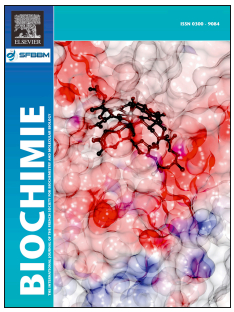


Gentisate 1,2-dioxygenase from the gram-positive bacteria *Rhodococcus opacus* 1CP: Identical active sites vs. different substrate selectivities

Natalya M. Subbotina, Alexey M. Chernykh, Anton I. Taranov, Anna D. Shebanova, Olga V. Moiseeva, Marta Ferraroni, Marina P. Kolomytseva

PII: S0300-9084(20)30274-1

DOI: https://doi.org/10.1016/j.biochi.2020.10.016

Reference: BIOCHI 6005

To appear in: *Biochimie*

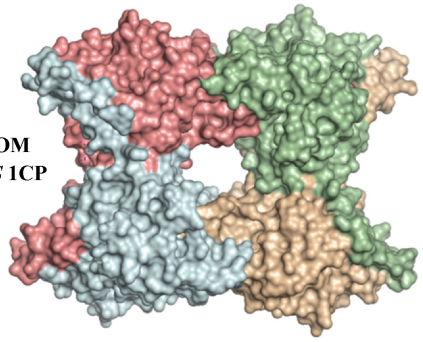
Received Date: 20 April 2020

Revised Date: 10 August 2020

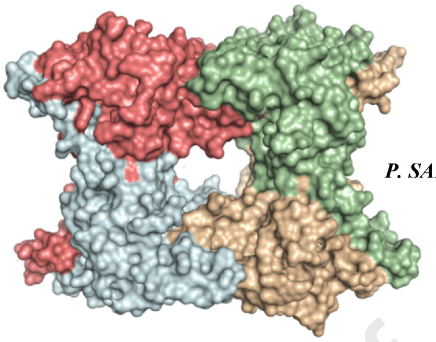
Accepted Date: 25 October 2020

Please cite this article as: N.M. Subbotina, A.M. Chernykh, A.I. Taranov, A.D. Shebanova, O.V. Moiseeva, M. Ferraroni, M.P. Kolomytseva, Gentisate 1,2-dioxygenase from the gram-positive bacteria *Rhodococcus opacus* 1CP: Identical active sites vs. different substrate selectivities, *Biochimie* (2020), doi: https://doi.org/10.1016/j.biochi.2020.10.016.

GENTISATE RING-CLEAVING DIOXYGENASES WITH DIFFERENT SUBSTRATE SELECTIVITY

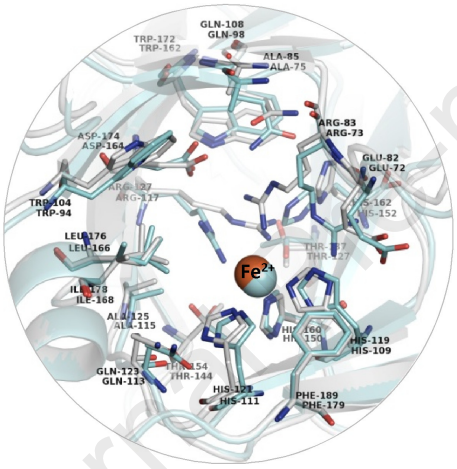


GDO FROM
***R. OPACUS* 1CP**



SDO FROM
P. SALICYLATOXIDANS

AND IDENTICAL AMINO ACID COMPOSITION OF THEIR ACTIVE SITES



Gentisate 1,2-dioxygenase from the gram-positive bacteria *Rhodococcus opacus* 1CP: identical active sites vs. different substrate selectivities

Natalya M. Subbotina[a,#], Alexey M. Chernykh[a,#], Anton I. Taranov[b], Anna D. Shebanova[a], Olga V. Moiseeva[a], Marta Ferraroni[c], Marina P. Kolomytseva[a,*]

[a]*Laboratory of enzymatic degradation of organic compounds, G.K. Skryabin Institute of Biochemistry and Physiology of Microorganisms of RAS, Federal Research Center, Pushchino Center for Biological Research of the Russian Academy of Sciences, Prospekt Nauki 5, Pushchino, Moscow Region, 142290, Russia*

[b]*Pushchino State Institute of Natural Sciences, Prospekt Nauki 3, Pushchino, Moscow Region, 142290, Russia*

[c]*Dipartimento di Chimica "Ugo Schiff", Università di Firenze, Via della Lastruccia 3, I-50019 Sesto Fiorentino, FI, Italy*

[#]These authors have contributed equally to the work

*Corresponding author.

*Corresponding author:

E-mail address: mkolomytseva@rambler.ru;

Keywords: *Rhodococcus*, 3-hydroxybenzoate, gentisate 1,2-dioxygenase, cupin, structure and function relationship, substrate specificity

Abstract

Gentisate 1,2-dioxygenases belong to the class III ring-cleaving dioxygenases catalyzing key reactions of aromatic compounds degradation by aerobic microorganisms. In the present work, the results of complete molecular, structural, and functional investigations of the gentisate 1,2-dioxygenase (*rho*-GDO) from a gram-positive bacterium *Rhodococcus opacus* 1CP growing on 3-hydroxybenzoate as a sole source of carbon and energy are presented. The purified enzyme showed a narrow substrate specificity. Among fourteen investigated substrate analogs only gentisate was oxidized by the enzyme, what can be potentially applied in biosensor technologies. The *rho*-GDO encoding gene was identified in the genomic DNA of the *R. opacus* 1CP. According to phylogenetic analysis, the *rho*-GDO belongs to the group of apparently most

recently acquired activities in bacterial genera *Rhodococcus*, *Arthrobacter*, *Corynebacterium*, *Nocardia*, *Amycolatopsis*, *Comamonas*, and *Streptomyces*. Homology modeling the *rho*-GDO 3D-structure demonstrates the composition identity of the first-sphere residues of the active site of *rho*-GDO and salicylate 1,2-dioxygenase from *Pseudaminobacter salicylatoxidans* (RCSB PDB: 2PHD), despite of their different substrate specificities. The phenomenon described for the first time for this family of enzymes supposes a more complicated mechanism of substrate specificity than previously imagined, and makes the *rho*-GDO a convenient model for a novel direction of structure-function relationship studies.

1. Introduction

Gentisate 1,2-dioxygenase (GDO, EC 1.13.11.4) is one of the key enzymes of the catabolism of a wide range of aromatic substrates such as: benzoate and its substituted derivatives, salicylate, 3-phenylpropionate, dibenzothiophene, naphthalene, phenanthrene, and anthracene (**Table 1**) [1-18]. GDOs catalyze aromatic ring fission between a hydroxyl and an adjacent carboxyl groups of substrate molecules, such as gentisate as well as methyl-, ethyl-, isopropyl-, halogen substituted gentisates, 1,4-dihydroxy-2-naphthoic acid and salicylate [1, 9, 11, 12]. GDOs belong to the class III bicupin dioxygenases, which are characterized by a six-stranded β-barrel fold and conserved amino acid motifs that provide the 3His ligand environment of a divalent metal ion, usually a Fe(II) ion [19, 20].

Table 1

Currently, about twenty enzymes with GDO activity have been isolated from different bacterial sources (**Table 1**) and partially characterized. Among them there is an atypical GDO (salicylate 1,2-dioxygenase (SDO)) isolated from the gram-negative bacterium *Pseudaminobacter salicylatoxidans* BN12 which is able to oxidize the ring fission not only of gentisate, but also of (substituted) salicylate(s) and a broad range of monohydroxylated substrates including 1-hydroxy-2-naphthoate [3, 4, 19, 21, 22]. Very specific GDOs with unusual selectivity were also purified and characterized. For example, GDO from the gram-negative bacterium *Xanthobacter polyaromaticivorans* 127W oxidized mainly 1,4-dihydroxy-2-naphthoate and gentisate [13]. 5-Aminosalicylate 1,2-dioxygenase from the gram-negative bacterium *Comamonas* sp. strain QT12 was able to cleave gentisate and predominantly 5-aminosalicylate as substrates [18].

Up to now the 3D-structures of only three GDO-like enzymes from gram-negative bacteria have been described. These are include the 3D-structures of GDOs from *Escherichia coli* O157: H7 (RCSB PDB: 2D40, [23]), *Ruegeria* (*Silicibacter pomeroyi* DSS3 (RCSB PDB: 3BU7, [24]),

and *P. salicylatoxidans* BN12 (RCSB PDB: 2PHD, [19, 21]). At the same time, there are no data on the full structure of GDOs of gram-positive bacteria.

Previously we have shown that the gram-positive bacterium *Rhodococcus opacus* 1CP actively utilized 3-hydroxybenzoate and gentisate as the sole source of carbon and energy using gentisate 1,2-dioxygenase [25]. This work focuses on the study of the biochemical, kinetic and structural properties of the novel gentisate 1,2-dioxygenase from this organism.

2. Materials and methods

2.1. Chemicals

3-Hydroxybenzoate and the reagents used to prepare culture media were from Reakhim (Russia). Gentisate, *β*-mercaptoethanol, 2,3-, 2,4- and 3,4-dihydroxybenzoates were from Fluka (Switzerland). Catechol, 4-methyl- and 4-chlorocatechol, protocatechuate, 2,6-dihydroxybenzoate, homogentisate, salicylate, hydroxyquinol, ammonium persulfate, Tris, protein markers for SDS-polyacrylamide electrophoresis (SDS-PAGE) and gel-filtration were obtained from Sigma (USA). Agarose was from Helicon (Russia), acrylamide and bis-acrylamide were from Dia-m (Russia), Coomassie brilliant blue R-250 was from Bio-Rad (USA). Ethanol was HPLC-purity from Sigma (USA).

2.2. Microorganism and growth conditions

R. opacus 1CP (VKM Ac-2638) biomass for gentisate 1,2-dioxygenase (*rho*-GDO) purification was obtained during the cell cultivation in a 10-l bioreactor (pO_2 = 80, under shaking 350 rpm at 29 ^{o}C) with 7 l of mineral media [26] and 3-hydroxybenzoate (up to 1.2 g/h) as a sole source of carbon and energy. Substrate exhaustion was monitored by the decrease in oxygen uptake. The cells were harvested at the exponential growth phase by centrifugation at 3000 g, washed twice with 50 mM Tris/HCl containing 10 % ethanol (v/v), pH7.2, and stored at -20 ^{o}C.

2.3. Preparation of crude extract and enzyme purification

Frozen cells were resuspended in an equal volume of 50 mM Tris/HCl with 10% ethanol (v/v), pH7.2, and disrupted by French press. After the treatment with 0.01 mg/mL DNAase (Sigma, USA), the cell extract was clarified by centrifugation for 30 min at 27000 g. The supernatant was used as a crude extract for further investigations.

All the procedures of the *rho*-GDO purification were performed in the presence of 50 mM Tris/HCl pH7.2 with 10% ethanol (v/v). *rho*-GDO was purified from crude extract of *R. opacus* 1CP cells using subsequent steps: anion-exchange chromatography on QAE-Toyopearl (30 mL volume, Supelco (Japanese)) with a linear gradient of NaCl (0.0-0.5 M), hydrophobic chromatography on Butyl Sepharose Fast flow (30 mL volume, Sigma-Aldrich (Germany)) with a linear gradient of $(NH_4)_2SO_4$ (0.6-0.0 M), and gel-filtration on Superdex 200 prep grade (Pharmacia Biotech (Switzerland)). During each step fractions with GDO activity were joined,

concentrated and desalted using ultrafiltration unit with Millipore membrane (10 kDa).

Protein concentration was determined by the Bradford method [27] with BSA as a standard.

2.4. Enzyme assay and analysis of the kinetic data

The *rho*-GDO activity was measured spectrophotometrically with a Shimadzu UV-160 spectrophotometer (Japan) in 1 cm quartz cuvettes at 25 ^{o}C with 0.33 mM gentisate as a substrate at 334 nm [28]. The reaction mixture contained also air-saturated 50 mM Clark and Lubs phosphate buffer at pH8.0. The reaction was started by adding the enzyme.

Apparent K_m and V_{max} values were calculated from 1/V *versus* 1/[S] plots [29]. Apparent k_{cat} values were determined using molecular masses of the enzymes. Specific rate of the *rho*-GDO with gentisate was determined using molar extinction coefficients for maleylpyruvate - 10.8 $mM^{-1}см^{-1}$ [28]. Activity of *rho*-GDO with the other substrate analogs was determined using the molar extinction coefficients of the corresponding reaction products [3, 13] as well as was concurrently investigated by monitoring spectrum dynamics of the reaction products in the range of 200-800 nm.

One unit of enzyme activity was defined as the amount of enzyme catalyzing the formation of 1 micromole of product per minute under the described conditions. Inhibition constants were determined graphically utilizing the method of Lineweaver and Burk [29] using three fixed gentisate concentrations: 40 μM, 80 μM and 160 μM. The enzyme activity was measured at least 3 times for each of 10 increasing concentrations of inhibitor for each gentisate concentration.

For calculation of each kinetic constant, the data of at least three experiments were used. All kinetic constants were obtained with an error <10%.

2.5. Enzyme characterization

The pH effect on the reaction rates of the *rho*-GDO was studied by determining the enzyme activities in 40 mM acetate-phosphate-borate/NaOH buffer (APB) in the range of pH5.6-10.0 [30], 50 mM tris/HCl buffer (pH7.2-9.0), 50 mM tris/H_2SO_4 buffer (pH7.2-9.0), and 50 mM Clark and Lubs phosphate buffer (pH6.0-8.0) [31]. The temperature effect on the enzymatic reaction rate was determined in the range of 15–60°C in 50 mM Clark and Lubs phosphate (pH8.0) using a Shimadzu TCC-controller (Japan).

Molecular mass of the *rho*-GDO was determined by 12% SDS-PAGE according to the modified Laemmli method [32]. Low molecular weight kit SigmaMarker (6,500-66,000 Da) was used as a standard.

For estimation of molecular mass under non-denaturing conditions, the protein was applied to a Superdex 200 column (110 mL, "Pharmacia Biotech", Switzerland) and eluted with 50 mM tris/HCl (pH7.2) buffer containing 10% ethanol (v/v) and 0,1 M NaCl. The molecular mass was calculated using a standard linear regression curve of the reference proteins (29,000-700,000 Da, Sigma (USA)).

2.6. Bacterial strains, plasmids and growth conditions

For DNA manipulations, *R. opacus* 1CP and *E. coli* strains were grown at 29°C and 37 ^{o}C, respectively, in LB media [33] under shaking at 200 rpm. *E. coli* DH5α and *E. coli* BL21(DE3) were obtained from Invitrogen (USA) and Novagen (USA), respectively. Culture growth was measured using UV-visible spectrophotometer Shimazu UV-160 at 545 nm.

Plasmid pGEM-T Easy (Promega, USA) was used for cloning the wild type of *rho*-GDO gene (*mhbD*). For expression of *mhbD* gene the pET22b(+) vector (Novagen, USA) was applied.

2.7. DNA manipulation, cloning, and sequencing

Genomic DNA from *R. opacus* 1CP was isolated by modified method using Genomic DNA Purification Kit (Fermentas). Cells were incubated in lysis solution for 1 h, following the DNA purification according to the manufacturer protocol. Purification of plasmid DNA was carried out using Plasmid Miniprep kit (Eurogene (Russia)). The quality and quantity of the extracted DNA were determined by the NanoDrop 2000 spectrophotometer (ThermoFisher).

The search of homological nucleotide sequences was performed by BLAST [34]. Alignment of the nucleotide sequences of 3-hydroxybenzoate degradation gene cluster of *R. jostii* RHA1 (GenBank: CP000431.1), *R. opacus* B4 (GenBank: AP011115.1), *Rhodococcus* sp. TFB (GenBank: JF754465.1) was performed using the MEGA6 software [35]. To design the primers GDO1f and GDO1r (Supplementary Tables S1) for highly conserved regions flanking the gentisate-1,2-dioxygenase genes of *R. jostii* RHA1, *R. opacus* B4, *Rhodococcus* sp. TFB (GenBank JF754465.1) the Primer3Plus web service was used [36].

Amplification of a part of gene cluster containing the complete sequence of the *mhbD* gene encoding the gentisate-1,2-dioxigenase from *R. opacus* 1CP was performed using the ScreenMix kit (Evrogen, Russia) and GDO1f and GDO1r primers. PCR mixture (25 µl) contained 0.8 µM of

each primer, 1 µl of the template genomic DNA, 5 µl ScreenMix kit solution (Evrogen, Russia). The following PCR program was used: an initial denaturation (95°C, 3 min) was followed by 33 cycles consisting of a denaturation step (94°C, 30 s), an annealing step (70°C, 30 s), and an extension step (72°C, 60 s). The final extension step was done at 72 °C for 10 min. Isolation of PCR-products from agarose gel was performed using Cleanup Standard kit (Evrogen, Russia). The resulting PCR product was ligated into pGEM-T Easy vector (Promega, USA) and sequenced using the Evrogen commercial service (Russia).

The GDO2f and GDO2r primers (Supplementary Tables S1) designed from the obtained genomic DNA sequence were used for amplification of full-length *mhbD* gene coding *rho*-GDO. The GDO2f primer contained a start codon and restriction site of SacI. The GDO2r primer contained a stop codon and HindIII restriction site.

PCR mixture (20 µl) contained 0.5 µM of each primer, 1 µl of genomic DNA, 10 µl KAPA HiFi HotStart ReadyMix (Roche). The following PCR program was used: an initial denaturation (95°C, 3 min) was followed by 30 cycles of a denaturation step (98°C, 20 s), an annealing step (69°C, 20 s), and an extension step (72°C, 60 s). The final extension step was done at 72°C for 5 min.

The pET22-*mhbD* plasmid was obtained by cloning the *mhbD* gene, coding *rho*-GDO into the pET22b (+) expression vector (Novagen) at the SacI and HindIII restriction sites.

Restriction, ligation, preparation of competent cells of *E. coli* strains and their subsequent transformation were processed according to standard protocols [37].

2.8. Expression procedure

The *rho*-GDO was expressed in *E. coli* strain BL21 (DE3) cells transformed with the pET22-*mhbD* plasmid. 10 mL of the cell culture was incubated overnight at 37°C with agitation in LB medium in the presence of 100 µg/mL ampicillin (Sigma-Aldrich). Then, 0.15 mL of the obtained inoculum was transferred to 100 mL of LB medium containing 100 µg/mL ampicillin and cultured at 37°C with agitation until the optical density OD_{545} was equal to 0.7. The expression of the *mhbD* gene was induced by 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 37 °C. After incubation for 3.5 h, the cells were pelleted by centrifugation at 3000 g, washed and resuspended in 5 mL of 50 mM Tris/HCl with pH7.2. Cell disruption was performed using an MSE ultrasonic disintegrator (UK) for 20 min at 200 kHz.

GDO-activity in the obtained crude extract was measured as described above. *E. coli* BL21(DE3) strain, not transformed with pET22-*mhbD*, was used as a control.

2.9. Nucleotide sequence accession number

The GenBank accession number of the sequence reported in this paper is KF806993 (AHH30829).

2.10. Phylogenetic analysis

All amino acid sequences of bacterial GDOs used in the present studies were obtained from GenBank, RCSB PDB, NCBI, and UniProtKB/Swiss-Prot databases. The alignment was performed with the MEGA6 software [35] using the ClustalW algorithm (default parameters). The tree was also constructed by MEGA6 using the Neighbor-Joining statistical method and the Poisson correction method with a number of bootstrap replications equal to 1000. All positions containing gaps and missing data were eliminated.

2.11. rho-GDO theoretical model

Homology modeling of *rho*-GDO was performed using Modeller software [38]. The crystal structures of salicylate 1,2-dioxygenase from *P. salicylatoxidans* (RCSB PDB: 2PHD) and salicylate 1,2-dioxygenase from *P. salicylatoxidans* W104Y mutant (PDB ID: 4FAG) were used as the templates for modeling *rho*-GDO. Tetrameric structure of 4FAG was modeled with UCSF Chimera [39] and Modeller [38] software. The best model was selected by means of DOPE assessment score implemented in the Modeller software [40]. Verification of the model was done with ProCheck v.3.4 [41], WHAT_CHECK [42] and UCSF Chimera [39]. In addition, the root mean square deviation (RMSD) between the main chain atoms of the model and the templates was calculated for the reliability of the model. Alignment of protein sequences was performed by MEGA6 software using Muscle algorithm (default parameters) [35].

The *rho*-GDO model data are available in the PMDB database under the accession number **PM0082625**.

3. Results and discussion

3.1. Purification and characterization of the native rho-GDO

Electrophoretically homogenous *rho*-GDO was isolated from the cells of *R. opacus* 1CP utilizing *m*-hydroxybenzoate as a sole source of carbon and energy, using anion exchange and hydrophobic chromatography, and gel-filtration step (Supplementary Tables S2 and Figure S3). The purified enzyme with specific activity of 108 U/mg had 177±2 kDa molecular mass and 41±1 kDa subunit mass. The enzyme was unstable in crude extract, but restored activity during subsequent purification (Supplementary Tables S2 and Figure S3).

Thus, *rho*-GDO was a tetramer like the majority of known GDOs from gram-negative and gram-positive strains **(Table 1**). *rho*-GDO belongs to the group of relatively high molecular mass enzymes, which property is not connected to their taxonomic affiliation (**Table 1**).

rho-GDO was the most active at 40-45^{o}C. pH optimum of the enzyme was observed in slightly alkaline conditions: in 50 mM Tris/HCl buffer – at pH8.0-8.3, in 50 mM Tris/H_2SO_4 buffer - at pH7.6-8.0, in 40 mM APB - at pH7.5-8.5, and in 50 mM Clark and Lubs phosphate buffer - at pH8.0 (Table 1). This is also typical to all known dioxygenases cleaving gentisate. In Clark and Lubs phosphate buffer, the enzyme activity was found 10 times higher than in other buffers. In subsequent experiments, this buffer was chosen to measure the enzyme kinetic parameters.

3.2. Kinetic properties of rho-GDO

The purified *rho*-GDO has a narrow substrate specificity (**Table 2**). Among all the investigated substrate analogs, containing a carboxyl- and adjacent hydroxyl-substituents, only gentisate was oxidized by the enzyme. It is interesting to note that *rho*-GDO has a different selectivity concerning the hydroxy-substituted salicylates as competetive inhibitors. Thus the absence of a hydroxyl-substituent at position 5 in salicylate led to a loss of catalytic activity and reduction of the strength of its binding to the *rho*-GDO active site (K_i=810 μM). Salicylate analogues with a hydroxyl-substituent at position 3 or 4 in the aromatic ring of the salicylate molecule, excepting position 5 (gentisate), were not substrates of the enzyme and had a reduced inhibitory effect (K_i values – 1069 and 1130 μM correspondingly). Nevertheless, the 6-hydroxyl substituted salicylate had inhibitory effects on enzyme catalysis and was more strongly bound to *rho*-GDO active site. A surprising effect was observed for 5-chlorosalicylate: this compound was

not a substrate (a competitive inhibitor), however it was bound strongly to the *rho*-GDO active site (K_i=57 μM) as compared with apparent gentisate binding (K_m=40 μM). Thus, the presence of a 5-hydroxy-substituent in the aromatic ring of salicylate (gentisate) is required for the performance of *rho*-GDO catalysis.

Table 2

Concurrently, another series of substrate analogs containing two adjacent hydroxyl substituents as reactive groups was tested: catechol, 4-methyl- and 4-chlorocatechols, hydroxyquinol (1,2,4-trihydroxybenzene) and protocatechuate (3,4-dihydroxybenzoate). With the exception of protocatechate, all tested compounds were competitive inhibitors. Catechol and especially its 4-methyl and 4-chlorosubstituted derivatives were not substrates but they were very strong inhibitors. The presence of hydroxyl- or carboxyl-substituent at 4-position of catechol ring led to a weak inhibitory effect up to disappearance of *rho*-GDO inhibition (protocatechuate). It can be concluded that the hydroxyl group at exactly position 5 in the salicylate-like molecule (gentisate) is important both for the catalytic process and for binding to the *rho*-GDO-active site, while at the same position in molecules containing two hydroxyls as adjacent reactive groups (substituted catechols) decrease the binding of the substrate analogues. Thus, the presence of the carboxyl and hydroxyl groups at positions 2 and 5, respectively, of the substrate molecule is mandatory for the *rho*-GDO catalysis.

Unfortunately, the kinetic and inhibition properties only of a few enzymes of all known GDOs (**Table 1**) have been investigated. The majority of the gentisate 1,2-dioxygenases isolated and characterized so far has a rather narrow substrate specificity (**Table 1**), which was determined using gentisate and its derivatives as substrates. For example, GDOs from *M. osloensis* [11], *P. (Comamonas) testosteroni* [9], *P. (Delftia) acidovorans* [9], *R. erythropolis* S1 [1], *Sphingomonas* sp. strain RW5 [7], *P. alcaligenes* NCIB 9867 and *P. putida* NCIB 9869 [12, 43] were active with gentisate and its methyl-, ethyl, bromo-, and chlorosubstituted derivatives. At the same time, the highest dioxygenase activity was observed with respect to gentisate. The exception was GDO from *P. alcaligenes* P25X which exhibited the highest activity with 3-methylgentisate as a substrate [12].

As a rule, GDOs are weakly active with structural analogues of gentisate: 2,4-, 2,3-, 2,6-, 3,5-dihydroxybenzoates and salicylates [1, 3, 4]. The exception is the GDO from *P. (Delftia) acidovorans* that is inhibited by hydroxybesoates like *rho*-GDO (Table 2) [44]. At the same time, this enzyme showed weak activity with hydroxyquinol, in contrast to *rho*-GDO (Table 2) [44].

Moreover, some GDOs make cleavage of hydroxynaphthoates. GDOs from *P. (Comamonas) testosteroni, P. (Delftia) acidovorans,* and *X. polyaromaticivorans* 127W were

active with 1,4-dihydroxy-2-naphthoate [9, 13], GDOs from *P. salicylatoxidans* and from *X. polyaromaticivorans* 127W – with 1-hydroxy-2-naphthoate [3, 13].

As well as *rho*-GDO the GDO from *Corynebacterium glutamicum* ATCC 13032 was not active with salicylate, 1-hydroxy 2-naphthoate, 3-hydroxysalicylate [6]. At the same time, this enzyme did not cleave methyl-, amino-, bromo- and fluorosalicylates [6]. Unfortunately, the inhibitory effects of the above compounds on the GDO from *C. glutamicum* ATCC 13032 were not evaluated, as well as catechol-like compounds, 5-chlorosalycilate, and 2-hydroxy-3-naphthoate were not tested as substrates, which greatly complicates the comparative analysis.

Recently, 5-aminosalicylate 1,2-dioxygenase from *Comamonas* sp. strain QT12 was purified and characterized. The enzyme was able to cleave gentisate and predominantly 5-aminosalicylate as substrates [18].

Unfortunately, there are no data concerning substrate specificity of GDOs with known 3D-structure from *E. coli* O157:H7: (*ec*-GDO; RCSB PDB: 2D40, [23]) and *R.* (*Silicibacter) pomeroyi* DSS3 (*rp*-GDO; RCSB PDB: 3BU7, [14, 24]) that complicates an insight into the nature of GDO function.

Thus, *rho*-GDO is a rather unusual enzyme among the GDOs possessing a relatively narrow spectrum of substrates.

3.3. Cloning and expression of rho-GDO

As a rule, the genes of rhodococci encoding the enzymes decomposing 3-hydroxybenzoate through gentisate are assembled into clusters. The *Rhodococcus* strains (*R. jostii* RHA1, *R. opacus* B4, and *Rhodococcus* sp. TFB) for which the complete nucleotide sequences of such gene clusters are known were selected as closest to *R. opacus* 1CP strain. Multiple alignment of their sequences was carried out and the highly conserved regions flanking the genes of gentisate 1,2-dioxygenases were found. Based on these regions, the primers GDO1f and GDO1r were designed to amplify and sequence a part of gene cluster (1253 bp) containing the complete *mhbD* gene of gentisate 1,2-dioxygenase from *R. opacus* 1CP.

The obtained open reading frame of 1077 bp and encoded a protein containing 358 amino acid residues. The predicted molecular weight of the protein was 39.6 kDa. BLAST analysis confirmed a high degree of homology of the obtained sequence with the known gene sequences of GDOs of rhodococci. The GDO genes from *R. opacus* PD630 (AHK33220, 94.26%), *R. jostii* RHA1 (ABG93677, 93.16%) and *R. opacus* R7 (AII05708, 92.42%) showed the highest identity with the *mhbD* gene of *R. opacus* 1CP. The nucleotide sequence of the *mhbD* gene coding *rho*-

GDO and the amino acid sequence of the *rho*-GDO from *R. opacus* 1CP (**Fig. 1**) have been deposited at GenBank accession numbers KF806993 and AHH30829, respectively.

Heterologous expression of the *mhbD* gene encoding the gentisate-1,2-dioxygenase of *R. opacus* 1CP in *E. coli* strain BL21(DE3) allowed to detect the GDO activity in the cell-free extract of *E. coli*. The typical spectral changes of gentisate transformation were observed. The peak at 334 nm indicated the formation of maleylpyruvate - the product of gentisate cleavage, which absorbs at this region of the spectrum (data not shown). In addition, the recombinant enzyme was not active with salicylate and 1-hydroxy-2-naphthoate. The activity of recombinant mhbD with gentisate in the cell-free extract was 0.0016 U/mL.

Fig. 1 (2-column fitting image)

3.4. Amino acid alignment of rho-GDO and evolutionary relationship

Phylogenetic analysis of the amino acid sequence of GDO from *R. opacus* 1CP obtained by translation of nucleic acid sequence of the identified *mhbD* gene (**Fig. 1**), as well as the currently characterized biochemically or structurally GDOs, and the GDOs from annotated genomes of other bacteria showed formal division into 5 separate groups. Each of them contained GDOs with confirmed properties, which indicates the belonging of the majority of the selected annotated proteins to the class of GDOs (**Fig. 2**).

Fig.2 (2-column fitting image)

The most distant group of GDOs are the enzymes of the archaea of the genus *Haloferax* and the actinobacteria of the genus *Mycobacterium* as well as 5-aminosalicylate 1,2-dioxygenase from proteobacteria *Comamonas* sp. strain QT12 (**Fig. 2**, group 1). The second group is formed by high molecular weight GDOs of gram-positive (*Rhodococcus*, *Bacillus*, *Mycobacterium*, and *Streptomyces* genera) and gram-negative (*Silicibacter* (*Ruegeria*) and *Sphingomonas* genera) bacteria (**Fig. 2**, group 2). The third branch of the tree is formed by low molecular weight GDOs of gram-negative bacteria of the genera *Delftia*, *Acinetobacter*, *Polaromonas*, and *Ruegeria* (**Fig. 2**, group 3).

The largest young branch of the phylogenetic tree is formed by two types of GDOs - low molecular weight gram negative GDOs (Figure 1, 4) and for the most part low molecular weight and high molecular weight gram-positive GDOs (Figure 1, 5). The exception is a few GDOs of gram-negative bacteria of the genera *Pseudoaminobacter* and *Comamonas*, which fall into the group of gram-positive GDOs (**Fig. 2**, group 5).

The *rho*-GDO described in this paper belongs to the group of the evolutionarily youngest enzymes (**Fig. 2**, group 5). The interesting fact is the presence of at least two types of GDOs in bacteria of the genus *Rhodococcus*: the evolutionarily earlier group of GDO-like enzymes with a large numbers of amino acids (365-382 aa) and the youngest group of smaller enzymes (356-359 aa) of the most gram-positive bacteria. Unfortunately, to date, there are no reports about the GDOs from bacteria of the *Rhodococcus* genus as well as their properties.

The enzyme, evolutionary most closely related to the high molecular weight GDO-like proteins of the *Rhodococcus* genus, is GDO with a known three-dimensional structure, isolated from the gram-negative bacterium *S. (Ruegeria) pomeroyi* DSS3 (RCSB PDB: 3BU7, [14, 24]) (**Fig. 2**, group 2).

Relatively close enzymes to the *rho*-GDO, isolated in this work, are salicylate 1,2-dioxygenase of known 3D-structure from *P. salicylatoxidans* (*ps*-SDO, RCSB PDB: 2PHD, 3NL1, [19, 22, 45]) and the purified and biochemical characterized GDO from gram-positive bacteria *C. glutamicum* ATCC13032 [5, 6], belonging to the group of evolutionarily youngest GDOs (group 5).

Alignment of the amino acid sequences of GDOs demonstrates high identity of *rho*-GDO and the low-molecular GDOs of *Rhodococcus* bacteria (> 96% identity), but the identity of the *rho*-GDO and GDOs of known crystal structure is less than 70% (**Fig. 1, 2**): with SDO from *P. salicylatoxydans* BN12 (RCSB PDB: 3NL1) - 66.5%, with GDO from *E. coli* O157: H7 (UniProtKB/Swiss-Prot: Q8X655) - 34.2%, and with GDO from *R. pomeroyi* DSS3 (RCSB PDB: 3BU7) - 29.9% identity. It is worth to note that the crystal structure is available only for GDOs from gram-negative bacteria.

According to the alignment the amino acid composition of *rho*-GDO active site is very similar to that of salycilate 1,2-dioxygenase from gram-negative *P. salicylatoxidans* (Fig. 1) which conversely shows a very broad substrate specificity [22]. This fact does not fit into the overall picture of the understanding of structural-functional relationships and encouraged us to perform further more rigorous investigations of *rho*-GDO structure using computational approaches.

3.5. Homology model of rho-GDO 3D-structure

According to the obtained homology modeling data the *rho*-GDO molecule was assembled from 4 subunits bound trough α-helixes, with each two subunits of the tetramer interacting through their N-terms forming a dimer (**Fig. 3ab** (1-3)). The dimer of the tetrameric structure is stabilized by two type of contacts: the first, between the N-terminal α-helixes of two adjacent

subunits, are formed by residues Ala11-Ser22 of one subunit and the α-helix formed by Ser52-Leu64 of another one (**Fig. 3ab** (1)); the other type of inter-subunit contacts are mediated through two nearest α-helixes of each subunit formed by Thr174-Ile168 and Asp340-Leu347 (**Fig. 3ab** (2)). The inter dimer contacts are made through α-helixes containing Trp225-Glu241 residues and Arg196-Thr202 residues of each subunit (**Fig. 3ab** (3)).

Fig. 3 (2-column fitting image)

Each subunit of the *rho*-GDO contains two cupin motifs with one catalytic ferrous located in the N-terminal cupin active site cavity (**Fig. 3ab**), which is oriented outwards in the tetrameric structure of GDOs and thus is accessible to the substrate molecule [46]. The secondary structure of the *rho*-GDO C-terminal cupin is similar to the N-terminal cupin domain. However, unlike the *R. pomeroyi* gentisate 1,2-dioxygenase structure (*rp*-GDO, RCSB PDB: 3BU7 [24]) the cavity formed by the ß-fold of the *rho*-GDO C-terminal cupin domain is unable to hold a ferrous ion because of the absence of coordinating residues (Asp287/His292, Arg285/His290, and Ser326/His330 in the structure of *rho*-GDO/*rp*-GDO, correspondingly) (**Fig. 4A and Fig. 1**). Moreover, the amino acid compositions of the C-cupin domain cavities of *rho*-GDO and *rp*-GDO do not match completely to each other, as is shown by the amino acid sequence alignment (**Figure 1, 4**). Thus, the *rho*-GDO, like the gentisate 1,2-dioxygenase (*ec*-GDO) from *E. coli* O157:H7 [23] and the salicylate 1,2-dioxygenase (*ps*-SDO) from *P. salicylatoxidans* [19], does not have a second C-cupin active site as in *rp*-GDO.

Fig. 4 (2-column fitting image)

3.6. Structure of the active site of rho-GDO and comparison with the active sites of other GDO-like enzymes

Ferrous ion of *rho*-GDO N-terminal cupin domain is coordinated by His111, His109, and His150 (**Fig. 1, 4C, 5**), what is common to all other GDOs [21, 23, 24, 46]. Arg73, Gln98, Arg117, His152, Trp162, and Asp164 residues are also conserved in all gentisate 1,2-dioxygenases (**Fig. 1, 4C, 5**) and are considered to participate in binding to a substrate molecule and in the subsequent acid-base catalysis [46]. His152 of *rho*-GDO active site corresponding to His162 of *ps*-SDO, His162 *rp*-GDO and His147 of *ec*-GDO is supposed to act as proton donor to the evolving peroxidate intermediate during catalysis (**Fig. 4C, 5**) [21, 23, 24, 46]. Asp164 and Gln98 residues of *rho*-GDO, which correspond to Asp174/Asp175/Asp159 and Gln108/Gln108/Gln93 in *ps*-SDO/*rp*-GDO/*ec*-GDO structures, accordingly, have been

implicated in the deprotonation of the 5-hydroxyphenyl group of gentisate [21, 23, 24, 46] (**Fig. 1, 4C, 5**). Arg73, Arg117, and Trp162 of the *rho*-GDO active site correspond to Arg83, Arg127, and Trp172 of the *ps*-GDO where they stabilize the substrate carboxyl group coordinated to the iron ion [19] (**Fig.** 5A).

Fig. 5 (2-column fitting image)

The other residues covering the inner surface of the *rho*-GDO active site cavity (Glu72, Ala75, Trp94, Gln113, Ala115, Thr144, Leu166, Ile168, Phe179) are identical to the relevant residues of salicylate 1,2-dioxygenase from *P. salicylatoxidans* (**Fig.** 5A, [19, 21]. This is also supported by amino acid sequence alignment (**Fig. 1**). It is also interesting that for the GDO from *C. glutamicum* ATCC 13032, which has catalytic properties similar to *rho*-GDO [6] and belongs to the same group of enzymes (Fig. 2), identical residues are in the relevant positions as the residues forming the cavity of the *rho*-GDO active center (Fig. 1). From the other side, the residues of the GDO from *P. (Delftia) acidovorans* that are similar to the *rho*-GDO (Table 2, [44]) and belongs to another phylogenetic group of GDOs (**Fig. 2**) do not completely coincide at the relevant positions (Fig. 1).

Residue Trp94 of *rho*-GDO corresponds to Trp104 of *ps*-SDO, which is able to form a hydrogen bond with the 5-phenolate of gentisate as discussed earlier [19, 45]. Thus, this residue is important for the specificity towards gentisate and 5-substituted substrate analogues. Salicylate is only a weak inhibitor of *rho*-GDO as well as catechol and its analogues (**Table 2**), whereas in the case of *ps*-SDO salicylate is a substrate even if the catalytic activity is sharply reduced compared to gentisate. Therefore, probably Trp is important for interaction with 5-hydroxyl of gentisate but is not determinant for the inability of *rho*-GDO to cleave substrate analogues, not containing hydroxyl-substituent at position 5.

3.7. Differences in the second-sphere residue composition of the active sites of rho-GDO and ps-SDO

As mentioned above, the residues covering the inner surface of the active site cavities of *rho*-GDO and *ps*-SDO are identical (**Fig. 5A**), whereas the enzymes have very different substrate specificity (**Table 2**). Thus, other factors determining difference in substrate specificity of the enzymes should be investigated.

In the second-sphere residues of the *rho*-GDO active site there are three positions differing from the *ps*-SDO structure, deserving consideration (**Fig. 6**). Ala96/Gly106, a residue not directly interacting with the substrate, but able to change conformation of the side chain of

Trp94/Trp104, is important for the substrate binding [22], Ala112/Ser122 interacting with His111/His121, one of the three ferrous ion coordinating residues, and His153/Met163 interacting with His152/His162, are important for acid-base catalysis [22].

Fig. 6 (1-column fitting image)

In the first case, the alanine residue of *rho*-GDO active site is at a position corresponding to Gly106 of *ps*-SDO. The structure of the Gly106Ala mutant of *ps*-SDO indicates that the mutation induces a change in orientation of the Trp104 side chain that causes an unproductive binding of salicylate and 1-hydroxy-2-naphthoate molecules to the active site [22]. Like the Gly106Ala mutant of *ps*-SDO, the investigated *rho*-GDO was not active with salicylate and 1-hydroxy-2-naphthoate. Unfortunately, there is no data about substrate specificity of the *rp*-GDO and *ec*-GDO of known structure, containing Gly at this position (**Fig. 4C** and **5B**). Nevertheless, the GDO from *X. polyaromaticivorans* 127W also containing Gly at this position is active with 1-hydroxy 2-naphthoate as the native *ps*-SDO (**Fig. 1**, [13]). At the same time, the relevant mutation Ala→Gly introduced into the GDO from *C. glutamicum* ATCC 13032 as well as Ala→Ser, Ala→Ile, and Ala→Asp mutations led to a weak ability of mutant enzyme to convert salicylate and several other monohydroxylated substrates [6].

The other interesting position is the neutral-polar residue Ser122 of *ps*-SDO, which takes part in the formation of an hydrogen bond network with Gln123 and Asp44 of another subunit maintaining the enzyme tetrameric structure (RCSB PDB:2phd), that is replaced by the small hydrophobic residue Ala112 in *rho*-GDO structure, which is unable to form a similar hydrogen bond network. This can be the reason of an increased conformational instability observed for *rho*-GDO. An Ala residue is also present in GDOs of other bacteria of the *Rhodococcus* genus, whereas a neutral-polar Ser, Thr, Asn or Met residues are generally located at this position in GDOs of other bacterial genus (**Fig. 1**). Thus, Ser is located in the relevant position in GDO from *C. glutamicum* ATCC 13032 that exhibits properties relatively close to those of *rho*-GDO [6], suggesting no significant role in substrate selectivity of GDOs.

The replacement of the polar Met163 in *ps*-SDO structure by the basic His153 in *rho*-GDO can also influence on the location and function of the substrate carboxyl group and can be involved in the subsequent acid-base catalytic role played by His152/His162 residue of *rho*-GDO and *ps*-SDO correspondingly. The alignment of the amino acid sequence of bacterial GDOs shows that the most of the low molecular GDO-like enzymes from genus *Rhodococcus* (**Fig. 1, 2**) as well as from gram-positive strains (data not shown) have a histidine in this position. Contrariwise, the majority of GDOs from gram-negative strains has a glycine residue or hydrophobic/neutral-polar residues at the same position (**Fig. 1,** data not shown). Unfortunately,

518 it is difficult to determine the contribution of one or another residue at this position to the
519 enzyme catalytic activity due to the lack of kinetic data and to clarify this point further
520 investigations are needed.
521
522 *3.8. Differences in the entrance of the rho-GDO and ps-SDO active sites*
523
524 Another interesting region in terms of differences between *rho*-GDO and *ps*-SDO substrate
525 specificity is the enzyme active site cavity entrance. In this region, two residues are different:
526 Ala77/Gly87 and Pro87/Ala97 in *rho*-GDO/*ps*-SDO structures correspondingly (**Fig. 6**).
527 Ferraroni and co-authors [45] previously showed that the side chain of Trp104 of the native *ps*-
528 SDO is oriented so that the hydrophobic part of its side chain is in an outside cavity position.
529 Binding of the substrate leads to a rotation of about 180^{o} around the CD-CB bond of Trp104 to
530 form a hydrogen bond with the 5-phenolate group of gentisate. The replacement of Gly87 and
531 Ala97 of *ps*-SDO by Ala77 and Pro87 in *rho*-GDO structure makes the surrounding of Trp94 of
532 *rho*-GDO (Trp104 in *ps*-SDO structure) more hydrophobic thus strengthening Trp94 residue in
533 the outside position and probably improving the affinity of the enzyme to binding of 5-
534 substituted substrates. According to this hypothesis, the obtained kinetic data of K_m/K_i for *rho*-
535 GDO with gentisate and 5-chlorosalicylate are lower than the corresponding values for *ps*-SDO
536 (**Table 2**).
537
538 *3.9. Differences in interchain interactions between rho-GDO and ps-SDO*
539
540 The 178-184 loops of each subunit of *ps*-SDO (residues 168-176 in *rho*-GDO) are the
541 regions which form close contacts between two adjacent subunits of each dimer in the enzyme
542 tetrameric structure(**Fig. 3 A2, Fig. 7**). In this loop, there are two noteworthy replacements:
543 His172/Gln182 and Tyr173/Gln183, in *rho*-GDO/*ps*-SDO structures correspondingly (**Fig. 7)**. In
544 the *ps*-SDO structure, Gln182 (**Fig. 7A**), takes part in the interactions between subunits (RCSB
545 PDB: 2PHD). This residue, located near the active site entrance, makes hydrogen bonds with the
546 Tyr98 side chain and with Asp44 of another subunit, contributing to the closure of the active site
547 upon substrate binding [19], thereby making the *ps*-SDO structure more compact (**Fig. 7A**).
548 **Fig. 7** (2-column fitting image)
549
550 In *rho*-GDO the replacement of the polar Gln182 by a histidine (His182), forming weaker
551 interactions with Asp34 than Tyr88, can make this region more unstable and affect the *rho*-GDO
552 activity.

The same applies to the Tyr173/Gln183 substitution in *rho*-GDO/*ps*-SDO, correspondingly (**Fig. 7**). In *ps*-SDO structure, the neutral-polar Gln183 residues of each subunit of the dimer make hydrogen bond with the same residue of another subunit. This interaction increases the compactness of the *ps*-SDO dimer (**Fig. 7A**), in contrast to *rho*-GDO, where in the same position there are two hydrophobic-aromatic Tyr173 residues (**Fig. 7B**).

4. Conclusions

Studies of gentisate oxidizing enzymes show that there is a curious phenomenon, consisting in the combined existence of different substrate specificity and the identity of the composition of their active sites, most likely obtained as a result of the exchange of a part of the genetic material between organisms of one taxonomic group and among different taxonomic groups.

This is the first report of a similar phenomenon for the class III ring-cleaving dioxygenases. A related effect was first discovered by us earlier on the example of the class I ring-cleaving (intradiol) dioxygenases. In fact, catechol 1,2-dioxygenase and 3-chlorocatechol 1,2-dioxygenase of *R. opacus* 1CP had the same composition of the active sites although showing different substrate specificities [47].

The results of our studies suggest a more complicated nature of the mechanism behind enzyme substrate specificity as commonly accepted, and should encourage researchers to more thoroughly investigate the relationships between enzymes and their substrates.

The unique gentisate 1,2-dioxigenase from the gram-positive strain *R. opacus* 1CP presented in this work, which has a narrow substrate specificity, has a high potential for use in biosensor technologies for monitoring the metabolism of tyrosine and salicylic acid by organisms, as well as a pollution of environment and different materials by polyaromatic compounds, the degradation of which is usually carried out through gentisate.

Acknowledgements

The work was supported by the RFBR (the research project №13-04-00880).

Conflict of interest

The authors declare that they have no conflict of interest with the contents of this article.

Author contributions

All authors contributed to this work by designing the study, conducting experiments, discussing the results and commenting on the manuscript. NS and MK developed the enzyme purification scheme and purified the enzyme; NS characterized the enzyme; NS and AT isolated and expressed the gene coding the enzyme; NS, AS and OM performed phylogenetic analysis; AC made homology modeling, AC, MF, and MK accomplished the enzyme structural analysis; MK conceptualized the study; MF and MK wrote the manuscript. All authors read and approved the manuscript for submission.

References

[1] A. Suemori, R. Kurane, N. Tomizuka, Purification and properties of gentasate 1,2-dioxigenase from *Rhodococcus erithropolis* S-1, Biosci. Biotech. Biochem. 57 (1993) 1781-1783. https://doi.org/10.1128/jb.178.12.3501-3507.1996

[2] P. Kiemer, B. Tshisuaka, S. Fetzner, F. Lingens, Degradation of benzoate via benzoyl-coenzyme A and gentisate by *Bacillus stearothermophilus* PK 1, and purification of gentisate 1,2-dioxygenase, Biol. Fertil. Soils 23 (1996) 307-313. https://doi.org/10.1007/BF00335959.

[3] J.-P. Hintner, C. Lechner, U. Riegert, A.E. Kunhm, T. Storm, T. Reemtsma, A. Stolz, Direct ring fission of salicylate by a salicylate 1,2-dioxygenase activity from *Pseudaminobacter salicylatoxidans*, J. Bacteriol. 183 (2001) 6936–6942. https://doi.org/10.1128/JB.183.23.6936-6942.2001.

[4] J.-P. Hintner, T. Reemtsma, A. Stolz, Biochemical and molecular characterization of a ring fission dioxygenase with the ability to oxidize (substituted) salicylate(s) from *Pseudaminobacter salicylatoxidans*, J. Biol. Chem. 279 (2004) 37250–37260. https://doi.org/10.1074/jbc.M313500200.

[5] X.-H. Shen, C.-Y. Jiang, Y. Huang, Z.-P. Liu, S.-J. Liu, Functional identification of novel genes involved in the glutathione-independent gentisate pathway in *Corynebacterium glutamicum*, Appl. Environ. Microbiol. 71 (2005) 3442–3452. https://doi.org/10.1128/AEM.71.7.3442-3452.2005.

[6] E. Eppinger, A. Stolz, Expansion of the substrate range of the gentisate 1,2-dioxygenase from *Corynebacterium glutamicum* for the conversion of monohydroxylated benzoates, Prot. Eng. Des. Sel. 30 (2017) 57–65. https://doi.org/10.1093/protein/gzw061.

[7] J. Werwath, H.-A. Arfmann, D.H. Pieper, K.N. Timmis, R.-M. Wittich, Biochemical and genetic characterization of a gentisate 1,2-dioxygenase from *Sphingomonas* sp. strain RW5, J. Bacteriol. 180 (1998) 4171–4176. PMCID:PMC107414.

[8] W. Fu, P. Oriel, Gentisate 1,2-dioxygenase from *Haloferax* sp. D1227, Extremophiles 2 (1998) 439-446. https://doi.org/10.1007/s007920050090.

[9] M.R. Harpel, J.D. Lipscomb, Gentisate 1,2-dioxygenase from *Pseudomonas*. Purification, characterization, and comparison of the enzymes from *Pseudomonas testosteroni* and *Pseudomonas acidovorans*, J. Biol. Chem. 265 (1990) 6301-6311. PMID:2156846.

[10] M. Suárez, E. Ferrer, M. Martín, Purification and biochemical characterization of gentisate 1,2-dioxygenase from *Klebsiella pneumoniae* M5al, FEMS Microbiol. Lett. 143 (1996) 89-95. https://doi.org/10.1111/j.1574-6968.1996.tb08466.x.

[11] K.L. Crawford, S.W. Hutton, P.J. Chapman, Purification and properties of gentisate 1,2-dioxygenase from *Moraxella osloensis*, J. Bacteriol. 121 (1975) 794-799. PMCID:PMC246005

[12] Y. Feng, H.E. Khoo, C.L. Poh, Purification and characterization of gentisate 1,2-dioxygenases from *Pseudomonas alcaligenes* NCIB 9867 and *Pseudomonas putida* NCIB 9869, Appl. Environ. Microbiol. 65 (1999) 946-950. PMCID: PMC91127.

[13] S. Hirano, M. Morikawa, K. Takano, T. Imanaka, S. Kanaya, Gentisate 1,2-dioxygenase from *Xanthobacter polyaromaticivorans* 127W, Biosci. Biotechnol. Biochem. 71 (2007) 192–199. https://doi.org/10.1271/bbb.60444.

[14] D. Liu, T. Zhu, L. Fan, J. Quan, H. Guo, J. Ni, Identification of a novel gentisate 1,2-dioxygenase from *Silicibacter pomeroyi*, Biotechnol. Lett. 29 (2007a) 1529-1535. https://doi.org/10.1007/s10529-007-9421-7.

[15] L. Huang, H. Hu, H. Tang, Y. Liu, P. Xu, J. Shi, K. Lin, Q. Luo, C. Cui Identification and characterization of a novel gentisate 1,2-dioxygenase gene from a halophilic *Martelella* strain, Sci. Reports 5 (2015) 14307. https://doi.org/10.1038/srep14307.

[16] D. Liu, S. Zhu, J. Ni, Purification and characterisation of a gentisate 1,2-dioxygenase from *Ralstonia solanacearum* GMI 1000, Ann. Microbiol. 57 (2007b) 307-312 https://doi.org/10.1007/BF03175065.

[17] H.J. Lee, J.M. Kim, S.H. Lee, M. Park, K. Lee, E.L. Madsen, C.O. Jeon, Gentisate 1,2-dioxygenase, in the third naphthalene catabolic gene cluster of *Polaromonas naphthalenivorans* CJ2, has a role in naphthalene degradation, Microbiology 157 (2011) 2891–2903. https://doi.org/10.1099/mic.0.049387-0.

[18] H. Yu, S. Zhao, L. Guo, 2018. Novel gene encoding 5-aminosalicylate 1,2-dioxygenase from *Comamonas* sp. strain QT12 and catalytic properties of the purified enzyme. J. Bacteriol. 200, e00395-17. https://doi.org/10.1128/JB .00395-17.

[19] M. Ferraroni, I. Matera, L. Steimer, S. Bürger, A. Scozzafava, A. Stolz, F. Briganti, Crystal structures of salicylate 1,2-dioxygenase-substrates adducts: A step towards the comprehension of the structural basis for substrate selection in class III ring cleaving dioxygenases, J. Struct. Biol. 177 (2012a) 431–438. https://doi.org/10.1016/j.jsb.2011.11.026.

[20] S. Fetzner, Ring-cleaving dioxygenases with a cupin fold, Appl. Environ. Microbiol. 78 (2012) 2505–2514. https://doi.org/10.1128/AEM.07651-11.

[21] I. Matera, M. Ferraroni, S. Bürger, A. Scozzafava, A. Stolz, F. Briganti, Salicylate 1,2-dioxygenase from *Pseudaminobacter salicylatoxidans*: crystal structure of a peculiar ring-

cleaving dioxygenase, J. Mol. Biol. 380 (2008) 856–868. https://doi.org/10.1016/j.jmb.2008.05.041.

[22] M. Ferraroni, I. Matera, S. Bürger, S. Reichert, L. Steimer, A. Scozzafava, A. Stolz, F. Briganti, The salicylate 1,2-dioxygenase as a model for a conventional gentisate 1,2-dioxygenase: crystal structures of the G106A mutant and its adducts with gentisate and salicylate, FEBS J. 280 (2013) 1643–1652. https://doi.org/10.1111/febs.12173.

[23] M.A. Adams, V.K. Singh, B.O. Keller, Z. Jia, Structural and biochemical characterization of gentisate 1,2-dioxygenase from *Escherichia coli* O157:H7, Mol. Microbiol. 61 (2006) 1469-1484. https://doi.org/10.1111/j.1365-2958.2006.05334.x.

[24] J. Chen, W. Li, M. Wang, G. Zhu, D. Liu, F. Sun, N. Hao, X. Li, Z. Rao, X.C. Zhang, Crystal structure and mutagenic analysis of GDOsp, a gentisate 1,2-dioxygenase from *Silicibacter pomeroyi*, Prot. Sci. 17 (2008) 1362–1373. https://doi.org/10.1110/ps.035881.108.

[25] N.M. Subbotina, M.P. Kolomytseva, L.A. Golovleva, 3-hydroxybenzoate and 2,5-dihydroxybenzoate metabolism in *Rhodococcus opacus* 1CP strain, Mikrobiologiia 81 (2012) 325-331. PMID:22880393.

[26] S.N. Gorlatov, O.V. Malítseva, V.I. Shevchenko, L.A. Golovleva, Degradation of chlorophenols by a culture of *Rhodococcus erythropolis*, Mikrobiologiya 58 (1989) 802–806.

[27] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding, Anal. Biochem. 72 (1976) 248–254. https://doi.org/10.1006/abio.1976.9999.

[28] L. Lack, The enzymic oxidation of gentisic acid, Biochim. Biophys. Acta 34 (1959) 117-123. https://doi.org/10.1016/0006-3002(59)90239-2.

[29] H. Lineweaver, D. Burk, The determination of enzyme dissociation constants, J. Am. Chem. Soc. 56 (1934) 658-666. https://doi.org/10.1021/ja01318a036.

[30] S.A. Kazakov, S.V. Balatskaya, B.H. Johnston, Ligation of the hairpin ribozyme in *cis* induced by freezing and dehydration, RNA. 12 (2006) 446-456. https://doi.org/10.1261/rna.2123506.

[31] V.E. Bower, R.G. Bates, pH values of the Clark and Lubs buffer solutions at 25^{o}C, J. Res. Nat. Bur. Stand. 55 (1955) 197-200. https://nvlpubs.nist.gov/nistpubs/jres/55/jresv55n4p197_A1b.pdf

[32] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (1970) 680-685. https://doi.org/10.1038/227680a0.

[33] G. Bertani, Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*, J. Bacteriol. 62 (1951) 293-300. PMCID: PMC386127.

[34] S. Altschul, W. Gish, W. Miller, E. Myers, D. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

[35] K. Tamura, G. Stecher, D. Peterson, A. Filipski, S. Kumar, MEGA6: molecular evolutionary genetics analysis version 6.0, Mol. Biol. Evol. 30 (2013) 2725–2729. https://doi.org/10.1093/molbev/mst197.

[36] A. Untergasser, I. Cutcutache, T. Koressaar, J. Ye, B.C. Faircloth, M. Remm, S.G. Rozen, 2012. Primer3 - new capabilities and interfaces. Nucleic Acids Res. 40, e115. https://doi.org/10.1093/nar/gks596.

[37] J. Sambrook, D.W. Russell, Molecular cloning: a laboratory manual, 3rd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, ISBN 978-1-936113-42-2, 2001.

[38] B. Webb, A. Sali, Comparative protein structure modeling using Modeller, Curr Prot Bioinf 54, John Wiley & Sons Inc, 5.6.1-5.6.37, 2016. https://doi.org/10.1002/0471250953.bi0506s15.

[39] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera - a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605-1612. https://doi.org/10.1002/jcc.20084.

[40] M.-Y. Shen, A. Sali, Statistical potential for assessment and prediction of protein structures, Prot. Sci. 15 (2006) 2507-2524. https://doi.org/10.1110/ps.062416606.

[41] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Cryst. 26 (1993) 283-291. https://doi.org/10.1107/S0021889892009944.

[42] R.W.W. Hooft, G. Vriend, C. Sander, E.E. Abola, Errors in protein structure, Nature 381 (1996) 272. https://doi.org/10.1038/381272a0.

[43] C.C. Yeo, C.L. Tan, X. Gao, B. Zhao, C.L. Poh, Characterization of hbzE-encoded gentisate 1,2-dioxygenase from *Pseudomonas alcaligenes* NCIMB 9867, Res. Microbiol. 158 (2007) 608-616. https://doi.org/10.1016/j.resmic.2007.06.003.

[44] M.R. Harpel, J.D. Lipscomb, Gentisate 1,2-dioxygenase from Pseudomonas. Substrate coordination to active site Fe^{2+} and mechanism of turnover, J. Biol. Chem. 265 (1990) 22187-22196. PMID: 2266121

[45] M. Ferraroni, L. Steimer, I. Matera, S. Bürger, A. Scozzafava, A. Stolz, F. Briganti, The generation of a 1-hydroxy-2-naphthoate 1,2-dioxygenase by single point mutations of salicylate 1,2-dioxygenase--rational design of mutants and the crystal structures of the A85H

732 and W104I variants, J. Struct. Biol. 180 (2012b) 563-571.
733 https://doi.org/10.1016/j.jsb.2012.08.007.
734 [46] D. Buongiorno, G.D. Straganz, Structure and function of atypically coordinated enzymatic
735 mononuclear non-heme-Fe(II) centers, Coord. Chem. Rev. 257 (2013) 541– 563.
736 https://doi.org/10.1016/j.ccr.2012.04.028.
737 [47] I. Matera, M. Ferraroni, M. Kolomytseva, L. Golovleva, A. Scozzafava, F. Briganti,
738 Catechol 1,2-dioxygenase from the gram-positive *Rhodococcus opacus* 1CP: quantitative
739 structure/activity relationship and the crystal structures of native enzyme and catechols
740 adducts, J. Struct. Biol. 170 (2010) 548-564. PMID: 20040374.
741
742

Figure legends

Fig. 1. The alignment of amino acid sequences of the gentisate oxidizing enzymes from different bacteria. AHH30829 *Rop* 1CP – GDO from *R. opacus* 1CP (GenBank accession number AHH30829), ELB94938 *Rw* IFP2016 – GDO from *R. wratislaviensis* IFP 2016 (GenBank accession number ELB94938), EHI40830 *Rhop* PD630 - *R. opacus* PD630 (GenBank accession number EHI40830), 3NL *Psal* BN12 – SDO from *P. salicylatoxidans* BN12 (RCSB Data Bank: 3NL1), BAC00417 *Cglu* – GDO from *C. glutamicum* ATCC 13032 (GenBank accession number BAC00417), EPD39425 *Dac* CCUG15835 – GDO form *P. (Delftia) acidovorans* CCUG15835 (GenBank accession number EPD39425), BAC98955 *Xpol* 127W - GDO from *X. polyaromaticivorans* 127W (GenBank accession number BAC98955), Q8X655 *Ecoli* O157:H7 – GDO from *E. coli* O157:H7 (UniProtKB/Swiss-Prot: Q8X655), 3BU7 *Spom* DSS3 – GDO from *S. (Ruegeria) pomeroi* DSS3 (RCSB Data Bank: 3BU7), EHI47270 *Rop* PD630 – GDO from *R. opacus* PD630 (GenBank accession number EHI47270), ELB89931 *Rw* IFP2016 – GDO from *R. wratislaviensis* IFP 2016 (GenBank accession number ELB89931), AAQ62856 *Hal* D1227 – GDO from *Haloferax* sp. D1227 (GenBank accession number AAQ62856). The GDOs which have the identical amino acid composition of the N-terminal cupin active site surfaces are marked by an asterisk. The GDOs containing a second C-terminal cupin active site bearing Fe^{2+} are marked by a double asterisk.

Fig. 2. Phylogenetic relationships of GDOs and related proteins based on the alignment of their amino acid sequences (GenBank, NCBI, UniProtKB/Swiss-Prot and RCSB PDB accession numbers of proteins are indicated). The bootstrap values (consensus support, %) are displayed on the tree.

Fig. 3. The overall structure of *rho*-GDO molecule calculated by homology modeling: a) homotetrameric native molecule and b) subunit structure: 1, 2, and 3 – the sites of contacts between subunits in *rho*-GDO tetramer (the description is presented in the text).

Fig. 4. The least-square superposition of the *rho*-GDO and *rp*-GDO (RCSB PDB: 3BU7) subunits (A) and their cupin domains with active site cavities (B – in C-terminal cupin domain and C – in N-terminal cupin domain). Atoms are colored as element type, i.e. carbon, in gray for *rho*-GDO and in pink for *rp*-GDO; oxygen, red; nitrogen, blue; iron, brown in *rho*-GDO and violet purple in *rp*-GDO structures. The names of *rp*-GDO residues are reported above the names of *rho*-GDO.

Fig. 5. The least-square superposition of the *rho*-GDO and *ps*-SDO subunits (A, RCSB PDB: 3NL1) and the *rho*-GDO and *ec*-GDO subunits (B, RCSB PDB: 2D40). Atoms are colored as element type, i.e. carbon, in gray for *rho*-GDO and in blue for *ps*-SDO; oxygen, red; nitrogen,

blue, iron, brown in *rho*-GDO, blue in *ps*-SDO, and violetpurple in *ec*-GDO structures. The names of *ps*-SDO and *ec*-GDO residues are reported above the names of *rho*-GDO.

Fig. 6. The different second-sphere residues (coloured in red) in the active sites of *rho*-GDO and *ps*-SDO (the least-square superposition of the *rho*-GDO and *ps*-SDO in complex with gentisate (coloured in green) (RCSB PDB: 3NL1)). Atoms colored as element type, i.e. carbon, in gray for *rho*-GDO and in blue for *ps*-SDO; oxygen, red; nitrogen, blue; iron, brown in *rho*-GDO and blue in *ps*-SDO structures. The names of *ps*-SDO residues are reported above the names of *rho*-GDO.

Fig. 7. The interchain interactions in *ps*-SDO (A) and *rho*-GDO (B). Atoms are colored as element type, i.e. oxygen, red; nitrogen, blue; iron, brown. Carbon elements are colored in green and blue in *ps*-SDO molecule, orange and pink - in *rho*-GDO structure.

Table 1 Physicochemical properties of known GDOs.

Strain	Growth substrate	GenBank accession number of aa seq.	Molecular mass, kDa	Subunit mass, kDa (subunit number□)	T optimum, ˚C°	pH optimum	Apparent K_m, μM (with gentisate)	State	References
Rhodococcus erythropolis S-1	salicylate, *m*-HBA	no info	328	43 (8)	40	8.5	53	native	[1]
Bacillus stearothermophilus PK1	benzoate	no info	260	40(6)	65-70	7.3	n.d.	native	[2]
Pseudoaminobacter salicylatoxydans BN12	naphthalene-sulfonate	3NL1	185	45(4)	n.d.	n.d.	79	native and recombinant	[3, 4]
Corynebacterium glutamicum ATCC 13032	gentisate, *m*-HBA	BAC00417	180	44(4)	n.d.	n.d.	n.d.	recombinant	[5, 6]
***Rhodococcus opacus* 1CP**	*m*-HBA	**AHH30829**	**177**	**41 (4)**	**40-45**	**7.6-8.5**	**40**	**native**	**the present work**
Sphingomonas sp. strain RW5	3,6-dichloro-2-methoxy-benzoate	CAA12267	177	38.5 (4)	n.d.	n.d.	15	native	[7]
Haloferax sp. D1227	3-phenylpropionate	AAQ62856	174	42(4)	45	7.2	95	native	[8]
Delftia (Pseudomonas) acidovorans	*m*-HBA	EPD39425	164	37.2-39.8(4)	n.d.	7.0-9.0	74	native	[9]
Pseudomonas testosteroni	*m*-HBA	no info	158	40.8(4)	n.d.	7.0-9.0	85	native	[9]
Klebsiella pneumoniae M5a1	*m*-HBA, gentisate	AAW63413	159	38(4)	30	8.0-9.0	52	native	[10]
Escherichia coli O157:H7	no info	Q8X655	154	38.9(4)	n.d.	7.5-8.5	11	recombinant	[2]
Moraxella osloensis OA3	salicylate	no info	154	40(4)	n.d.	7.0-8.5	7.1	native	[11]
Pseudomonas alcaligenes NCIB 9867	*m*-HBA	AAD49427	154	39(4)	50	8.0	92	native	[12]
Xanthobacter polyaromaticivorans 127W	dibenzothiophene	BAC98955	146	40.0 (4)	n.d.	n.d.	18.6	recombinant	[13]
Ruegeria (Silicibacter) pomeroyi DSS-3	no info	3BU7	132	45(3)	40	8.0	12	recombinant	[14]
Martelella strain AD-3	phenanthrene, anthracene	AMM86058	120	38(3)	30	7.0	26.6	recombinant	[15]
Pseudomonas putida NCIB 9869	*m*-HBA	no info	82	41(2)	50	8.0	143	native	[12]
Ralstonia solanacearum GMI 1000	no info	WP_011001037	n.d.	38	30	8.0	56	recombinant	[16]
Polaromonas naphthalenivorans CJ2 NagI2	naphthalene	AAZ93402	80	39(2)	50	8.0	31	recombinant	[17]
NagI3,		no info	80	38.7(2)	65	7.8	10	recombinant	[17]
Comamonas sp. strain QT12	*m*-aminobenzoate	ARB18233	n.d.	45	10	8.0	823	recombinant	[18]

n.d. – not determined; *m*-HBA – *m*-hydroxybenzoate, no info – no information, □ - predicted subunit number

Table 2 Kinetic properties of the *rho*-GDO in comparison to other known GDOs. All kinetic constants were obtained with an error <10% (the data of at least three independent experiments were used for each kinetic parameter calculation).

Substrate	Structure	***rho*-GDO** (the present work)		***ps*-SDO** [4, 22]		***Pseudomonas acidovorans* GDO** [44]		***Rhodococcus erythropolis* S-1** [1]		***Xanthobacter polyaromaticivorans* 127W** [13]	
		K_m or $K_i^{\#}$, µM	**V_{rel}, %**	**K_m or $K_i^{\#}$, µM**	**V_{rel}, %**	**K_m or $K_i^{\#}$, µM**	**V_{rel}, %**	**K_m or $K_i^{\#}$, µM**	**V_{rel}, %**	**K_m or $K_i^{\#}$, µM**	**V_{rel}, %**
gentisate (5-hydroxysalicylate)		40.00	100.00	79.00*	100.00*	80	100	n.d.	100.00	n.d.	100.00
salicylate		810.00#	0.00	12.00*	1.02*	2300.00-3000.00#	0.00	n.d.	5.00	n.d.	n.d.
2,3-dihydroxybenzoate (3-hydroxysalicylate)		1069.00#	0.00	67.00*	0.22*	3300.00#	0.00	n.d.	4.00	n.d.	n.d.
2,4-dihydroxybenzoate (4-hydroxysalicylate)		1130.00#	0.00	39.00*	0.09*	6400.00#	0.00	n.d.	3.00	n.d.	n.d.
2,6-dihydroxybenzoate (6-hydroxysalicylate)		220.00#	0.00	n.d.	n.d.	2000.00#	0.00	n.d.	n.d.	n.d.	n.d.
5-chlorosalicylate		57.00#	0.00	85.00*	0.22*	1400.00#	0.00	n.d.	n.d.	n.d.	0.00
homogentisate		444.00#	0.00	n.d.	n.d.	6800.00#	0.00	n.d.	n.d.	n.d.	n.d.
1-hydroxy-2-naphthoate		n.d.	0.00	140.00**	18.00**	n.d.	n.d.	n.d.	n.d.	n.d.	2.90
2-hydroxy-3-naphthoate		n.d.	0.00	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
catechol		44.90#	0.00	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	0.00
4-methylcatechol		6.70#	0.00	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
4-chlorocatechol		1.45#	0.00	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
hydroxyquinol		330.00#	0.00	n.d.	n.d.	n.d.	0.5	n.d.	n.d.	n.d.	n.d.
3,4-dihydroxybenzoate (protocatechuate)		not inhibitor and not substrate	0.00	n.d.	n.d.	n.d.	n.d.	n.d.	6.00	n.d.	n.d.

n.d. – not determined; * - the data from the paper of Hinter et al., 2004, the V_{rel} values are given as percentage from the enzyme activity with gentisate [4]; ** - the data from the paper of Ferraroni et al., 2013, the V_{rel} values are given as percentage from the enzyme activity with gentisate [22]

Fig. 1

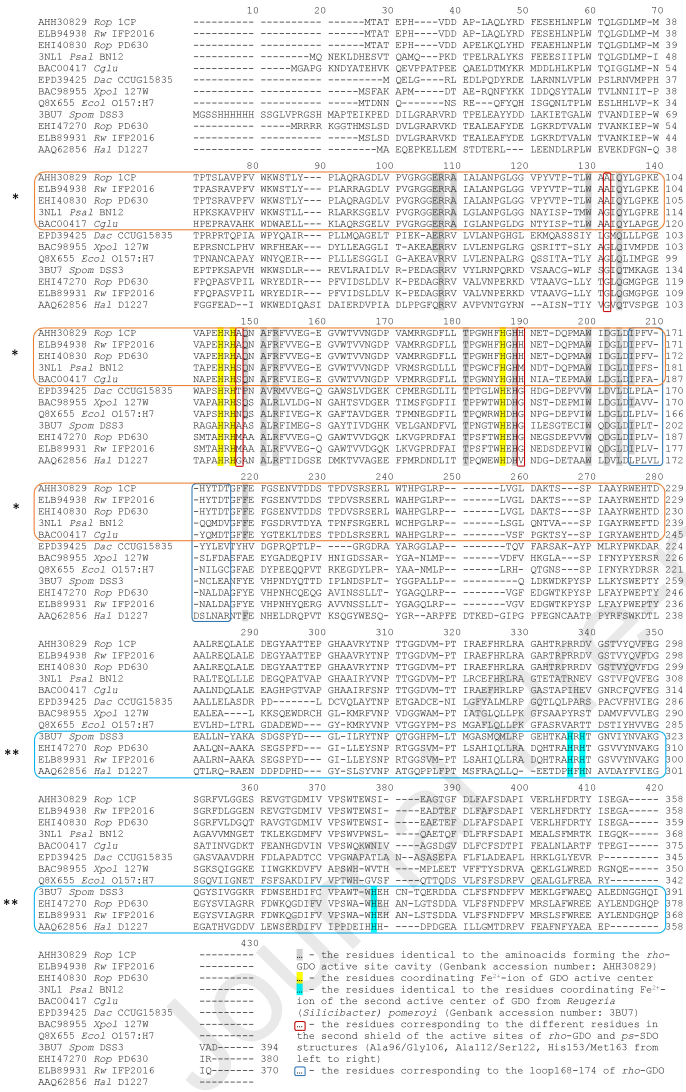


Fig. 2

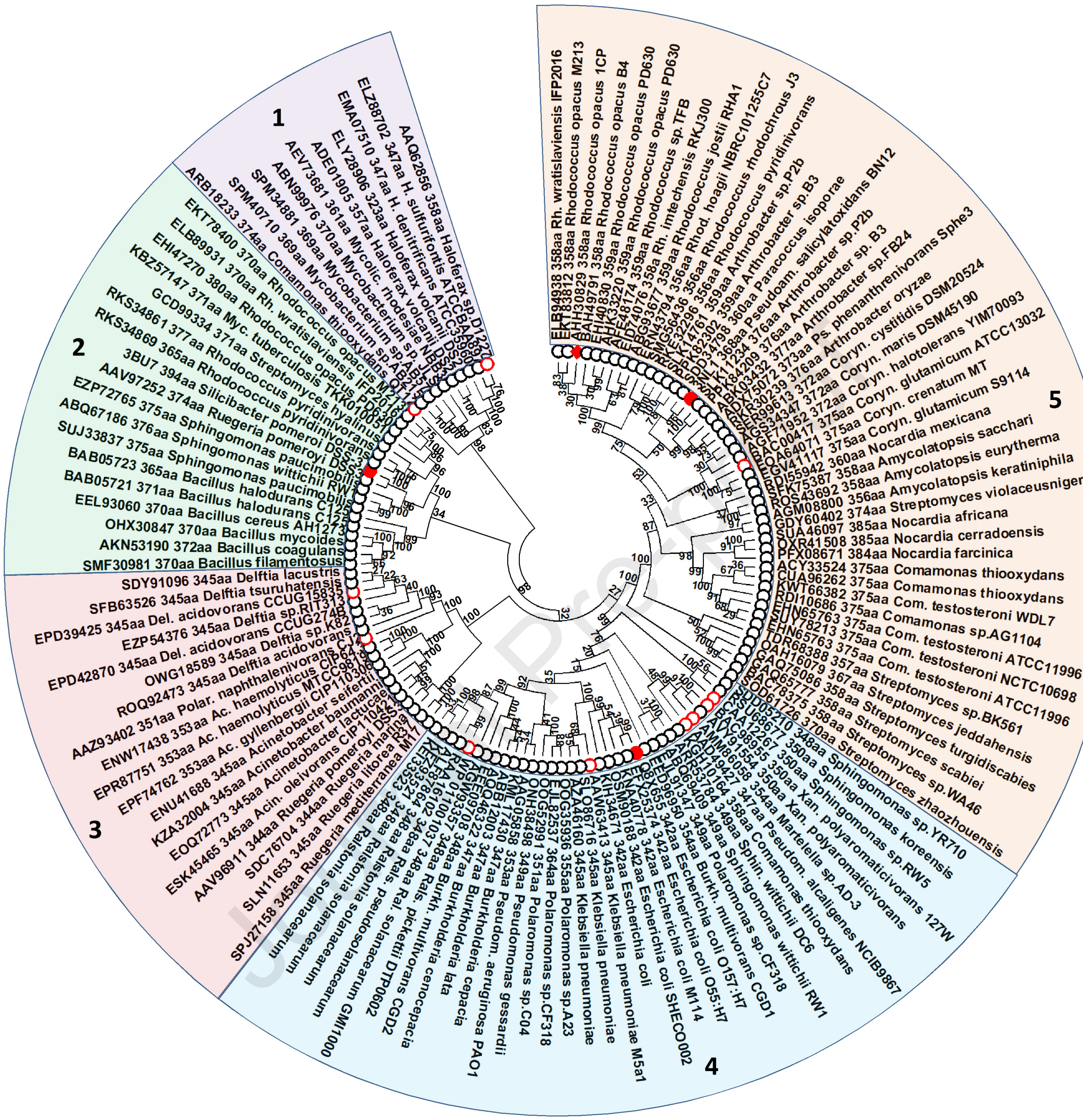


Fig. 3.

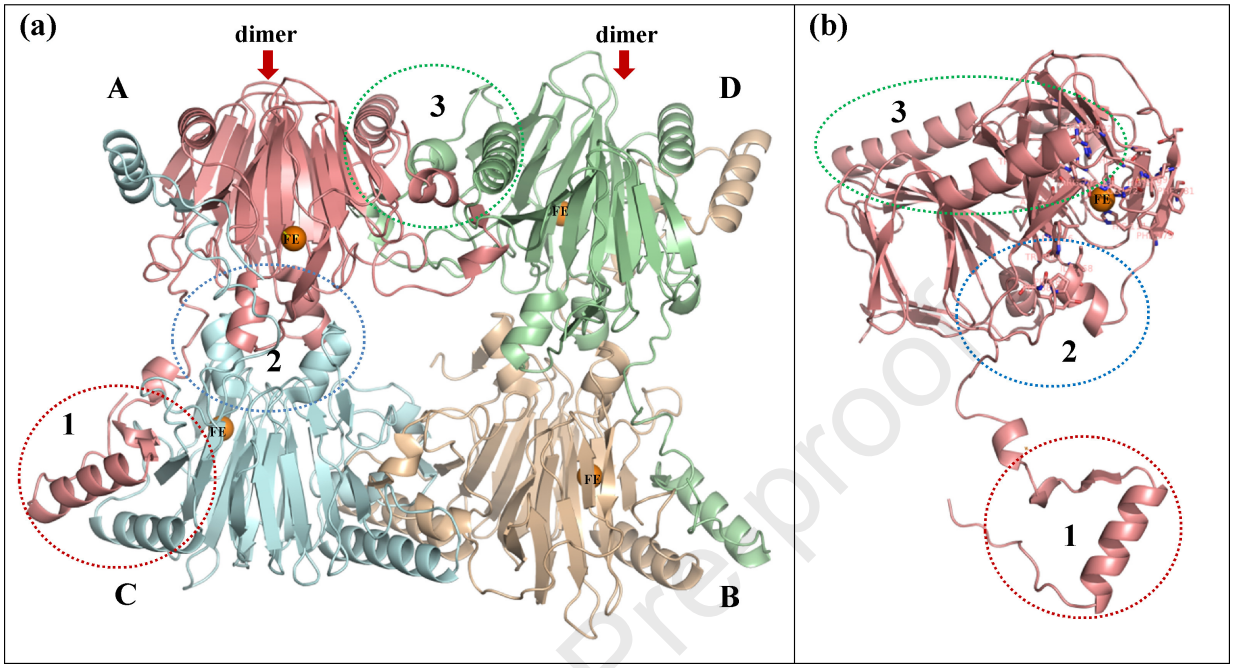


Fig. 4.

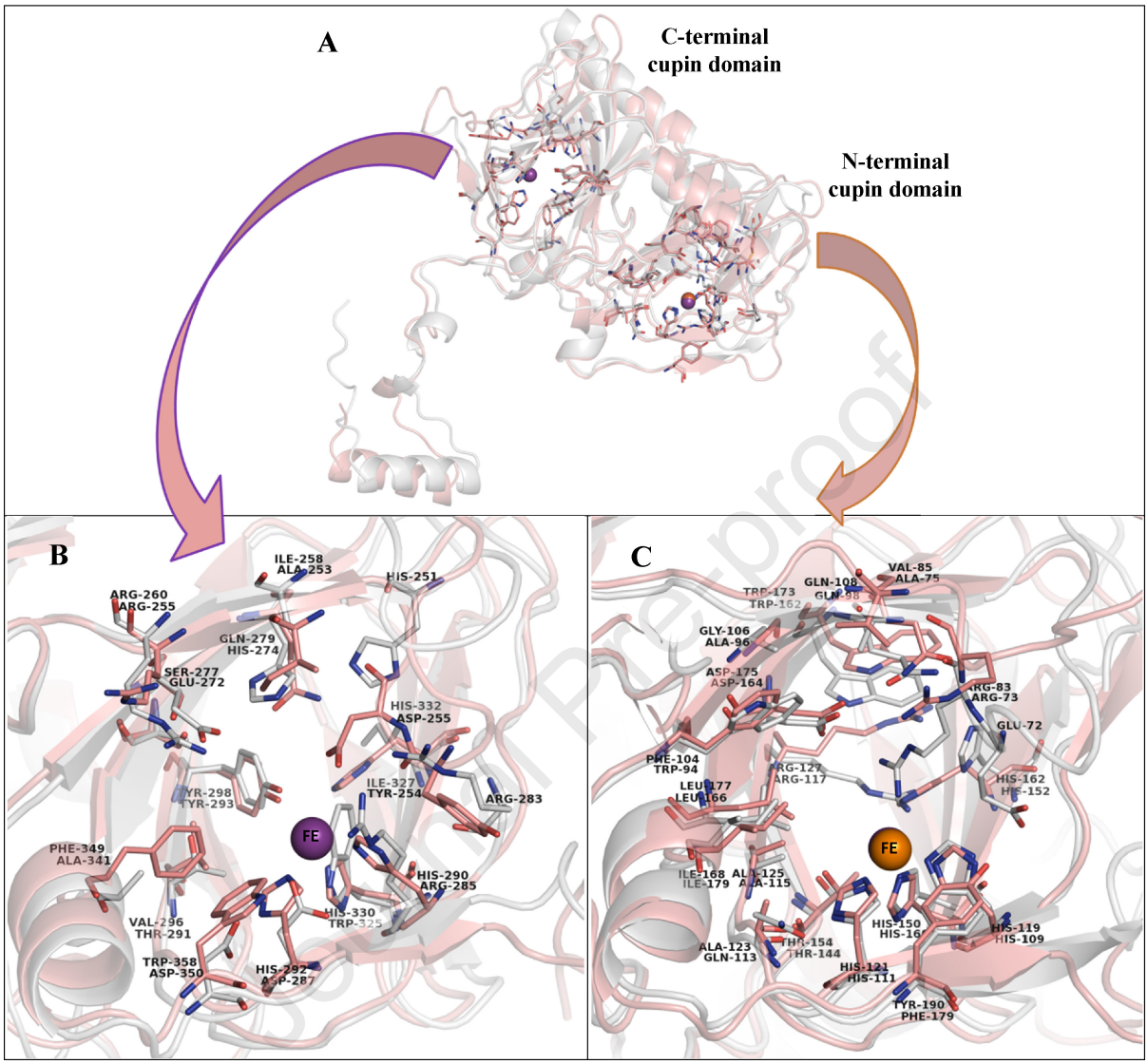

A
C-terminal
cupin domain
N-terminal
cupin domain
B
ILE-258
ALA-253
HIS-251
ARG-260
ARG-255
GLN-279
HIS-274
SER-277
GLU-272
HIS-332
ASP-255
TYR-298
TYR-293
ILE-327
TYR-254
ARG-283
FE
PHE-349
ALA-341
HIS-290
ARG-285
HIS-330
TRP-325
VAL-296
THR-291
TRP-358
ASP-350
HIS-292
ASP-287
C
VAL-85
ALA-75
TRP-173
TRP-162
GLN-108
GLN-98
GLY-106
ALA-96
ASP-175
ASP-164
ARG-83
ARG-73
GLU-72
PHE-104
TRP-94
ARG-127
ARG-117
HIS-162
HIS-152
LEU-177
LEU-166
FE
ILE-168
ILE-179
ALA-125
ALA-115
HIS-150
HIS-16
HIS-119
HIS-109
ALA-123
GLN-113
THR-154
THR-144
HIS-121
HIS-111
TYR-190
PHE-179

Fig. 5.

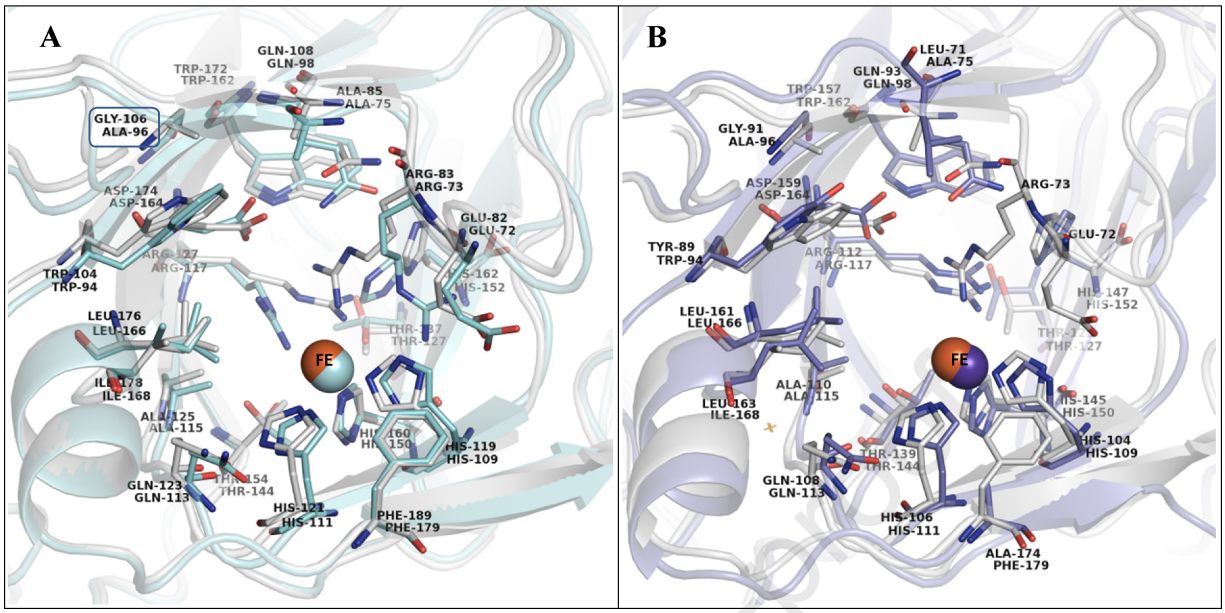


Fig. 6.

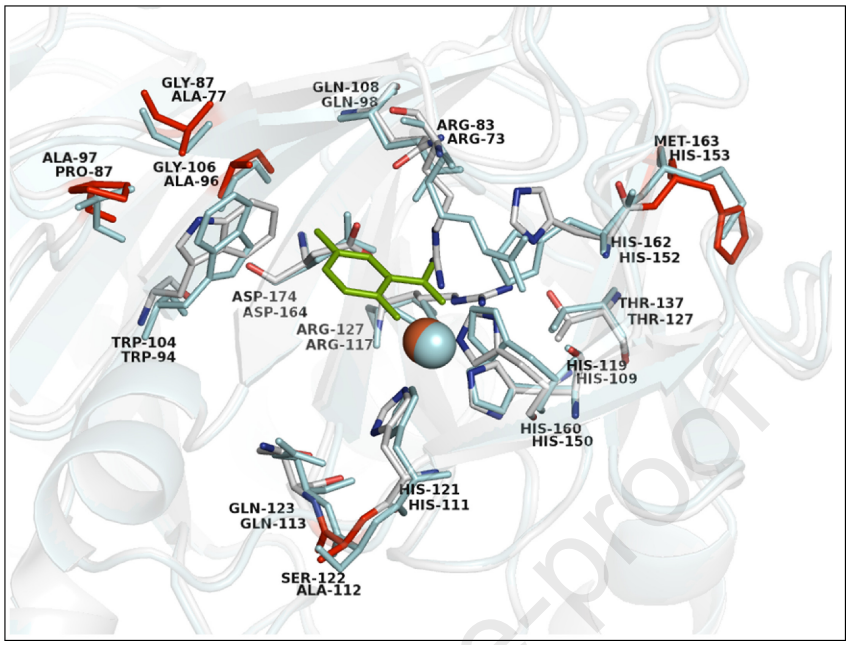
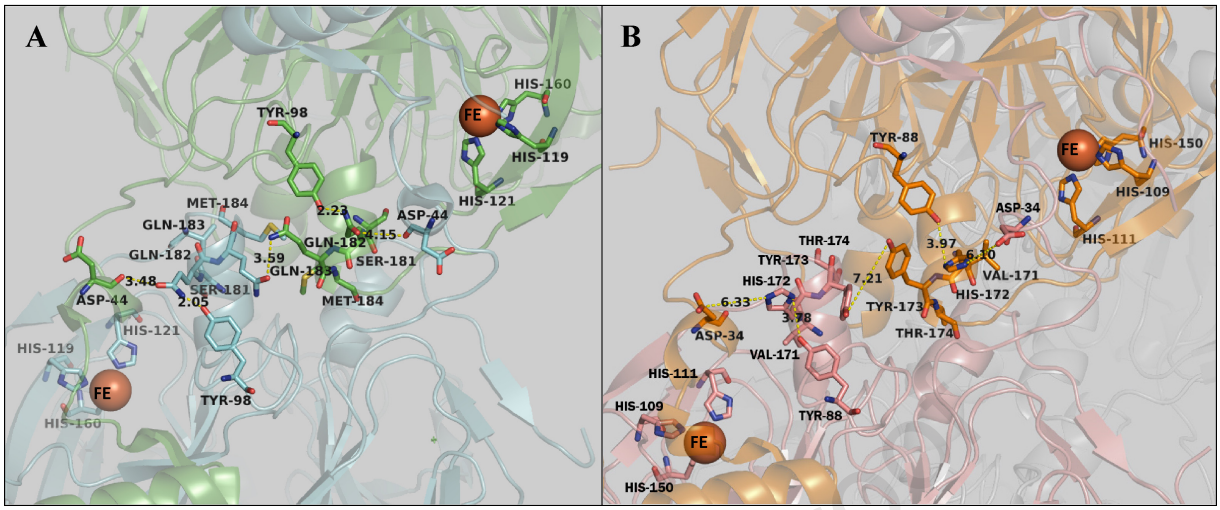


Fig. 7.



Highlights

- Gentisate 1,2-dioxygenase of *Rhodococcus opacus* 1CP was purified and characterized
- Gene coding the *rho*-GDO was identified and expressed
- Phylogenetic relationship among gentisate 1,2-dioxyhenases (GDOs) was studied
- A homology model of the 3D-structure of the first gram-positive GDO was presented
- Phenomenon of the active site identity of diverse substrate selectivity GDOs was shown

Conflict of Interest

The authors declare that they have not known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. All authors have not and will have not any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations.