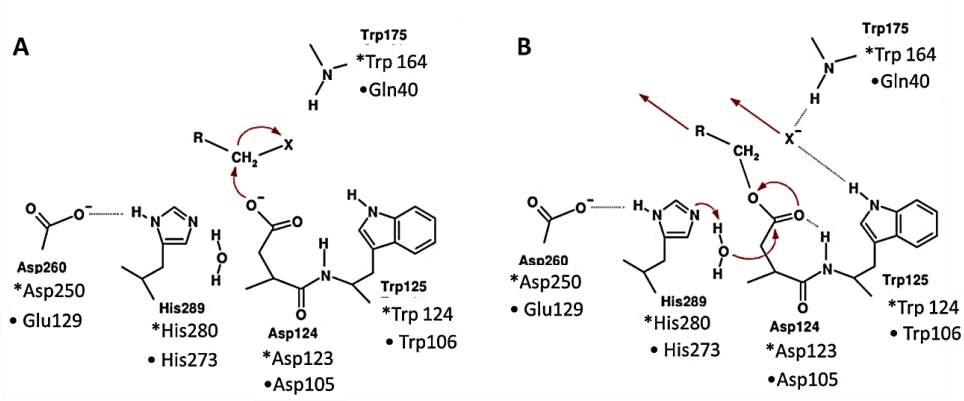

Figure 4. General catalytic mechanism of haloalkane dehalogenases. **A**. Formation of the covalent intermediate by S_N2 substitution. **B**. Hydrolysis of the intermediate. Residue numbers refer to DhlA, *DpcA, •DmxA, respectively. X represents halogen. Adapted from (Bosma *et al.*, 2003, Janssen, 2004).

The positive charge on the histidine is stabilized by its hydrogen bonding with the side-chain oxygen of catalytic acid: aspartate (is inherent for HLD-I, Asp 260/250) or glutamic acid (is inherent for HLD- II: DhaA, LinB and DbjA enzymes), or, DmxA is an exclusion, glutamine (Glu129). It possibly also helps to preserve the correct position of the catalytic acid and the base. The alcohol product releases and the halogen ion, which determines the overall reaction rate, is removed from the active site (Ridder *et al.*, 1999, Verschueren, Franken, *et al.*, 1993, Lewandowicz *et al.*, 2001, Janssen, 2004, Franken *et al.*, 1991, Krooshof *et al.*, 1998, Pries *et al.*, 1995).

2.3.4. DpcA from *Psychrobacter cryohalolentis* K5 and DmxA from *Marinobacter* sp ELB17

The enzymes were identified by phylogenetic analysis (Chovancová *et al.*, 2007). *Psychrobacter cryohalolentis* K5 was isolated from a cryopeg (saline water lens) within forty-thousand-year-old Siberian permafrost where the *in situ* temperature is -9 to -12°C and cryopeg salinity – 12–14%. Cryopeg samples were obtained from the Kolyma Lowland region of Siberia, Russia, a tundra zone with a harsh climate and continuous permafrost that has remained frozen for at least 40,000–50,000 years. The samples were recovered from depths of 11–24 m below the surface (Novototskaya-Vlasova *et al.*, 2012, Bakermans *et al.*, 2006).

DpcA. It has overall moderate level of the enzymatic activity, however, in comparison to other biochemically characterized HLDs, DpcA possesses one of the narrowest substrate specificity profiles. Among 30 examined substrates, the enzyme exhibits the highest activity toward 1-bromobutane, 1- bromohexane and 1,3-dibromopropane, while zero activity is observed toward short chlorinated substrates.

DpcA enzyme shows better activities toward longer substrates with the preference to brominated, in lesser degree it demonstrates activity to iodinated and the minimum of activity - with chlorinated substrates (Drienovska *et al.*, 2012).

DmxA. The bacteria *Marinobacter* sp ELB17 was originally isolated from the depth of 17 m from the permanently ice-covered east lobe of Lake Bonney in the Taylor Valley, in Antarctica. As was discovered later, both east and west lobes of this Antarctic lake possess significant abundance of ELB17 (Ward & Priscu 1997). The optimal growth conditions are ranged in the temperatures 12-15 °C and the salinity of 18-35 g/l (Ward & Priscu 1997).

The activity of DmxA towards thirty examined halogenated substrates was rather low. The enzyme shows slight preference towards brominated and bromo-chlorinated substrates. It's highest specific activity was observed in the reaction with 1,3-dibromopropane. Temperature and pH profile of DmxA was measured with 1,3-diiodopropane. The maximum activity of DmxA toward 1,3-diiodopropane was detected at 55 °C and pH 9.0. Experiments on measuring of enantioselectivity revealed that DmxA is one of the most enantioselective haloalkane dehalogenase (Chrast et. al, 2019, not published).

2.4. Extremozymes

Temperature is one of the major environmental factors that affects the evolution and activities of organisms (Brock, 2012). However, the proteins that can be active in extreme temperatures are of special interest, because they can be advantageous for biotechnological applications (Sarmiento *et al.*, 2015, Turner *et al.*, 2007). The possession of information on the structural features of such enzymes, and the application sitedirected mutagenesis of less stable enzymes, are useful for the improvement of the properties of enzymes (Turner *et al.*, 2007).

2.4.1. Adaptation of the enzyme to cold

A colder environment is prevailing on the Earth. Various adaptive features enable the cell to adapt to the cold (Feller & Gerday, 2003). The understanding of cold-adaptation is not clear. Several comparative sequence and structure analyses have been made (Gianese *et al.*, 2001, Metpally & Reddy, 2009, Adekoya *et al.*, 2006), and the properties of the enzyme change to cold-adaptation have been discussed previously as well (D'Amico *et al.*, 2002, Feller & Gerday, 2003, 1997); thus far an unambiguous universal formula does not exist. However, there is an established relationship: the improving of conformational flexibility of appropriate parts of the molecular structure, at the same time, protein shows a low stability, and compensation of activity at low temperatures -– being is the main common opinion through all the viewpoints on the enzymes adapted to cold (Feller & Gerday, 1997, Georlette *et al.*, 2003).

This is a challenge for cold-adaptation, so the enzymes require an increase in flexibility of the structural components, which are involved in the catalytic cycle. Other protein regions, if not involved in catalysis, may or may not maintain a high rigidity (Georlette *et al.*, 2004).

In the opinion of Miyazaki et al. (Miyazaki et al., 2000) the obtaining of a high activity at low temperature and stability at high temperatures is achievable simultaneously. And the cold-adaptation conforms the effects of evolution, rather than any intrinsic physical-chemical limitations on proteins.

There is a thought, that cold-adapted enzymes require an improved flexibility of the structural moiety involved in the catalytic activity, whereas other regions are not implicated in catalysis can retain a certain rigidity (D'Amico *et al.*, 2002) and it was proved on the psychrophilic elastase, which is characterized by a reduced flexibility locally in some zones, that are distant from the functional sites. When compared to the mesophilic enzyme, it was in agreement with the hypothesis, suggesting that such local rigidity can be beneficial for the catalytic activity of psychrophilic enzymes (Papaleo *et al.*, 2006).

Feller and Gerday discuss that structural factors such as a reduced number of proline (on loops) and arginine residues can influence the cold-activity of an enzyme. These residues restrict backbone rotations, and can form multiple hydrogen bonds and salt bridges, respectively. While clusters of glycine residues enable localized chain mobility, all weak interactions (ion pairs, aromatic

interactions, hydrogen bonds and helix dipoles) are less abundant, and non-polar core clusters have a weaker hydrophobicity, making the protein interior less compact. Frequently, stabilizing cofactors bind weakly, and loose or relaxed protein extremities seem to favour unzipping. The disappearance of solvent-exposed ion pairs, a larger proportion of non-polar groups which are exposed to the surrounding medium and an excess of negative charges that favour interactions with the solvent – generally characterize the surface of a protein. However, each protein family has it's own strategy in order to decrease stability by using one or a combination of these structural alterations (Feller & Gerday, 2003).

The analysis from Adekoya et al (Adekoya *et al.*, 2006), presented a comparative sequence and structure analysis that reveals features of cold adaptation of an enzyme in the thermolysin family (and a comparison of DpcA versus DhlA, in parentheses, the number of residues, is presented):

1) fewer arginines (17 vs 15),

2) a lower Arg / (Lys + Arg) ratio (0.56 vs 0.55),

3) a lower fraction of large aliphatic side chains, expressed by the

(Ile + Leu) / (Ile + Leu + Val) ratio, (0.75 vs 0.73),

4) more methionines (6 vs 11),

5) more serines (16 vs 13),

6) more of the thermolabile amino acid asparagine (27 vs 26).

Other scientists (Metpally & Reddy, 2009) found that residues such as Ala, Asp, Ser and The are much more often observed in psychrophiles when compared to mesophiles. On the other hand, residues Glu and Leu are significantly less favoured in psychrophile proteomes. They also observed that tiny/small and neutral amino acid groups are notably preferred in psychrophiles where, as charged, basic, aromatic and hydrophilic groups are notably less favoured. They also noted that psychrophilic proteins avoid including the amino acids Glu, Phe, Lys, Asn and Tyr, in comparison to more preferable residues, such as Ala, Asp, Gly, Ser, and Thr as compared to mesophile proteins (Metpally & Reddy, 2009). The agreement with the finding is observed in the comparison of DpcA sequence with that of DhlA: fewer numbers of Glu, Phe and Tyr. The number of residues of Asp, Gly, Ser and Tyr prevail in DpcA.

The results obtained from the analysis by Gianese (Gianese *et al.*, 2001) strongly suggest that the charged residues Glu, Arg and Lys are to be replaced in psychrophilic enzymes. Three residue exchanges are indicated: Glu → Ala, Val → Ala, Arg → Lys. Substitutions involving charged residues are: Lys → Ser, Lys → Asn, Arg → Ser, and hydrophobic residues (Val → Ile)); and these substitutions are mostly present at exposed sites: α-helices or coil regions. Only the substitution Lys → Asn is favoured at the subunit interface. Asn is more frequent in psychrophilic than in meso/thermophilic enzymes.

As an example, Ala12 is in DpcA and DppA instead of Ser13 of DhlA. Thr 12 of DmrA and Glu12 of DccA- are on the α-helice and exposed to the surface. The position in other enzymes is quite shifted even if it is Ala (DppA, DmrA) or Ser (DhlA), Lys (DccA). And this region is near to the Mg^{2+} ion binding site of DccA. Ala 96 from DpcA and Ala 91 from DccA which is placed on the loop are different from Thr of DmrA and DppA and Val of DhlA. Val 102 of the helix is in DpcA, DmrA and DccA, while Arg 102 is in DhlA and DppA.

Other authors, (Feller et al., 1996; Feller and Gerday, 1997; Marshall, 1997) on the basis of comparative single-family analysis, state that for the adaptation to cold the decreased Arg content or decreased Arg/(Lys + Arg) molar ratio is a feature. Gianese et al. have the point of view that one of the mechanisms of low-temperature adaptation is the substitution of charged residues Glu, Arg and Lys. It is shared by most families. As an example, on the loop Lys86 in the DhlA structure is substituted by Arg 86 of DpcA.

Ala is much more often observed on α-helical regions, while Val shows a tendency for β-strands. Psychrophiles avoid Val in buried regions of β-sheets in comparison with the corresponding regions of thermo/mesophiles (for example, Val49 in double conformation of DhlA (PDBID 1B6G) is substituted by Pro 48 in DpcA in the beta-sheet, Val 121 of DhlA versus Phe120 of DpcA, Val 122 of DhlA versus Cys121 of DpcA). However, DpcA and DhlA have Val 77 and Val76, respectively.

Val 156 of DhlA (belong to α-helice, exposed to the surface) vs Asp154 in DpcA (and within the position of the residues it is highly different from the others: DppA (Leu154), DccA (153 Gly), excluding DmrA – at this position Asp154 as well). Ser107 of the helix in DpcA is the only, in comparison to Ala, of the rest of the residues in a similar position. Ala134 is on the helix of DpcA

and DccA versus substituted residue in DmrA (by 134 Gly) and in DppA and DhlA by 133Pro and 134 Pro, respectively. In the position of Ala 209 of DpcA and of DmrA, and 208 of DccA, DppA has Val 207 and DhlA – 219 Val. Numerous of Ala on the helices and loops, exposed to the surfaces are observed in DpcA 3D structure.

The overall effect of the two mutations considered, Val → Ala, Val → Ile, is the decrease in the number of carbon atoms in the hydrophobic core (Gianese *et al.*, 2001).

In conclusion, a deeper understanding on the structural basis of cold-adaptation neded as more structures of cold-adapted proteins and especially enzymes are determined.

2.4.2. The characteristics affecting the thermostability

Thermostable enzymes are studied extensively. They are applied for various industrial applications: in the food and paper industry, for pharmaceuticals, detergents, removal of toxic wastes and drilling for oil (Haki & Rakshit, 2003). The molecular mechanism of thermostability and thermophilicity vary among the enzymes (Vieille *et al.*, 1996), this intrinsic property is determined by the primary structure of the protein (Ward & Moo-Young, 1988).

However, there are some common features, which are identified in these enzymes, such as an increase of interactions (hydrogen bonds, electrostatic interactions, hydrophobic interactions, metal binding) (Li *et al.*, 2005, Vogt *et al.*, 1997). Disulfide bonds play a significant role in an improvement of the protein stabilization (Matsumura *et al.*, 1989, Sheng *et al.*, 2016, Betz, 1993).

The studies of comparison of thermophilic enzymes and their counterparts have shown, that the thermostable enzymes are characterized by a higher content of hydrophobic amino acids and a decreased number glycine residues, and polar and charged amino acids (Maras *et al.*, 1994). However the higher content of charged residue was also observed for thermophilic proteins in the amino-acid composition analysis done by Dechert *et al.*, as well as the preference of Arg over Lys residues (Deckert *et al.*, 1998), and the substitution of Gly by Ala and the substitution of Lys by Arg (Vieille & Zeikus, 2001).

Thermophilic enzymes are characterized by shortened loop regions (Thompson & Eisenberg, 1999), by the decreased number of specific amino acids at helical ends (e. g., Pro), Cβ-branched amino acids (as Ile, Thr, Val) in the

helices of thermophilic proteins (Vieille & Zeikus, 2001, Facchiano *et al.*, 1998). Nevertheless, the larger amount of such amino acids in the active site of thermostable phosphotriesterase-like lactonase from *Geobacillus stearothermophilus* was observed (Hawwa *et al.*, 2009). Hawwa et al. also noticed the advantageous number of prolines at the beginning of the helices, and the number of negatively charged residues at the beginning of the enzyme helices, and the number of aromatic-aromatic interactions of the overall structure (Hawwa *et al.*, 2009).

It was found, that helices of thermophilic proteins are more stable than those of mesophilic counterparts (Facchiano *et al.*, 1998). Overall, thermostable enzymes are characterized by increased protein rigidity and by displaying fewer internal motions in comparison to their mesophilic homologues (Feller & Gerday, 2003, Vieille & Zeikus, 2001). The molecular stability of thermophilic and hyperthermophilic enzymes is connected mostly with the temperature of growth of the parent organisms (Cowan, 1997).

2.5. Methods of protein structure determination

The dry mass of a cell is mostly composed of protein molecules. Their location inside of a cell is not random; we now know that nearly every major process in a cell is carried out by assemblies of 10 or more protein molecules, carrying biological functions. Each of these protein assemblies interacts with several other large complexes of proteins. We can observe the entire cell as a factory, a cooperative net of large protein machines and smaller molecules, realizing organized and challenging works in order to sustain life (Alberts, 1998). To understand the molecular function of a protein a three-dimensional structure is necessarily needed. There are several developed structural biology methods, which are aimed at the visualization of the architecture of macromolecular assemblies, and they are invaluable to our understanding and comparing of mechanism of a biomolecular function (Nogales, 2015, Alberts, 1998).

3D-structure are inestimable source of information, which provides: a) the shape and the domain structure, b) classification of the proteins, c) prediction of function for uncharacterized proteins, d) allosteric modulation and interaction with other macromolecules, interactions with small ligands: metal ions, nucleotides, substrates, cofactors and inhibitors, e) structural information for structure-based drug development, f) evidence for the mechanism of enzymatic

reactions, g) evidence for presence and function of posttranslational modifications: disulfide bonds, N-glycosylation and etc., h) experimental evidence for presence of transmembrane domains.

The common structure determination process for different methods includes several stages: 1) gene cloning, 2) protein purification, 3) sample preparing, 4) protein structure identification experiment, 5) structure solution, 6) validation and deposition (Rupp, 2010). *In silico* analysis with the help of bioinformatics resources (as ExPASy, ProtParam, ESPript and etc.) should be done prior and during the study.

Since the first structure of protein haemoglobin (Perutz *et al.*, 1968) (PDB ID: 2DHB) was determined, and the structural biology has been developed, we are able to use several primary methods to determine the protein structure:

I. Protein X-ray crystallography
II. Cryo-electron microscopy (cryo-EM)
III. Nuclear magnetic resonance (NMR) spectroscopy
IV. Small-angle X-ray and neutron scattering (SAXS and SANS).

Different information is obtained from the experiment from each of the method that is used to solve the structure: X-ray diffraction pattern (for X-ray crystallography), distances between close atoms and information on the local conformation (for NMR spectroscopy), an image of the overall shape of molecule (for electron microscopy).

Thus each method has it's advantages and disadvantages. And it should be chosen adequately to the characteristic of the protein which is under study and addressing to a particular problem.

2.5.1. X-ray crystallography

Since 1985 Wilhelm Conrad Röntgen in 1895 discovered X-rays it become an indispensable tool for structure discovery.

X-rays are high energy electromagnetic radiation with the range of the wavelength from about 10 to 10^{-3} nm. And it is between ultraviolet (UV) and γ-rays in electromagnetic spectrum. Max von Laue realized that wavelength of the X-rays is similar to the spasing of atomic planes in a crystal and this enables to the crystal diffracts X-rays and, consequently, obtaining the information about atom arrangement in the crystal (Waseda *et al.*, 2011, Suryanarayana & Norton, 2013). In 1914 he was awarded the Nobel Prizes the discovery of the diffraction

of X-rays on a periodic arrangement of particles, crystals. What was the contribution to the later work of Sir William Henry Bragg and Sir Lawrence Bragg, father and son, rewarded for "their service in analysis of crystals structure by means of X-rays". And the following honours during the last 100 years were mostly given to crystallographers (Jaskolski *et al.*, 2014).

The most numerous atomic structures, which were deposited in the protein data bank (PDB, Berman *et al.*, 2003, it was founded in 1971) were determined by X-ray crystallography. X-ray crystallography is still the best method for accurate and complete crystal structure determination (Zou *et al.*, 2011). Determining the structure of proteins and other biomolecules at the atomic level is the central issue behind understanding many aspects of biology. X-ray crystallography is the best-known method for structural biology, however, it is suitable only for samples that can be crystallized (Bai *et al.*, 2013).

2.5.1.1. Protein crystallization principle

The meaning of the crystallization is in the kinetically controlled phase separation in a thermodynamically metastable supersaturated system. The process results in the crystal formation. In this case thermodynamic parameters (extensive: protein or reagent concentrations and intensive: temperature, pH) can be controlled, but there are limited influence and knowledge about the kinetic parameters (equilibrium rates, nucleation, reassembly, kinetics of crystal growth) (Chayen & Line, 2007).

Crystallization takes place in a few steps: 1) nucleation, when molecules realize the first-order phase transition from a fully disordered state self-assembly to an ordered one, periodic three-dimensional array – critical nuclei, 2) crystal growth, when periodic three-dimensional array from critical nuclei is build 3) cessation of growth (Juárez-Martínez *et al.*, 2001, Feher, 1986). The achievement of the supersaturated state is the crucial importance challenge, which is required to crystallize the protein.

The crystallization process can be described by the phase diagram (Fig. 5).

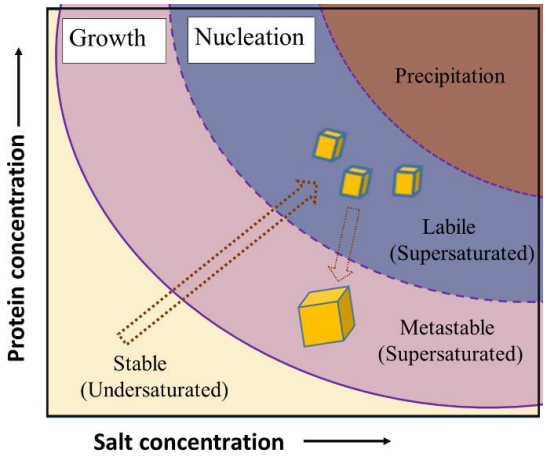

Figure 5. Macromolecular crystallization phase diagram.

The undersaturated and supersaturated zones are segregated by a sharp line on the solubility diagram. The line marks the maximum solubility at specific concentrations of a precipitant and represents the equilibrium, which exists between the free-molecule phase and the solid phase. The others subdivisions (metastable and labile zones) by dashed line of the supersaturation zone are more uncertain. The crystal will develop from the nuclei in the metastable zone, however nucleation here will not occur. In the labile zone is the possibility for both events. The final zone is at a very high supersaturation, where the precipitation might be most probable. As crystals can be grown from a supersaturated solution only, the creation of such a supersaturated solution for the protein of interest underlines the issue in growing protein crystals. Adapted from (McPherson & Gavira, 2013).

Nucleation

The understanding of the nucleation process during the crystallization process includes the understanding of the intramolecular forces in solution. The interactions between the protein molecules includes interactions of amino acid residues, carrying charged and uncharged regions, and in solution are affected, for example, by screened repulsive Coulomb interactions and attractive van der Waals interactions. Additionally to this, protein–ion dispersion forces and hydrophobic and hydration forces are also important (Gunton *et al.*, 2007, Israelachvili, 2010).

At supersaturated conditions, accidental collisions of solute particles lead to the formation of critical-size nuclei. It is an unstable cluster, which can then become a seed, initiating crystal growth. The critical nucleus is smaller at the higher supersaturation and there is the higher probability for nucleation (Rosenberger *et al.*, 1996).

Nucleation energy. During the crystallization process the kinetic barrier exists. To achieve the phase separation, nucleation must overcome it. In case of homogeneous nucleation, high supersaturation is needed to overcome the free energy barrier $\Delta G_n^{\#}$ with the help of formation of critical nuclei to which additional collisions of other molecules will lead their attaching. At heteronucleation a foreign material is introduced. It allows to form a nucleus at lower supersaturation with a lower free energy barrier $\Delta G_n^{\#}$ (Durbin & Feher, 1996, Rupp, 2010).

Entropic contributions. Thermodynamically, entropic contributions drive the crystallization process.

The hydrophobic side chains of protein are liable to the tendency to association in polar medium (water) in such a way that a hydrophobic protein core is formed. Hydrogen-bonded water molecules form a clathrate cage, surrounding hydrophobic residues, that decreases the entropy of the solvent.

When the protein molecules reorganize, move from the liquid state to form the crystal state, these water molecules shall be undergone of rearrangement. Thereby, the releasing of ordered water molecules from the hydration shell around hydrophobic residues (resulting the fewer molecules need to be ordered) is the main force that leads to the folding of a protein. This results in fewer water molecules need to be ordered – the gain in the entropy of solvent ($\Delta S_{solvent}$ is increasing), and to a decrease in free energy, and a stabilizing of interactions results (Gunton *et al.*, 2007, Rupp, 2010).

The protein self-assembling into crystals requires decreasing of the free energy of crystallization, $\Delta G_{crystallization}$ ($\Delta G_{crystallization}$ must be negative).

The fundamental equation of the free energy of protein crystallization from solution ($\Delta G_{crystallization}$), where is the total entropy change is separated into the entropy change of the solute (protein, $\Delta S_{protein}$), that is decreasing and the entropy change of the solvent ($\Delta S_{solvent}$), that is increasing, *Eq.*2 (Rupp, 2010):

Equation 2

$$\Delta G_{crystallization} = \Delta H_{crystallization} - T\,(\Delta S_{protein} + \Delta S_{solvent})_{crystallization}$$

Where $\Delta G_{crystallization}$ is Gibbs free energy of crystallization, $\Delta H_{crystallization}$ is the standard enthalpy of crystallization; the enthalpy change is weakly negative for crystal formation. $\Delta S_{protein}$ and $\Delta S_{solvent}$ are entropy contributions. $\Delta S_{protein}$ is the entropy change followed the attachment of the protein molecules to the crystal. The next term $\Delta S_{protein}$ to be negative. In case of $\Delta S_{solvent}$ is positive, the condition for the phase transition is favourable, when it is negative, is favourable for the free protein molecules in solution.

This can explain that targeted surface residue mutation will improve the crystallization (Gunton *et al.*, 2007, Rupp, 2010, Shiryayev *et al.*, 2005, Derewenda & Vekilov, 2006).

Crystal growth and its cessation

It is possible to observe the crystal growth with the help of atomic force microscopy (AFM). The two modes of growth exist: a) normal (or crystal face) growth and b) tangential, layer growth. The normal growth occurs by initiating of new layers or islands, which have step edges. The new molecules can be added to them. The tangential growth is an easier process: the new molecules continually join the step edges of layers and extends them.

New steps can be generated by two-dimensional nucleation or by a screw dislocation crossing the crystal face. The supersaturation will influence which mechanism will be prevailing (Chernov, 2003, McPherson *et al.*, 2003).

The faster crystals grow, the smaller are their sizes in the end of experiment.

After a certain time, the growth cessation occurs. It can be promoted by blocking of favourable growth sites on the surface of a crystal with the impurities (Feher, 1986).

It also can be explained by the the natural disorders of a crystal due to increased weakness of bonds between adjacent protein molecules, caused by the addition of precipitating salts (Falcón Rodríguez *et al.*, 2000).

2.5.1.2. Crystallization techniques

Numerous crystallization techniques have been developed to crystallize proteins. They are schematically represented on Fig.6. The selection of the technique depends on the specific purpose: initial screening, growth optimization, ease of automation and harvesting.

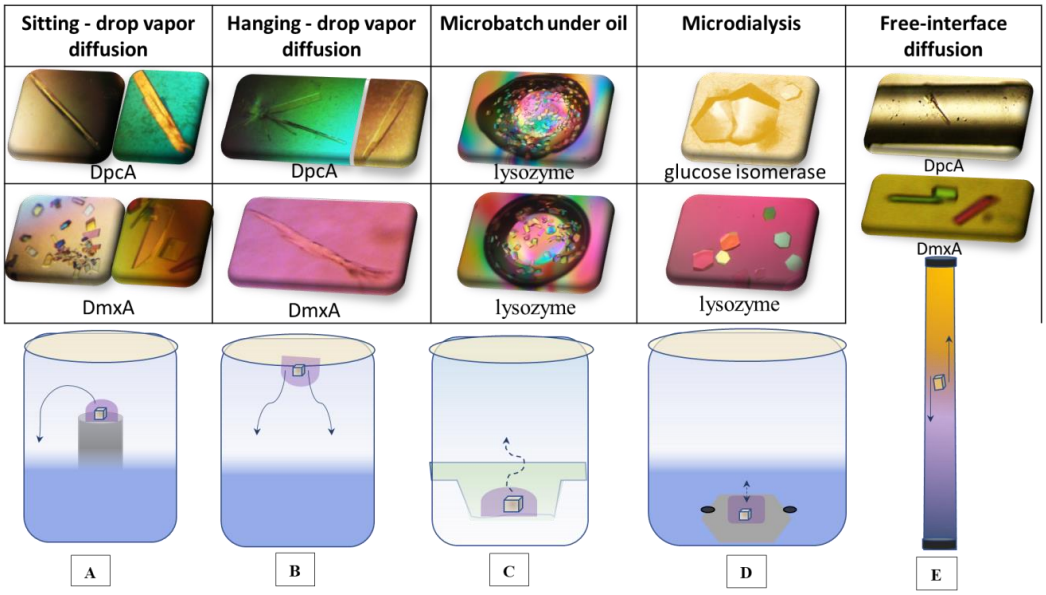

Figure 6. Schematic representation of common crystallization techniques and their result at the end of the experiment: **A** - hanging-drop vapor - diffusion, **B** - sitting-drop vapor - diffusion, **C**- microbatch under oil, **D** – microdialysis (picture of glucose isomerase is adapted from (S. J. Lee & Cudney, 2004)), **E** - free-interface diffusion. DpcA and DmxA crystals presented for illustration, obtained by A, B and E techniques, lysozyme crystals from our laboratory – C and D. Adapted from (Rupp, 2010, Chayen & Line, 2007).

The studied macromolecule is in solution at the beginning of all crystallization techniques experiments. When the solution is brought into supersaturation, the protein comes out of solution and form crystals. The concentration of protein in the solute drops when the nuclei is formed, followed by the appearing system in the metastable zone, where growth should occur without the formation of further nuclei (Chayen, 1998).

Vapor diffusion techniques. They are the most popular among crystallographers and resultful for structure uploading in the protein data bank (PDB). The term “diffusion” is applied in the meaning that water molecule

transport occurs because of evaporation from the drop to the reservoir solution (Bergfors, 2009).

In the hanging-drop vapor diffusion technique (Fig.6, A), a small aqueous drop of protein is placed on the cover slide, and is mixed with the equal amount of the crystallization solution from a reservoir. The cover slide is flipped over and seals the reservoir or well with the greased rim. We obtain a closed system with the twice higher concentration of precipitant in the reservoir than that on the cover side. Because of the difference in concentrations, water vapor from the hanging drop diffuses into the reservoir, protein and precipitant concentrations increase proportionally until equilibrium is reached (Chayen, 1998, Rupp, 2010).

The sitting-drop vapor diffusion (Fig.6, B) has the same principle, except for the drop support setup geometry. In this case, the drop is placed directly on the shelf or post and the system is sealed by tape. The equilibration here proceeds at a slower rate than in the similar hanging-drop experiment.

The method with the drop which is hanging – is good for manual small-scale setups, while the sitting-drop method is preferred, when the crystallization is robotized. There are also other advantages for each of the vapor diffusion technique. For hanging-drop application are inherent following benefits: reducing crystals sticking to the plastic and, as a consequence, it is easier for crystal harvesting. While for sitting-dop vapor diffusion are beneficial due to it's higher mechanical stability and temperature change resistance, easier for carrying out seeding, using in automation and application in high-throughput screening (HTS) and it is without the consequences of the low surface tension (Chayen, 1998, Bergfors, 2009, Rupp, 2010).

The speed, simplicity, small amount of sample and plenty of different plates and consumables on the market – are the common advantages over the rest of the techniques, and put vapor diffusion in the first line of choice (Bergfors, 2009, Rupp, 2010).

Microbatch crystallization under oil. This technique has been used for many years and it is also adapted for robotization. In batch crystallization (Fig.6, C) protein and precipitant are mixed together in order to obtain a supersaturated protein solution. The drop is dispensed on the bottom of the microplate well, which is covered by the surface of the oil (water-impermeable: paraffin, water – permeable: silicone, oils mixture), that acts as a barrier between the reservoir and

the crystallization drop. During the incubation water vapor diffuses slowly through the layer of oil. The type of oil, its thickness and the presence of alcohols, detergents, lipids (substances that can diffuse into oil) will influence the rate of vapor diffusion and consequently the speed of crystallization (Chayen, 1997, Rupp, 2010, D'Arcy *et al.*, 1996, Rayment, 2002, Chayen, 1998).

Dialysis. Microdialysis is the technique (Fig.6, D) (Lagerkvist *et al.*, 1972, Zeppezauer *et al.*, 1968), where the sample is placed in the plastic chamber (button) and is sealed with the semi-permeable membrane that is held by a secured O-ring. This blocks protein (and high molecular PEGs), but let to penetrate to small molecules (like salts, ligands and other compounds smaller than the size of the pores of a dialysis membrane), and the protein slowly reaches the equilibrium with the precipitant solutions. The obtaining crystal by the use of a salting-out scenario is more often, where high concentrations of salt or other small membrane-permeable compounds decrease the protein solubility. Exceptional salting – in a scenario to remove solutes is used, where dialysis against distilled water takes place.

The advantage is that the button with the protein can be exposed to different precipitants, while consuming a large amount of the protein is a disadvantage. The technique is more preferable for growing large crystals (Rupp, 2010, Luft *et al.*, 2014).

Free-interface diffusion. The basic idea of the free-interface diffusion (Salemme, 1972) is to provide the contact of precipitant with the protein without premixing in a glass capillary (Fig.6, E). The higher density solution (usually precipitant) is placed beneath the lower density solution (usually protein) and the capillary is held vertically. The liquids will equilibrate against each other freely, by diffusion only. The high supersaturation region, suitable for nucleation, will be observed at the point of contact of liquids from the start. Within the time of the experiment, the gradient will become smoothed with the lower supersaturation and become more feasible for the controlled growth. It is designed for a single-nucleation event at a height supersaturation.

The extension of the free-interface diffusion method is counter-diffusion that takes advantage of the difference in the diffusion speed between protein and small molecules, and looks for numerous events of precipitation. The process is driven by diffusive mass transport: as the precipitating agent at a very high

concentration, (salt), will move like a "wave" into the protein solution, while the protein molecules, stay at the same place, because of their slower diffusion as they are much larger. The convective mixing must be limited by the microgravity, gels, internal diameter of the capillary (equal or narrower to 0.2 mm). That endorses variations of the technique, adding low-melting agarose (Biertümpfel *et al.*, 2002), and using microfluidics chips (Dhouib *et al.*, 2009). The drawback of the technique is the consuming of a larger amount of the sample, and the expenses of the chips (García-Ruiz, 2003, Bergfors, 2009, Luft *et al.*, 2014, Otalora *et al.*, 2009, Rupp, 2010, Moreno *et al.*, 1999)

The change in the technique can affect the vapor-phase equilibration, respectfully the outcome.

2.5.1.3. Growing derivative crystals

It is needed to obtain the derivatives for phase determination and for studying protein-ligand interactions. It can be possible by: a) co-crystallization of a protein and ligand, and b) soaking of the protein crystals in mother-liquor solution that contains ligand. The co-crystallization method is directed to produce the crystal complexes of proteins with large ligands (such as nucleic acids or other proteins). For the smaller molecules soaking can be used, by diffusion of ligands to the active sites through channels of water in the crystal. Soaking is more favourable for the crystallographer for the comparison of a pure protein structure with that of a complex of protein-ligand (Rhodes, 2006).

Another possibility – to produce selenomethionine derivatives by gene expression and purification of the protein with incorporated selenomethionine instead of methionine residues.

For metalloproteins the phasing is possible due to their introduced heavy-atom compounds. For example, haemoglobin containing iron (II) ions that can be used for the obtaining of phase information (Rhodes, 2006).

2.5.1.4. Crystallization screening and optimization strategy

Almost all crystallization experiments start from the screening of the initial condition of crystallization. There are numerous of companies commercially providing crystallization screens, which are used for trials, unless the researcher uses home-made prepared stock solutions as well.

It has been estimated that for the number of proteins for 288 trials there exists a high probability of finding successful crystallization hit condition (Rupp, 2010).

It is a highly rare outcome when a good diffraction crystal is at once obtained. Observed crystalline precipitate, phase separations and small crystals are worth pushing. After the carried out initial crystallization screening experiments where the main parameters were determined, the optimization is following (Chayen & Saridakis, 2008).

To obtain high quality diffraction data good crystal is needed exceptionally. For this reason, few methods can be applied: 1) fine-tuning-is variation of chemical parameters (protein or precipitant concentration, pH, and protein/precipitant) and physical parameters (temperature, vibration) of initial crystallization hit condition (Fig. 7); 2) using additives (commercially available screen of 96); 3) performing seeding.

Different tricks are also used by scientists to optimize crystals: using cross-crystallization method (Tomcová *et al.*, 2006) and to obtain protein crystals: crystallization in pores, that is effective in the inducing of nucleation (Chayen *et al.*, 2001, Nanev *et al.*, 2017).

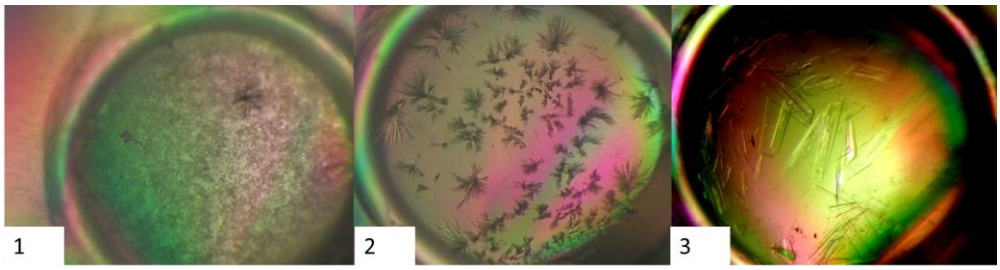

Figure 7. An example of the result of DpcA enzyme optimization. The protein used at the same concentration. **1.** 0.2 M Calcium chloride dihydrate, 0.1 M Sodium acetate trihydrate pH 4.6, 20% v/v 2-Propanol (prot/prec=1/1); **2.** Changed cocktail composition: 0.2 M Ammonium acetate, 0.1 M Sodium acetate trihydrate, pH 4.6, 30% w/v Polyethylene glycol 4,000 (prot/prec=1/1); **3.** Increased pH, added additive, protein to precipitant ratio: 0.2 M ammonium acetate, 30% (w/v) PEG 4.000 and 0.1 M sodium acetate trihydrate pH 5.88, (prot/prec=2/1).

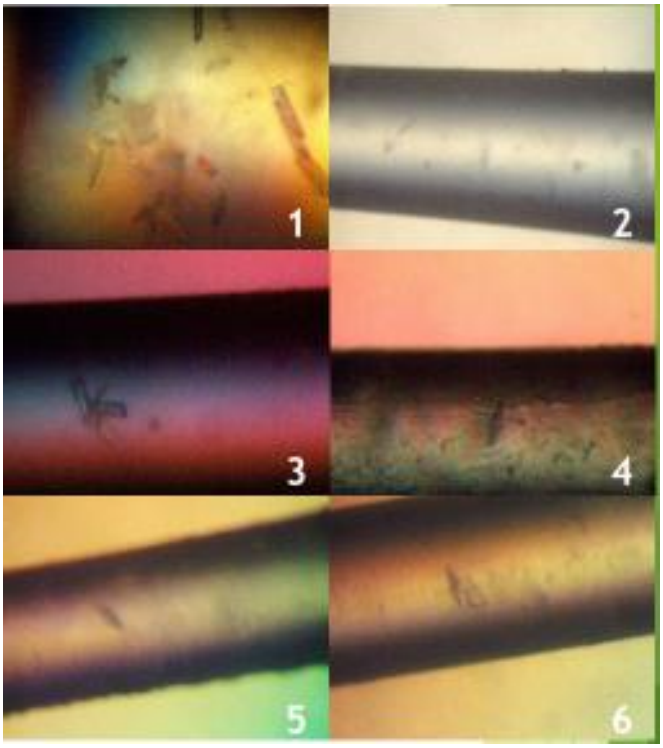

Figure 8. The optimization results of two batches of DmxA, purified in 0.5 mM Tris-HCl buffer (pH 7.5). Experiment is performed by counter diffusion with two additives from Hampton research (1,3- sarcosine; 2,4,5,6- L-proline). Better crystal formation is observed in the cases of 1 and 3 with the addition of 0.1M sarcosine.

2.5.1.5. The protein crystal structure determination

2.5.1.5.1. Crystal and its properties

The crystal with its symmetry is a "nucleus" of a crystallography. Since the 18th century, it has been the assumption that the form of a crystal symmetry is the expression of the regularity of the internal organization.

The protein crystal is an ordered 3D arrangement of molecules that are held by weak noncovalent interactions (Fig 9). The simplest and identical repeating building blocks that constitute the crystal are referred to as its unit cells (Rhodes, 2006). In three dimensions, a unit cell is characterized by length of basis vectors (**a, b** and **c**) and by three angles between them (α, β and γ). The angle α lays between the basis vectors **a** and **b,** β is between **a** and **c**, *and vectors* **a** *and* **b** *enclose angle* γ (Massa, 2000, Rupp, 2010). In regard to the symmetry in protein

crystallography it is possible to apply three operations to the structure that is the rotation, translation and screw axes (combination of the first two).

The lattice of the structure – is the array of identical points (equivalent to each other by translation symmetry). In each unit cell is contained the equivalent of one lattice point (one repeat unit). When the operations to symmetry (symmetry of the space group) is applied and the smallest portion of a crystal structure can generate the whole crystal – we speak about the asymmetric unit (Valvoda *et al.*, 1992).

Only the atoms coordinate in the asymmetric unit are deposited in the various crystallographic databases. The generation of the remainder of the contents of the unit cell is set of operations which depend on Space Group (Prince, 2004).

The internal symmetry of the unit cell and its content is described by its space group. Space groups belong to seven crystal systems: triclinic, monoclinic, orthorhombic, trigonal, tetragonal, hexagonal and cubic. There are 230 possible space groups existing and in 65 among them can be exploited by biological macromolecules, because of their chirality (Hejátko & Hakoshima, 2018). The knowledge of the space group and dimensions of the unit cell are essential for the interpretation of the diffraction data (Rhodes, 2010)

To define the position of atoms in the crystal structure, the information about the unit cell, asymmetric unit and the space group is needed (Rhodes, 2006).

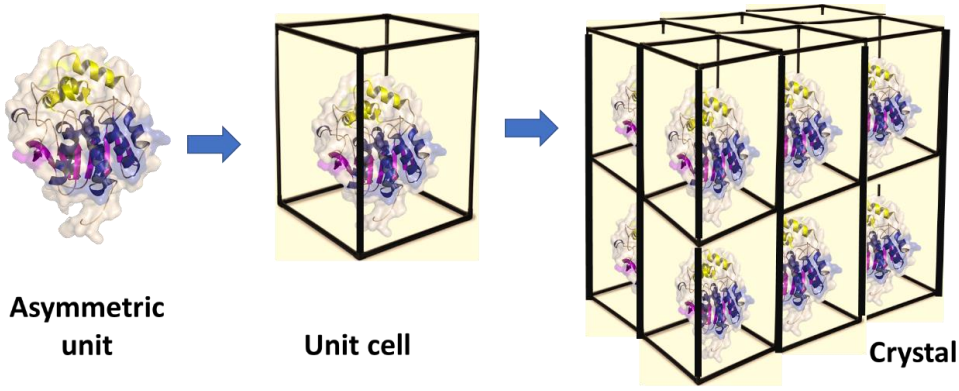

Figure 9. Relationship between an asymmetric unit, unit cell and a whole crystal. The asymmetric unit is represented by HLD. Transformed from (Rupp, 2010).

Important properties of the protein crystal are mentioned below:

a. Protein crystals are highly fragile in comparison to inorganic crystals and are easily crushed by gentle pressure (e.g. a needle crash test).
b. Is coloured by the Izit dye (Hampton Research, CA, USA) that will penetrate the solvent channels of the protein crystals. It will also have weak birefringent under cross polarizers.
c. Real macroscopic crystals are not perfect. They posses mosaicity, mutual misorientation of smaller domain crystal. Arrays of crystal are unevenly aligned with each other, the reflections that come up from the crystal are cone-shaped, because of this they must be measured over a small angular range, but not at a single angle.
d. Different forms of crystals vary in diffraction quality, in how easily they will be reproduced or grown, but the diffraction quality is the most important criterion.
e. Molecular structure of the protein is not altered by crystallization (conclusion is resulted from the structures from different types of crystals) (Rhodes, 2006).
f. The solvent content in the crystal is from 27% to 65%, with an average of 43% (Matthews, 1968). Water molecules, occupying sites inside crystallized biological macromolecules or at their surface, play an important role in their functional and structural properties (Vuister *et al.*, 1991).
g. Some enzymes retain their activity in the form of crystal (Rhodes, 2006).

2.5.1.5.2. X-ray data collection

When the crystal is obtained (Fig. 10, 1), it has to be fished out and prepared for the diffraction experiment on the synchrotron or in-house diffraction source. Cryo flash freezing in liquid nitrogen is needed. In most cases, different cryoprotectants are used to test on the crystal (glycerol, ethylene glycol, propylene glycol, PEG 400, MPD, sucrose and others). After the suitable cryoprotectant is found, the crystal is frozen and is kept on a loop in a vial in Dewar with liquid nitrogen. Than a Dewar container, filled with pucks of vials with crystals is taken on the synchrotron (Fig.10, 2). The vials are transferred to the local container with the liquid nitrogen and one by one are examined.

Carefully and quickly the loop with a crystal is mounted onto a goniometer (Fig.10, 4.A) that is between an X-ray source and X-ray detector (Fig. 10, 3-5). The goniometer is placed on the gonistat, all together; the system provides precise automatic rotation of the crystal with respect to the X - ray beam and the detector, which in its turn changes the position with respect to X-ray beam. The stream of nitrogen gas is directed on the crystal to prevent crystal radiation damage and decrease the propagation of free radicals.

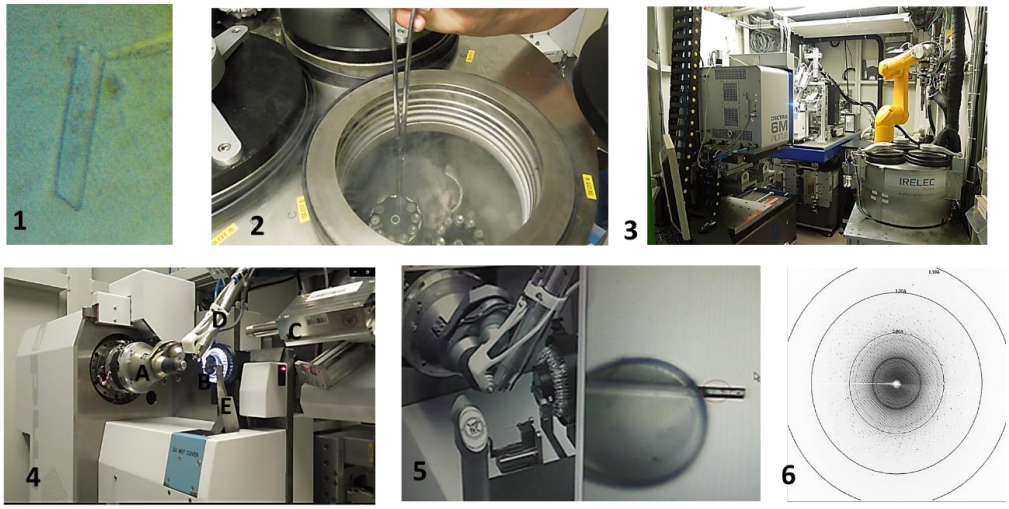

Figure 10. Crystal preparation for acquiring diffraction data. **1.** Protein crystal of DpcA is in the crystallization drop. **2.** Containers with the liquid nitrogen on the synchrotron, where the vials in pucks are kept during the experiment. **3.** BESSY (HZB, Berlin, MX BL14.1 (Mueller *et al.*, 2012)). Pilatus 6M pixel-detector (DECTRIS, Switzerland) enabling rapid data collection and CATS sample changer robot (IRELEC, France) are shown. **4.** A loop with the crystal is placed on the goniometer head Mini-Kappa (Arinax, France) (A) and is tested by the beamline via microdiffractometer MD2 (B), the crystal is seen in through the camera. X-fluorescence detector (C) can be additionally used (for the determination of the presence of chemical elements of interest in the sample). The crystal is under the cryostream (D) continually. To prevent the excessive radiation the direct beam is blocked by the beamstop (E). **5**. The frozen crystal before the start of the experiment, camera view. **6.** Resulted diffraction image. *Photos are prepared by K. Doleželová.*

X-ray beam leaving the collimator are reduced by the focusing mirrors (a system of metal plates), and directed at the crystal to irradiate it (Rhodes, 2006). The X-rays are scattered by the electrons of the atoms in the crystal and the

reflections are recorded on the detector area, displaying the image with black spots (Fig. 10.6). The spots of the reflections have different intensities. The darker the reflection, the greater the intensity, which is the measure of the strength of the beam that diffracted and produced the spot. The reflection position in its turn can provide the direction in which the particular beam was diffracted by the crystal. The data collection results are a set of diffraction images that contain the information of the intensities for each single reflection and its position, which includes the information on the molecular arrangement of atoms in protein crystal. Such diffraction images (frames of data) are captured within each rotation of a crystal through a 0.1º-1.5º angle (Rupp, 2010, Rhodes, 2006).

The particle storage rings, as a synchrotron radiation source, are associated with the accelerators of the particles (Rhodes, 2006). Their concept is complex, Vladimir Veksler (Veksler, 1945) in the USSR, and Edwin McMillan (McMillan, 1945) in the USA, proposed, independently, the design of the synchrotron.

A spin-polarized beam, which is produced by a particle source, is injected into a linear accelerator (linac). The electron bunches are accelerated there to be then transferred to a booster synchrotron, which is intended to overcome the energy gap between the linac and followed by the main ring (Mane *et al.*, 2005). In the storage rings, electrons or positrons circulate at velocities approaching the speed of light under an ultra-high vacuum. The circular path of electrons is maintained by powerful bending magnets, which are one of the synchrotron light sources. X-rays are emitted, when an electron moves in a curved motion passing through a magnetic field. Additionally, wiggler and undulator magnets are sources of the generation of radiation and control the beam parameters of emitted photons (Brown *et al.*, 1983, Vinokurov & Levichev, 2015, Rhodes, 2006, Diener *et al.*, 2019, Mobilio *et al.*, 2014). Finally, the radiation at the beamline can be used for various experiments. During the experiment the scientists are outside of the experimental hutch, and with the help of computer software, change the crystal and it's orientation and position of the crystal to an X-ray beam, as well as other parameters, and decide on the strategy for data collection in order to obtain diffraction images, process the data (Fig. 10.6).

2.5.1.5.3. X-ray scattering, Bragg's Law and the phase problem

The electrons of an atom, surrounding its nucleus, form an electron cloud that scatter X-rays. The photons of the X-ray travel as an electromagnetic wave and when they hit electrons, they cause electron oscillations coherent with impacting photons. The electrons emit waves that have an identical frequency and in fixed phase relation to others, which will result in their superimposed scattered wave.

Concerning the scattering of all of the electrons of the atom, the scattered wave amplitude emanating from an atom depends on the electron distribution in a given position and is described by electron density. This enables the calculation of the atomic scattering factor. However, there is a phase difference between the partial waves that are emitted from the electrons of atoms in the "back" from those in the "front", when the photon electric field hits all of the electrons in an atom. This difference results in backward scattering weakening (Rupp, 2010).

The obtained diffraction spots were interpreted by William Henry Bragg (Bragg & Bragg, 1913) and interpreted as being due to the X-ray reflection of atoms on planes (Fig. 11). And crystal diffraction can occur when Bragg's Law is satisfied:

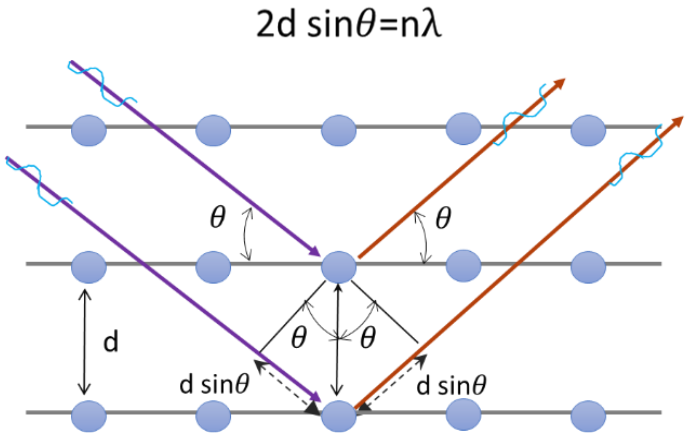

Figure 11. Schematic representation of the incoming incident beam (violet colour) and scattered (dark red) scattered X-rays on the electrons of atoms (blue dots) arranged in a crystal on atomic planes (grey). Where d is the distance between parallel planes, θ is the angle between incident (scattered) wave and a plane, λ is wavelength of the incident X-ray, n is an integer, and $2d\sin\theta$ is the path difference between waves. And the angle of incident have to be equal to angle of diffraction $\theta=\theta$ (Rupp, 2010).

Thus, the radiation that is emitted by the unit cell is the summation of the waves of all atoms scattering from the unit cell (Rupp, 2010). The reflections obtained on the detector can be described by a structure factor (F_{hkl}), and it can be transformed to the electron density by Fourier summation (Rhodes, 2006).

In the diffraction experiment, as was mentioned above, we obtain the measured intensities of waves scattered from the crystal planes and they are designated by lattice indices or Miller indices (*hkl*). They specify how many numbers of planes exist per unit cell in *x*, *y*, *z* directions, respectfully (Rupp, 2010, Taylor, 2003). The intensity of (*hkl*) reflection (I_{hkl}) is proportional to square modulus $|F_{hkl}|^2$. Correspondingly, this enables the determination of the structure factors absolute value $|F_{hkl}|$ experimentally, if taking the square root of the reflection intensity $(I_{hkl})^{1/2}$, measured on the detector. As the structure factor (F_{hkl}) is the complex number $|F_{hkl}|e^{i\varphi}$, where $|F_{hkl}|$ is the structure factor amplitude and $\varphi(hkl)$ is the phase angle between the diffracted waves, then we know I_{hkl} from the diffraction data, and we know structure factor (F_{hkl}), however the information on the $\varphi(hkl)$ is lost (Atkins & De Paula, 2006, Rhodes, 2006, Rupp, 2010).

The desired electron density distribution (ρ (*x*, *y*, *z*), electron density map) per unit cell volume (V) provides the opportunity to obtain the atomic coordinates (*x*, *y*, *z*) in the studied molecule of a protein. And mathematically, its relationship with the structure factor expressed through Fourier transformation (Rupp, 2010, Rhodes, 2006):

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

Amplitudes — Phases - ?

2.5.1.5.4. Structure solution

There are several techniques for solving the phase problem: the direct method, molecular replacement, isomorphous replacement, and anomalous dispersion.

Direct method

The direct reconstruction of the mathematical relationship between the structure factors and the phase information is used to obtain the missing component (phases) as they are obtained from the observed diffraction pattern.

The calculation method is the following: firstly, choosing phases for a few strong reflections. Then phases for other reflections with the help of the phase relationships among the strong reflections are calculated. The electron density can be determined and explained, when the sufficient phases have been calculated (Cowtan, 2001).

This method requires high resolution data (<1.2 Å) and is limited to about 1,000 atoms of protein (Taylor, 2003).

Molecular replacement (MR)

This is the most frequently used method, when the phases from structural factors of a homology model are available to estimate the phases of a new protein (Atkins & De Paula, 2006). The sequence identity should be >25% and the r.m.s. deviation between the α C atoms of the model and the final structure should be <2.0 Å (Taylor, 2003).

Isomorphous replacement (IR)

The structure of haemoglobin was determined by using this method (Green *et al.*, 1954), and until the mid-1990s it played the main role in structure determination. For the successful application of this method, the diffraction data from the native protein crystal are required within the data of diffraction, obtained from the crystal, in which the heavy atom was introduced. For this purpose, heavy atoms by crystal soaking or co-crystallization are incorporated in the molecular structure without the change of the crystal's unit-cell parameters or orientation of the protein in the cell (Sanderson & Skelly, 2007). In the case of a single heavy atom derivative, the method is referenced to SIR; for multiple – MIR.

Determining the position localizations of the extra atoms and the scattering from them can be calculated both in magnitude and in phase. The structure factors from the measured diffraction patterns of the native crystal, $|F_P|$, and that of the derivative, $|F_{PH}|$, are compared and the isomorphous difference, $|F_H|$, is estimated: $|F_H| \backsimeq |F_{PH}| - |F_P|$ (Sanderson & Skelly, 2007, Crick & Magdoff, 1956, Taylor, 2003, Cowtan, 2001).

Anomalous dispersion, single and multiple (SAD and MAD)

This method uses the advantage heavy atoms have to absorb the X-ray at the specific wavelength, "absorption edges". Atoms such as carbon, nitrogen and oxygen do not contribute to anomalous scattering, thus mutated seleno-methionine instead of methionines, or metal-containing proteins, or soaking of

heavy atoms into a protein sequence enable to use anomalous scattering (Rhodes, 2006). This anomalous scattering (or anomalous dispersion) breaks Friedel's Law: $|F(hkl)| \neq |F-(hkl)|$, which let to indicate the anomalous scatterers, while using the dispersive difference between the data recorded at different wavelength (Rhodes, 2006, Rupp, 2010, Taylor, 2003).

A single wavelength use is referred for SAD, while for MAD, the data are collected at several wavelengths (usually three) to maximize the absorption and dispersive effects (Taylor, 2003).

Single and multiple isomorphous replacement with anomalous scattering (SIRAS and MIRAS)

For this method, the combination of IR with anomalous dispersion is inherent.

Firstly, the data with the native crystal are collected with the denoting of $|F_P|$ for each of the reflections is followed by determination of the heavy atom-derivative, and then the second data collection happens also at the same wavelength. Obtaining of $|F_{PH}|$. The third data collection – is at another wavelength for the maximization of anomalous scattering (Rhodes, 2006).

2.5.1.5.5. Structure refinement, validation and deposition

This is the last stage of the structure determination, while adjusting parameters that describe the structure (for example, positions of atoms). By iteration of cycles of model building and calculation of a map, it is intended to give the best fit between the observed intensities and the calculated from the model structure that was deduced from the diffraction pattern (Atkins & De Paula, 2006).

The very final stage of the research on such a problem is the deposition of the atomic coordinates and structure factors, and the publication of the pertinent paper (Wlodawer et al., 2013).

2.5.1.5.6. Perspectives

The synchrotron beamlines are continuously developing and improving, applying new devices and software. The implementation of microfocus beamlines and ultrabrilliant femtosecond X-ray pulses of X-ray free-electron lasers (XFELs) raise new challenges in structure determination. XFEL requires

an electron beam with a high energy that will be obtained from the path through an extremely long wiggler (Rupp, 2010) and for microcrystallography this allows the collection of data without radiation damage (Yamamoto *et al.*, 2017).

2.5.2. NMR spectroscopy

Since the 1940's, when physical chemists (Isidor Rabi and Felix Bloch with Edward Mills Purcell) invented NMR spectroscopy for studying the properties of atomic nuclei, the method became useful in determining the molecular structure of organic compounds and protein molecules.

The spectroscopy field is associated with the interaction between electromagnetic radiation and matter. NMR is based on the conception of nuclear spin which can applied to atomic nuclei such as ^{1}H, ^{13}C, ^{15}N, ^{19}F and ^{13}P, which possess an overall spin. This spin induces the nuclei adoption of preferred orientation in an electromagnetic field. The nuclear spins can change the orientation under influence of the electromagnetic radiation in the radio-frequency range and each orientation has a different energy (de Graaf, 2013). Transitions between states with certain energies are permitted according to the postulates of quantum mechanics and, when we apply pulses of electromagnetic radiation at frequencies that precisely match these energy gaps, we are able to observe transitions that give rise to NMR signals. Nuclei in different chemical environments (e.g. the different ^{1}H nuclei in a protein) will resonate at different frequencies and a plot of intensity against resonance frequency is known as a 1D NMR spectrum.

Resonance frequencies are typically reported as "chemical shifts" in units of p.p.m., which corrects for the fact that the raw frequencies (usually in units of MHz) scale with the size of the magnet (Kwan *et al.*, 2011)

^{1}H or proton NMR spectroscopy was the first established technique for determining the structure of small molecules, up to 100. With the increasing number of residues appear the resonance overlap and the line width problem that cause limitations for this technique. It becomes possible to overcome the issues after the introduced pulsed NMR, combined with Fourier transformation (Ernst *et al.*, 1987). This one-dimensional NMR shows signals for each hydrogen atom. The next step, that expanded the opportunities – the establishment of a second frequency dimension in NMR spectroscopy – ^{15}N-labelled heteronuclear single-

quantum coherence (HSQC) spectrum, that detects chemical shifts and coupling of nuclei (^{1}H - ^{15}N or ^{1}H - ^{13}C) (Jeener *et al.*, 1979); and additionally of three- and four-dimensional NMR, using isotopic labelling of the protein by most commonly ^{1}H, ^{13}C, ^{15}N (Oschkinat *et al.*, 1988, Vuister *et al.*, 1991) (Bax & Grzesiek, 1993).

NMR spectroscopy study structural models of proteins in near physiological conditions and is suitable for defining flexible parts of proteins and their possible conformation. NMR is applicable for protein dynamics, kinetics and thermodynamics. However, the atomic structure determination has a limitation for the proteins, containing no more, than approximately 300 amino acids (and about 25 kDa). It is connected with the losses of sensitivity (due to NMR relaxation) and increasing of spectral complexity in large proteins (Rhodes, 2006).

The advantage of NMR spectroscopy (Fig. 12) is the determination of the structure of the protein in solution and defining flexible parts of proteins and their possible conformation. The disadvantage is requiring large amounts of material and in addition, the stability of studied protein at room temperature under a rather lengthy period for data acquisition.

However, NMR analysis is applied to protein domains, which tend to be small enough for this technique and stable structures can be obtained (Lodish *et al.*, 2007).

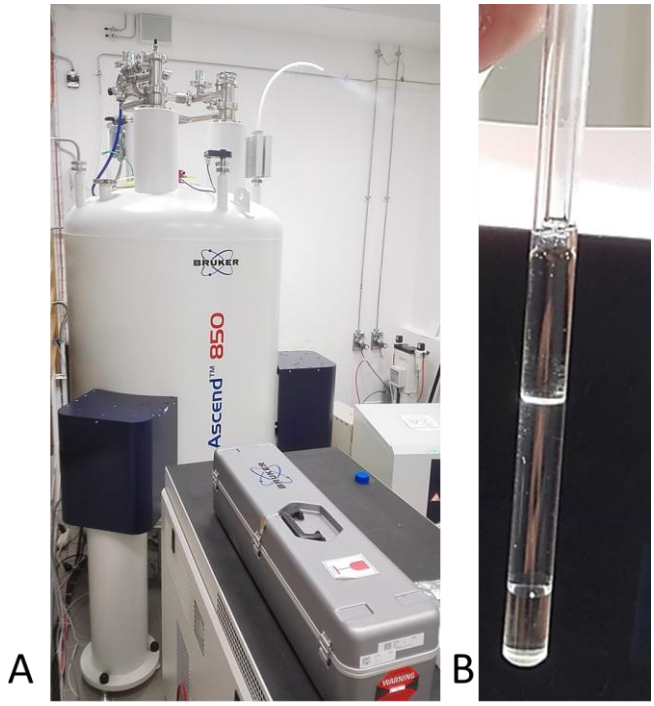

Figure 12. A. Institute of Organic Chemistry and Biochemistry of the CAS, Prague. Magnet: Bruker superconducting 850 MHz, Ascend™, console: Avance III HD 850. **B.** The sample of DmxA enzyme in the glass tube. *Pictures prepared by K. Doleželova.*

2.5.3. Cryo-electron microscopy

At the period of the development of various phasing methods applied to X-ray crystallography to uncover the lost phase information in electron diffraction, a new method of reconstruction of 3D objects from a set of images obtained from electron microscope was introduced by DeRosier and Klug in 1968 (De Rosier & Klug, 1968). The 3D reconstruction method is based on the extracted information from Fourier transform from images that contain both: phase and amplitude information presented in electron microscopy (EM) by digitized image processing. They compared the 3D reconstruction method with the structure analysis method in X-ray diffraction and concluded, that the difference is in a phase (Zou *et al.*, 2011). The phase together with the amplitudes of Fourier components are preserved and by the help of focusing of the diffracted electron beam into an image allows the reconstruction of 3D structure, while the phase problem exists during the recording of the X-ray diffraction data (De Rosier &

Klug, 1968). The method brought with it the opportunity to obtain hundreds of macromolecular structures.

Later, Richard Henderson discovered the structure of bacteriorhodopsin (Henderson *et al.*, 1990) (PDB ID:1BRD), the first protein structure solved with the new method at atomic resolution, which was implemented with the help of flash- freezing (vitrification) technology by Jacques Dubochet (J. & A.W., 1981), allowing to molecules to retain their natural shape during imaging. The molecular structure obtaining was thankfully to the algorithms and image-processing software, which was developed by Joachim Frank (Frank, 1975, Penczek *et al.*, 1994, Nogales, 2015).

The structural information from transmission electron microscopy can be obtained from the level of cells to the macromolecules of the biological specimens. The resolution of the obtained data from the EM depends on the sample, the method it was prepared, technical specifications of the devices used for the experiment, parameters of the imaging (Thompson *et al.*, 2016).

Currently there are several ways of obtaining 3D structural information from the collected 2D EM data:

a) electron tomography (with or without subtomogram averaging) (Briggs, 2013, Lucic *et al.*, 2005) – for "unique" assemblies (organelles and cells),

b) single particle analysis (Cheng *et al.*, 2015) – for purified, "homogenous" protein complexes (Ex. ribosome (Bai *et al.*, 2013)),

c) helical reconstruction (Egelman, 2015) – for protein assemblies with helical symmetry,

d) ED crystallography (Schenk *et al.*, 2013, Arheit *et al.*, 2013) – from 2D crystals of proteins, significantly smaller than 150 kDa that form ordered 2D arrays (such as bacteriorhodopsin) or the so-called Micro-ED (micro-crystal electron diffraction) from sufficiently thin 3D crystals of proteins or small molecules.

The structure determination begins from the appliance of the purified protein on the grid. Then the sample is plunged in liquid ethane (around −180 °C), flash-freezing, creating a thin layer of vitreous ice that ideally contains identical copies of the protein particles in different orientations. The sample is transferred to the electron microscope for imaging and data collection. The microscope column is kept under high vacuum, because of the strong interaction

of electrons with matter, the sample must be imaged in a solid state. The obtained 2D electron micrographs of protein particles, embedded in vitreous ice, are then selected by hand or by automated algorithms, and aligned and averaged (Milne *et al.*, 2013, Carroni & Saibil, 2016).

The method includes the following steps (Fig.13):

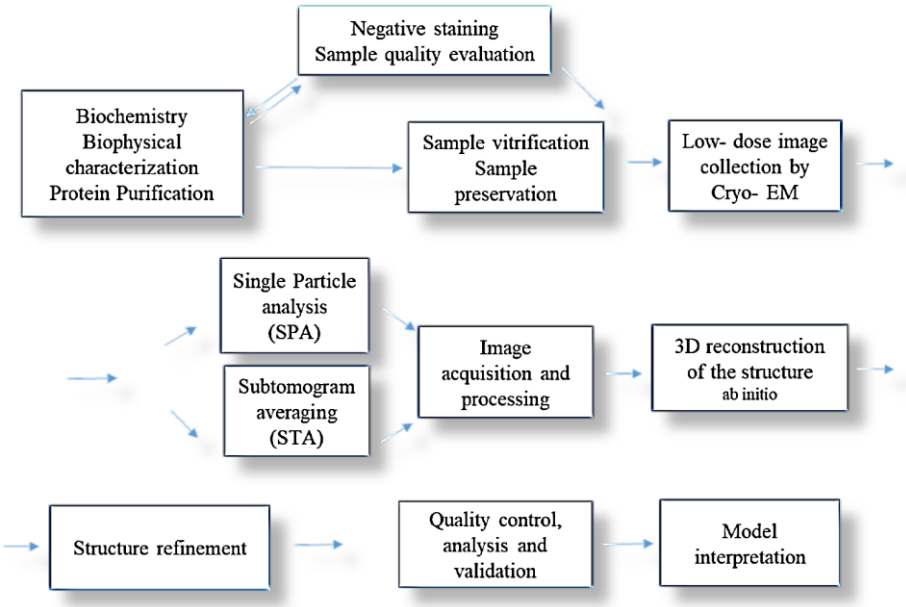

Figure 13. Steps of 3D structure determination by cryo-electron microscopy. Transformed from references: (Thompson *et al.*, 2016, Dubochet *et al.*, 1988, Grassucci *et al.*, 2007, Cheng *et al.*, 2015, Murata & Wolf, 2018)

Cryo-electron microscopy single particle analysis (cryo-EM SPA) is suitable for investigation: dynamic biological process (investigate a broad spectrum of drug-target interactions and dynamic conformational states), protein structures that are difficult to solve using other methods, protein complexes, large virus assemblies and aggregates.

In some cases, it is difficult to obtain 3D crystals large enough for X-ray diffraction of macromolecular structures however such 2D crystals are suitable for single-particle reconstruction EM and ED (Zou *et al.*, 2011).

2.6. SAXS

Small-angle X-ray scattering (SAXS) is a biophysical method for studying the structural transitions of biological macromolecules in solution, their overall

shape and the structural properties of materials at the nanoscale and for the quantitative analysis of flexible systems, including proteins, which are intrinsically disordered. The proteins and complexes in broad molecular size range (from few kDa up to GDa) can be used for the study and under different experimental conditions varying from the extreme to the nearly native (Kikhney & Svergun, 2015, Blanchet & Svergun, 2013).

The combination of results obtained from SAXS with the data of the other structural information provides an insight into conformation of a protein in solution (Tsutakawa et al., 2007).

It is needed from 10 to 100 µl of sample per measurement (total of 1–2 mg of purified protein including an obligatory measurement of concentration series (e.g. 1, 2, 5 and 10 mg/ml). The SAXS experiment includes following steps: a quartz capillary is filled in with the sample solution of particles and it is illuminated by a collimated monochromatic X-ray beam. This results the recorded intensity of the scattered X-rays by an X-ray detector. The intensity of the pure solvent is collected as well and then it is used for the substraction from the sample. The signal from the particles of interest is only left. The outcoming pattern of the scattering is related to the overall shape and size of the studied particles (Kikhney & Svergun, 2015).

2.7.Cross-linked enzyme crystals and their activity

The development of highly active and stable biocatalysts is often a key parameter in the technological process (Burton *et al.*, 2002). The development of the stabilized enzymes without diminishing and diluting their activity plays a highly important issue for their use as industrial catalysts (Govardhan, 1999).

There are few methods to immobilize the enzymes:

a) Cross-linking of the protein using a bifunctional reagent (such as glutaraldehyde).
b) Enzyme adsorption onto an inert carrier.
c) Immobilization in a polymerized gel (synthetic and non-synthetic).
d) A reactive insoluble support is used to which the enzyme is bound covalently (Norouzian, 2003).

Enzymes are well-known tools for achieving industrially important chemical reaction in region-, strereo- and chemoselective ways (Jegan Roy &

Emilia Abraham, 2004). There is a need of using enzymes with the extended stability under harsh conditions: enchanced pH stability, in an organic solvent environment, at elevated (or low) temperatures, and improved activity for various industrial applications (Jegan Roy & Emilia Abraham, 2004).

Cross-linking crystals (CLECs) is a technology that corresponds to these requirements. CLECs are extremely stable, not only with respect to temperature, but also to other inactivating agents like organic solvents. CLEC are also highly resistant to proteolysis. The specific activity for CLEC is significantly higher in comparison to immobilized enzymes, however, activity can be severely restricted by the substrate size exclusion and low molecular flexibility of the enzyme (Illanes, 1999, Noritomi *et al.*, 1998, Khalaf *et al.*, 1996, Govardhan, 1999).

The first data have shown that crystals can exhibit catalytic activity without disruption of the lattice reported on the histidyl residues in myoglobin and the tyrosyl residues in haemoglobin. It was shown that the enzymes ribonuclease-S and chymotrypsin, carboxypeptidase-A are catalytically active in the crystalline state 1960th (Quiocho & Richards, 1964). The first CLECs were obtained for thermolysin (Persichetti *et al.*, 1995). Concerning the HLDs, none of CLEC was previously mentioned, except of obtained immobilized enzyme LinB *from Sphingobium japonicum* UT26 (Stepankova, Bidmanova, *et al.*, 2013). Other examples of cross-linked enzymes are: lipase from *Burkholderia cepacia* (Rajan & Emilia Abraham, 2008), laccase from *Trametes versicolor* as a biosensor for the determination of phenols (Roy *et al.*, 2005) and other enzymes.

Hight specific activity is one of the most essential characteristics of a synthetically useful biocatalyst. The enzymatic activity in the crystalline state is controlled by the size of the crystal, size of the substrate and conformation of the enzyme (Margolin, 1996). The process requires extensive optimization in order to ensure high activity and stability.

New industrial processes can be based on the fact that a new interesting enzyme is available, or the desired products are identified, after this the selection of a biocatalyst that permits conversion of available reactants is following. Such an enzyme might be available commercially, or it might have been described in the literature (Arnold, 2001). Nowadays the interest of CLECs technology is raising broadening the scope of biocatalysis.

3. Materials and methods

3.1. Crystallization

For the first crystallization trials DpcA from *Psychrobacter cryohalolentis* K5 and DmxA from *Marinobacter* ELB17 were purified and dialyzed over 50 mM phosphate buffer (pH 7.5). The enzymes were prepared at concentrations: DpcA of 12.71 mg ml^{-1}, DmxA of 12.71 mg ml^{-1}. Searching for crystallization conditions that are suitable for the crystal growth of DpcA and DmxA proteins was started with an application of the crystallization conditions previously reported for DhaA (Stsiapanava, et al., 2011). Crystallization drops were performed manually by sitting-drop vapor diffusion (Ducruix & Giegé, 1999) with the different volume protein-to-precipitant ratios (1/1, 1/2, 1/3, 2/1, respectfully) with the commercially available crystallization screens: Crystal Screen and Crystal Screen 2 (Hampton Research, Aliso Viejo, California, USA), several screens from JBScreen Classic (Jena Bioscience, Jena, Germany). However, the experiments resulted few crystals, that were confirmed as crystals of salt by diffraction on the laboratory diffractometer Nonius FR 591 equipped with a detector MarReserch 345 at the temperature 120 K (Prague, IOCB).

Subsequenty purified enzymes DpcA and DmxA in 50 mM Tris–HCl buffer pH 7.5 at a concentration of 10 mg ml^{-1} were used for futher crystallization screening by sitting-drop vapor-diffusion technique (Ducruix & Giegé, 1999) in CombiClover crystallization plates (Emerald BioSystems, Bainbridge Island, USA Combi Clover crystallization plate, Emerald Biostructures) with the screens as mentioned above and additionally with MD-2 (Molecular Dimensions Ltd, Suffolk, England) and counterdiffusion screening kits (Granada Crystallization Box (GCB), 24 conditions and PEG screening; Triana Science and Technology, Granada, Spain). Drops contained the same ratios as in the previous experiments. Experiments were set up at 277, 285, 293 and 298 K.

The crystallization conditions of DpcA and DmxA were further optimized using varying compound concentrations, temperature and pH. The the Additive Screen kit (Hampton Research, Aliso Viejo, California, USA) with 96 unique reagents was used additionally for both enzymes. Crystallization was carried out by sitting-drop and hanging- drop vapor diffusions, counter diffusion (using capillaries, GCB, 3D capillaries with low-melting agarose, crystallization in gel TMOS).

The solutions for optimization and the final composition of the optimised crystallization conditions were prepared using chemicals from Sigma-Aldrich (USA).

The rystallized protein identity was checked by 12% SDS–PAGE staining of a gel by Coomassie Blue wColorBurst Electrophoresis Marker MW 8–220 kDa (Sigma–Aldrich/Merck, USA).

Co-crystallizaion with the ligand. For the co-crystallization of DpcA with 10% (v/v) 1-bromohexane in a solution of 0.1 M glycine buffer pH 8.6 was used at optimized conditions by the sitting drop vapor diffusion with different ratios of protein/precipitant solution/substrate/additive.

Soaking with the ligands. 50 µl of 1,2-dichloroethane was added into the reservoir well containing 800 µl of the precipitating solution with the crystals of DpcA or DmxA.

3.2. Diffraction data collection

The obtained crystals were first tested by the in-house diffractometer Nonius FR 591 equipped with a detector MarReserch 345 at the temperature 120 K (Prague, IOCB). The compete diffraction data sets of DpcA were collected on MX14.2 beam line operated by the Helmholtz-Zentrum Berlin (HZB) at the BESSY II electron storage ring (Berlin-Adlershof, Germany) (Mueller *et al.*, 2015) at the wavelength 0.978 and 100 K using a Rayonics MX-225. The crystal of DmxA were mounted into nylon loops (Hampton Research, USA) and MicroLoops (MiTeGen; Jena Bioscience GmbH, Germany) and directly flash-cooled in a liquid-nitrogen stream without additional cryoprotection. The diffraction data were collected at beamline ID29 at the ESRF, European Synchrotron Radiation Facility, France (de Sanctis *et al.*, 2012) at the wavelength of 0.972 Å and 100 K using a Pilatus 6M-F detector.

3.3. Enzyme structure determination, model validation and deposition

Data diffraction for DpcA crystal were indexed, integrated and scaled by HKL-3000 (Minor *et al.*, 2006). The diffraction data for DmxA were automatically processed by XDS program package (Kabsch, 1993, 2010a, 2010b). The structure of DpcA and DmxA were solved by molecular

replacement with MOLREP (Vagin & Teplyakov, 2010) from the CCP4 software suite (Winn *et al.*, 2011), where the coordinates of homologue haloalkane dehalogenases were used. For the structure refinement REFMAC5, version 5.8.0155 (Murshudov *et al.*, 2011) software as part of the CCP4 suite (Collaborative Computational Project, Number 4, 1994) was used. Additionally several cycles of refinement of DpcA were carried using SHELXL (Sheldrick & Schneider, 1997) and Phenix.refine (Adams *et al.*, 2010). The models of the enzyme were also optimised by PDB_REDO web server (Joosten *et al.*, 2014). The manual structure building steps were performed in Coot (Emsley *et al.*, 2010); other data and coordinate manipulations were performed in the CCP4 program package (Collaborative Computational Project, Number 4, 1994). The structure of DmxA was refined by restrained isotropic and TLS refinement (Painter & Merritt, 2006, Winn *et al.*, 2001, Winn *et al.*, 2003) using 2 TLS groups and local NCS refinement with 1 NCS group carried out by REFMAC5 (Murshudov et al., 2011).

The quality of the structure models with respect to the experimental data was controlled by assessed using the the internal tools of Coot and the MolProbity service (Chen *et al.*, 2010), (Hintze *et al.*, 2016).

The structure validation and analyses were performed using MOLPROBITY servic, SFCHECK (Vaguine *et al.*, 1999), wwPDB Validation Server (Berman *et al.*, 2003) , and PISA server (Krissinel & Henrick, 2007). Atomic coordinates and experimental structure factors of DpcA and DmxA were deposited in the RCSB Protein Data Bank under PDB codes 6F9O and 5MXP, respectfully.

Active site volumes were calculated with CASTp (Dundas et al., 2006) using a 1.4 Å radius probe. All figures were prepared using PyMOL 1.5 (Schrodinger, 2015). Structure similarity search of the PDB was performed with the program Dali (Holm & Rosenstrom, 2010).

4. Results and discussion

4.1 Crystallographic analysis of new psychrophilic haloalkane dehalogenases: DpcA from *Psychrobacter cryohalolentis* K5 and DmxA from *Marinobacter* sp. ELB17

This chapter is based on paper:

K. Tratsiak, O. Degtjarik, I. Drienovska, L. Chrast, P. Rezacova, M. Kuty, R. Chaloupkova, J. Damborsky and I. Kuta Smatanova. Crystallographic analysis of new psychrophilic haloalkane dehalogenases: DpcA from Psychrobacter cryohalolentis K5 and DmxA from Marinobacter sp. ELB17. Acta Cryst. (2013). F69, 683-688.

ABSTRACT

Haloalkane dehalogenases are hydrolytic enzymes with a broad range of potential practical applications such as biodegradation, biosensing, biocatalysis and cellular imaging. Two newly isolated psychrophilic haloalkane dehalogenases exhibiting interesting catalytic properties, DpcA from *Psychrobacter cryohalolentis* K5 and DmxA from *Marinobacter sp.* ELB17, were purified and used for crystallization experiments. After the optimization of crystallization conditions, crystals of diffraction quality were obtained. Diffraction data sets were collected for native enzymes and complexes with selected ligands such as 1-bromohexane and 1,2-dichloroethane to the resolution ranging from 1.05 Å to 2.49 Å.

Publication omitted for copyright reasons

4.2 Crystal structure of cold-adapted haloalkane dehalogenase DpcA from *Psychrobacter cryohalolentis* K5

This chapter is based on paper:

K. Tratsiak, T. Prudnikova, I. Drienovska, J. Damborsky, J. Brynda, P.Pachl, M. Kuty, R. Chaloupkova, P. Rezacova and I. Kuta Smatanova. Crystal structure of cold-adapted haloalkane dehalogenase DpcA from *Psychrobacter cryohalolentis* K5. Accepted for publication in Acta Cryst. F (2019). (IF $_{2017}$ = 0.989)

ABSTRACT Haloalkane dehalogenases (HDLs) convert halogenated aliphatic pollutants to less toxic compounds by a hydrolytic mechanism. Owing to their broad substrate specificity and high enantioselectivity, haloalkane dehalogenases can function as biosensors to detect toxic compounds in the environment or can be used for the production of optically pure compounds. Here we present structural analysis of the haloalkane dehalogenase DpcA isolated from the psychrophilic bacterium *Psychrobacter cryohalolentis* K5 at atomic resolution of 1.05 Å. This enzyme exhibits a low temperature optimum, making it attractive for environmental applications such as biosensing at the subsurface environment where the temperature typically does not exceed 25 ºC. Our structure revealed that among structural homologues of HLD-I subfamily, DpcA possesses the shortest access tunnels and one of the most widely open main tunnels. Comparative analysis revealed major differences in the region of the α4 helix of the cap domain, which is one of the key determinants of tunnels' anatomy. The crystal structure of DpcA contributes towards better understanding of structure-property relationships of cold-adapted enzymes.

Keywords: haloalkane dehalogenase; α/β-hydrolase; X-ray diffraction; psychrophile; structural analysis

Publication omitted for copyright reasons

4.3 Deciphering of paradoxically high thermostability of dehalogenase from psychrophilic bacterium

This chapter is based on paper:

L.Chrast, **K. Tratsiak,** L. Daniel, T. Prudnikova, J. Brezovsky, I. Kuta Smatanova, R. Chaloupkova, J. Damborsky. Deciphering of paradoxically high thermostability of dehalogenase from psychrophilic bacteriumStructural basis of paradoxically high thermostability of dehalogenase from psychrophilic bacterium. Prepared for publication in FEBS Journal (2019). (IF 2017 = 4.530).

ABSTRACT

Haloalkane dehalogenases are enzymes with a broad application potential in biocatalysis, bioremediation, biosensing and cell imaging. Novel haloalkane dehalogenase DmxA originating from the psychrophilic bacterium *Marinobacter* sp. ELB17 surprisingly possesses the highest thermal stability (melting temperature T_m = 65.9 C) of all biochemically characterized wild type dehalogenases. The enzyme was successfully expressed and the crystal structure was solved with the resolution of 1.45 Å. DmxA structure contains several features distinct from known members of haloalkane dehalogenase family: (i) a unique composition of catalytic residues, (ii) a dimeric state mediated by a disulphide bridge and (iii) narrow tunnels connecting the enzyme active site with the surrounding solvent. Crucial role of narrow tunnels in the paradoxically high stability of DmxA enzyme was confirmed by *in silico* analysis and mutagenesis experiments.

Publication omitted for copyright reasons

5. Conclusions

The new biologically attractive haloalkane dehalogenases DpcA from *Psychrobacter cryohalolentis* K5 and DmxA from *Marinobacer* sp. ELB17 were characterized crystallographically. The structures of these enzymes, which are different in their temperature optima (DpcA is a cold-adapted enzyme, and DmxA is highly thermostable), allow us to gain insights into their structural properties of specificity and stability.

Haloalkane dehalogenases are bacterial enzymes that realize cleavage of the carbon-halogen bond of halogenated aliphatic compounds by hydrolysis (Janssen et al. 1994). This catalytic reaction makes DpcA and DmxA attractive targets for the enzymes' catalytic efficiency and broadens their substrate specificity towards important environmental pollutants (Copley 1998, Otyepka & Damborský, 2002).

This work combines standard techniques used to study the structures of DpcA and DmxA enzymes. X-ray crystallographic analysis is employed to study structures, for the further construction of biotechnologically effective enzymes and their combinations. This is the precious source of the structural information for the understanding of the functions of the enzymes. The project was subdivided into subprojects: the crystallization of the enzymes, crystallographic experiments on the enzymes' crystals, the analysis of the structures of DpcA and their comparison with homologous enzymes, the analysis of the DmxA enzyme structure.

The first part of the thesis was focused on **crystallization, co-crystallization with ligands, and the collection and processing of the diffraction data of DpcA and DmxA crystals.** The crystallization of proteins is the essential step of the research in order to obtain high quality diffraction data to study a structure of a protein by X-ray crystallography and to solve the protein structure.

Searching for crystallization conditions suitable for the crystal growth of DpcA and DmxA in 0.5 mM phosphate buffer, pH 7.5 was started with the application of the previously reported crystallization conditions for DhaA wild type (Stsiapanava *et al.*, 2011), however it was not successful as well as further crystallization screening with the commercially available screens. Exchanging the enzyme purification buffer, screening pH, temperature, precipitants and salt varying were the factors that led to crystal formation following its optimization. Additionally, screening with 96 additives was used to improve the result of

crystallization. The crystals suitable for the X-ray diffraction experiment were obtained in a newly discovered precipitant composition. Almost all crystallization techniques were successfully used to grow crystals of the selected enzymes (Tratsiak *et.al.*, 2013).

Diffraction data for DpcA were collected at the beamline of 14.2 operated by the Helmholtz-Zentrum Berlin (HZB) (Germany) at the BESSY II electron storage ring (Mueller *et al.*, 2012) to the 1.05 Å resolution.

The second part of this thesis described an important contribution to the understanding of the molecular basis of cold adaptation in proteins. We presented the **X-ray structural analysis of DpcA from *Psychrobacter cryohalolentis* K5 as a base for understanding its enzymatic properties that can guide modification of this cold-adapted enzyme for various biotechnological applications.** DpcA is HLD that was biochemically characterized as that of the first of psychrophilic origin (Drienovska *et al.*, 2012). DpcA exhibits a low temperature optimum, making it attractive for environmental applications such as biosensing at the subsurface environment where the temperature typically does not exceed 25 ºC.

We determined the structure of DpcA by a molecular replacement with the use of a search model of DhlA from *Xanthobacter autotrophicus* (PDB ID 1B6G) (Ridder *et al.*, 1999). Crystallographic analysis of the DpcA crystal revealed the structural features of cold-adaptation of this enzyme. We identified DpcA as a monomer, representing the biologically active unit. We were able to model all protein residues (1-309) onto the electron density map within the C-terminal histidine tag (residues 304-309). We observed the highly conserved main domain (residues 1-154, 221-309) that consists of an eight-stranded β-sheet, where β2 is lying in an antiparallel orientation with respect to the direction of the β-sheet. The β-sheets are surrounded by six α-helices. The cap domain (residues 155-220), is inserted between β-strand β6 and α-helix α8 of the main domain and is represented by four α-helices and connective loops. We identified and interpreted the non-protein electron density in the active site as water molecules and a chloride anion.

We carried out structural and bioinformatic assessment of the differences between DpcA and other homologues of the HLD subfamily and identified substantial differences in the architecture of the active site and access tunnels of DpcA. We discovered that the studied structure possesses the shortest access

tunnels (the length is 7.9 Å) and one of the most widely open main tunnels (its radius is 1.7 Å), which explains the preference of DpcA for the smaller substrates. We denoted the DpcA active site cavity volume is 435 $Å^3$, and is comparable to that of the DppA from *Plesiocystis pacifica* (Hesseler *et al.,* 2011).

Furthermore, the comparative analysis of relative structures revealed the major differences in the region of the α4 helix of the cap domain (residues 162-173, including halide-stabilizing residue Trp164). The α4 helix covers the enzyme active site. We indicated that the region of residues that includes residues 155-159 has an increased flexibility and forms a loop with the high dynamic disorder. The α4 helix is important in determining the tunnel's properties and the flexibility in this region thus directly affects that of the active site. This corresponds to the previous observation that adaptive mutations favouring the active-site flexibility are located outside the catalytic site (Feller and Garday, 2003).

The third part of the thesis is dedicated to the **diffraction, data collection and processing of DmxA from *Marinobacer* sp. ELB17 and its structural analysis in order to explain its extraordinary stability.** Based on the obtained diffraction data that were collected at beamline ID29 at the ESRF, European Synchrotron Radiation Facility, France (de Sanctis *et al.*, 2012) up to 1.45 Å resolution and the solved crystal structure, we elucidated the paradoxically high thermostability of this enzyme from psychrophilic bacterium with X-ray crystallographic study and site-directed mutagenesis. We solved the structure by molecular replacement using the structure of DhaA from *Rhodococcus rhodochrous,* PDB ID 4E46 (Stepankova, Khabiri, *et al.,* 2013) as a search model.

We elucidated that the DmxA structure is composed of two domains. The core domain comprised of typical α/β-hydrolase structure consisting of an eight-stranded β-sheet (with β2 lying in an antiparallel orientation) is covered by six α-helices. The cap domain, inserted between β-strand β6 and α-helix α8, consists of five short α-helices connected by six loop insertions. The active site is situated between domains. We identified and interpreted the electron density in a vicinity of the active site as a water molecule and acetate ion with an occupancy of 1.0 in both molecules present in the asymmetric unit. The water molecule (HOH338) is situated in a canonical halide-binding pocket of the enzyme.

Our findings of the crystallographic analysis of DmxA revealed three unique features of this enzyme that can contribute to its thermostability.

(1) A unique composition of catalytic pentad. We detected that the enzyme structure possesses a unique halide-stabilizing residue, glutamine 40, which is a part of a composition of catalytic machinery (instead of the conventional asparagine typical of other members of the subfamily). Glutamine 40 is not involved in contact with HOH338, its Nε2 atom is situated 5.42 Å far from the water molecule, and the side chain of the residue faces away from the halide-binding site making polar interactions with the side chain hydroxyl of tyrosine 68 and main chain carbonyl of leucine 203. This led us to suppose that glutamine (Q) 40 residue cannot be involved in the stabilization of the substrate nor releasing a halide anion in the DmxA active site. Thanks to this, further construction of Q40N supported these data.

(2) We recognized that the enzyme dimerization is due to a disulfide bridge. We showed that the enzyme forms a dimeric structure through a disulfide bridge by cysteine 294 residues from each monomer of C-terminal helices (residues 281-296). DmxA represents the first example of oligomerization via disulfide bridge formation reported in the HLD enzyme family so far.

(3) The 3D structure of DmxA allowed us to find the unusually narrow tunnels connecting the buried enzyme active site with the surrounding solvent. We succeeded in identifying the important residues in the bottleneck of the main tunnel that influence the size of access tunnels, which was shown to be a crucial feature determining the high stability of DmxA. It was revealed that the DmxA variants possess a shift in substrate specificity and dramatically modified thermostability, thereby providing insight into the role of access tunnels for the paradoxical robustness of this enzyme.

The results published in this thesis extend our knowledge about the structures of haloalkane dehalogenases and the structural features on adaptation to cold and the factors influencing the thermostability.

6. References

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Crystallographica Section D* **66**, 213-221.

Adekoya, O. A., Helland, R., Willassen, N.-P. & Sylte, I. (2006). *Proteins: Structure, Function, and Bioinformatics* **62**, 435-449.

Akers, K. S., Sinks, G. D. & Schultz, T. W. (1999). *Environmental toxicology and pharmacology* **7**, 33-39.

Alberts, B. (1998). *Cell* **92**, 291-294.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. & Walter, P. (2002). *Molecular Biology of the Cell, Fourth Edition*. Garland Science.

Alberts, B., Johnson, A., Lewis, J., Walter, P., Raff, M. & Roberts, K. (2002). *Molecular Biology of the Cell 4th Edition: International Student Edition*. Routledge.

Arheit, M., Castano-Diez, D., Thierry, R., Gipson, B. R., Zeng, X. & Stahlberg, H. (2013). *Methods Mol Biol* **955**, 171-194.

Arnold, F. H. (2001). *Nature* **409**, 253.

Atkins, P. W. & De Paula, J. (2006). *Atkins' Physical Chemistry*. W.H. Freeman.

B L Blaney, a. & Ewing, G. E. (1976). *Annual Review of Physical Chemistry* **27**, 553-584.

Bai, X.-c., Fernandez, I. S., McMullan, G. & Scheres, S. H. W. (2013). *eLife* **2**, e00461.

Bakermans, C., Ayala-del-Rio, H. L., Ponder, M. A., Vishnivetskaya, T., Gilichinsky, D., Thomashow, M. F. & Tiedje, J. M. (2006). *Int J Syst Evol Microbiol* **56**, 1285-1291.

Bax, A. & Grzesiek, S. (1993). *Accounts of Chemical Research* **26**, 131-138.

Belkin, S. (1992). *Biodegradation* **3**, 299-313.

Berg JM, T. J., Stryer L. (2002).

Berg, J. M., Tymoczko, J. L., Gatto, G. J. & Stryer, L. (2015). *Biochemistry*, Eighth ed. New York: W.H. Freeman & Company, a Macmillan Education Imprint.

Bergfors, T. M. (2009). *Protein crystallization*. La Jolla, Calif.: International University Line.

Bergmann, J. G. & Sanik, J. (1957). *Analytical Chemistry* **29**, 241-243.

Berman, H., Henrick, K. & Nakamura, H. (2003). *Nature Structural Biology* **10**, 980.

Betz, S. F. (1993). *Protein Science* **2**, 1551-1558.

Bhatt, P., Kumar, M. S., Mudliar, S. & Chakrabarti, T. (2007). *Biodegradation of Chlorinated Compounds—A Review*.

Biertümpfel, C., Basquin, J., Suck, D. & Sauter, C. (2002). *Acta Crystallographica Section D* **58**, 1657-1659.

Bláha, L., Damborský, J. & Němec, M. (1998). *Chemosphere* **36**, 1345-1365.

Bogdanovic, X., Hesseler, M., Palm, G. J., Bornscheuer, U. T. & Hinrichs, W. (2010). *Acta crystallographica. Section F, Structural biology and crystallization communications* **66**, 828-830.

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

Boysen, G., Pachkowski, B. F., Nakamura, J. & Swenberg, J. A. (2009). *Mutation research* **678**, 76-94.

Bragg, W. H. & Bragg, W. L. (1913). *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **88**, 428-438.

Briggs, G. E. & Haldane, J. B. (1925). *The Biochemical journal* **19**, 338-339.

Briggs, J. A. G. (2013). *Current Opinion in Structural Biology* **23**, 261-267.

Brock, T. D. (2012). *Thermophilic Microorganisms and Life at High Temperatures*. Springer New York.

Brown, G., Halbach, K., Harris, J. & Winick, H. (1983). *Nuclear Instruments and Methods in Physics Research* **208**, 65-77.

Brunger, A. T. (1992). *Nature* **355**, 472-475.

Burton, S. G., Cowan, D. A. & Woodley, J. M. (2002). *Nature Biotechnology* **20**, 37.

Campbell, D. W., Müller, C. & Reardon, K. F. (2006). *Biotechnology letters* **28**, 883-887.

Carlucci, L., Zhou, E., Malashkevich, V. N., Almo, S. C. & Mundorff, E. C. (2016). *Protein science : a publication of the Protein Society* **25**, 877-886.

Carroni, M. & Saibil, H. R. (2016). *Methods (San Diego, Calif.)* **95**, 78-85.

Chaloupkova, R., Prudnikova, T., Rezacova, P., Prokop, Z., Koudelakova, T., Daniel, L., Brezovsky, J., Ikeda-Ohtsubo, W., Sato, Y., Kuty, M., Nagata, Y., Kuta Smatanova, I. & Damborsky, J. (2014). *Acta crystallographica. Section D, Biological crystallography* **70**, 1884-1897.

Chayen, N. (1997). *Journal of Applied Crystallography* **30**, 198-202.

Chayen, N. (1998). *Acta Crystallographica Section D* **54**, 8-15.

Chayen, N. E. & Line, I. U. (2007). *Protein Crystallization Strategies for Structural Genomics*. International University Line.

Chayen, N. E. & Saridakis, E. (2008). *Nat Methods* **5**, 147-153.

Chayen, N. E., Saridakis, E., El-Bahar, R. & Nemirovsky, Y. (2001). *Journal of Molecular Biology* **312**, 591-595.

Chen, V. B., Arendall, W. B., III, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Crystallographica Section D* **66**, 12-21.

Cheng, Y., Grigorieff, N., Penczek, Pawel A. & Walz, T. (2015). *Cell* **161**, 438-449.

Chernov, A. A. (2003). *Journal of Structural Biology* **142**, 3-21.

Chovancová, E., Kosinski, J., Bujnicki, J. M. & Damborský, J. (2007). *Proteins: Structure, Function, and Bioinformatics* **67**, 305-316.

Cooper, G. M. (2000). *The Cell: A Molecular Approach*. ASM Press.

Copeland, R. A. (2000). *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis*. Wiley.

Copley, S. D. (1998). *Current Opinion in Chemical Biology* **2**, 613-617.

Cowan, D. A. (1997). *Comparative biochemistry and physiology. Part A, Physiology* **118**, 429-438.

Cowtan, K. (2001). *e LS*.

Crick, F. & Magdoff, B. S. (1956). *Acta Crystallographica* **9**, 901-908.

D'Amico, S., Claverie, P., Collins, T., Georlette, D., Gratia, E., Hoyoux, A., Meuwis, M.-A., Feller, G. & Gerday, C. (2002). *Philosophical Transactions of the Royal Society B: Biological Sciences* **357**, 917-925.

D'Arcy, A., Elmore, C., Stihle, M. & Johnston, J. E. (1996). *Journal of Crystal Growth* **168**, 175-180.

Davidson, J. R. & William, S. (1958). Method of operating a refrigeration system using a chlorine containing halo-alkane as a refrigerant: Google Patents.

de Graaf, R. A. (2013). *In Vivo NMR Spectroscopy: Principles and Techniques*. Wiley.

De Rosier, D. J. & Klug, A. (1968). *Nature* **217**, 130-134.

de Sanctis, D., Beteva, A., Caserotto, H., Dobias, F., Gabadinho, J., Giraud, T., Gobbo, A., Guijarro, M., Lentini, M., Lavault, B., et al. (2012). ID29: a high-intensity highly automated ESRF beamline for macromolecular crystallography experiments exploiting anomalous scattering. *Journal of Synchrotron Radiation* **19**: 455-461.

Deckert, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L., Graham, D. E., Overbeek, R., Snead, M. A., Keller, M., Aujay, M., Huber, R., Feldman, R. A., Short, J. M., Olsen, G. J. & Swanson, R. V. (1998). *Nature* **392**, 353.

Derewenda, Z. S. & Vekilov, P. G. (2006). Entropy and surface engineering in protein crystallization, Vol. 62, pp. 116-124. Wiley-Blackwell.

Dhouib, K., Khan Malek, C., Pfleging, W., Gauthier-Manuel, B., Duffait, R., Thuillier, G., Ferrigno, R., Jacquamet, L., Ohana, J., Ferrer, J.-L., Théobald-Dietrich, A., Giegé, R., Lorber, B. & Sauter, C. (2009). *Lab on a Chip* **9**, 1412-1421.

Diener, R., Dreyling-Eschweiler, J., Ehrlichmann, H., Gregor, I.-M., Kötz, U., Krämer, U., Meyners, N., Potylitsina-Kube, N., Schütz, A. & Schütze, P. (2019). *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **922**, 265-286.

Drienovska, I., Chovancova, E., Koudelakova, T., Damborsky, J. & Chaloupkova, R. (2012). *Appl Environ Microb* **78**, 4995-4998.

Dubochet, J., Adrian, M., Chang, J. J., Homo, J. C., Lepault, J., McDowall, A. W. & Schultz, P. (1988). *Q Rev Biophys* **21**, 129-228.

Ducruix, A. & Giegé, R. (1999). *Crystallization of Nucleic Acids and Proteins: A Practical Approach*. Oxford University Press.

Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y. & Liang, J. (2006). *Nucleic Acids Research* **34**, W116-W118.

Durbin, S. D. & Feher, G. (1996). *Annual Review of Physical Chemistry* **47**, 171.

Egelman, E. H. (2015). *Archives of Biochemistry and Biophysics* **581**, 54-58.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Crystallographica Section D* **66**, 486-501.

Ernst, R. R., Bodenhausen, G. & Wokaun, A. (1987). *Principles of nuclear magnetic resonance in one and two dimensions*. Clarendon Press Oxford.

Facchiano, A. M., Colonna, G. & Ragone, R. (1998). *Protein Eng* **11**, 753-760.

Falcón Rodríguez, C., Aguilera Morales, S. & Falcón Rodríguez, F. (2000). *Journal of Crystal Growth* **208**, 651-656.

Feher, G. (1986). *Journal of Crystal Growth* **76**, 545-546.

Feller, G. & Gerday, C. (1997). *Cellular and molecular life sciences : CMLS* **53**, 830-841.

Feller, G. & Gerday, C. (2003). *Nature reviews. Microbiology* **1**, 200-208.

Fetzner, S. (1998). *Applied microbiology and biotechnology* **50**, 633-657.

Fortova, A., Sebestova, E., Stepankova, V., Koudelakova, T., Palkova, L., Damborsky, J. & Chaloupkova, R. (2013). *Biochimie* **95**, 2091-2096.

Foundation, N. (1999). *Nobel Lectures in Chemistry*. World Scientific.

Frank, J. (1975). *Ultramicroscopy* **1**, 159-162.

Franken, S. M., Rozeboom, H. J., Kalk, K. H. & Dijkstra, B. W. (1991). *The EMBO journal* **10**, 1297-1302.

Fried, S. D. & Boxer, S. G. (2017). *Annual review of biochemistry* **86**, 387-415.

Fung, H. K. H., Gadd, M. S., Drury, T. A., Cheung, S., Guss, J. M., Coleman, N. V. & Matthews, J. M. (2015). *Molecular microbiology* **97**, 439-453.

García-Ruiz, J. M. (2003). *Methods in Enzymology*, pp. 130-154: Academic Press.

Gehret, J. J., Gu, L., Geders, T. W., Brown, W. C., Gerwick, L., Gerwick, W. H., Sherman, D. H. & Smith, J. L. (2012a). *Protein science : a publication of the Protein Society* **21**, 239-248.

Gehret, J. J., Gu, L. C., Geders, T. W., Brown, W. C., Gerwick, L., Gerwick, W. H., Sherman, D. H. & Smith, J. L. (2012b). *Protein Sci* **21**, 239-248.

Georlette, D., Blaise, V., Collins, T., D'Amico, S., Gratia, E., Hoyoux, A., Marx, J.-C., Sonan, G., Feller, G. & Gerday, C. (2004). *FEMS Microbiology Reviews* **28**, 25-42.

Georlette, D., Damien, B., Blaise, V., Depiereux, E., Uversky, V. N., Gerday, C. & Feller, G. (2003). *Journal of Biological Chemistry* **278**, 37015-37023.

Gianese, G., Argos, P. & Pascarella, S. (2001). *Protein Engineering, Design and Selection* **14**, 141-148.

Gomes, J. & Steiner, W. (2004). *Food technology and Biotechnology* **42**, 223-235.

Grassucci, R. A., Taylor, D. J. & Frank, J. (2007). *Nature protocols* **2**, 3239-3246.

Green, D., Ingram, V. M. & Perutz, M. F. (1954). *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **225**, 287-307.

Gribble, G. W. (2003). *Chemosphere* **52**, 289-297.

Gunton, J. D., Shiryayev, A. & Pagan, D. L. (2007). *Protein Condensation: Kinetic Pathways to Crystallization and Disease*. Cambridge University Press.

Haki, G. D. & Rakshit, S. K. (2003). *Bioresource Technology* **89**, 17-34.

Hasan, K., Fortova, A., Koudelakova, T., Chaloupkova, R., Ishitsuka, M., Nagata, Y., Damborsky, J. & Prokop, Z. (2011). *Appl Environ Microbiol* **77**, 1881-1884.

Hawwa, R., Aikens, J., Turner, R. J., Santarsiero, B. D. & Mesecar, A. D. (2009). *Archives of biochemistry and biophysics* **488**, 109-120.

Hejátko, J. & Hakoshima, T. (2018). *Plant Structural Biology: Hormonal Regulations*. Springer International Publishing.

Hemminki, K. (1983). *Archives of toxicology* **52**, 249-285.
Henderson, R., Baldwin, J. M., Ceska, T. A., Zemlin, F., Beckmann, E. & Downing, K. H. (1990). *Journal of Molecular Biology* **213**, 899-929.
Hesseler, M., Bogdanovic, X., Hidalgo, A., Berenguer, J., Palm, G. J., Hinrichs, W. & Bornscheuer, U. T. (2011). *Applied microbiology and biotechnology* **91**, 1049-1060.
Hintze, B. J., Lewis, S. M., Richardson, J. S. & Richardson, D. C. (2016). *Proteins: Structure, Function, and Bioinformatics* **84**, 1177-1189.
Ho, B., Baryshnikova, A. & Brown, G. W. (2018). *Cell Systems* **6**, 192-205.e193.
Holloway, P., Trevors, J. T. & Lee*, H. (1998). *Journal of Microbiological Methods* **32**, 31-36.
Holmquist, M. (2000). *Current Protein and Peptide Science* **1**, 209-235.
Illanes, A. (1999). *Electronic Journal of Biotechnology* **2**, nd-nd.
Isidorov, V., Povarov, V. & Prilepsky, E. (1993). *J. Ecol. Chem* **1**, 19-25.
Israelachvili, J. N. (2010). *Intermolecular and Surface Forces*. Saint Louis, UNITED STATES: Elsevier Science & Technology.
J., D. & A.W., M. (1981). *Journal of Microscopy* **124**, 3-4.
Janssen, D. B. (2004). *Curr Opin Chem Biol* **8**, 150-159.
Janssen, D. B., Gerritse, J., Brackman, J., Kalk, C., Jager, D. & Witholt, B. (1988). *European journal of biochemistry* **171**, 67-72.
Janssen, D. B., Pries, F., van der Ploeg, J., Kazemier, B., Terpstra, P. & Witholt, B. (1989). *Journal of bacteriology* **171**, 6791-6799.
Janssen, D. B., Scheper, A., Dijkhuizen, L. & Witholt, B. (1985). *Applied and environmental microbiology* **49**, 673-677.
Jaskolski, M., Dauter, Z. & Wlodawer, A. (2014). *The FEBS Journal* **281**, 3985-4009.
Jeener, J., Meier, B., Bachmann, P. & Ernst, R. (1979). *The Journal of chemical physics* **71**, 4546-4553.
Jegan Roy, J. & Emilia Abraham, T. (2004). *Chemical Reviews* **104**, 3705-3722.
Jesenska, A., Monincova, M., Koudelakova, T., Hasan, K., Chaloupkova, R., Prokop, Z., Geerlof, A. & Damborsky, J. (2009). *Appl Environ Microbiol* **75**, 5157-5160.
Jesenska, A., Pavlova, M., Strouhal, M., Chaloupkova, R., Tesinska, I., Monincova, M., Prokop, Z., Bartos, M., Pavlik, I., Rychlik, I., Mobius, P., Nagata, Y. & Damborsky, J. (2005). *Appl Environ Microb* **71**, 6736-6745.
Joosten, R. P., Long, F., Murshudov, G. N. & Perrakis, A. (2014). *IUCrJ* **1**, 213-220.

Jordan, A., Harnisch, J., Borchers, R., Le Guern, F. & Shinohara, H. (2000). *Environmental Science & Technology* **34**, 1122-1124.

Juárez-Martínez, G., Garza, C., Castillo, R. & Moreno, A. (2001). *Journal of Crystal Growth* **232**, 119-131.

Khalaf, N., Govardhan, C. P., Lalonde, J. J., Persichetti, R. A., Wang, Y.-F. & Margolin, A. L. (1996). *Journal of the American Chemical Society* **118**, 5494-5495.

Koshland Jr, D. (1953). *Biological reviews* **28**, 416-436.

Kotik, M., Vanacek, P., Kunka, A., Prokop, Z. & Damborsky, J. (2017). *Applied microbiology and biotechnology* **101**, 6385-6397.

Kikhney, A. G. & Svergun, D. I. (2015). *FEBS Letters* **589**, 2570-2577.

Krooshof, G. H., Ridder, I. S., Tepper, A. W. J. W., Vos, G. J., Rozeboom, H. J., Kalk, K. H., Dijkstra, B. W. & Janssen, D. B. (1998). *Biochemistry* **37**, 15013-15023.

Kwan, A. H., Mobli, M., Gooley, P. R., King, G. F. & Mackay, J. P. (2011). *Febs j* **278**, 687-703.

Lagerkvist, U., Rymo, L., Lindqvist, O. & Andersson, E. (1972). *J Biol Chem* **247**, 3897-3899.

Lesk, A. M. (2001). *Introduction to Protein Architecture: The Structural Biology of Proteins*. Oxford University Press.

Lewandowicz, A., Rudziński, J., Tronstad, L., Widersten, M., Ryberg, P., Matsson, O. & Paneth, P. (2001). *Journal of the American Chemical Society* **123**, 4550-4555.

Li, W. F., Zhou, X. X. & Lu, P. (2005). *Biotechnol Adv* **23**, 271-281.

Lichtenthaler, F. W. (1995). *Angewandte Chemie International Edition in English* **33**, 2364-2374.

Liu, Y., Yvon-Lewis, S. A., Thornton, D. C. O., Butler, J. H., Bianchi, T. S., Campbell, L., Hu, L. & Smith, R. W. (2013). *Journal of Geophysical Research: Oceans* **118**, 3950-3965.

Lodish, H., Berk, A., Kaiser, C. A., Krieger, M. & Scott, M. P. (2007). *Molecular Cell Biology / Iclicker*. Macmillan Higher Education.

Lodish, H. F. & Lodish, H. (1999). *Molecular cell biology / Harvey Lodish ... [et al.]*, 4th ed. ed. New York

Los, G. V. & Wood, K. (2007). *High Content Screening*, pp. 195-208: Springer.

Lucic, V., Forster, F. & Baumeister, W. (2005). *Annual review of biochemistry* **74**, 833-865.

Luft, J. R., Newman, J. & Snell, E. H. (2014). *Acta crystallographica. Section F, Structural biology communications* **70**, 835-853.

Lusty, C. (1999). *Journal of Applied Crystallography* **32**, 106-112.

Mane, S. R., Shatunov, Y. M. & Yokoya, K. (2005). *Reports on Progress in Physics* **68**, 1997-2265.

Maras, B., Valiante, S., Chiaraluce, R., Consalvi, V., Politi, L., De Rosa, M., Bossa, F., Scandurra, R. & Barra, D. (1994). *Journal of protein chemistry* **13**, 253-259.

Marek, J., Vevodova, J., Smatanova, I. K., Nagata, Y., Svensson, L. A., Newman, J., Takagi, M. & Damborsky, J. (2000). *Biochemistry* **39**, 14082-14086.

Margesin, R. & Schinner, F. (2001). *Applied microbiology and biotechnology* **56**, 650-663.

Margolin, A. L. (1996). *Trends in Biotechnology* **14**, 223-230.

Mase, T., Yabuki, H., Okai, M., Ohtsuka, J., Imai, F. L., Nagata, Y. & Tanokura, M. (2012). *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* **68**, 652-654.

Massa, W. (2000). *Crystal Structure Determination*. Springer.

Matsumura, M., Signor, G. & Matthews, B. W. (1989). *Nature* **342**, 291-293.

Matthews, B. W. (1968). *J Mol Biol* **33**, 491-497.

Mazumdar, P. A., Hulecki, J. C., Cherney, M. M., Garen, C. R. & James, M. N. (2008). *Biochim Biophys Acta* **1784**, 351-362.

McMillan, E. M. (1945). *Physical Review* **68**, 143.

McPherson, A. & Gavira, J. A. (2013). *Acta crystallographica. Section F, Structural biology communications* **70**, 2-20.

McPherson, A., Kuznetsov, Y. G., Malkin, A. & Plomp, M. (2003). *Journal of Structural Biology* **142**, 32-46.

Metpally, R. P. R. & Reddy, B. V. B. (2009). *BMC Genomics* **10**, 11.

Michaelis, L. & Menten, M. L. (2007). *Die kinetik der invertinwirkung*. Universitätsbibliothek Johann Christian Senckenberg.

Milne, J. L. S., Borgnia, M. J., Bartesaghi, A., Tran, E. E. H., Earl, L. A., Schauder, D. M., Lengyel, J., Pierson, J., Patwardhan, A. & Subramaniam, S. (2013). *The FEBS journal* **280**, 28-45.

Milo, R. (2013). *Bioessays* **35**, 1050-1055.

Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. (2006). *Acta crystallographica. Section D, Biological crystallography* **62**, 859-866.

Miyazaki, K., Wintrode, P. L., Grayling, R. A., Rubingh, D. N. & Arnold, F. H. (2000). *J Mol Biol* **297**, 1015-1026.

Mobilio, S., Boscherini, F. & Meneghini, C. (2014). *Synchrotron Radiation: Basics, Methods and Applications*. Springer Berlin Heidelberg.

Moore, R. (1996). *J. Geophys. Res* **101**, 20,899-820,908.

Moreno, A., Antonio González-Ramírez, L., de los Angeles Hernández-Hernández, M. a., Oliver-Salvador, C., Soriano-García, M. & Rodríguez-Romero, A. (1999). *Journal of Crystal Growth* **196**, 587-594.

Mueller, U., Darowski, N., Fuchs, M. R., Forster, R., Hellmig, M., Paithankar, K. S., Puhringer, S., Steffien, M., Zocher, G. & Weiss, M. S. (2012). *Journal of Synchrotron Radiation* **19**, 442-449.

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). *The European Physical Journal Plus* **130**, 141.

Murata, K. & Wolf, M. (2018). *Biochimica et Biophysica Acta (BBA) - General Subjects* **1862**, 324-334.

Murrey, H. E., Judkins, J. C., Am Ende, C. W., Ballard, T. E., Fang, Y., Riccardi, K., Di, L., Guilmette, E. R., Schwartz, J. W., Fox, J. M. & Johnson, D. S. (2015). *J Am Chem Soc* **137**, 11461-11475.

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. (2011). *Acta Crystallographica Section D* **67**, 355-367.

Nakai, C., Tsuchihara, K., Mimaki, S., Esumi, H., Arakawa, H., Goto, M., Totsuka, Y., Suzuki, Y., Ochiai, A., Kojima, M., Kinoshita, M., Takemura, S., Tanaka, S., Kubo, S., Marubashi, S., Matsuda, T., Shibata, T., Nakagama, H. & Nakamori, S. (2016). *Carcinogenesis* **37**, 817-826.

Nanev, C. N., Saridakis, E. & Chayen, N. E. (2017). *Scientific reports* **7**, 35821-35821.

Nardini, M. & Dijkstra, B. W. (1999). *Curr Opin Struct Biol* **9**, 732-737.

Newman, J., Peat, T. S., Richard, R., Kan, L., Swanson, P. E., Affholter, J. A., Holmes, I. H., Schindler, J. F., Unkefer, C. J. & Terwilliger, T. C. (1999). *Biochemistry-Us* **38**, 16105-16114.

Niehaus, F., Bertoldo, C., Kähler, M. & Antranikian, G. (1999). *Applied microbiology and biotechnology* **51**, 711-729.

Nogales, E. (2015). *Nature Methods* **13**, 24.

Noritomi, H., Koyama, K., Kato, S. & Nagahama, K. (1998). *Biotechnology Techniques* **12**, 467-469.

Norouzian, D. (2003). *Iranian Journal of Biotechnology* **1**, 197-206.

Novak, H. R., Sayer, C., Isupov, M. N., Gotz, D., Spragg, A. M. & Littlechild, J. A. (2014). *FEBS Letters* **588**, 1616-1622.

Novototskaya-Vlasova, K., Petrovskaya, L., Yakimov, S. & Gilichinsky, D. (2012). *FEMS Microbiology Ecology* **82**, 367-375.

Ohana, R. F., Encell, L. P., Zhao, K., Simpson, D., Slater, M. R., Urh, M. & Wood, K. V. (2009). *Protein expression and purification* **68**, 110-120.

Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., Schrag, J., Sussman, J. L., Verschueren, K. H. G. & Goldman, A. (1992). *Protein Engineering, Design and Selection* **5**, 197-211.

Oschkinat, H., Griesinger, C., Kraulis, P. J., Sørensen, O. W., Ernst, R. R., Gronenborn, A. M. & Clore, G. M. (1988). *Nature* **332**, 374.

Otalora, F., Gavira, J. A., Ng, J. D. & Garcia-Ruiz, J. M. (2009). *Progress in biophysics and molecular biology* **101**, 26-37.

Otyepka, M. & Damborský, J. (2002). *Protein Science* **11**, 1206-1217.

Painter, J. and Merritt, E.A. (2006). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallographica Section D* **62**: 439-450.

Painter, J. and Merritt, E.A. (2006). TLSMD web server for the generation of multi-group TLS models. *Crystallographica Section* **39**: 109-111.

Papaleo, E., Riccardi, L., Villa, C., Fantucci, P. & De Gioia, L. (2006). *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1764**, 1397-1406.

Pavelka, A., Sebestova, E., Kozlikova, B., Brezovsky, J., Sochor, J. & Damborsky, J. (2016). *IEEE/ACM transactions on computational biology and bioinformatics* **13**, 505-517.

Pavlova, M., Klvana, M., Jesenska, A., Prokop, Z., Konecna, H., Sato, T., Tsuda, M., Nagata, Y. & Damborsky, J. (2007). *J Struct Biol* **157**, 384-392.

Penczek, P. A., Grassucci, R. A. & Frank, J. (1994). *Ultramicroscopy* **53**, 251-270.

Perutz, M. F., Miurhead, H., Cox, J. M., Goaman, L. C., Mathews, F. S., McGandy, E. L. & Webb, L. E. (1968). *Nature* **219**, 29-32.

Pieters, R. J., Spelberg, J. H. L., Kellogg, R. M. & Janssen, D. B. (2001). *Tetrahedron Letters* **42**, 469-471.

Pikkemaat, M. G., Linssen, A. B. M., Berendsen, H. J. C. & Janssen, D. B. (2002). *Protein Engineering, Design and Selection* **15**, 185-192.

Poelarends, G. J., Wilkens, M., Larkin, M. J., van Elsas, J. D. & Janssen, D. B. (1998). *Appl. Environ. Microbiol.* **64**, 2931-2936.

Pries, F., Kingma, J., Krooshof, G. H., Jeronimus-Stratingh, C. M., Bruins, A. P. & Janssen, D. B. (1995). *Journal of Biological Chemistry* **270**, 10405-10411.

Pries, F., Kingma, J., Pentenga, M., van Pouderoyen, G., Jeronimus-Stratingh, C. M., Bruins, A. P. & Janssen, D. B. (1994). *Biochemistry* **33**, 1242-1247.

Prince, E. (2004). *International Tables for Crystallography,Volume C: Mathematical, Physical and Chemical Tables*. Springer Netherlands.

Prokop, Z., Opluštil, F., DeFrank, J. & Damborský, J. (2006). *Biotechnology Journal: Healthcare Nutrition Technology* **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). *Angewandte Chemie* **49**, 6111-6115.

Prudnikova, T., Mozga, T., Rezacova, P., Chaloupkova, R., Sato, Y., Nagata, Y., Brynda, J., Kuty, M., Damborsky, J. & Kuta Smatanova, I. (2009). *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* **65**, 353-356.

Quiocho, F. A. & Richards, F. M. (1964). *Proceedings of the National Academy of Sciences* **52**, 833-839.

Rayment, I. (2002). *Structure* **10**, 147-151.

Rhodes, G. (2006). *Crystallography Made Crystal Clear : A Guide for Users of Macromolecular Models*. Burlington, UNITED STATES: Elsevier Science & Technology.

Rhodes, G. (2010). *Crystallography Made Crystal Clear: A Guide for Users of Macromolecular Models*. Elsevier Science.

Ridder, I. S., Rozeboom, H. J. & Dijkstra, B. W. (1999). *Acta Crystallographica Section D* **55**, 1273-1290.

Rittié, L. & Perbal, B. (2008). *Journal of Cell Communication and Signaling* **2**, 25-45.

Roldan-Arjona, T., Garcia-Pedrajas, M. D., Luque-Romero, F. L., Hera, C. & Pueyo, C. (1991). *Mutagenesis* **6**, 199-205.

Rosenberger, F., Vekilov, P. G., Muschol, M. & Thomas, B. R. (1996). *Journal of Crystal Growth* **168**, 1-27.

Rupp, B. (2010). *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*. Garland Science.

S. J. Lee, S. & Cudney, R. (2004). *A modified microdialysis button for use in protein crystallization*.

Salemme, F. R. (1972). *Archives of Biochemistry and Biophysics* **151**, 533-539.

Sanderson, M. R. & Skelly, J. V. (2007). *Macromolecular Crystallography: conventional and high-throughput methods*. OUP Oxford.

Sarmiento, F., Peralta, R. & Blamey, J. M. (2015). *Frontiers in bioengineering and biotechnology* **3**, 148-148.

Sasaki, Y. F., Saga, A., Akasaka, M., Ishibashi, S., Yoshida, K., Su, Y. Q., Matsusaka, N. & Tsuda, S. (1998). *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **419**, 13-20.

Sato, Y., Monincova, M., Chaloupkova, R., Prokop, Z., Ohtsubo, Y., Minamisawa, K., Tsuda, M., Damborsky, J. & Nagata, Y. (2005). *Appl Environ Microb* **71**, 4372-4379.

Sato, Y., Natsume, R., Tsuda, M., Damborsky, J., Nagata, Y. & Senda, T. (2007). *Acta crystallographica. Section F, Structural biology and crystallization communications* **63**, 294-296.

Scarratt, M. & Moore, R. (1996). *Marine Chemistry* **54**, 263-272.

Schenk, A. D., Philippsen, A., Engel, A. & Walz, T. (2013). *Journal of Structural Biology* **182**, 173-185.

Schrodinger, LLC (2015). The PyMOL Molecular Graphics System, Version 1.8.

Sheldrick, G. M. & Schneider, T. R. (1997). *Methods in Enzymology* **277**, 319-343.

Sheng, J., Ji, X., Zheng, Y., Wang, Z. & Sun, M. (2016). *Biotechnology letters* **38**, 1809-1815.

Shiryayev, A., Pagan, D. L., Gunton, J. D., Rhen, D. S., Saxena, A. & Lookman, T. (2005). *J Chem Phys* **122**, 234911.

Simoni, R. D., Hill, R. L. & Vaughan, M. (2002). *Journal of Biological Chemistry* **277**, e23.

Slater, J. H., Bull, A. T. & Hardman, D. J. (1995). *Biodegradation* **6**, 181-189.

Sobol, Z., Engel, M. E., Rubitski, E., Ku, W. W., Aubrecht, J. & Schiestl, R. H. (2007). *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **633**, 80-94.

Soper, R., Green, N. P. O., Stout, G. W. & Taylor, D. J. (1990). *Biological Science*. Cambridge University Press.

Stepankova, V., Bidmanova, S., Koudelakova, T., Prokop, Z., Chaloupkova, R. & Damborsky, J. (2013). *ACS Catalysis* **3**, 2823-2836.

Stepankova, V., Khabiri, M., Brezovsky, J., Pavelka, A., Sykora, J., Amaro, M., Minofar, B., Prokop, Z., Hof, M., Ettrich, R., Chaloupkova, R. & Damborsky, J. (2013). *Chembiochem : a European journal of chemical biology* **14**, 890-897.

Stsiapanava, A., Chaloupkova, R., Fortova, A., Brynda, J., Weiss, M. S., Damborsky, J. & Kuta Smatanova, I. (2011). *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* **67**, 253-257.

Stucki, G. & Thueer, M. (1995). *Environmental science & technology* **29**, 2339-2345.

Sumner, J. B. (1926). *Journal of Biological Chemistry* **69**, 435-441.

Suryanarayana, C. & Norton, M. G. (2013). *X-Ray Diffraction: A Practical Approach*. Springer US.

Tait, V. K. & Moore, R. M. (1995). *Limnology and Oceanography* **40**, 189-195.

Taylor, G. (2003). *Acta Crystallographica Section D: Biological Crystallography* **59**, 1881-1890.

Thompson, M. J. & Eisenberg, D. (1999). *J Mol Biol* **290**, 595-604.

Thompson, R. F., Walker, M., Siebert, C. A., Muench, S. P. & Ranson, N. A. (2016). *Methods (San Diego, Calif.)* **100**, 3-15.

Tilley, R. J. D. (2006). *Crystals and Crystal Structures*. Wiley.

Tomasi, A., Albano, E., Bini, A., Botti, B., Slater, T. F. & Vannini, V. (1984). *Toxicologic pathology* **12**, 240-246.

Tomcová, I., Branca, R. M. M., Bodó, G., Bagyinka, C. & Smatanová, I. K. (2006). *Acta crystallographica. Section F, Structural biology and crystallization communications* **62**, 820-824.

Tratsiak, K., Degtjarik, O., Drienovska, I., Chrast, L., Rezacova, P., Kuty, M., Chaloupkova, R., Damborsky, J. & Kuta Smatanova, I. (2013). *Acta Crystallogr Sect F Struct Biol Cryst Commun* **69**, 683-688.

Turner, P., Mamo, G. & Karlsson, E. N. (2007). *Microbial Cell Factories* **6**, 9.

Ugazio, G., Burdino, E., Danni, O. & Milillo, P. (1973). Hepatotoxicity and lethality of halogenoalkanes: Portland Press Limited.

Vagin, A. & Teplyakov, A. (2010). *Acta crystallographica. Section D, Biological crystallography* **66**, 22-25.

Vaguine, A. A., Richelle, J. & Wodak, S. (1999). *Acta Crystallographica Section D: Biological Crystallography* **55**, 191-205.

Valvoda, V., Polcarová, M. & Lukáč, P. (1992). *Základy strukturní analýzy*. Karolinum.

van den Burg, B. (2003). *Current Opinion in Microbiology* **6**, 213-218.

van Pée, K.-H. & Unversucht, S. (2003). *Chemosphere* **52**, 299-312.

Veksler, V. I. (1945). *J. Phys.* **9**, 153-158.

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

Verschueren, K. H. G., Franken, S. M., Rozeboom, H. J., Kalk, K. H. & Dijkstra, B. W. (1993). *Journal of Molecular Biology* **232**, 856-872.

Vieille, C., Burdette, D. S. & Zeikus, J. G. (1996). *Biotechnology annual review* **2**, 1-83.

Vieille, C. & Zeikus, G. J. (2001). *Microbiology and Molecular Biology Reviews* **65**, 1-43.

Vinokurov, N. A. & Levichev, E. B. (2015). *Physics-Uspekhi* **58**, 850-871.

Vogt, G., Woell, S. & Argos, P. (1997). *Journal of Molecular Biology* **269**, 631-643.

Vuister, G. W., Boelens, R., Kaptein, R., Hurd, R. E., John, B. & Van Zijl, P. C. M. (1991). *Journal of the American Chemical Society* **113**, 9688-9690.

Wang, H., Liu, S. & Du, S. (2013). *Organic Pollutants-Monitoring, Risk and Treatment*: IntechOpen.

Ward, O. P. & Moo-Young, M. (1988). *Biotechnology Advances* **6**, 39-69.

Ward, B.B. and Priscu, J.C. (1997). Detection and characterization of denitrifying bacteria from a permanently ice-covered Antarctic lake. *Hydrobiologia* **347**: 57-68.

Waseda, Y., Matsubara, E. & Shinoda, K. (2011). *X-Ray Diffraction Crystallography: Introduction, Examples and Solved Problems*. Springer Berlin Heidelberg.

Weber, L. W., Boll, M. & Stampfl, A. (2003). *Critical reviews in toxicology* **33**, 105-136.

Westerbeek, A., Szymanski, W., Feringa, B. L. & Janssen, D. B. (2011). *ACS Catalysis* **1**, 1654-1660.

Winn, M., Isupov, M. & Murshudov, G. N. (2001). *Acta Crystallographica Section D: Biological Crystallography* **57**, 122-133.

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 Crystallographica Section D: Biological Crystallography* **67**, 235-242.

Winn, M. D., Murshudov, G. N. & Papiz, M. Z. (2003). *Methods in enzymology*, pp. 300-321: Elsevier.

Wlodawer, A., Minor, W., Dauter, Z. & Jaskolski, M. (2013). *The FEBS Journal* **280**, 5705-5736.

Yamamoto, M., Hirata, K., Yamashita, K., Hasegawa, K., Ueno, G., Ago, H. & Kumasaka, T. (2017). *IUCrJ* **4**, 529-539.

Yokouchi, Y., Noijiri, Y., Barrie, L. A., Toom-Sauntry, D., Machida, T., Inuzuka, Y., Akimoto, H., Li, H. J., Fujinuma, Y. & Aoki, S. (2000). *Nature* **403**, 295.

Zeppezauer, M., Eklund, H. & Zeppezauer, E. S. (1968). *Archives of Biochemistry and Biophysics* **126**, 564-573.

Zou, X., Hovmöller, S. & Oleynikov, P. (2011). *Electron Crystallography. Electron Microscopy and Electron Diffraction*. Oxford University Press.

7. Supplementary materials

Table S1. List of experimentally characterized HLDs.

HLD	Organism	PDB ID	Subfamily	Reference
DhlA	*Xantobacter autotrophicus*	2HAD	HLD-I	(Franken *et al.*, 1991)
DppA	*Plesiocystis pacifica* SIR-1	2XT0	HLD-I	(Hesseler *et al.*, 2011)
DpcA	*Psychrobacter cryohalolentis* K5	6F9O	HLD-I	(Tratsiak *et al.*, 2019 accepted)
DccA	*Caulobacter crescentus*	5ESR	HLD-I	(Carlucci *et al.*, 2016)
DmrA	*Mycobacterium rhodesiae JS60*	4MJ3	HLD-I	(Fung *et al.*, 2015)
DmbB	*Mycobacterium tuberculosis*		HLD-I	(Jesenska *et al.*, 2005)
DhmA	*Mycobacterium avium* N85		HLD-I	(Pavlova *et al.*, 2007)
eHLD-B	*metagenome-derived*		HLD-I	(Kotik *et al.*, 2017)
DmxA	*Marinobacter sp.* ELB17	5MXP	HLD-II	(Chrast *et al*., not published)
DbjA	*Bradyrhizobium japonicum*	3A2M	HLD-II	(Prokop *et al.*, 2010)
DbeA	*Bradyrhizobium elkanii* USDA94	4K2A	HLD-II	(Chaloupkova *et al.*, 2014)
DhaA	*Rhodococcus rhodochorus*	1BN6	HLD-II	(Newman *et al.*, 1999)
LinB	*Sphingomonas paucimobilis* UT26	1CV2	HLD-II	(Marek *et al.*, 2000)
DmbA (Rv2579)	*Mycobacterium tuberculosis*	2QVB	HLD-II	(Mazumdar *et al.*, 2008)
DatA	*Agrobacterium tumefaciens* C58	3WI7	HLD-II	(Hasan *et al.*, 2011)
DmmA	*Moorea producta*	3U1T	HLD-II	(Gehret *et al.*, 2012b)
HanR	*Rhodobacteraceae family bacterium*	4BRZ	HLD-II	(Novak *et al.*, 2014)
DspA	*Strongylocentrotus purpuratus*		HLD-II	(Fortova *et al.*, 2013)
DmlA	*Mesorhizobium loti* MAFF303099		HLD-II	(Sato *et al.*, 2005)

eHLD-B	*metagenome-derived*	HLD-II	(Kotik *et al.*, 2017)
DrbA	*Rhodopirellula baltica* SH1	HLD-III	(Jesenska *et al.*, 2009)
DmbC	*Mycobacterium tuberculosis*	HLD-III	(Jesenska *et al.*, 2009)

ktratsiak@gmail.com

Crystallographic study of biotechnologically attractive haloalkane dehalogenases DpcA and DmxA
Ph.D. Thesis Series, No. 7. 2019, 165 pp.

Printed in the Czech Republic by Typodesign
Edition of 20 copies

University of South Bohemia in České Budějovice
Faculty of Science
Branišovská 31
CZ-37005 České Budějovice, Czech Republic

Phone: +420 387 776 201
www.prf.jcu.cz, e-mail: sekret@prf.jcu.cz

www.prf.jcu.cz, e-mail: sekret@prf.jcu.cz