Heterogeneous & Homogeneous & Bio- & Nano-
CHEMCATCHEM
CATALYSIS

# Accepted Article

**Title:** Rhodococcus erythropolis Oleate Hydratase: a New Member in the Oleate Hydratase Family Tree - Biochemical and Structural Studies

**Authors:** Jan Lorenzen, Ronja Janke, Ayk Waldow, Farah Qoura, Bernhard Loll, and Thomas Brück



**To be cited as:** *ChemCatChem* 10.1002/cctc.201701350

**Link to VoR:** http://dx.doi.org/10.1002/cctc.201701350

A Journal of
ChemPubSoc Europe

WILEY-VCH

www.chemcatchem.org

# *Rhodococcus erythropolis* Oleate Hydratase: a New Member in the Oleate Hydratase Family Tree - Biochemical and Structural Studies

Jan Lorenzen[a,d], Ronja Janke[b,d], Ayk Waldow[b], Farah Qoura[a], Bernhard Loll[b,c,e] and Thomas Brück[a,e]

[a] Professorship for Industrial Biocatalysis, Technical University Munich, Lichtenberg Str. 4, 85748 Garching, Germany

[b] Institute of Chemistry and Biochemistry, Freie Universität Berlin, Takustr. 6, 14195 Berlin, Germany

[c] moloX GmbH, Takustr. 6, 14195 Berlin, Germany

[d] These authors contributed equally to this work

[e] To whom correspondence should be addressed:

Prof. Dr. Thomas Brück
Professorship for Industrial Biocatalysis
Department of Chemistry
Technical University Munich
Lichtenberg Str. 4, 85748 Garching, Germany
Tel.: +49-(0) 89 289-13253
E-mail: brueck@tum.de

Dr. Bernhard Loll
Institute of Chemistry and Biochemistry
Structural Biochemistry
Freie Universität Berlin
Takustr. 6,
14195 Berlin, Germany
Phone: +49 (0) 30 838-57348
E-mail: loll@chemie.fu-berlin.de

## Abstract

Recently, the enzyme family of oleate hydratases (OHs: EC 4.2.1.53) have gained a rising scientific and economic interest, as these FAD-binding bacterial enzymes do not require cofactor recycling and feature high thermal and pH stability. Their products, hydroxy fatty acids, are used in specialty chemical applications including surfactant and lubricant formulations. The “oleate hydratase engineering database”, established by Schmid et al. (2017), divides all OHs into 11 families (HFam1 to 11). To date, only two crystal structures of homodimeric OHs from the families HFam2 and HFam11 have been reported. In this study, we biophysically characterized an OH belonging to the HFam3 family, originating from the marine bacterium *Rhodococcus erythropolis*, for the first time. The crystal structure revealed that this new OH (OhyRe) surprisingly is a monomer in its active form. This particular feature provides new avenues for enzyme engineering and recycling via immobilization.

## Introduction

Unsaturated fatty acids are toxic to many bacteria due to their surfactant properties, destabilization of membranes and due to the unsaturated fatty acids inhibitory effect on the fatty acid biosynthesis.[1] Hence, mechanisms have evolved to counteract these effects when organisms are exposed to unsaturated fatty acids, such as linoleic acid or oleic acid. In consequence, unsaturated fatty acids are saturated by hydrogenation. However, an intermediate step in this hydration process is catalyzed by fatty acid hydratases.[2] Hydration of unsaturated fatty acids has so far only been identified in Gram-negative and Gram-positive bacteria [3] and is mediated by hydratases.

Hydratases are a group of lyases that catalyze hydration and dehydration of a substrate[4]. Even though many hydratases are known, only few OHs have been described.[2] OHs belong to the group of fatty acid hydratases (EC 4.2.1.53) and convert oleic acid into (R)-10-hydroxystearic acid (10-HSA).[5] To date all characterized OHs bind an FAD cofactor,[2] albeit the role of FAD has remained enigmatic.

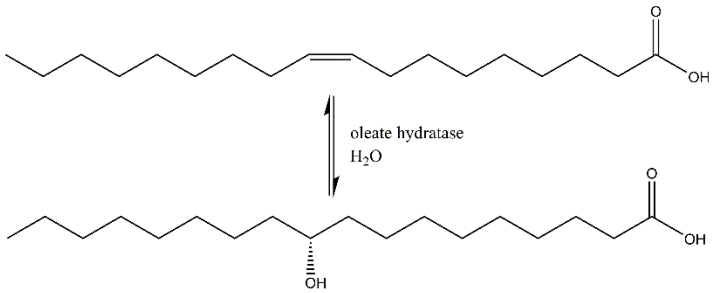


**Figure 1:** Oleate hydratase (OH) reaction scheme.

Merely two structures of OHs have been reported. One of the OH from *Lactobacillus acidophilus* (LAH), crystalized in the apo state as well as bound to the preferred substrate linoleic acid.[6] The second structure of an OH (OhyA) has been determined from *Elizabethkingia meningoseptica*, bound to its cofactor FAD.[7] Even though LAH and OhyA convert different substrates, both utilize FAD as a cofactor. Interestingly the bound FAD was oxidized in both hydratases, but reduced FAD led to significantly higher activity in OhyA.[7]

Most recently, Schmid et al. assembled a “hydratase engineering database” including a total of 2046 protein sequences.[8] Sequence comparison revealed 11 distinct families (HFam1 to 11) with a sequence identity of 62 % within the respective family. The only two structurally characterized OHs, LAH and OhyA fall into two different families, HFam2 and HFam11. For all other families structural information has remained elusive. Remarkably, in regions attributed to decorate the active site, amino acid sequences are not strictly conserved within these families.

Hydratases are in particular interesting for commercial applications since their products, hydroxyl fatty acids, have a wide range of industrial and pharmaceutical applications. Hydroxyl fatty acids can be utilized in applications for the food industry as precursors of lactones,[9] as well as an emollient for cosmetics. Further applications are the employment as a surfactant or as additives in lubricant formulation and as a green material in polymer science.[10]

In this study, we aimed at obtaining novel insights into the reaction mechanism of the OhyRe enzyme of *R. erythropolis*, using a combined structural, biophysical and biochemical approach. The OhyRe enzyme was biochemically characterized in terms of pH and temperature optimum as well as substrate specificity. The kinetic constants were determined for the conversion reaction of oleic acid to 10-HSA. Furthermore, we were able to solve the crystal structure of OhyRe at 2.64 Å resolution. The structure of OhyRe provides structural information for the OH HFam3 family of OHs for the first time. We have compared the overall architecture as well as structural features of the OhyRe active site to a set of reported OHs. As we identified two key residues, which differed in the OhyRe structure, these were replaced by site directed mutagenesis to residues conserved in the other family members. Most interestingly, these conservative amino acid replacements resulted in a strong decrease of hydration activity, indicating that OhyRe may follow a distinct reaction mechanism.

## Results & Discussion

### Isolation & characterization of OhyRe derived from *R. erythropolis*

*R. erythropolis* is a Gram-positive, non-motile representative of the nocardiaceae family.[11] *R. erythropolis* is a wide spread soil bacterium which is mainly found in marine habitats, from coastal to deep sea sediments [12] and exhibits a broad range of biochemically interesting enzymes.[13] As the genome of *R. erythropolis* has recently been deciphered,[14] we have performed sequence specific genome mining, targeting alternative OH activities in this extremophile organism.

To this end we have identified the 1.68 kB sequence related to known OHs albeit at a very low sequence identity level of max. 38 %. The gene for the putative OH was termed *ohyRe*, according to the nomenclature used by Bevers et al.[5a] for the *Elizabethkingia meningoseptica* OH ohyA. The corresponding OhyRe protein is composed of 559 amino acids with a calculated molecular weight of 66.3 kDa. The N-terminally $His_6$-tagged protein was successfully expressed in *Escherichia coli BL21(DE3)* cells. Initial activity tests with the crude extract resulted in high conversion rates (approx. 85%) from oleic acid to 10-HSA, confirming that OhyRe indeed displayed OH activity. To further characterize the enzyme, we have devised a $Ni^{2+}$-affinity purification protocol. After final His-Trap HP affinity chromatography, the purified enzyme was colorless and lost its hydration activity (Fig. S1). This observation suggested a loss of the FAD

10.1002/cctc.201701350

cofactor, which has been uniformly reported to be essential for the hydration activity in all known OHs. Moreover, the only reported loss of FAD during purification was also associated with a strong decrease in hydration activity.[6] Therefore, the addition of various FAD concentrations (5 µM to 100 µM) to the purified enzyme preparation has been evaluated to reconstitute the activity of purified OhyRe. The addition of FAD to a final concentration of 20 µM lead to the recovery of the enzymatic activity of OhyRe and was applied for all following experiments. After establishing an efficient purification protocol, OhyRe was biochemically characterized.

Initial tests demonstrated that purified OhyRe has a temperature optimum for the hydration activity at 28 °C (Fig 2). Interestingly, OhyRe exhibits a relatively wide temperature tolerance with high activity in the low temperature region (72 % of max. at 20 °C) and still moderate activity in the high temperature area (21 % of max. at 45 °C). This wide temperature tolerance might reflect the diversity of the occurrence of *R. erythropolis* from costal to deep sea habitats.

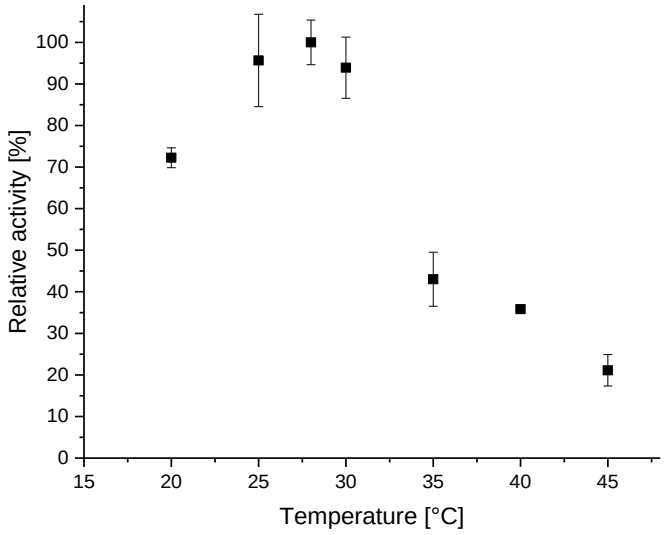


**Figure 2:** Determination of the temperature optimum for the oleate hydratase (OH) OhyRe from *R. erythropolis* in 20 mM Tris-HCl buffer at pH 7.2. Relative activity is given to the temperature optimum at 28 °C (100 %). All experiments were carried out in biological duplicates and technical triplicates.

The determination of the pH optimum (Fig. 3) showed a high pH tolerance in a range from pH 5.0 to 8.0 for OhyRe, the highest activity was reached at pH of 7.2.

10.1002/cctc.201701350

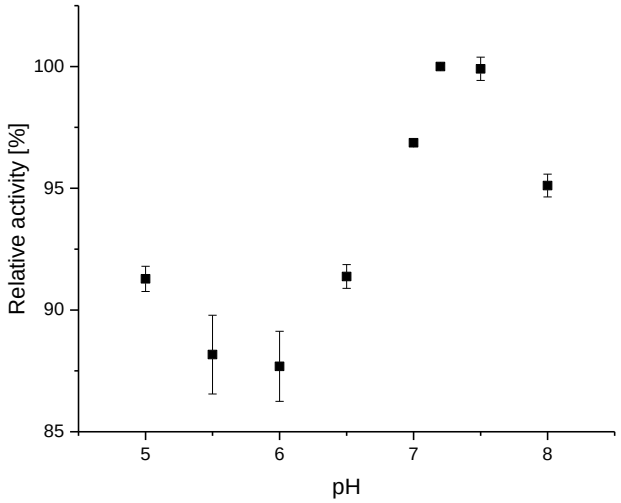


**Figure 3:** Determination of the pH optimum for the oleate hydratase (OH) OhyRe from *R. erythropolis* at 28°C. Relative activity is given to the optimum at pH 7.2 (100 %). All experiments were carried out in biological duplicates and technical triplicates.

These results correlate with the recommended cultivation conditions for *R. erythropolis* published by the German Collection of Microorganisms and Cell Culture (Leibniz Institute DMSZ).

The substrate spectrum of OhyRe was experimentally determined in reactions with the purified enzyme and different free fatty acids (FFAs) (Table 1). All experiments were carried out in biological duplicates and technical triplicates. The enzyme showed neither detectable activity for any saturated fatty acid tested (C14:0 to C18:0), nor for myristoleic acid, vaccenic acid, eicosatrienoic acid and eicosatetraenoic acid, respectively. Low conversion rates could be observed for linoleic acid, α- and γ-linolenic acid. The highest affinity, under the given conditions, was detected for oleic acid, followed by palmitoleic acid. Complex lipids, like the tested triolein or microalgae derived oils were not converted by the enzyme. These results correspond to published data for other known OHs [6] and confirm the preference of OHs for the 9-Z double bond in the fatty acid carbon chain. The kinetic parameters were experimentally defined for OhyRe, with oleic acid as the preferred substrate, and resulted in a $K_M$ value of 0.49 ± 0.1 mM and a $k_{cat}$ value of 34 ± 5 $min^{-1}$.

**Table 1:** Tested substrates, identified products and corresponding specific activity of conversion reaction with OhyRe. All experiments were carried out in biological duplicates and technical triplicates.

| substrate | | product | specific activity [nmol mg-enzyme$^{-1}$ min$^{-1}$] |
|---|---|---|---|
| C16:1 ω7 | palmitoleic Acid | 10-hydroxypalmitic acid | 1205 ± 70.6 |
| C18:1 ω9 | oleic Acid | 10-hydroxystearic acid | 1266 ± 30.3 |
| C18:2 ω6 | linoleic Acid | 10-hydroxy-12(Z)-octadecenoic acid | 156 ± 3.3 |

| | | | |
|---|---|---|---|
| C18:3 ω3 | α-linolenic Acid | 10-hydroxy-12,15(Z,Z)-octadecadienoic acid | 150 ± 1.7 |
| C18:3 ω6 | γ-linolenic Acid | 10-hydroxy-6,12(Z,Z)-octadecadienoic acid | 615 ± 36.0 |

## OhyRe mutants

Engleder et al. [7] demonstrated by multiple sequence alignment of 10 OHs that the catalytically crucial residues are highly conserved within the OH family. The alignment of *OhyRe* with 14 different OHs (Fig. S2) predominantly confirmed this hypothesis, but also depicted two exceptions. Interestingly, two amino acid residues that are highly conserved in all other OHs differ in the sequence of *OhyRe*, namely the residues M77 and V393. Further structural comparison of OhyRe and OhyA (Fig. S3) showed that these residues correlate with the residues E122 and T436 from OhyA, which are responsible for the activation of the water molecule or the binding of the carboxylate, [7] respectively. A conservative substitution of the two identified residues to the familial conserved residues (M77E, V393T) was executed by site directed mutagenesis and resulted in the effect depicted in Fig. 4.

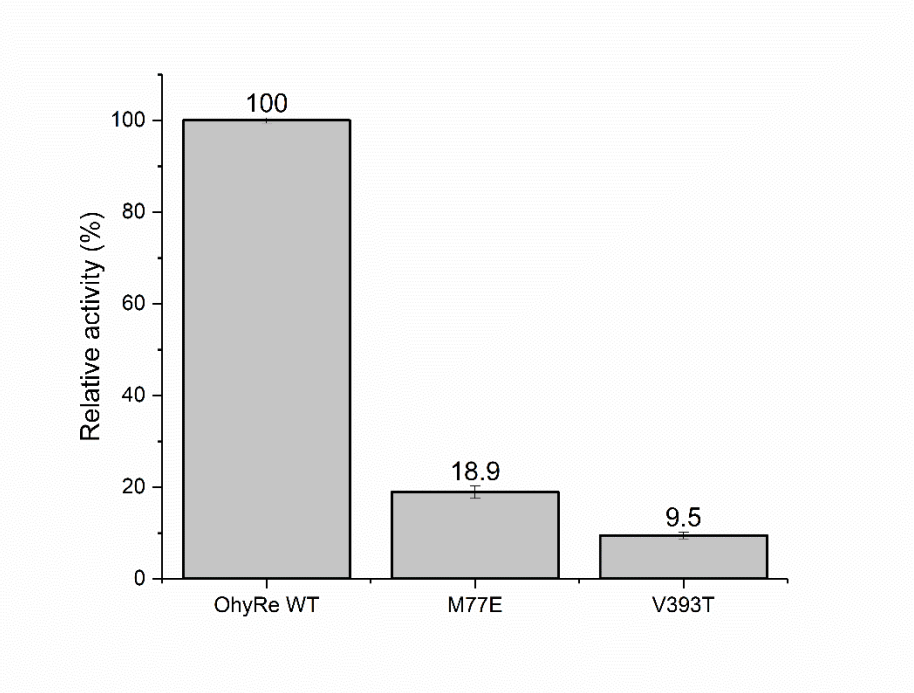


**Figure 4:** Relative conversion activity from oleic acid to 10-HSA from OhyRe wild type enzyme (OhyRe WT), set as 100 %, and its respective mutants (M77E, V393T). All experiments were carried out in biological duplicates and technical triplicates.

Fig. 4 shows the relative hydration activity of the OhyRe wild type enzyme compared to the two constructed mutants M77E and V393T with oleic acid as substrate. The conservative substitution of both residues resulted in a strong decrease in the conversion efficiency from oleic acid to 10-HSA of 81 % and 90 %, respectively. This strong decrease in hydration activity supports the hypothesis that OHs from the HFam3 family might follow a distinct catalytic mechanism.

In comparison to the two other structurally characterized OHs LAH [6] and OhyA [7], OhyRe shows a sequence identity of 35 % and 34 %, respectively. Beside the low sequence identity, both the LAH and OhyA have large N- and C-terminal amino acid sequence extensions (Fig. S4 and 6A,B), which are absent in OhyRe. To understand whether those extensions are common for all other OHs, we performed a sequence alignment (Fig. S2). Interestingly, OH sequences can be roughly divided in sequences with amino acid extensions at the N- and the C-termini present or absent. Consequently, this raised the question, whether OhyRe might be structurally different compared to the members of the HFam2 and HFam11 families. Hence we tried to crystalize the protein for structure determination by X-ray crystallography.

## Overall structure

We obtained crystals of OhyRe that belong to space group $P6_522$. The diffraction of the OhyRe crystals was anisotropic, resulting in complete data set to a resolution of 2.64 Å. The structure of OhyRe was determined by molecular replacement with two protein chains in the asymmetric unit corresponding to a very high solvent content of about 75%. The structure was refined with good stereochemistry to good R factors (Table S1, supporting information). Both polypeptide chains are practically identical with a root mean square deviation (rmsd) of 0.1 Å for 522 pairs of Cα-atoms. The structure is complete, except for the 5 N-terminal residues and a loop region ranging from residue 42 to 67 due to missing electron density caused by its high flexibility (Fig. 5A). No redox cofactor could be identified in the electron density maps. Hence our structure represents the apo-state of OhyRe from *R. erythropolis*.

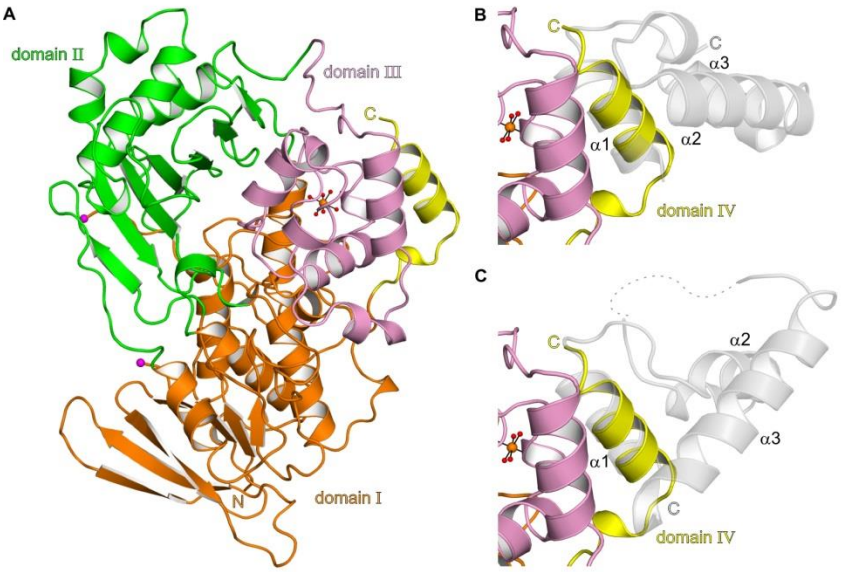


**Figure 5:** Overall structure of OhyRe in cartoon representation. **(A)** Domain organization of OhyRe: Domain I in orange, domain II in green, domain III in pink and domain IV in yellow. The bound $Mg^{2+}$ and coordinating waters are shown as orange and red spheres, respectively. Magenta spheres indicate the boarders of a flexible linker, ranging from residue 61 to 86, which could not be modelled due to its flexibility. **(B)** Zoom onto domain IV with a superposition of OhyRe and

LAH, domain IV of LAH is shown in gray. **(C)** Superposition of OhyRe and OhyA, domain IV of OhyA is shown in gray. Dashed line indicates an un-modelled loop region.

The closest related structures, as analyzed by the DALI software,[15] are the OHs LAH (PDB-id 4ia6) and OhyA (PDB-id 4uir). OhyRe superimposes with a rmsd of 1.8 Å for 500 pairs of Cα-atoms and 1.6 Å for 501 pairs of Cα-atoms, respectively. In analogy to the structures of LAH and OhyA, OhyRe is composed of three core domains (residues 1–540; Fig. 5A and 6) that are related to other FAD-dependent enzymes. Domain I (residues 6–115, 223–300, 322–344, and 478–540) is a mixed α/β domain composed of a parallel five-stranded β-sheet packed between two α-helices on one side and a three-stranded antiparallel β-sheet on the other side (Fig. 5A), resembling a variant of the Rossmann fold. Domain II (residues 116–131, 301–321, and 345–477) consists of an antiparallel β-sheet (Fig. 5A) flanked by three α-helices defining the cofactor- and substrate-binding site in conjunction with domain I. Domain III (residues 132-222) is exclusively α-helical (Fig. 5A) and its fold is structurally related to monoamine oxidases.[16] Domain IV (residues 541-559) of OhyRe is composed of merely one single α-helix. In contrast to the structures of LAH and OhyA the C-terminal domain is truncated by 32 or 48 residues, respectively. As a consequence two α-helices are absent in domain IV of OhyRe (Fig. 5B and C). The structure of LAH has been determined in apo and product-bound state. Upon substrate recognition and binding the two C-terminal α-helices of domain IV undergo a large conformational change and thereby allow the substrate to enter a cavity that ends at a cleft between domain I and III.[6]

For both the LAH and OhyA a homodimeric quaternary structure in solution as well as *in crystallo* has been reported.[6-7] For other OHs such as the one from *L. fusiformis*,[17] *S. maltophilia* [18] and *M. caseolyticus* [19] dimers in solution were reported as well. As analyzed by the PISA server,[20] OhyRe is a monomer *in crystallo*. To analyze the situation in solution and to confirm our structural analysis, we performed multi-angle light scattering (MALS) experiments. Our experiments clearly revealed that OhyRe is monomeric in solution (Fig. S5). The latter finding is in agreement to our observations for OhyRe assembly in the crystal and strengthens our hypothesis that OHs, lacking N- and C-terminal extensions, are monomers. Additionally we examined whether the extensions might have functional implications. As discussed above, in the structure of LAH, domain IV reflects the C-terminal extension, suggested to be involved in substrate gating. Since the two reported and our crystal structures cover nearly the complete amino acid sequence of the respective protein, we superimposed the structures and inspected in particular the monomer-monomer interfaces of dimeric LAH and OhyA. Most of the dimer interface is established by residues that reside within the N- or C-terminal extension, including domain IV of LAH (Fig. 6A) and OhyA (Fig. 6B). In summary, we propose that the presence or absence of N- and C-terminal extensions within the amino acid sequence directly influences the oligomeric state of OHs. In case of OhyRe the extensions are absent and hence the protein is monomeric in solution.

Notably, the crystal structure revealed a coordinated $Mg^{2+}$ ion close to the putative active site of OhyRe (section S1). However, experimental data showed that the $Mg^{2+}$ ion has no influence on the hydration activity of OhyRe.

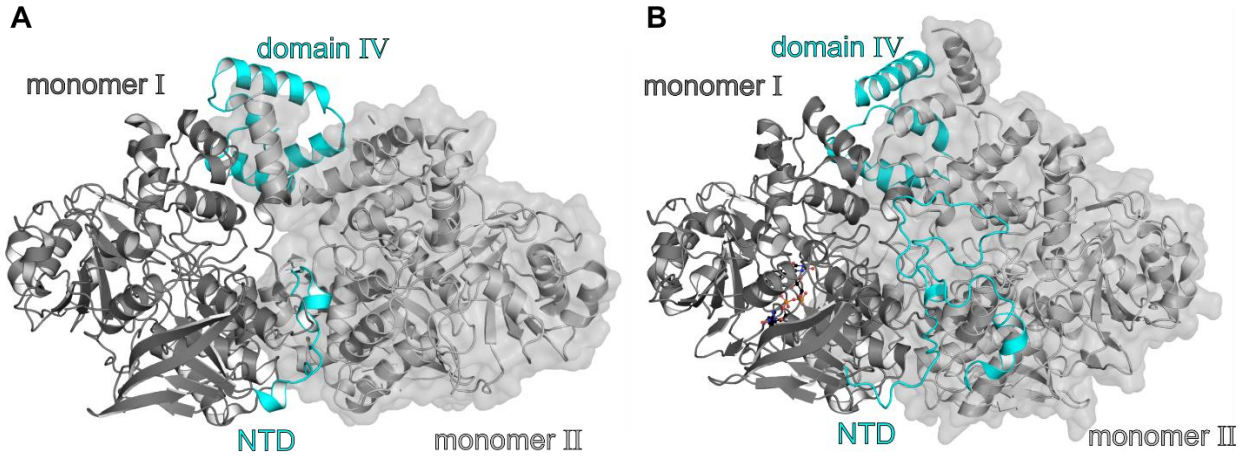


**Figure 6:** Oligomeric state of oleate hydratases (OHs). **(B)** LAH monomer I is shown in dark gray. Monomer II in surface and cartoon representation (light gray). The N-terminal domain (NTD) and domain IV in cyan. **(C)** The NTD OhyA of each monomer meanders into the other monomer.

## FAD binding

All characterized OHs utilize FAD as cofactor.[2] Different roles of the FAD cofactor have been proposed either being directly involved in catalysis or indirectly having merely a structural role to stabilize the enzyme. In the crystal structure of LAH, no FAD could be located in its binding site.[6] In contrast, in the crystal structure of OhyA, FAD is bound, but only to one of the two polypeptide chains.[7] Furthermore a $K_d$ of 1.8 x $10^{-6}$ M was determined by isothermal calorimetry.[7] All indicate a rather low binding affinity of FAD to studied OHs. For our structural studies, we attempted to add FAD prior to the final purification step or to incubate OhyRe with FAD prior to the crystallization experiments. Moreover, we tried to soak FAD in our apo OhyRe crystals. All our attempts to gain structural information of OhyRe bound to FAD remained unsuccessful so far. A primary sequence analysis reveals the presence of a characteristic recognition sequence $GXGXXGX_{21}E/D$ for nucleotide binding, also reflected by the Rossmann fold of the protein. A superposition of the structures of OhyA and OhyRe (Fig. 7A and B) clearly reveals that the cofactor binding site is large enough to accommodate a FAD molecule. Furthermore in the crystal structure of OhyA a PEG molecule was identified that might mimic a bound oleate molecule. For latter substrate enough space would be in the active site of our OhyRe structure (Fig. 7A and B).

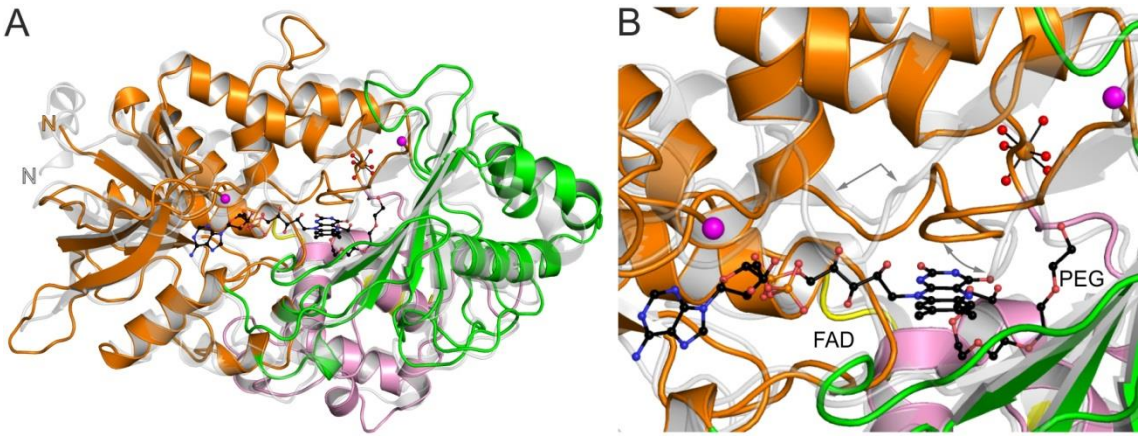


**Figure 7**: Superposition of OhyRe and OhyA. Same color coding as in Figure 5A. **(A)** OhyRe is colored as in Figure 1 and OhyA in gray. View on the active site of OhyA. Magenta spheres indicate the boarders of a flexible linker in OhyRe, ranging from residue 42 to 67, which could not be modelled due its flexibility. In contrast upon FAD binding in OhyA this loop region is getting structured, indicated by two gray arrows. Moreover a short loop fragment rearranges upon cofactor/substrate binding (gray arrow-headed arc). In general, there would be enough space in OhyRe to accommodate the cofactor as well as the substrate. Solvent exposed on the back of the active site, the $Mg^{2+}$ binding site located. **(B)** Same view in (A) zoomed on the FAD cofactor binding site of OhyA.

### Putative active site with catalytically important residues

As mentioned above OHs can be, based on their amino acid sequences, sub-divided in 11 families.[8] Based on the structure of OhyA, that belongs to the family HFam1,[8] E122 was proposed to be important for cofactor binding and being involved in catalysis by activating a water molecule for the attack on the partially charged double bond of the substrate.[7] Latter residue resides within a conserved loop region $^{118}$RGGREM$^{123}$, almost strictly conserved within all OHs. In the sequence of the OhyRe, that falls into the family HFam3, the ultimate and penultimate amino acids are different, but characteristic for members of the largest HFam1 as well HFam3 family.[8] In the respective loop region OhyRe has the sequence $^{73}$RGGRML$^{78}$. The exchange of OhyA E122 to OhyRe M77, will drastically alter the chemistry and can certainly not activate a water molecule. To investigate the influence of amino acid at position 77 in OhyRe, we generated a point mutation M77E. This single mutation led to a drastic decrease in hydration activity of OhyRe. OhyA Y241 was proposed to be as well important for catalysis, by protonation of the double bond [7] and is conserved with OhyRe Y205. Hence, it is tempting to propose that OhyRe might have a different catalytic mechanism as proposed for OhyA.

Interestingly, the fatty acid double-bond hydratase from *Lactobacillus plantarum*, LPH,[21] also exhibits the HFam3 family specific motive RGGREM. In contrast to LPH, which solely converts linoleic acid, OhyRe exhibits hydration activity for both, C16 and C18 unsaturated fatty acids. In addition, LPH was shown to be a homotrimer in its active form, whereas OhyRe is a monomer.

The exact binding site of the substrates of OHs could not yet be resolved. In the crystal structure of LAH, a cavity was observed running from the surface of the protein towards the interface formed by domain I and III. This site is known to be a substrate binding site in other related enzyme families. Volkov *et al.* identified a linoleic acid (LA) molecule bound in this cavity with its carboxylate oriented towards the protein surface and the hydrophobic tail pointing to the protein

interior (Fig. 8). In contrast, a PEG molecule has been identified close to the FAD cofactor in the structure of OhyA, possibly mimicking a bound substrate molecule (Fig. 8). Strictly this binding site is fundamental different to the “interdomain”-binding site as described for LAH (Fig. 7B and Fig. 8). To date, there is no structural information on an OH with a real substrate bound to the enzyme. Therefore it is very speculative which amino acid residues are crucial for catalysis. A definitive answer could only provide a structure with bound substrate. This would allow the identification of key amino acids, important for catalysis. These residues could then be targeted by structure-based mutagenesis. Moreover, the given resolution of 2.64 Å of our OhyRe structure, does not allow the identification of water molecules within the active site. The location of the water molecules represents the second bottle neck for the identification of catalytically relevant residues, as the precise position of the water molecules might be as well crucial for catalysis. Based on the structure of OhyA, the T436A mutant had been generated, showing a dramatic decrease in activity. Latter observation was interpreted as the hydroxyl of T436 being important for recognition and binding of the carboxylate function of the substrate. Taken together, both structures not only display distinct substrate binding sites, but also differ in the orientation of the carboxylate. In order to increase the enzymatic activity of OhyRe, we speculated that a mutation V393T (equivalent to T436 in OhyA) could potentially improve recognition and binding of the substrate. But unexpectedly, the activity was drastically reduced (Fig 4).

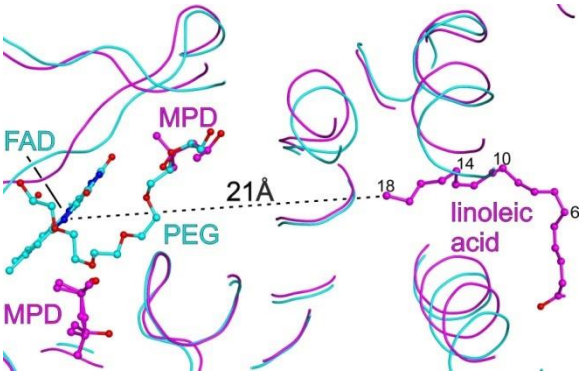


**Figure 8:** View on the potential nucleotide and substrate binding site. Superposition of LAH (magenta) and OhyA (cyan) as ribbon. FAD and PEG bound to LAH are shown in stick representation. The linoleic acid substrate of LAH and 2-methyl-2,4-pentanediol (MPD) molecules bound to LAH are shown in stick representation. Notably a 2-methyl-2,4-pentanediol (MPD) molecule in the LAH structure superimposed partially with the bound PEG molecule in the OhyA structure. The terminal carbon atom of linoleic acid is in distance of 21 Å to FAD, indicated by a dashed black line.

## Conclusion

The enzyme family of OHs offers the sustainable conversion and valorization of biogenic fatty acids into high value lubricant additives. In contrast to heme dependent P450 hydroxylases, the conversion of oleic acid to 10-hdroxysteric acid by OHs can be accomplished without the need for co-factors. In this study a new OH from *R. erythropolis* could be identified by targeted genome mining. The biochemical characterization showed that the enzyme, OhyRe, exhibits a high pH-tolerance as well as high hydration activity over a wide temperature range. The crystal structure of OhyRe was solved to a resolution of 2.64 Å, surprisingly revealing a monomeric structure for

the enzyme, which contrasts all other reported members of the OH family. These results were ultimately confirmed by MALS experiments of the protein in solution. To our knowledge, OhyRe is the first strictly monomeric OH and the first member of the HFam3 family of OHs that was characterized and crystallized so far. The HFam3 family of OHs exhibits high potential for further biocatalytical applications, as the monomeric enzyme provides a variety of options for immobilization strategies. The free N- and C-terminal ends are not involved in dimer association and could e.g. be feasible targets for the fusion of anchor-proteins. Furthermore, two mutants of OhyRe, M77E and V393T, were generated to clarify the importance of these two residues in the catalytic mechanism of OhyRe. These conservative substitutions, to residues being highly conserved within in the OH family, resulted in a drastic decrease of hydration activity of 80 to 90 %. This suggests that OhyRe might follow a distinct reaction mechanism, which has to be clarified in future experiments. Data in part presented in this manuscript are included in a patent application, regarding the technical process for 10-HSA production (date of submission to the EPO: 29.09.2017).

## Materials and Methods

### Chemicals

All chemicals utilized in this work were purchased from Sigma-Aldrich (Munich, Germany) and Carl Roth (Karlsruhe, Germany), at the highest purity grade available.

### Bacterial strains & plasmids

The *Escherichia coli* strains XL-1 Blue and DH5α were used for cloning whereas the strains BL21 (DE3) and Rosetta2 were employed for protein expression. All plasmids and *E. coli* strains were obtained from Novagene/Merck Millipore.

### Cloning for characterization experiments

*The myosin-cross-reactive antigen coding gene O5Y_00450* from *Rhodococcus erythropolis CCM2595*, was taken as a template for a codon-optimized gene-synthesis (Life Technologies, Regensburg), for an *E. coli* host strain. The obtained synthetic gene was sub-cloned in a pET28a expression plasmid and transformed into chemically competent *F* cells.

### Cloning for crystallization experiments

For protein production in *E. coli*, the *OhyRe* ORF was amplified with the forward primer 5'-TATACCATGGGAATGAGCAGCAATCTGAGCC-3' and reverse primer 5'-TATAAAGCTTTTAACGAAACATGGTAACTGCTGC-3' containing *Hind*III and *Nco*I recognition sites, respectively) and cloned into the pETM11 vector, resulting in a construct with N-terminal hexa-histidine-tag that could be cleaved off by TEV protease.

### OhyRe mutant construction

The mutants *OhyRe-M77E* and *OhyRe-V393T* were constructed by overlap-extension-PCR using primer sets listed in Table S2 (supporting information).

**Protein expression and protein purification for characterization experiments**

The expression of the native OhyRe or respective mutants thereof was carried out in *E. coli BL21DE3* cells, grown in Laure Broth (LB) medium. Pre-cultures were inoculated from a cryo-stock and grown over night in 100 ml LB in a 500 ml baffled flask at 120 rpm at 37°C. Main cultures were grown up to an optical density, measured at 600 nm ($OD_{600}$), of 0.6 – 0.8, before the expression was induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM. After 16 h of incubation at 16°C the cells were harvested and the resuspended (20 mM Tris-HCl buffer and 25 mM imidazole pH 7.2) cell pellets were disrupted by high pressure homogenization (EmulsiFlex-B15, AVESTIN). A subsequent centrifugation step at 20.000 g for 40 min at 4°C (Beckmann coulter J-20 XP) was applied for the separation of the cell debris from the soluble protein fraction. The soluble protein fraction was utilized for affinity chromatography via a $Ni^{2+}$-NTA His-trap column (HisTrap FF, GE Healthcare; flow rate 1 ml/min). The purified protein solution was desalted using HiPrep 26/10 desalting column (GE Healthcare). Protein amounts were quantified using 2-D quant kit (GE Healthcare) according to manufacturer's instructions.

**Protein expression and protein purification for crystallization experiments**

OhyRe fused to an N-terminal hexa-histidine-tag in pET-M11 vector was transformed in *E. coli* Rosetta2. Protein was expressed in auto-induction media at 37 °C until an $OD_{600}$ ~ 0.8 was reached and subsequently cooled down to 20 °C.[22] Cells grew over night and were harvested by centrifugation (6 min, 6'000 rpm at 4 °C). For resuspension of the cell pellet, buffer A was used (50 mM Tris/HCl pH 7.5, 500 mM NaCl, 30 mM imidazole and 1 mM DTT). Cells were lysed by homogenization at 4 °C and the lysate was cleared by centrifugation (1 h, 21'000 rpm at 4 °C). $Ni^{2+}$-NTA beads (cv ~ 1 ml; GE Healthcare) were equilibrated with buffer A. OhyRe was loaded on the column and washed with 10 cv of buffer A. OhyRe was eluted in a linear gradient to buffer A supplemented to 500 mM imidazole. Pooled fractions were dialyzed over night against buffer B (20 mM Tris/HCl pH 7.5, 150 mM NaCl and 1 mM DTT) supplemented with TEV protease to cleave off the N-terminal His-tag. Uncleaved protein was separated from cleaved protein via an additional $Ni^{2+}$-NTA affinity chromatography. Size exclusion chromatography was performed with a HighLoad Superdex S200 16/60 column (GE Healthcare), equilibrated with buffer B. Pooled protein fractions were concentrated with Amicon-Ultra 30'000 to 19.2 mg/ml as measured by the absorbance at 280 nm.

**Enzyme characterization**

The determination of the enzymatic properties of the OH from *Rhodococcus erythropolis CM2595* was (unless otherwise stated) executed under the following standard reaction conditions. The tests were carried out in a reaction volume of 200 µl in 20 mM Tris-HCl buffer (pH 7.2), containing the purified enzyme (final conc. of 5 µM), FAD (final conc. 20 µM) and 720 µM OA as substrate at 28°C for 15 min. **Thermo-stability** was tested in a temperature range from 20 – 45°C. **pH**

**tolerance** was monitored in a range from pH 5 – 8. To determine the **substrate specificity** of the myosin-cross-reactive antigen from *Rhodococcus erythropolis CM2595*, the purified enzyme was tested on myristic acid (C14:0), myristoleic acid (C14:1), palmitic acid (C16:0), palmitoleic acid (C16:1), stearic acid (C18:0), vaccenic acid (C18:1 *trans*-11) oleic acid (C18:1, *cis*-9), linoleic acid (C18:2), linolenic acid (C18:3), eicosatrienoic acid (C20:3), eicosatetraenoic acid (C20:4) and triolein as potential substrates. For the determination of the **kinetic parameters** of OhyRe towards oleic acid, different substrate concentrations (90 µM to 1.44 mM) and reaction times (1 to 15 min) were tested under standard conditions. The reactions were stopped by the addition of an equal volume of ethyl acetate (EtOAc) and instant, intensive vortexing. The resulting solvent phase was separated from the water phase and applied for silylation and subsequent analysis.

**Lipid analysis**

The preparation of the extracted lipid fractions for GC measurements was performed according to Volkov et al.[23] Hydroxylated fatty acids were identified by GC-MS by injecting 1 µl into a Thermo Scientific™ TRACE™ Ultra Gas Chromatograph coupled to a Thermo DSQ™ II mass spectrometer and the Triplus™ Autosampler injector. Column: Stabilwax ® fused silica capillary (30 m × 0.25 mm, film thickness 0.25 µm) Restek. (Program: initial column temperature 150 °C, increasing (4 °C/min) up to a final temperature of 250 °C. Carrier gas: hydrogen (flow rate 3.5 ml/min). Peaks were identified by comparison to fatty acid standards or by specific molecular masses detected. Extracts resulting from kinetic experiments were analyzed with the Shimadzu™ GC-2025 system equipped with a flame ionization detector. Column: Zebron ZB-WAX (30 m x 0.32 mm, film thickness 0,25 µm) Phenomenex. Carrier gas: hydrogen (3.00 ml/min). Program: initial column temperature 150 °C for 1 min; increasing 5°C/min to 240°C, hold for 6 min. Peaks were identified by comparison to the respective standards.

**Crystallization**

Crystals were obtained by the sitting-drop vapor-diffusion method at 18 °C with a reservoir solution composed of 100 mM Bis-Tris/HCl pH 5.5 to pH 6.0, 255 mM to 300 mM magnesium formate and 5 % (v/v) glycerol. Crystals were cryo-protected with 25% (v/v) MPD supplemented to the reservoir resolution and subsequently flash-cooled in liquid nitrogen.

**Diffraction data collection, structure determination and refinement**

Synchrotron diffraction data were collected at the beamline 14.1 of the MX Joint Berlin laboratory at BESSY (Berlin, Germany) or beamline P14 of Petra III (Deutsches Elektronen Synchrotron, Hamburg, Germany). X-ray data collection was performed at 100 K. Diffraction data were processed with XDS [24] (Table S1). The structure was solved via molecular replacement in PHASER [25] by using the structure of the hydratase from *Lactobacillus acidophilus* (PDB 4ia5) [6] as a search model. Crystals of OhyRe belong to the space group $P6_522$, with two molecules in the asymmetric unit. For the refinement 2.8% of the reflections were set aside for the calculation of the $R_{free}$. Model building and water picking was performed with COOT.[26] The structure was initially refined by applying a simulated annealing protocol and in later refinement cycles by maximum-likelihood restrained refinement using PHENIX.refine.[27] Model quality was evaluated

with MolProbity [28] and the JCSG validation server.[29] Secondary structure elements were assigned with DSSP [30] and ALSCRIPT [31] was used for secondary structure based sequence alignments. Figures were prepared using PyMOL.[32]

**Multi-angle light scattering (MALS)**

MALS experiment was performed at 18°C. OhyRe was loaded onto a Superdex 200 increase 10/300 column (GE Healthcare) that was coupled to a miniDAWN TREOS three-angle light scattering detector (Wyatt Technology) in combination with a RefractoMax520 refractive index detector. For calculation of the molecular mass, protein concentrations were determined from the differential refractive index with a specific refractive index increment (*d*n/*d*c) of 0.185 ml/g. Data were analyzed with the ASTRA 6.1.4.25 software (Wyatt Technology).

**Accession numbers**

The atomic coordinates and structure factor amplitudes have been deposited in the Protein Data Bank under the accession code 5ODO.

## Acknowledgement

We gratefully acknowledge the financial support by the German Federal Ministry of Education and Research (Project: Advanced Biomass Value, 03SF0446C). We also want to thank Martina Haack for her excellent technical support in the analytics and Tom Schuffenhauer for his excellent experimental support. We are grateful to C. Alings for excellent technical support. We are grateful to M. Wahl for continuous encouragement and support. R. Janke is supported by the Elsa-Neumann stipend of the state of Berlin and the Nüsslein-Volhard foundation. We acknowledge beamtime and support at beamline P14 of PETRA III (Deutsches Elektronen Synchrotron, Hamburg, Germany). We accessed beamlines of the BESSY II (Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung II) storage ring (Berlin, Germany) via the Joint Berlin MX-Laboratory sponsored by the Helmholtz Zentrum Berlin für Materialien und Energie, the Freie Universität Berlin, the Humboldt-Universität zu Berlin, the Max-Delbrück Centrum, and the Leibniz-Institut für Molekulare Pharmakologie.

**Keywords:** Lyases • Oleate Hydratase • Fatty Acids • Hydration • Protein Structure

## References


[1] a) D. L. Greenway, K. G. Dyke, *Journal of General Microbiology* **1979**, *115*, 233-245; b) C. J. Zheng, J. S. Yoo, T. G. Lee, H. Y. Cho, Y. H. Kim, W. G. Kim, *FEBS Lett* **2005**, *579*, 5157-5162.

[2] A. Hiseni, I. W. C. E. Arends, L. G. Otten, *ChemCatChem* **2015**, *7*, 29-37.

[3] D. L. Greenway, K. G. Dyke, *J Gen Microbiol* **1979**, *115*, 233-245.

[4] J. Jin, U. Hanefeld, *Chem Commun (Camb)* **2011**, *47*, 2502-2510.

[5] a) L. E. Bevers, M. W. Pinkse, P. D. Verhaert, W. R. Hagen, *J Bacteriol* **2009**, *191*, 5010-5012; b) E.-Y. Jeon, J.-H. Lee, K.-M. Yang, Y.-C. Joo, D.-K. Oh, J.-B. Park, *Process Biochemistry* **2012**, *47*, 941-947.

[6] A. Volkov, S. Khoshnevis, P. Neumann, C. Herrfurth, D. Wohlwend, R. Ficner, I. Feussner, *Acta Crystallogr D Biol Crystallogr* **2013**, *69*, 648-657.

[7] M. Engleder, T. Pavkov-Keller, A. Emmerstorfer, A. Hromic, S. Schrempf, G. Steinkellner, T. Wriessnegger, E. Leitner, G. A. Strohmeier, I. Kaluzna, D. Mink, M. Schurmann, S. Wallner, P. Macheroux, K. Gruber, H. Pichler, *Chembiochem* **2015**, *16*, 1730-1734.

[8] J. Schmid, L. Steiner, S. Fademrecht, J. Pleiss, K. B. Otte, B. Hauer, *Journal of Molecular Catalysis B: Enzymatic* **2017**.

[9] Y. Wache, M. Aguedo, A. Choquet, I. L. Gatfield, J. M. Nicaud, J. M. Belin, *Appl Environ Microbiol* **2001**, *67*, 5700-5704.

[10] a) C. T. Hou, *N Biotechnol* **2009**, *26*, 2-10; b) J. W. Song, E. Y. Jeon, D. H. Song, H. Y. Jang, U. T. Bornscheuer, D. K. Oh, J. B. Park, *Angew Chem Int Ed Engl* **2013**, *52*, 2534-2537.

[11] M. Goodfellow, G. Alderson, *Journal of General MIcrobiology* **1977**, *100*, 25.

[12] a) B. R. Langdhal, P. Bisp, K. Ingvorsen, *Microbiology* **1996**, *142*; b) S. C. Heald, P. F. B. Brandao, R. Hardicre, A. T. Bull, *Antonie van Leeuwenhoek* **2001**, *80*.

[13] C. C. de Carvalho, M. M. da Fonseca, *Appl Microbiol Biotechnol* **2005**, *67*, 715-726.

[14] H. Strnad, M. Patek, J. Fousek, J. Szokol, P. Ulbrich, J. Nesvera, V. Paces, C. Vlcek, *Genome Announc* **2014**, *2*.

[15] L. Holm, P. Rosenstrom, *Nucleic Acids Res* **2010**, *38*, W545-549.

[16] L. De Colibus, M. Li, C. Binda, A. Lustig, D. E. Edmondson, A. Mattevi, *Proc Natl Acad Sci U S A* **2005**, *102*, 12684-12689.

[17] B. N. Kim, Y. C. Joo, Y. S. Kim, K. R. Kim, D. K. Oh, *Appl Microbiol Biotechnol* **2012**, *95*, 929-937.

[18] Y. C. Joo, E. S. Seo, Y. S. Kim, K. R. Kim, J. B. Park, D. K. Oh, *J Biotechnol* **2012**, *158*, 17-23.

[19] Y. C. Joo, K. W. Jeong, S. J. Yeom, Y. S. Kim, Y. Kim, D. K. Oh, *Biochimie* **2012**, *94*, 907-915.

[20] E. Krissinel, K. Henrick, *J Mol Biol* **2007**, *372*, 774-797.

[21] J. Ortega-Anaya, A. Hernandez-Santoyo, *Biochim Biophys Acta* **2015**, *1848*, 3166-3174.

[22] F. W. Studier, *Protein Expr Purif* **2005**, *41*, 207-234.

[23] A. Volkov, A. Liavonchanka, O. Kamneva, T. Fiedler, C. Goebel, B. Kreikemeyer, I. Feussner, *J Biol Chem* **2010**, *285*, 10353-10361.

[24] W. Kabsch, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 125-132.

[25] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *Journal of applied crystallography* **2007**, *40*, 658-674.

[26] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 486-501.

[27] a) P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 213-221;
b) P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart, P. D. Adams, *Acta Crystallogr D Biol Crystallogr* **2012**, *68*, 352-367.
[28] V. B. Chen, W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 12-21.
[29] H. Yang, V. Guranovic, S. Dutta, Z. Feng, H. M. Berman, J. D. Westbrook, *Acta Crystallogr D Biol Crystallogr* **2004**, *60*, 1833-1839.
[30] W. Kabsch, C. Sander, *Biopolymers* **1983**, *22*, 2577-2637.
[31] G. J. Barton, *Protein Eng* **1993**, *6*, 37-40.
[32] W. DeLano, *The PyMOL Molecular Graphics System* http://www.pymol.org . **2002**.