ChemCatChem

Research Article
doi.org/10.1002/cctc.202300847

Chemistry Europe
European Chemical Societies Publishing

www.chemcatchem.org

Special Collection

Biochemical and Structural Characterization of a Uronic Acid Oxidase from *Citrus sinensis*

Alessandro Boverio,[a, b] Hugo L. van Beek,[c] Simone Savino,[c] Adeline Ranoux,[d] Wouter J. J. Huijgen,[d] Harry W. C. Raaijmakers,[d] Marco W. Fraaije,*[a] and Nikola Lončar*[c]

Aldaric acids are attractive diacids that can be prepared by selective oxidation of carbohydrates. For this, effective biocatalysts are in demand. This work reports on the discovery, biochemical and structural characterization of a VAO-type flavin-containing carbohydrate oxidase from *Citrus sinensis*: $URAO_{Cs3}$. $URAO_{Cs3}$ could be overexpressed using prokaryotic and eukaryotic expression systems. Extensive biochemical characterization revealed that the enzyme displays a high thermostability and an exquisite selectivity for uronic acids, galacturonic acid and glucuronic acid. The enzyme was further investigated by determining the crystal structure. The selective oxidation of D-galacturonic acid in a complex mixture was demonstrated, showing how $URAO_{Cs3}$ was found to be highly effective in selectively producing galactaric acid while leaving other carbohydrates untouched. In addition to the specific discovery of $URAO_{Cs3}$, these findings suggest that plant proteomes can be an interesting source for new biocatalysts.

Introduction

Aldaric acids are diacids of sugars which represent highly valuable building blocks for a vast repertoire of applications.[1–3] Despite their huge potential as renewable building blocks, the current production of aldaric acids mainly depends on chemical processes.[4] It would be attractive to develop a biotechnological process for producing aldaric acids. Metabolic pathways leading to aldaric acids are known and typically involve as last biocatalytic step the oxidation of uronic acids such as galacturonic acid. Galacturonic acid is the main monomer of the polysaccharide pectin, a key component of the plant cell wall and, for example, abundantly present in fruit.[5] In addition to fruit, pectin can be isolated from agricultural residues such as sugar beet pulp and citrus pulp. For hydrolysis of pectin, various enzymes have been reported. Yet, for the enzymatic selective oxidation of uronic acids, only a limited number of enzymes has been described. Some bacteria have been shown to produce oxidative enzymes acting on uronic acids,[6,7] employing $NAD(P)^+$ as oxidizing cofactor. While dehydrogenases have inherent drawbacks when used as isolated enzymes, several attempts have been made to produce galactaric acid employing whole cells that contain such dehydrogenases.[8,9] In plants, instead, oxidases acting on uronic acids have been identified in the past decades.[10,11] Recently, the sequence of a uronic acid oxidase from bitter orange, *Citrus aurantium*, ($URAO_{Ca}$) was identified and characterized.[12] This revealed that it contains a flavin cofactor as redox prosthetic group to assist in catalysis. In fact, the FAD cofactor was found to be an integral part of the protein as it was covalently attached to a cysteine and histidine residue.[13] Such bicovalent binding of flavin cofactors has only been observed in a very small number of flavoenzymes. It has been suggested to be linked to flavoenzymes acting on relatively bulky compounds, such as polysaccharides.[14] $URAO_{Ca}$ can be regarded a carbohydrate oxidase and its sequence also shows homology with known carbohydrate oxidases.[15] Flavin-dependent carbohydrate oxidases form a class of enzymes that can be divided in two different families according to their structural fold: the Glucose-Methanol-Choline (GMC) family and Vanillyl Alcohol Oxidase (VAO) family.[15–17] While, in the GMC family, the FAD cofactor is typically noncovalently bound and deeply buried in the active site, the carbohydrate oxidases from the VAO family often harbor a covalently attached FAD cofactor.

VAO-type oxidase structures can be divided in two major domains: a substrate binding domain (S-domain) formed by the C-terminal part of the protein sequence and a flavin binding domain (F-domain) at the N-terminus[18] Even though, VAO-type carbohydrate oxidases are mostly known from fungi,[14,19,20] a few examples have also been found in plants.[21]

In this work, we report on the discovery, characterization, and crystal structure of a VAO-type flavin-containing carbohy-

[a] *A. Boverio, Prof. Dr. M. W. Fraaije*
Molecular Enzymology group, University of Groningen
Nijenborgh 4, 9747AG Groningen (The Netherlands)
E-mail: m.w.fraaije@rug.nl

[b] *A. Boverio*
Department of Biology and Biotechnology
University of Pavia, via Ferrata 9, 27100 Pavia (Italy)

[c] *Dr. H. L. van Beek, Dr. S. Savino, Dr. N. Lončar*
GECCO Biotech, 9747 AG Groningen (The Netherlands)
E-mail: n.loncar@gecco-biotech.com

[d] *Dr. A. Ranoux, Dr. W. J. J. Huijgen, Dr. H. W. C. Raaijmakers*
Cosun RD&I, Cosun Innovation Center,
Kreekweg 1, 4671 VA Dinteloord (The Netherlands)

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cctc.202300847

This publication is part of a joint Special Collection with ChemBioChem published dedicated to the conference Biotrans 2023. Please see our homepage for more articles in the collection

drate oxidase from sweet orange, *Citrus sinensis*: $URAO_{Cs3}$. Except for a high degree of sequence homology with $URAO_{Ca}$, this newly characterized oxidase also shares a similar substrate acceptance with displaying high activity on uronic acids. $URAO_{Cs3}$ could be overexpressed using prokaryotic and eukaryotic expression systems (*Escherichia coli* and *Pichia pastoris*). With having access to purified recombinant $URAO_{Cs3}$, we could study the catalytic and structural aspects of this oxidase in more detail. This revealed that the enzyme displays a high thermostability and is very selective in recognizing uronic acids as substrates. By solving its crystal structure, we could assign the substrate binding pocket and catalytic machinery that explains the exquisite selectivity for uronic acids. These are attractive features when considering $URAO_{Cs3}$ as biocatalyst for the selective oxidation of uronic acids into their respective diacids. The established expression systems and elucidated crystal structure provide a good basis for developing this plant oxidase into a valuable biotechnological tool. To demonstrate its value as biocatalysts, the selective oxidation of galacturonic acid in a complex mixture was tested. Remarkably, $URAO_{Cs3}$ was found to be highly effective in selectively producing galactaric acid while leaving other carbohydrates untouched. Except for the specific discovery of $URAO_{Cs3}$, our study also demonstrates that plant proteomes can be an interesting source for new biocatalysts.

Results

A parallel approach in enzyme identification: isolation from orange peel and bioinformatic analysis

To identify the enzyme that catalyzes the oxidation of uronic acids in orange peel, both an experimental and a computational approach was carried out in parallel, to increase the chance of success. In a more classical approach, we set out to isolate the uronic acid oxidase from fresh plant material, orange peels. It was previously shown that orange peels contain uronic acid oxidase activity,[10,12] and that the respective enzyme(s) could be extracted from the peels.

Yet, the exact identity was not established at that time. Upon extracting a protein fraction and subjecting it to separation based on hydrophobic interaction chromatography, a colored protein fraction was obtained. This sample was tested for oxidase activity on 16 different sugars (D-glucose, D-galactose, D-mannose, D-xylose, L-arabinose, fructose, N-acetyl galactosamine, rhamnose, L-fucose, D-maltose, D-glucosamine, D-melezitose, D-melibiose, D-glucuronic acid, D-galacturonic acid, and polygalacturonic acid) using the HRP-based assay which relies on the oxidase-catalyzed formation of hydrogen peroxide. A substrate concentration of 10 mM was used in all the cases. This revealed clear oxidase activity for galacturonic and glucuronic acid, and some residual oxidase activity for polygalacturonic acid. No oxidase activity was observed for any other tested carbohydrate. This confirmed the presence of one or more oxidases active on uronic acids. The protein sample was then analyzed by SDS-PAGE. After incubating the polyacrylamide gel in 5% acetic acid for 2 minutes, the gel was analyzed for the presence of proteins carrying a covalently attached flavin cofactor, as it is known that VAO-type carbohydrate oxidases typically contain a covalent FAD cofactor (Figure S1). Such in-gel flavin fluorescence detection is a convenient approach to identify covalent flavoproteins as it is similar in sensitivity when compared with regular protein staining methods and can be followed up by protein staining of the same gel or proteomic analysis. Indeed, the oxidase-enriched protein sample was found to contain a dominant flavin-fluorescent protein with a molecular mass of around 70 kDa (Figure S1). Such molecular weight matches well with previously reported results for a URAO from *C. aurantium*. Next, MS analysis was performed on the protein band that displayed fluorescence (Figure S1). After tryptic digestion and subsequent analysis using an FT-ICR/Orbitrap, various peptide sequences were identified. The resulting peptide sequences were analyzed using the predicted proteome of *C. sinensis*. A clear hit was found corresponding to sequence KDO75192.1, identified by 29 unique peptides (Figure S2). In UniProt, it is annotated as an FAD-binding PCMH-type domain-containing protein. The full-length protein (570 residues) has a predicted molecular mass of 63678 Da while it is also predicted to be glycosylated, potentially leading to a higher molecular mass as observed by SDS-PAGE analysis (Figure S1). Yet, the first 33 N-terminal residues are predicted to be signal peptide for secretion. A BLAST search in the PDB revealed that the protein shares highest sequence identity with an VAO-type oxidase, AtBBE15, from the plant *Arabidopsis thaliana* (PDB: 4UD8) that harbors a bicovalently bound FAD cofactor and acts on monolignols.[22]

In parallel, we also embarked on a bioinformatic approach to identify putative URAOs based on known and inferred sequence features. Such genome mining approach[23] can be very effective and complementary to an experimental approach. We hypothesized that URAOs are probably homologs of other known VAO-type carbohydrate oxidases, similar to the recently identified URAO from *C. aurantium*. Using the sequence of glucooligosaccharide oxidase (GOOX) from *Acremonium strictum* (Uniprot accession code: Q6PW77), we first performed a PHI-BLAST search in the predicted proteome of *C. sinensis*. This resulted in a large set of sequences (>100) displaying $>20\%$ sequence identity. As filtering pattern, a sequence motif was used that included the two conserved Cys and His residues that are known to covalently couple the FAD cofactor. In the next step, all retrieved sequences were inspected upon alignment with GOOX for the presence of a positively charged residue (Arg or Lys) downstream from the flavinylated Cys and at a position where in GOOX the residue could be close to the C6 of the bound monosaccharide. The presence and position of such residue was inferred from a structural inspection of the structure of GOOX in which a substrate analog is bound (PDB:2AXR). This suggested that, for accommodating a carboxylate group at the 6-position of a uronic acid substrate, a lysine or arginine should be present at or close to the position equivalent to Ala351. Through this targeted bioinformatic trawl of the proteome of *C. sinensis*, four protein sequences were identified: KDO75192.1 ($URAO_{Cs1}$), KDO48902.1 ($URAO_{Cs2}$), XP_

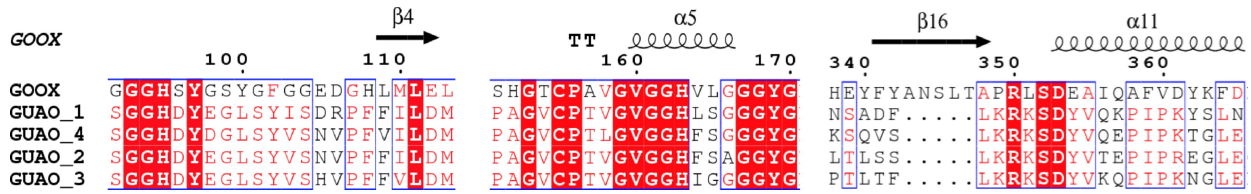

Figure 1. Multiple sequence alignment. Multiple sequence alignment of the four sequences of URAOs with GOOX that was used as query. In position 95 and 155 the histidine and cysteine residues responsible for the bicovalent linkage of the FAD are conserved. In position 351 the motif responsible for the second positive charge in the active site is conserved among the four chosen candidates while it was absent in the query (GOOX).

006468466.1 ($URAO_{Cs3}$), KDO77595.1 ($URAO_{Cs4}$) (Figure 1). The four proteins were very similar in sequence displaying >82% sequence identity when compared with each other. This also revealed that the enzyme discovery approaches converged: one of the identified sequences ($URAO_{Cs1}$) was the protein identified via the above-mentioned proteomic analysis. Based in these results, we continued with setting up recombinant expression for these four putative URAOs.

Expression and characterization

After successfully cloning the four URAO-encoding genes in a pBAD-His-SUMO vector, the expression of the respective SUMO-fused enzymes was tested in *E. coli* ORIGAMI2 cells. After induction, cell free extracts (CFE) obtained from the cultures corresponding to the four candidates were tested against galacturonic and glucuronic acid (10 mM final concentration).

Both $URAO_{Cs1}$ and $URAO_{Cs3}$ showed activity with two uronic acids, with $URAO_{Cs3}$ showing a higher activity judged on the rate of color formation. Therefore, it was decided to study $URAO_{Cs3}$ in more detail. Expression and purification from *E. coli* cells yielded only very small amounts of enzyme. Therefore, *P. pastoris* was chosen as alternative expression host. For this, a codon-optimized gene of $URAO_{Cs3}$ was cloned in a pPICZα vector, and the resulting product was transformed in *P. pastoris* X-33 cells. 16 colonies were picked and grown according to the protocol described in the materials and methods section and the resulting filtrates were subsequently tested for oxidase activity. Two transformants showed a relatively high galacturonic acid oxidase activity (~30 mU/mL). Successful expression was also confirmed through SDS-PAGE analysis revealing, after incubation in acetic acid, a single flavin-fluorescent band at the expected molecular weight. One transformant was subsequently used for growth and expression at a larger scale to prepare sufficient amount of purified enzyme for a detailed biochemical and structural characterization.

Substrate scope, kinetic properties, and thermostability

After a single step of IMAC purification, ~20 mg URAO could be purified from 1 L of culture broth. SDS-PAGE analysis revealed a single band of about 70 kDa (Figure S3). Oxidase activity of purified $URAO_{Cs3}$ was tested with 16 potential carbohydrate substrates. This resulted in very similar results (Table 1) when compared with that obtained with the URAO obtained from orange peels: the highest activity was observed with galacturonic acid and glucuronic acid.

Table 1. $URAO_{Cs3}$ substrate scope.

Substrate	Relative activity (%)[a]
arabinose	0.1
galactose	0.3
mannose	0.2
galacturonic acid	59.1
glucuronic acid	100.0
D-glucosamine	0.3
D-melezitose	0.2
D-melibiose	1.0
polygalacturonic acid	4.2

[a] Activity was measured using the HRP-coupled assay using 20 mM of substrate at pH 6.5, 24 °C. Rhamnose, glucose, xylose, fructose, fucose, N-acetyl-D-galactosamine and maltose did not show any activity.

Only with polygalacturonic acid a modest activity was observed which may be due to trace amounts of short oligomers or monomers. $URAO_{Cs3}$ did not show significant activity on all other tested carbohydrates (<1% activity). It shows that $URAO_{Cs3}$ is highly specific for uronic acids. Based on the substrate screening results, the steady state kinetic parameters were determined for galacturonic and glucuronic acid (Table 2).

Initial reaction rates were determined in triplicate and could be fitted successfully using the Michaelis-Menten formula. From the analysis, it emerged that the enzyme is most efficient with galacturonic acid, exhibiting a relatively low K_M (35.5 μM) and a k_{cat} of 4.2 s^{-1}. The K_M for glucuronic acid was 4-fold higher, while its k_{cat} was 2-fold higher. Clearly, $URAO_{Cs3}$ is an efficient biocatalyst for oxidizing both uronic acids with a slightly higher catalytic efficiency for galacturonic acid. The pH optimum of

Table 2. Steady state parameters of $URAO_{Cs3}$.[a]

Substrate	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M (mM^{-1} s^{-1})
galacturonic acid	35.5 ± 0.6	4.2 ± 0.1	120
glucuronic acid	160 ± 20	8.9 ± 0.5	55

[a] The kinetic parameters were measured at 25 °C in 50 mM KPi, pH 7.0.

$URAO_{Cs3}$ activity was determined in the range of pH 3.0 to 9.0. The enzyme displayed its highest activity at pH 7.0. At more acidic pH values, activity rapidly declined with no activity recorded at pH 3.0 (Figure 2A). The activity was also monitored at different salt concentrations to test the ionic strength effect. A clear inhibitory effect was observed when testing NaCl (Figure 2B). With 500 mM NaCl, the activity decreased to only 50% of the activity when compared to activity in the absence of salt. Testing other salts revealed that the chloride anion seems to cause a specific inhibition of $URAO_{Cs3}$ activity. Furthermore, it was found that some sulfate salts boost the activity of the oxidase. Inhibition of by chloride has been observed for other oxidases[24,25] and may be due to its ability to act as dioxygen mimic. The thermostability of $URAO_{Cs3}$ was probed using the ThermoFluor method which provides an indication of thermal unfolding (Figure 3A). Interestingly, the enzyme was found to display a relatively high thermostability judged from the measured apparent melting temperatures (T_m). Over a range of pH 5.0 to pH 8.0, the T_m varied only mildly (80–84 °C) with the lowest stability at the highest pH value. Due to the deleterious effect on activity by chloride, thermostability was also measured with a gradient of NaCl at pH 7.0. Yet, it was observed that the salt did not affect the thermostability of $URAO_{Cs3}$. The effect of substrate, galacturonic acid, was also probed and revealed a positive effect on thermostability. The presence of 1.0 M substrate increased the T_m by 5 °C (Figure 3B). This result is in line with a stabilizing effect of substrate-enzyme interactions.

After having established the catalytic properties of $URAO_{Cs3}$, the biocatalyst was tested for the selective conversion of galacturonic acid in a complex mixture (Figure 4A). For this, an enzyme-hydrolyzed sugar beet pulp pectin sample was used which contained a mixture of carbohydrates. Except for galacturonic acid (0.32 wt.%), also various other carbohydrates were present such as arabinose, galactose, and glucose (0.11 wt.%). Incubation of this mixture with $URAO_{Cs3}$ convincingly showed that the enzyme is highly selective in only oxidizing galacturonic acid into galactaric acid. No conversion of glucose, arabinose and galactose, present in the complex mixture, was observed. The enzyme was also tested with a 0.35 wt.% solution of only D-galacturonic acid (Figure 4B) which revealed a similar rate of conversion. This shows that the biocatalyst also does not seem to suffer from inhibition by other components present in a pectin hydrolysate. Finally, a sample of sugar beet pulp hydrolysate (containing 11.4 wt.% D-galacturonic acid of 45.7 wt.% total dry matter) was diluted to 0.36 wt.% galacturonic acid and tested which resulted in selective and nearly complete conversion of galacturonic acid into galactaric acid (data not shown). Clearly, the oxidase performs well in complex mixtures. These features render $URAO_{Cs3}$ a highly efficient biocatalyst to produce galactaric acid starting from plant mass (pectin) derived galacturonic acid.

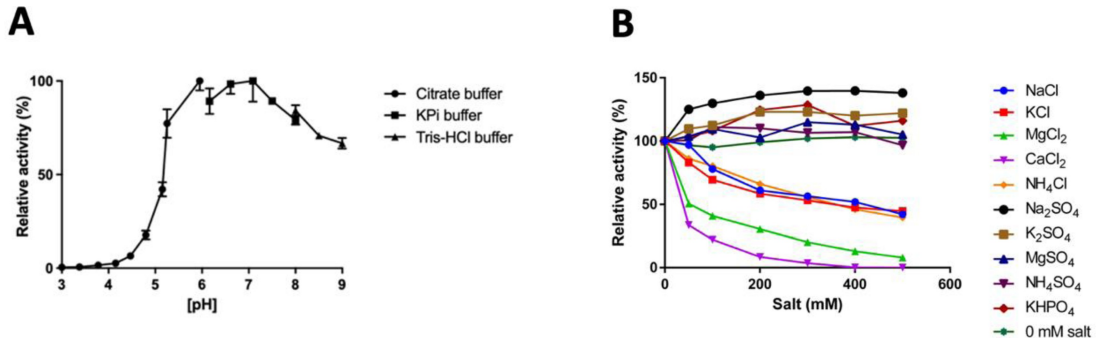

Figure 2. Effect of pH and salts on $URAO_{Cs3}$ activity. All reactions were performed using 20 mM galacturonic acid and 0.10 μM enzyme at 25 °C. A) pH-dependent oxidase activity using three different buffer systems (sodium citrate, potassium phosphate, TRIS/Cl). B) Effects of various salts on $URAO_{Cs3}$ activity.

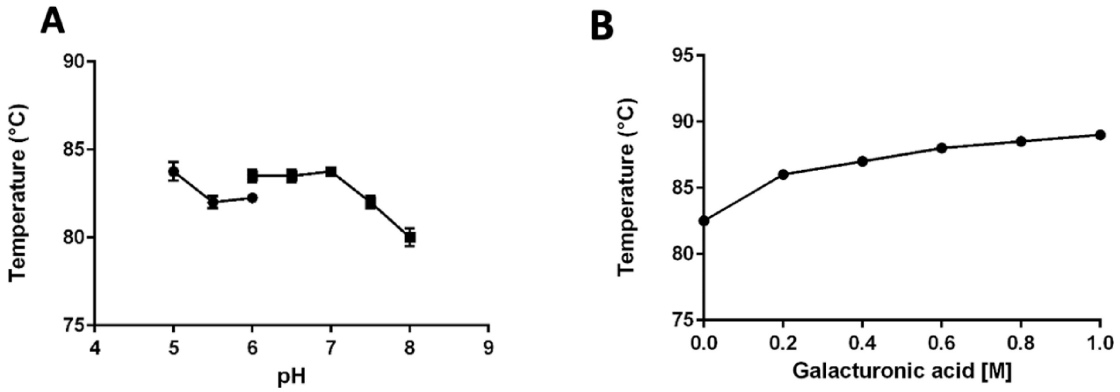

Figure 3. Thermal stability of $URAO_{Cs3}$. Apparent melting temperatures were measured using the Thermofluor method. A) Thermostability at pH values ranging from 5 to 8 (pH 5–6: sodium citrate; pH 6–8: potassium phosphate). B) Effect of galacturonic acid (from 0 to 1.0 M) on thermostability (50 mM potassium phosphate, pH 7.0). Temperature was increased by 0.5 °C every 30 seconds.

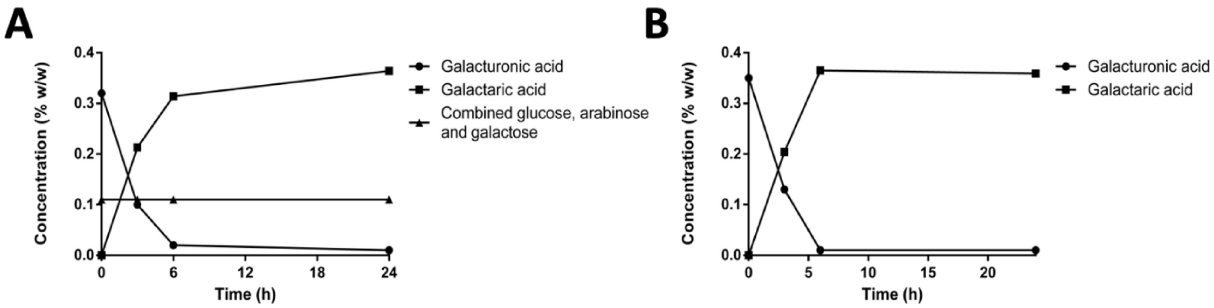

Figure 4. Conversion of galacturonic acid into galactaric acid. A) Reaction profile over time in a complex mixture. B) Reaction profile over time of pure galacturonic acid in solution.

Crystal structure of $URAO_{Cs3}$

To have a better understanding of the catalytic features of $URAO_{Cs3}$ by which it selectively accepts uronic acids, we set out to elucidate its crystal structure. $URAO_{Cs3}$ formed rectangular shaped yellow crystals in 20% PEG3350 and 0.20 M sodium iodine. These crystals allowed elucidation of its crystal structure at 1.9 Å resolution. One enzyme molecule was present in the asymmetric unit. This result is in line with size exclusion chromatography experiments which hinted to a monomeric conformation of the enzyme in solution. The overall structure of $URAO_{Cs3}$ is similar to other known carbohydrate oxidoreductases belonging to VAO-like family. Structure alignment with AtBBE15 (PDB: 4UD8)[22] and with GOOX (PDB: 2AXR) resulted in RMSD values of 0.7 Å and 1.4 Å for the Cα atoms and 51% and 26% of sequence identity, respectively. As for all VAO-type enzymes, two major domains can be recognized in the $URAO_{Cs3}$ structure: a substrate binding domain (S-domain) and a flavin binding domain (F-domain). The F-domain comprises residues from the N-terminus until residue 228 together with residue 454 to the C-terminus. The S-domain is composed of residues from 229 to 453 (Figure 5A). The S-domain is composed by α/β folds. Three parallel β strands (β1-β3) and five antiparallel β strands (β4-β8) are present together with nine α-helices (α1-α9). The S-domain, on the other hand, contains seven antiparallel β strands (β9-β15) which are placed above the isoalloxazine ring and are flanked by four α-helices (α10-α13) on the outside and two α-helices (α14-α15) on the inside. Inspection of the crystal structure revealed that the isoalloxazine ring of the FAD cofactor is bicovalently bound via 8α-N1-histidy and 6-S-cysteinyl linkages with residues His89 and Cys151, respectively. The redox active part of the flavin cofactor is embedded between the two domains (Figure 5B). The bicovalent attachment of the FAD has been observed in other VAO-type eukaryotic carbohydrate oxidases, with GOOX[20] as first example. A single *N*-linked GlcNAc moiety could be fitted in the electron density, attached to Asn49. Between the first and the second α-helix a disulfide bridge is present (Cys12-Cys74) to potentially contributing to the thermostability of the enzyme. The substrate binding groove of $URAO_{Cs3}$ that leads to the isoalloxazine ring of the FAD cofactor is similar compared to other carbohydrate oxidases such as xylooligosaccharide oxidase (XylO) and chitooligosaccharide oxidase (ChitO).[19,26] Beyond the overall architecture of the binding groove, a different composition of residues is present in the active site that dictates specific binding of uronic acids. Interestingly, the crystal structure of $URAO_{Cs3}$ also revealed a chloride ion bound in the active site, providing a potential hint about how the enzyme interacts with its uronic acid substrate (Figure 5C). The chlorine ion is positioned close to the Tyr91-OH (3.2 Å) and Arg351-NH2 (3.5 Å). In fact, the arginine stood out already when doing the sequence comparison with other carbohydrate oxidases (vide sopra) and expected to be directly involved in binding the carboxylate moiety of the uronic acids. Arg351 is part of a sequence motif (K-R-K) that is present in the other three homologs from *C. sinensis* (Figure 1) while the Arg is not conserved in other VAO-type carbohydrate oxidases (for example, in GOOX the Arg is replace by an Ala.[22] Furthermore, the binding of the chloride ion in that specific area could be the reason why an inhibition occurs using high concentration of salt. Other residues present in the active site adjacent to the redox-active flavin moiety are: Tyr91, Tyr165, Arg265, Asn383, Lys408, Gln410, Tyr452.

Despite several soacking attempts, no crystal structure was obtained with a substrate bound. For this reason, an *in silico* approach was used to understand which residues are involved in binding uronic acids. Docking of galacturonic acid resulted in a pose where the carboxylic moiety is placed in close contact with the Arg351 (3.1 Å) (Figure 6). Such binding mode is in line with the observed bound chloride ion. Except for a hint for the specific recognition of the carboxylate moiety of galacturonic acid, the docked pose of the substrate also revealed a possible role of Tyr452 as catalytic base: the hydroxyl moiety at the C1 of galacturonic acid point towards the Tyr452 (3.0 Å) which could act as a base to promote the hydride transfer from C1 to the N5 of the FAD cofactor. It is worth noting that this Tyr is conserved in other VAO-type carbohydrate oxidases, such as GOOX, XylO and ChitO, confirming a conserved role in catalysis.

Conclusions

Aldaric acids represent a precious potential resource for the (chemical) industry with multifold applications. Nevertheless, the synthesis of these compounds is driven almost entirely by chemical processes leaving an attractive space for biotechno-

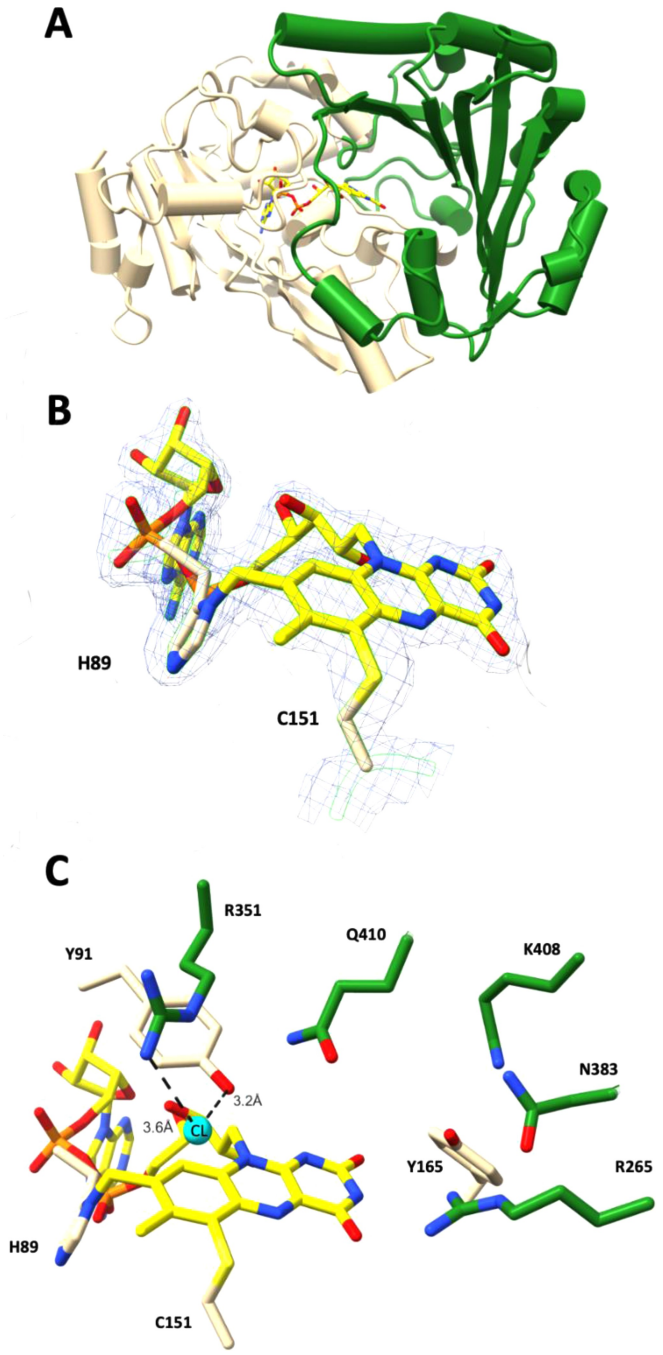

Figure 5. Structure analysis of $URAO_{Cs3}$. A) F-domain is represented in wheat and the S-domain is in forest green, the FAD cofactor is in yellow. B) The bicovalently bound FAD, linked via 8α-N1-histidyl and 6-S-cysteinyl linkages, is highlighted with a weighted 2Fo-Fc electron density map. The contour level of the map is 1.0 σ. C) Active site of $URAO_{Cs3}$ with a chloride ion in pink and the FAD cofactor in yellow. Distances are indicated in Å.

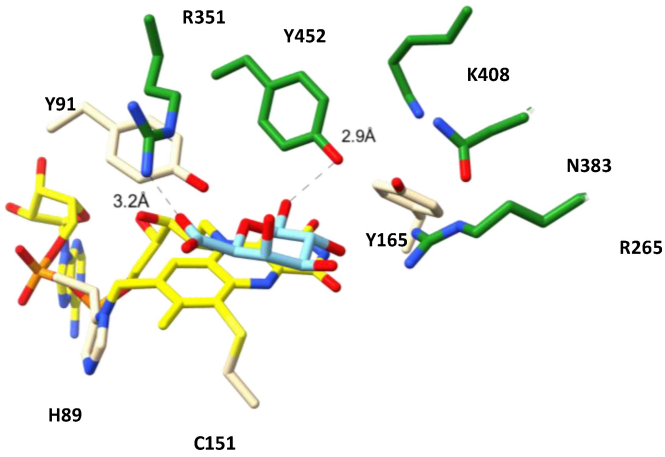

Figure 6. Docked galacturonic acid in the active site of $URAO_{Cs3}$. Galacturonic acid is in light blue, and the FAD cofactor is in yellow. Distances are indicated in Å.

logical applications. In nature, aldaric acids are the product of the oxidation of uronic acids. Flavin-dependent carbohydrate oxidases are a valuable tool in biocatalysis with a high substrate specificity and regioselectivity.[14] The recent identification of uronic acid oxidases in plants[12,13] inspired us to further investigate these types of enzymes. We identified the major URAO in orange peels and could identify three more putative URAOs in the predicted proteome of *C. sinensis*. The URAOs are VAO-type oxidases, all predicted to contain a bicovalent FAD cofactor. One of the URAOs ($URAO_{Cs3}$) could be expressed recombinantly with two different expression hosts: *E. coli* and yeast. The enzyme displayed a very high thermostability and maximal activity at neutral pH. Its substrate scope revealed a high selectivity towards uronic acids. This selective oxidation of galacturonic acid into galactaric acid could be demonstrated by incubating a complex mixture, hydrolyzed pectin, which resulted in a full and selective oxidation of galacturonic acid without inhibition by or conversion of the other carbohydrates. To identify the structural features that dictate the unique substrate profile, the crystal structure of $URAO_{Cs3}$ was determined. The crystal structure revealed also a solvent exposed active site allowing uronic acid to bind. A chloride ion was bound in the structure giving an indication about the role of Arg351 in the stabilization of the carboxylic moiety of the substrate.

The location of the chloride ion could also be the reason of the observed enzyme's inhibition at high concentration of salt in the buffer. Docking of galacturonic acid confirmed our hypothesis about the crucial role of a positive charge (Arg351) in the active site to allow uronic acid to bind. Due to the high thermostability and substrate specificity against uronic acids, $URAO_{Cs3}$ represents a valuable biocatalyst to produce aldaric acids.

Materials and Methods

Enzyme isolation from orange peels

For purifying $URAO_{Cs}$ from plant material, 2 kilograms of fresh oranges were purchased from a local supermarket. The obtained peels (470 grams) were homogenized in cold buffer (50 mM Tris-HCL, 150 mM NaCl, 5% glycerol, pH 7.5) using a blender. A further homogenization step was performed using an UltraTurrax rotor-stator homogenizer (14,000 rpm). The

resulting mass was left at 4 °C for 30 min, followed by a centrifugation step (3,430 g, 1 h, 4 °C). The supernatant was filtered over cotton, filter paper and 1 μm glass fiber filter. To remove the pigments that were still present in the extract, 200 mL of sample was loaded on a 50 mL Phenyl-Sepharose column (GE healthcare).

The resulting flow-through was collected and concentrated using a centrifugal ultrafiltration unit (30 kDa cut-off). The obtained sample was then analyzed by SDS-PAGE and subsequent detection for in-gel flavin fluorescence in 5% acetic acid using an UV bench.

Cloning, transformation, expression, and purification in *E. coli*

Synthetic genes encoding for four putative uronic acid oxidases ($URAO_{Cs1\text{-}4}$) codon optimized for *E. coli* with BSAI sites at 5′ and 3′ (Integrated DNA Technologies) were cloned with the Golden Gate methodology in a pBAD His-SUMO vector. For transformation, 3.5 μL of the plasmid were added to 100 μL of *E. coli NEB10β* RbCl competent cells and incubated on ice for 30 minutes. Cells were subsequently heat-shocked at 42 °C for 30 s and incubated again on ice for 3 minutes. 900 μL of LB-SOC was added and cells were incubated 1 h at 37 °C and plated on LB-agar supplemented with 50 $\mu g.mL^{-1}$ ampicillin and incubated overnight at 37 °C. Cloning was verified through sequencing. pBAD-SUMO vectors carrying the four different URAOs were subsequently transformed in *E. coli* ORIGAMI2 cells which is known from previous study to be able to express covalently bound flavoenzymes[26] A pre-inoculum of 5 mL of LB with ampicillin (50 $\mu g.mL^{-1}$) was grown overnight at 30 °C and diluted in 2 L baffled flasks containing 400 mL of Terrific Broth medium supplemented with ampicillin (50 $\mu g.mL^{-1}$) and L-arabinose (0.02%). Cultures were left at 17 °C for 72 hours before harvesting. Cells were harvested by centrifugation (4,950 g, 20 min, 4 °C) and flash frozen at −20 °C. Cell pellets were subsequently resuspended in buffer A (50 mM KPi, 150 mM NaCl, pH 7.7). PMSF (0.10 mM), lysozyme (0.5 mg/mL) and DNAse (1 mg/mL) were added to the lysis solution to prevent protein degradation. Cells were disrupted by sonication (5 s on, 5 s off, 70% amplitude for a total of 15 min) and then centrifuged (19,800 g, 1 h, 4 °C). Resulting cell free extract was loaded on a gravity column containing 4 mL of TALON resin previously equilibrated with buffer A. After a washing step with buffer A (5 CV), and one with buffer B (50 mM Kpi, 150 mM NaCl, 5 mM imidazole, pH 7.7), the protein was eluted with buffer C (50 mM Kpi, 150 mM NaCl and 150 mM imidazole, pH 7.7). Elution buffer was then exchanged against storage buffer (50 mM Kpi, 150 mM NaCl, 10% glycerol, pH 7.7).

Cloning, transformation, expression in *P. pastoris*

A synthetic DNA fragment containing a gene encoding for $URAO_{Cs3}$, codon optimized for *P. pastoris*, was ordered at Twist Bioscience and cloned into a pPICZα fused with an export tag using the Gibson cloning methodology.[27] Since the vector lacked suitable restriction sites, both genes and vector were amplified with the corresponding pair of primers listed in Table S*1* using a standard PCR protocol and Pfu Ultra II. The resulting product was transformed in *E. coli* NEB10β RbCl competent cells. Cells were then plated on low-salt LB agar plates with zeocin (25 μg/mL) as antibiotic. Cloning success was verified through sequencing. 2 μg of cut vector DNA using SacI in CutSmart buffer, was transformed subsequently in freshly made electrocompetent *P. pastoris* X-33 cells using the Bio-Rad MicroPulser electroporator. Cells were plated on YPDS-agar with an increasing amount of zeocin as antibiotic (100, 200 and 500 μg/mL) to potentially discover variants with multiple insertion. 16 colonies were picked and grown in rich medium (YPD) for 48 h at 28 °C. Induction of expression was done by adding 1% methanol using 2% methanol solution in YP (doubling the culture volume), and subsequent addition of 1% methanol as 50% methanol mix with YP at 56, 72 and 80 h. Cells were harvested at 96 h from start of the main culture by centrifugation (4,950 g, 20 min, 4 °C) and the supernatant was tested for activity.

Substrate scope, steady state kinetics, and pH optimum

The concentration of the enzyme was measured using the extinction coefficient of 6-S-cysteinyl bound FMN ($\varepsilon_{445} = 11.6\ mM^{-1}\ cm^{-1}$).[20] Oxidase activities were measured based on the hydrogen peroxide produced by the oxidase. Hydrogen peroxide formation was coupled to the activity of horse radish peroxidase (HRP) and chromogenic peroxidase substrates (AAP and DCHBS).[28] Absorbance measurements were performed using a Synergy H1 microplate reader (Biotek) at 515 nm ($\varepsilon_{515} = 26\ mM^{-1}\ cm^{-1}$).

Thermostability

Due to the quenching of the fluorescence of the FAD cofactor because of its bicovalently attachment to the protein, the ThermoFAD method[29] could not be used to monitor enzyme unfolding. Instead, the thermostability of $URAO_{Cs3}$ was determined using the Thermofluor method.[30] For this, 2.5 μL of concentrated enzyme was mixed with 17.5 μL buffer or substrate solution and added to 5 μL of 100x diluted SYPRO orange dye solution. Fluorescence was measured using the FRET filter set on a CFX96 RT-PCR machine.

Selective conversion of galacturonic acid

For conversion of galacturonic acid, a 25 mL solution of 0.35 wt.% galacturonic acid in 50 mM potassium phosphate buffer was prepared. Using NaOH, the pH was adjusted to pH 6.0. Then, 1.0 μL of catalase (135 kU/mL) and 131 μL of $URAO_{Cs3}$ (10.1 U/mL) was added to the reaction mixture. The concentration of galacturonic acid used was on purpose below the solubility of galactaric acid in order to be able to study the

enzymatic conversion in solution. The pH was monitored to ensure the pH remained around 6. The resulting solution was incubated in a shaking water bath at 30 °C and 110 rpm. Samples were taken at 3-, 6- and 24-hour time intervals. Each sample was deactivated by heating for 10 minutes at 95 °C.

Secondly, enzymatically hydrolyzed pectin syrup (1 g, 19.4% w/w galacturonic acid at 52.6 Bx % (refractive index corresponding to amount of dry matter) was diluted in 54.4 mL demineralized water resulting in targeted 0.35% w/w galacturonic acid and the pH was adjusted to pH 6.0 using NaOH (and HCl if needed). The ratios of galacturonic acid to glucose and galacturonic acid to the sum of arabinose, galactose and glucose of the starting material were 32:1 and 2.9:1 w/w respectively. A 25 mL aliquot of the pectin syrup solution was removed and 1.0 μL of catalase (135 kU/ml) and 131 μL of $URAO_{Cs3}$ (10.1 U/mL) was added. The resulting solution was incubated in a shaking water bath at 30 °C and 110 rpm. Samples were taken at 3-, 6-, and 24-hour time intervals. Samples for analysis were inactivated by heating for 10 minutes at 95 °C. The samples were analysed by HPLC and HPAEC-PAD (see Supporting Information for analytical details).

Protein crystallization, structure elucidation and docking

Purified protein from the *P. pastoris* cultivation was loaded on a Superdex200 10/300 column (Cytiva) using an ÄKTA pure. Two wavelengths (280 nm and 450 nm) for monitoring protein elution were used during the purification. The central fractions of the peak were pooled together and concentrated until 18 mg/mL. Different crystallization conditions were tested using a Mosquito crystallization robot (TTP LabTech, Melbourn, UK). Large rectangular shaped yellow crystals appeared in different conditions. The best condition was obtained with the sitting drop vapor diffusion methodology using 20% PEG3350 and 0.20 M sodium iodine. 25% glycerol was used as cryoprotectant. Crystals were subsequently flash frozen in liquid nitrogen and sent to ESRF for data collection. The best crystal diffracted at 1.9 Å resolution using MASSIF-1 as beamline.[31] Data were scaled using XDS. Molecular replacement was done using Phaser.[32] Model building was performed using Buccaneer.[33] Structure refinement was done using COOT[34] and REFMAC5[35] of the CCP4 package.[36] Data collection statistics are summarized in Table 3. Docking simulations were performed using the refined structure of $URAO_{Cs3}$. For the docking, Yasara[37] was used, applying the AMBER IPQ force field with a cell of 20×20×20 Å which included the whole active site, and docking the substrate using AutoDock Vina:[38] 100 runs were performed while using the default settings.

Conflict of Interests

The authors declare no conflict of interest.

Table 3. Data collection and refinement statistics of $URAO_{Cs3}$.

Diffraction Data	wild type
PDB code	8ATE
Space group	$P3_22_1$
Unit cell axes (Å)	101.65, 101.65, 108.75
Unit cell angles (°)	90, 90, 120
Resolution (Å)	108.8 (1.9)
R_{merge}	0.082 (2.769)
R_{pim}	0.030 (1.1017)
$CC_{1/2}$	0.999 (0.628)
Completeness (%)	90.0 (99.4)
Unique reflections	45258 (3312)
Multiplicity	15.8 (16.2)
Overall I/σ (I)	16.9 (1.2)
Protein residues	502
FAD molecules	1
Water molecules	84
R/R_{free} (%)	20.02 /24.31
Rms bond length	0.0115
Rms bond angle	1.56
Ramachandran outliers (%)	0.0

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords: aldaric acid · biocatalysis · flavin · galacturonic acid · oxidase

[1] N. van Strien, S. Rautiainen, M. Asikainen, D. A. Thomas, J. Linnekoski, K. Niemelä, A. Harlin, *Green Chem.* **2020**, *22*, DOI 10.1039/D0GC02293D.
[2] S. Rautiainen, P. Lehtinen, J. Chen, M. Vehkamäki, K. Niemelä, M. Leskelä, T. Repo, *RSC Adv.* **2015**, *5*, DOI 10.1039/c5ra01802a.
[3] T. Mehtiö, M. Toivari, M. G. Wiebe, A. Harlin, M. Penttilä, A. Koivula, *Crit. Rev. Biotechnol.* **2016**, *36*, DOI 10.3109/07388551.2015.1060189.
[4] C. L. Mehltretter, C. E. Rist, *J. Agric. Food Chem.* **1953**, *1*, DOI 10.1021/jf60012a005.
[5] N. C. Carpita, M. C. McCann, *J. Biol. Chem.* **2020**, *295*, DOI 10.1074/jbc.REV120.014561.
[6] S. H. Yoon, T. S. Moon, P. Iranpour, A. M. Lanza, K. J. Prather, *J. Bacteriol.* **2009**, *191*, DOI 10.1128/JB.00586-08.
[7] J. W. Ahn, S. Y. Lee, S. Kim, S. J. Cho, S. B. Lee, K. J. Kim, *Acta Crystallogr. F Struct. Biol. Cryst. Commun.* **2011**, *67*, DOI 10.1107/S1744309111012437.
[8] D. Mojzita, M. Wiebe, S. Hilditch, H. Boer, M. Penttilä, P. Richard, *Appl. Environ. Microbiol.* **2010**, *76*, DOI 10.1128/AEM.02273-09.
[9] T. Paasikallio, A. Huuskonen, M. G. Wiebe, *Microb. Cell Fact.* **2017**, *16*, DOI 10.1186/s12934-017-0736-3.
[10] J. Riov, *Plant Physiol.* **1975**, *55*, DOI 10.1104/pp.55.4.602.
[11] D. Cantu, L. C. Greve, S. Lurie, J. M. Labavitch, *Phytochemistry* **2006**, *67*, DOI 10.1016/j.phytochem.2005.10.003.
[12] R. Pressey, *Phytochemistry* **1993**, *32*, DOI 10.1016/0031-9422(93)85141-D.
[13] Y. Wei, Y. L. Tan, E. L. Ang, H. Zhao, *ChemBioChem* **2020**, *21*, DOI 10.1002/cbic.201900546.
[14] D. P. H. M. Heuts, D. B. Janssen, M. W. Fraaije, *FEBS Lett.* **2007**, *581*, DOI 10.1016/j.febslet.2007.09.019.

[15] S. Savino, M. W. Fraaije, *Biotechnol. Adv.* **2021**, *51*, DOI 10.1016/j.biotechadv.2020.107634.
[16] E. W. van Hellemond, N. G. H. Leferink, D. P. H. M. Heuts, M. W. Fraaije, W. J. H. van Berkel, *Adv. Appl. Microbiol.* **2006**, *60*, DOI 10.1016/S0065-2164(06)60002-6.
[17] N. G. H. Leferink, D. P. H. M. Heuts, M. W. Fraaije, W. J. H. van Berkel, *Arch. Biochem. Biophys.* **2008**, *474*, DOI 10.1016/j.abb.2008.01.027.
[18] M. W. Fraaije, W. J. H. Van Berkel, J. A. E. Benen, J. Visser, A. Mattevi, *Trends Biochem. Sci.* **1998**, *23*, DOI 10.1016/S0968-0004(98)01210-9.
[19] A. R. Ferrari, H. J. Rozeboom, J. M. Dobruchowska, S. S. Van Leeuwen, A. S. C. Vugts, M. J. Koetsier, J. Visser, M. W. Fraaije, G. Hart, *J. Biol. Chem.* **2016**, *291*, DOI 10.1074/jbc.M116.741173.
[20] C. H. Huang, W. L. Lai, M. H. Lee, C. J. Chen, A. Vasella, Y. C. Tsai, S. H. Liaw, *J. Biol. Chem.* **2005**, *280*, DOI 10.1074/jbc.M506078200.
[21] M. Toplak, G. Wiedemann, J. Ulićević, B. Daniel, S. N. W. Hoernstein, J. Kothe, J. Niederhauser, R. Reski, A. Winkler, P. Macheroux, *FEBS J.* **2018**, *285*, DOI 10.1111/febs.14458.
[22] B. Daniel, T. Pavkov-Keller, B. Steiner, A. Dordic, A. Gutmann, B. Nidetzky, C. W. Sensen, E. Van Der Graaff, S. Wallner, K. Gruber, P. Macheroux, *J. Biol. Chem.* **2015**, *290*, DOI 10.1074/jbc.M115.659631.
[23] N. Ziemert, M. Alanjary, T. Weber, *Nat. Prod. Rep.* **2016**, *33*, DOI 10.1039/c6np00025h.
[24] E. De Jong, W. J. H. Van Berkel, R. P. Van Der Zwan, J. A. M. De Bont, *Eur. J. Biochem.* **1992**, *208*, DOI 10.1111/j.1432-1033.1992.tb17231.x.
[25] R. Baron, C. Riley, P. Chenprakhon, K. Thotsaporn, R. T. Winter, A. Alfieri, F. Forneris, W. J. H. Van Berkel, P. Chaiyen, M. W. Fraaije, A. Mattevi, J. A. McCammon, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, DOI 10.1073/pnas.0903809106.
[26] S. Savino, S. Jensen, A. Terwisscha van Scheltinga, M. W. Fraaije, *FEBS Lett.* **2020**, *594*, DOI 10.1002/1873-3468.13854.
[27] D. G. Gibson, L. Young, R. Y. Chuang, J. C. Venter, C. A. Hutchison, H. O. Smith, *Nat. Methods* **2009**, *6*, DOI 10.1038/nmeth.1318.
[28] V. Vojinović, A. M. Azevedo, V. C. B. Martins, J. M. S. Cabral, T. D. Gibson, L. P. Fonseca, *J. Mol. Catal. B* **2004**, *28*, DOI 10.1016/j.molcatb.2004.02.003.
[29] F. Forneris, R. Orru, D. Bonivento, L. R. Chiarelli, A. Mattevi, *FEBS Journal* **2009**, *276*, DOI 10.1111/j.1742-4658.2009.07006.x.
[30] M. D. Cummings, M. A. Farnum, M. I. Nelen, *J. Biomol. Screening* **2006**, *11*, DOI 10.1177/1087057106292746.
[31] M. W. Bowler, D. Nurizzo, R. Barrett, A. Beteva, M. Bodin, H. Caserotto, S. Delagenière, F. Dobias, D. Flot, T. Giraud, N. Guichard, M. Guijarro, M. Lentini, G. A. Leonard, S. McSweeney, M. Oskarsson, W. Schmidt, A. Snigirev, D. von Stetten, J. Surr, O. Svensson, P. Theveneau, C. Mueller-Dieckmann, *J. Synchrotron Radiat.* **2015**, *22*, DOI 10.1107/S1600577515016604.
[32] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.* **2007**, *40*, DOI 10.1107/S0021889807021206.
[33] K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62*, DOI 10.1107/S0907444906022116.
[34] P. Emsley, K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, DOI 10.1107/S0907444904019158.
[35] O. Kovalevskiy, R. A. Nicholls, F. Long, A. Carlon, G. N. Murshudov, *Acta Crystallogr. D Struct. Biol.* **2018**, *74*, DOI 10.1107/S2059798318000979.
[36] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, DOI 10.1107/S0907444910045749.
[37] E. Krieger, G. Vriend, *Bioinformatics* **2002**, *18*, DOI 10.1093/bioinformatics/18.2.315.
[38] O. Trott, A. J. Olson, *J. Comput. Chem.* **2009**, DOI 10.1002/jcc.21334.

Manuscript received: July 7, 2023
Revised manuscript received: August 25, 2023
Accepted manuscript online: August 28, 2023
Version of record online: September 26, 2023