Mechanistic Insights into Glycoside 3-Oxidases Involved in *C*-Glycoside Metabolism in Soil Microorganisms

André Taborda

   Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av da República, 2780-157 Oeiras, Portugal

Tomas Frazão

   Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av da República, 2780-157 Oeiras, Portugal

Miguel V. Rodrigues

   Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av da República, 2780-157 Oeiras, Portugal

Xavier Fernández-Luengo

   Department of Chemistry, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

Ferran Sancho

   Zymvol Biomodeling, C/ Pau Claris, 94, 3B, 08010 Barcelona, Spain

Maria Fátima Lucas

   Zymvol Biomodeling, C/ Pau Claris, 94, 3B, 08010 Barcelona, Spain

Carlos Frazão

   Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av da República, 2780-157 Oeiras, Portugal

M. Rita Ventura

   Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av da República, 2780-157 Oeiras, Portugal

Laura Masgrau

   Department of Chemistry, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

Patrícia T. Borges

   Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Av da República, 2780-157 Oeiras, Portugal

Ligia O Martins (✉ lmartins@itqb.unl.pt)

   1Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa
   https://orcid.org/0000-0003-0082-9591

Research Article

Abstract

C-glycosides are natural products with important biological activities but are recalcitrant to degradation. Glycoside 3-oxidases (G3Oxs) are newly identified bacterial flavo-oxidases from the glucose-methanol-coline (GMC) superfamily that catalyze the oxidation of *C*-glycosides with the concomitant reduction of O_2 to H_2O_2. This oxidation is followed by C-C acid/base-assisted bond cleavage in two-step *C*-deglycosylation pathways. Soil and gut microorganisms have different oxidative enzymes, but the details of their catalytic mechanisms are largely unknown. Here, we report that *Ps*GO3x oxidizes at 50,000-fold higher specificity (k_{cat}/K_m) the glucose moiety of mangiferin to 3-keto-mangiferin than free D-glucose to 2-keto-glucose. Analysis of *Ps*G3Ox X-ray crystal structures and *Ps*GO3x in complex with glucose and mangiferin, combined with mutagenesis and molecular dynamics simulations, revealed distinctive features in the topology surrounding the active site that favors catalytically competent conformational states suitable for recognition, stabilization, and oxidation of the glucose moiety of mangiferin. Furthermore, their distinction to pyranose 2-oxidases (P2Oxs) involved in wood decay and recycling is discussed from an evolutionary, structural, and functional viewpoint.

Introduction

C-glycosides represent a large group of natural products in which the anomeric carbon of glucose is directly connected via carbon-carbon bonding to an aglycone moiety (anthraquinone, flavone, terpenoids, phenols, among others). These compounds are secondary metabolites produced by plants and microorganisms and exhibit great structural diversity, wide natural distribution, and significant biological activities, including antioxidant, anti-inflammatory, antibacterial, antiviral, and antitumor[1,2]. *C*-glycosides have shown increasing importance in the pharmaceutical, agricultural, and food industries, and a great effort has been focused on their synthesis[3,4]. Several *C*-glycosides like puerarin (daidzein 8-C-β-D-glucoside) have been the precursor of clinical drugs, and biotechnological strategies have been optimized for large-scale production of plant *C*-glycosides via heterologous expression systems[2,5].

Gut microbiota that performs bioconversion of *C*-glycosides to aglycones with beneficial health effects has been identified over the past decade[6−8]. However, very recently, microbial screenings and biochemical studies suggest a ubiquitous process of C-C bond cleavage reactions in nature[9,10]. Compared with other glycosides (*O*-, *N*- and *S*-glycosides), *C*-glycosides are considerably more stable against chemical and enzymatic treatments. Because of this, *C*-glycosides are not deglycosylated by glycoside hydrolases, the so-called glycosidases. Instead, their microbial catabolic pathway includes enzymes that catalyze an oxidative step followed by the C-C bond cleavage by a complex of deglycosylation enzymes [9]. Soil and intestinal microorganisms have similar C−C bond-cleaving enzymes[9]. In contrast, the initial oxidation step by *C*-glycoside-3-oxidases that oxidize the C3 of the sugar moiety is catalyzed by NAD(H) anaerobic oxidoreductase in intestinal microorganisms[6,7,11] and by FAD-dependent oxidoreductases in soil microorganisms[8,10]. These latter enzymes display ~ 60% similarity to pyranose oxidases (POx, pyranose:

oxygen 2-oxidoreductase; EC 1.1.3.10) from the glucose-methanol-choline (GMC) superfamily of enzymes[12].

G3Oxs from soil bacteria such as *Microbacterium sp.* 5-2b CarA, *Arthrobacter globiformis* NBRC12137 *Ag*CarA, and *Microbacterium trichothecenolyticum* NBRC15077 *Mt*CarA do not exhibit detectable activity for glucose and instead, oxidize *C*-glycosides such as carminic acid, mangiferin, and C6-glycosylated flavonoids, as well as *O*-glycosides but at a significantly lower rate, at C3 to form the corresponding 3-keto *C*-glycosides[10]. An analogous substrate preference was observed in bacterial *Sc*P2Ox from *Streptomyces canus* that displayed 100 to 1000-fold higher enzymatic activity towards the oxidation of *C*-glycoside puerarin compared to monosaccharide[13]. It was suggested that FAD-dependent G3Oxs from the POx family are ancestors of enzymes that oxidize glucose at the C2 position, the pyranose 2-oxidases (P2Oxs)[10,13]. The fungal P2Oxs, have most likely been acquired by horizontal gene transfer from bacteria and could have evolved and specialized over time to oxidize lignocellulose-derived sugars such as D-glucose, D-xylose, or D-galactose. Fungal P2Oxs are secreted to the extracellular space and are involved in wood decay and recycling. They are the most extensively studied POx, particularly from *Trametes multicolor*, *Peniophora sp*, and *Phanerochaete chrysosporium*[14−16]. They comprise a highly conserved flavin-binding domain with a Rossman-like-fold where FAD covalently binds and a substrate-binding domain. P2Oxs are homotetrameric, and the access to the active site is restricted by four channels that route the substrate from the enzyme surface to the active site cavity[17−19]. The structural characterization of bacterial GO3x, *Mt*CarA, and *Sc*P2Ox revealed a few structural and functional aspects of these enzymes[10,13]. However, many fundamental questions remain, particularly the mechanisms behind the diverse substrate specificities.

Here, we investigated the bacterial enzyme *As*P2Ox from *Arthrobacter siccitolerans* (now *Pseudoarthrobacter siccitolerans*[20,21]), renamed for *Ps*P2Ox, using a combination of experimental and computational approaches that reveal functional and structural details that explain these enzymes ability to bind and oxidize larger glycoside substrates. This work contributed to unveiling the catalytic mechanism of a critical catabolic enzyme involved in the degradation of recalcitrant C-glycosides in nature that remain to be fully disclosed and advancing our understanding of the structure-function relationships among members of the POxs family of enzymes.

Results And Discussion

Biochemical and kinetic characterization. The pseudo-second-order constants measured by transient state kinetics for the reductive and oxidative-half reactions (0.15 ± 0.02 M^{-1} s^{-1} and (0.76 ± 0.06) × 10^6 M^{-1} s^{-1}) indicate that the rate-limiting step of *Ps*P2Ox is the reductive-half reaction, *i.e.*, the oxidation of D-Glc to 2-keto-D-Glc, similarly to other studied POx (**Supplementary Figures 1 and 2**)[22,23]. The compound used to measure *Ps*P2Ox enzymatic activity in the coupled assay with horseradish peroxidase (HRP)[20], 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS) is after oxidation to a green cation radical, reduced by the enzyme to the dicationic form, leading to an underestimation of the enzymatic activity

(**Supplementary Table 2**). Therefore we have optimized the assay using substrates 4-aminophenazone (AAP) and 3,5-dichloro-2-hydroxybenzene-sulfonic acid (DCHBS), which originate after oxidation of a pink chromogen, N-(4-antipyryl)-3-chloro-5-sulfonate-p-benzoquinone-monoimine[24], which is not reduced by PsP2Ox (see Supplementary Information). This assay allowed re-estimate the catalytic parameters for monosaccharides D-Glc, D-Xyl, D-Gal, D-Ara, and D-Rib (**Table 1 and Supplementary Figure 3**). Still, the obtained results confirmed that the enzyme is a poor biocatalyst for the oxidation of monosaccharides. In contrast, and notably, the K_m of PsP2Ox for molecular oxygen is 1 to 3 orders of magnitude lower than the described for other POxs (**Supplementary Table 4**). Following the substrate preference of recently characterized CarA enzymes[10] and ScP2Ox[13] towards the oxidation of C- and O-glycosides, the activity of PsP2Ox was tested for mangiferin (Mang), carminic acid, and rutin (**Supplementary Figure 4**); PsP2Ox is inactive towards carminic acid and rutin; however, Mang is oxidized at a catalytic efficiency (k_{cat}/K_m) four orders of magnitude higher than D-Glc, displaying a k_{cat} around 40-fold higher and a K_m 1000-fold lower (**Table 1**), revealing a striking selectivity for the oxidation of the sugar moiety of the C-glycoside. To identify the product(s) of the reaction, enzymatic reactions with Mang were monitored by TLC (**Supplementary Figure 5**) and characterized using a combination of 1D and 2D-NMR (**Supplementary Figures 6, 7, and 8 and Supplementary Table 5**). The NMR data showed that oxidation must occur at the 3-OH, affording the 3-keto mangiferin (**Fig 1a**). Similarly to Mang ^{1}H NMR, the anomeric proton of the reaction product was easily identified at 5.17 ppm (**Fig. S6**). Using this peak as a starting point, COSY NMR (**Supplementary Figure 8a**) was used to assign neighbor protons: the anomeric H1 correlates to the H2 (4.69 ppm), and the H5 (3.27 ppm) is correlated to H4 (4.14 ppm) and H6 (3.55 ppm). The peaks of H2 and H4 are doublets (**Supplementary Table 5 and Supplementary Figure 8a**), indicating the absence of the coupling constants with a vicinal H3. The chemical shift values of H2 and H4 also appear to have shifted downfield, as well as their carbons (correlated through HMQC in **Supplementary Figure 8b**), indicating the existence of a more vicinal electronegative group due to the oxidation at C3. Due to the significantly higher specificity of the enzyme towards the C3 position of Mang compared to the oxidation of the C2 position of D-Glc, we renamed this enzyme from PsP2Ox to PsG3Ox. The phylogenetic analysis of bacterial and fungal POxs (**Fig 1b**;[13]) corroborates a clear division between those with a higher specificity towards the oxidation of D-Glc and other monosaccharides at the C2 position, represented by well-characterized fungal P2Oxs (**Fig. 1b**, green clade), and those with lower specificity for D-Glc and higher specificity towards the oxidation of C-glycosides at the C3-position, the bacterial G3Ox (**Fig. 1b**, blue clade). Interestingly, the bacterial KaP2Ox, the closest bacterial member to fungal counterparts, shows a high specificity for D-Glc[25]; (**Supplementary Table 4**). To investigate the presumable role of PsG3Ox in C-glycosides catabolism, a Basic Local Alignment Search Tool (BLAST) was performed in the genome of the drought-tolerant $P.$ $siccitolerans$ to find putative C-deglycosylases in the neighborhood of PsG3Ox coding gene. The PsG3Ox belongs to a gene cluster that has a similar organization to other soil bacteria[9]. Downstream of the $psg3ox$ gene, two homologs, PsG3OxC1, and PsG3OxC2, were found showing 37% and 33% identity when compared to CarC enzyme and two other ORFs, PsG3OxB1 and PsG3OxB2, shows 49% and 38% identity to CarB, which is C-glycoside deglycosylases (CGDs) that catalyze the cleaving of the C-C bond between the sugar and the aglycone moiety ([9]; **Fig 1c**). These

results hint at the role of *Ps*G3Ox as part of the *C*-deglycosylation catabolic pathway with a putative biological function in, e.g., sugar uptake as a carbon source from natural glycosylated compounds.

***Ps*G3Ox overall structure.** The crystal structure of *Ps*G3Ox was refined to 2.01 Å resolution, and the crystal structures of *Ps*G3Ox-Glc and *Ps*G3Ox-Mang complexes were refined at 2.35 and 2.60 Å resolutions (**Supplementary Table 2**). All crystals belong to space group C2221, showed similar cell dimensions, and contained one molecule in the asymmetric unit (*a.u.*). The overall structure of *Ps*G3Ox is composed of a flavin- and a substrate-binding domain, comparable to other POx structures (**Fig. 2a**). The flavin domain has a Rossmann fold-like structure of class α/β and contains one non-covalently-bond-FAD molecule. The substrate-binding domain comprises six-stranded central β-sheets (β8-β13) and six α-helices (α2-α6 and α9). *Ps*G3Ox shows an *r.m.s.d.* value of 0.84 Å compared with homologous Cα positions of the recently characterized bacterial *Mt*CarA (PDB 7DVE)[10]. This is in contrast to the significantly higher *r.m.s.d.* values (1.70-1.81 Å) of structures of P2Oxs from the fungal origin, *T. multicolor* (PDB 1TT0), *Peniophora sp*. (PDB 1TZL) and *P. chrysosporium* (PDB 4MIF). *Ps*G3Ox shows, when compared to the fungal enzymes, i) a significantly more solvent-exposed FAD cavity (**Fig. 2b,c**), ii) smaller monomers' size (~ 500 residues instead of the ~ 600 residues) (**Fig. 2f**), iii) a monomeric, instead of the tetrameric fungal state (**Fig. 2b,c**), and, iv) significantly higher <a.d.p.>'s values (52 Å^2) than fungal enzymes (12-30 Å^2). The structural model of *Sc*P2Ox predicted using RoseTTaFold[13] shows a monomeric enzyme and a similar overall fold to *Mt*CarA and *Ps*GO3x.

The structural alignment revealed that *Ps*G3Ox (as well as *Mt*CarA) display nine deletions and four insertions when compared to fungal enzymes (**Fig. 2f**). Two of the deletions (**boxes 1 and 3**) correspond to the oligomerization loop and arm regions involved in the fungal P2Ox inter-subunit interactions (**Fig. 2a**)[17]. An additional deletion (**boxes 9, 10, and 11**) corresponds to a region in fungal enzymes known as the head domain. A shorter insertion 319-321 (**box 12 in Fig. 2f**) is located near the flavin cofactor, while the insertions 186-190 (**box 5**) and 506-509 (**box 13**) are more distant from the active site (**Supplementary Figure 9**). The major insertion in *Ps*G3Ox is the insertion-1 segment containing 33-residues (**box 2**), forming two α-helices, α2 (60-71) and α3 (83-88), together with two loops (72-82 and 89-93), which are close to the substrate-binding domain and in the neighboring of the FAD access (**Fig. 2d**). Interestingly, insertion-1 is located in an equivalent structural position to the oligomerization domains of fungal enzymes (**Supplementary Figure 10a**), i.e., on the interaction interfaces; mainly, the region 60-70 of *Ps*G3Ox is close to the fungal oligomerization loop, while region 71-93 (the most flexible part of insertion-1, from now on referred as the insertion-1 loop) is located near the fungal oligomerization arm domain. Insertion-1 is also present in bacterial homologs (**Supplementary Figure 10b-f**), except for the dimeric bacterial *Ka*P2Ox, which contains two oligomerization domains comparable to those in fungal enzymes (**Supplementary Figure 10g**). The segments that differentiate monomers from oligomers tend to be located on the interaction interfaces, where they mediate or disrupt oligomerization and are usually loops[26]. It has been claimed that homooligomerization in glycosyltransferases and other proteins might be crucial for their function[27]. We speculate that the evolutionary mechanism of POxs

homooligomerization can hypothetically occur through the deletion and insertion of segments in the region where insertion-1 locates, promoting the stabilization of dimers and tetramers [28].

Catalytic center in *PsG3Ox*. The FAD cofactor in *Ps*G3Ox is non-covalently bonded, similar to other characterized bacterial enzymes, except *Ka*P2Ox, and in contrast to fungal enzymes where FAD is covalently linked through its 8α-methyl group, *e.g.,* H167[NE2] in *Tm*P2Ox (PDB 2IGK). *Ps*G3Ox's equivalent residue, H127, is approximately 6.5 Å away from the FAD[C8M] atom (**Fig. 2e**). *In vitro* deflavinylation (**Supplementary Figure 11**) followed by FAD incorporation, revealed that Apo-*Ps*G3Ox binds exogenous FAD with an estimated dissociation constant $K_D = (2.0 \pm 0.7) \times 10^{-7}$ M restoring the enzymes' total activity (**Supplementary Figure 12**). The estimated K_D value indicates a high affinity for FAD, comparable to other flavoenzymes (for more details, see Supplementary Information)[29-31]. The access to the FAD cavity is made through a cavity that contains (i) residues [125]AAHW[128] that is at the same structural position of fungal flavinylation motif, [165]STHW[168] in *Tm*P2Ox, that covalently binds FAD, (ii) the substrate loop ([346]ASPVPLADD[354]), which in fungal enzymes, is reportedly a dynamic gating segment that fine-tunes the enzymes' reactivity [15,32], and (iii) the insertion-1 segment (60-93 residues) (**Fig. 2d**).

We have attempted to trigger a covalent attachment of FAD to H127 in *Ps*G3Ox by replacing residues A125 and A126 with serine and threonine to mimic the fungal flavinylation motif (**Fig. 2f**). The single variants A125S and A126T still show non-covalently bonded FAD (**Supplementary Figure 13**) and 2- to 3-fold decreased catalytic efficiency than wild-type (**Supplementary Table 6**). The double variant S125A-T126A was produced in the apo-form. Considering that the N120 side-chain interacts through a hydrogen bond with H127, which may preclude the covalent binding to FAD (**Fig. 2e**), site-directed mutagenesis was used to construct N120V, since a valine is present in the fungal *Tm*P2Ox at that position. Still, this replacement also impaired FAD incorporation leading to the production of an apo-enzyme. The site-saturation mutagenesis of N120 resulted in two active variants, N120C and N120Y, displaying a FAD non-covalently bond (**Supplementary Figure 13**). The replacement of conserved H440 and N484 residues by alanine resulted in inactive enzymes, confirming their key catalytic role (**Supplementary Table 6**); H440 is expected to act as a proton acceptor for the C2-OH group of sugar substrates, with the support of N484, that stabilizes the protonated intermediate through a hydrogen bond, after the hydride transfer of C2 hydrogen to the FAD[N5] atom[32,33]. *Ps*G3Ox's substrate loop, [346]ASPVPLADD[354], has three polar and six hydrophobic residues and shows comparable structural flexibility to the fungal *Tm*P2Ox substrate loop ([452]DAFSYGAVQ[460]) (**Supplementary Table 7**). In the primary sequence alignment, this region is very well conserved among bacterial enzymes (**Supplementary Figure 10h**), except in the case of *Ka*P2Ox, which exhibits an amino acid composition ([369]DAFHYGDVP[377]) comparable to the fungal enzymes. To assess the role of the loop in the catalytic properties of *Ps*G3Ox, a substrate loop truncated variant (Δloop 345-359) was constructed and characterized. This variant shows a comparable catalytic efficiency (k_{cat}/K_m) for D-Glc and Mang to the full-length protein, even though the productive binding is negatively affected: whereas the affinity for both substrates increases, the turnover number (k_{cat}) is 3 to 5-fold lower (**Supplementary Table 6**). The insertion-1 segment in *Ps*G3Ox lowers the substrate loop and active site

exposure to solvent (**Supplementary Table 8**). Deleting the insertion-1 loop between residues 73-93 (its most flexible region) resulted in an inactive, FAD-depleted variant. After *in vitro* flavinylation, a fully loaded FAD enzyme was obtained, but enzymatic activity was not recovered (**Table 1**). These results suggest that insertion-1 is involved in FAD incorporation and proper stabilization at the active site and plays a crucial role in catalysis.

Conformational loop changes upon substrate binding and catalysis. The structures of complexes *Ps*G3Ox-Glc and *Ps*G3Ox-Mang are similar to the native structure with *r.m.s.d.* values of 0.40 Å and 0.42 Å. (**Fig. 3a-c; Supplementary Figure 14a**) but display significantly (almost 2-fold) higher FAD cavity volumes (**Supplementary Table 9**). Furthermore, these structures show higher *a.d.p.s* values and lack visible electron density maps of residues close to the insertion-1 and substrate loop regions (**Supplementary Table 10; Fig. 3b,c; Supplementary Figure 14c,d**). The conformation of these non-visible regions was modeled using Rosetta, and the loop candidates were scored based on the lowest possible Rosetta energies (**Fig. 3d-f**). In the *Ps*G3Ox substrate-free, similarly to *Mt*CarA, the substrate loop displays a closed conformation, and the active site has a smaller dimension and higher content of hydrophobic residues (**Fig. 3d,g**). In *Ps*G3Ox-Glc and *Ps*G3Ox-Mang structures, the substrate loop adopts a semi-open and open conformation, respectively. The insertion-1 loop, accordingly to the Rosetta best models, also displays a closed and an open conformation in the substrate-free structure and substrate-complexes, respectively (**Fig. 3h-i**). In fungal P2Oxs, the substrate loop follows open-to-closed conformational transitions that discriminate between electron-donor and electron-acceptor substrates and probe the regioselective oxidation at the C2 or C3 position of monosaccharides (**Supplementary Figure 15**)[34-36]. Therefore, to provide additional insights into the conformational landscape of the substrate loop and insertion-1 loop along the catalytic cycle of *Ps*G3Ox conventional MDs (**Fig. 4k-m; Supplementary Figure 16-18**) and GaMDs (**Fig. 4a-j; Supplementary Figure 19-21**) simulations were performed. The A352-FAD^{N5} and G84-FAD^{N5} distances were used to discriminate between closed ($\leq$ 20 Å) and open ($\geq$ 25 Å) conformations of the loops. For the substrate-free Model I, the substrate loop and the insertion-1 loop remain closed during simulations (**Supplementary Figure 18**), even when an enhanced sampling is pursued with the GaMD simulations (**Fig. 4a,e; Supplementary Figure 21**). This observation suggests a clear preference for a closed active site in the absence of substrates, in agreement with crystallographic data. The hydrophobic composition of the substrate loop hypothetically favors this conformation. Contrary, when D-Glc is bound in the active site, the simulations suggest that the semi-open substrate loop conformation (with closed insertion 1, Model II) is the only one capable of retaining D-Glc properly bound, with stable interactions with catalytic and non-catalytic residues (**Fig. 4b,f, and Supplementary Figures 16-22**); in a semi-open substrate loop conformation there is more available space in the A346-P348 zone to accommodate D-Glc (purple triangle in **Fig. 3h**). In the first Model II simulation, the A352-FAD^{N5} and G84-FAD^{N5} distances remain ~15-20 Å (**Supplementary Figures 18 and 21**). However, in the second GaMD simulation, the loops adopt a more open form (distances A352-FAD^{N5} > 20 Å and G84-FAD^{N5} ~ 20 Å, which may relate to D-Glc movement inside the binding cavity, losing some of the required catalytic interactions and finally leaving the active site (**Supplementary Figures 21 and 22**). Supporting the idea that the open loop conformation cannot properly retain D-Glc, in Model III simulations (open

loops, with initial A352-FAD^{N5} and G84-FAD^{N5} distances ~ 30 Å), D-Glc goes away from the active site after 300 ns in cMD and before 200 ns in GaMD (**Supplementary Figure 22**). Interestingly, when long GaMD simulations are run starting from open loops conformations in Model III*, without D-Glc in the active site, substrate loop, and insertion-1 loop follow open to closed transition (**Fig. 4c,g; Supplementary Figures 19 and 21**). This agrees with the preference mentioned above for closed states in substrate-free *Ps*G3Ox. The open-to-closed transitions (**Fig 4i**) are also clear by the main movements (PC1) of the protein in this trajectory, as seen in the Principal Component Analysis (PCA) (**Fig. 4d,h**). Significant higher *r.m.s.f.* values were found for insertion-1 and substrate loop that started with open conformations (**Fig. 4j**). Other simulations, e.g., starting from the closed models, showed much lower fluctuations (**Supplementary Figures 17 and 20**), corroborating the higher stability of the enzyme in a closed state. Overall, the simulations performed with D-Glc suggest that the substrate loop adopts a narrow range of conformations that allow effective oxidation of this substrate, affecting productive binding and explaining the poor catalytic parameters obtained for D-Glc. *In silico* molecular dynamics using Mang as substrate (Model IV) indicates that the substrate and insertion-1 loops adopt an open conformation (**Fig. 4k,l**) which is required to enlarge the active site pocket and allow bulkier substrates such as glycosides to bind (**Fig. 3f**). During the 400 ns of MD simulations of Model IV (open loops), both loops remain open (distance A352-FAD^{N5} > 30 Å and G84-FAD^{N5} > 20-25 Å) (**Fig. 4k,l; Supplementary Figure 18**), and Mang remains correctly oriented for ~280 ns, suggesting a stable complex. However, at the end of the simulation, Mang leaves its oxidation position in the active site. Interestingly, it moves towards the substrate loop, where the mangiferin's aglycon motif is accommodated between the loop's hydrophobic residues P348 and V349 and the residue F431 (**Fig. 4m**). This suggests that the aglycon motif of substrates such as Mang can interact with the substrate loop during substrate recruitment and/or product release. The dynamic transitions of the substrate loop and insertion-1 loop hint at a joint function of these two regions, not only in the access to the active site but also in the proper accommodation and oxidation of substrates.

Enzyme-substrate interactions. In the crystal structure of the *Ps*G3Ox-Glc complex, D-Glc^{O2} is located at 4.2 and 4.7 Å of the catalytic residues H440^{NE2} and N484^{ND2}, and at 4.2 Å from FAD^{N5} (**Fig. 5a**). The hydrogen atom at the C2 atom of D-Glc points to the flavin N5 atom, which might facilitate a hydride transfer and support D-Glc oxidation at the C2 position (Mendes et al., 2016). In fungal P2Oxs, D-Glc establishes shorter distances (2.5-3.0 Å) to the FAD^{N5} and catalytic residues[34,37]. Molecular docking of D-Glc into *Ps*G3Ox revealed a wider network of interactions than those observed in the crystal structure (**Fig. 5a,b; Supplementary Figures 23, 24, and 25**). The binding of D-Glc positioned for C2, and C3 oxidation (based on the corresponding catalytic distances) was abundant (**Supplementary Figure 23**). Remarkably, in the semi-open system, a higher C2/C3 ratio was obtained, suggesting a preference for oxidation of D-Glc at the C2 oxidation (**Supplementary Figure 23**). This observation supports that a semi-open substrate loop conformation is essential to bring and retain D-Glc close to the FAD and catalytic residues and properly orient D-Glc for C2 oxidation. In the *Ps*G3Ox-Mang crystal complex (and dockings), more hydrogen bonds were observed, which may explain the higher affinity and catalytic efficiency of *Ps*G3Ox for this substrate compared with D-Glc (**Fig 5c,d; Supplementary Figures 26 and 27**). Notably, in both

experimental and computational dockings, a clear preference for the positioning of Mang oriented towards the C3 oxidation was observed in contrast to the C2 positioning followed for D-Glc (**Fig. 5c,d**; **Supplementary Figure 26**). Computational dockings were performed to explore the structural reasons behind the lack of activity toward carminic acid and rutin (**Table 1**; **Supplementary Figure 28a**). Carminic acid binds in a non-catalytically competent manner: H2/C2 is relatively well oriented towards FAD^{N5}, but it is the OH3 that interacts with catalytic H440; no interaction with the catalytic residue N484 was observed. In the case of rutin, the dockings predict binding in an orientation compatible with catalysis at C3 (**Supplementary Figure 28b**). Inhibition assays show a 5-fold lower inhibition constant (K_i) for rutin (0.1 mM) as compared to carminic acid (**Supplementary Figure 29**), suggesting that rutin can bind more strongly to the active site, supporting the dockings' data; FAD reduction was not observed in enzymatic assays performed in anaerobic conditions in the presence of rutin (data not shown). Therefore we hypothesize that the absence of activity with rutin might be associated with its intrinsic properties, e.g., a hypothetic redox potential that impairs electron transfer. However, this possibility was not investigated here.

Alanine mutagenesis was used to investigate the importance of residues K55, R94, T129, Q297, and Q340 in substrate binding and catalysis (**Fig. 5e; Supplementary Table 6**). In general, the replacement to Ala resulted, with a few exceptions, in decreased k_{cat} and comparable K_m values to the wild type. The absence of Q297 and, particularly, of Q340 has a significant detrimental effect in catalysis for both substrates, even if slightly higher for D-Glc. On the other hand, the oxidation of Mang is particularly affected in R94A and T129A variants, whereas the oxidation of D-Glc is primarily unaffected. These two variants, along with K55A and Q340A, display higher K_m values for Mang and are thus expected to be involved in Mang's binding and productive pose in the active site. K55A shows a similar k_{cat} for Mang following its furthest location from the FAD and catalytic residues in the docking simulation models (**Fig. 5d**), contrary to the complex crystal structure where K55 is located nearby the aglycon motif of the mangiferin (**Fig. 5b**). The results obtained with R94 and Q297 are by the MDs that suggest a role in anchoring bulky substrates. In *Ps*G3Ox, the residue T129 establishes hydrogen bonds with $P348^N$ and FAD^{N5} (**Supplementary Figure 14**), similar to residue T169 in *Tm*P2Ox (**Supplementary Figure 15**); however, whereas T169 was suggested to trigger the transition between the different conformations of the substrate loop [34-37], T129 shows the same configuration independently of the substrate loop conformation (**Supplementary Figure 14**).

Concluding Remarks

This work establishes the distinct structural, functional, and mechanistic features between two groups within the POx family of enzymes. These enzymes are known to have a clear preference for the oxidation of D-Glc, at the C2 or C3 positions, yielding the respective ketoaldoses as products. However, it is clear the existence of two distinct phylogenetic groups with different activity profiles in this family: those with a higher specificity towards the oxidation of the monosaccharide D-Glc, the P2Oxs, and those with a higher specificity towards the oxidation of the D-Glc moiety of C-glycosides, the G3Ox group. These latter are

primarily inactive towards D-Glc oxidation. At the same time, the first group is most likely inactive to the bulkier C-glycosides, considering their overall tetrameric fold with buried active sites, in contrast with monomeric G3Ox with a more accessible active site. This work revealed that insertion-1, a striking structural segment of G3Ox's active site, plays a vital role in catalysis. Insertion-1 locates in bacterial GO3xs at the same region where the oligomerization loop and domain locates in dimeric and tetrameric P2Ox. Therefore we speculate that structural changes in these (non-conserved) regions may contribute to the formation of different functional oligomeric states that might be important in the regulation of enzyme activity; i.e., we propose that the catalytic mechanisms and binding specificities within the POx family of enzymes are modulated through different homooligomerization states. All POx members show, at their active site, a substrate loop that plays a prominent role in regioselectivity control by modulating its conformation depending on the type/size of the substrate. In *Ps*G3Ox and *Mt*CarA, the substrate loop is projected into the active site cavity, sterically occluding the binding of substrates, contrary to P2Oxs fungal enzymes. This propensity to adopt a closed conformation probably relates to its hydrophobic composition and articulated interaction with the insertion-1 segment, which contributes to lowering the substrate loop and active site exposure to solvent. In these enzymes, the dynamic transitions of substrate loop and insertion-1 hint at a joint function in the access to the active site and in allowing proper accommodation of the substrates. In the presence of small substrates such as D-Glc, a semi-open conformation must be adopted to create enough space for substrate binding and orientation towards C2 oxidation while simultaneously avoiding an excessive enlargement of the active-site cavity that would result in the loss of substrate, impairing the catalysis. Most likely, the difficulties in establishing this equilibrium, either by a lack of trigger(s) for the formation of the semi-open state or its maintenance, are reflected in the low specificity of G3Ox to D-Glc. In contrast, an open conformation is mandatory for the binding of Mang. In *Ps*G3Ox, hydrogen-bond interactions between the substrate loop and the substrate were not observed, in opposition to fungal P2Oxs; however, the *Ps*G3Ox substrate loop can create a "hydrophobic clamp" that interacts with the aglycone moiety of mangiferin, suggesting a role of the substrate loop to assist the substrate recruitment or product release. Active-site residues also play a role in anchoring the substrate/product to the enzyme. Overall, a combined experimental and computational investigation of *Ps*G3Ox allowed mapping the relationships between the enzyme sequence-structure-function and elucidating functional transitions that accompany substrate binding and release. Such changes highlight the fine control of access to the catalytic site required by the enzyme mechanism and, in turn, the specificity offered by the enzyme towards different substrates. Work is ongoing to explore the molecular determinants for substrate specificity among bacterial POxs. This undertaking is essential for advancing fundamental biochemical insights in protein science and shedding light on the diversity of microbial catabolic pathways of natural compounds.

Declarations

Author contributions

L.O.M. conceived this study; A.T. performed the mutagenesis, kinetic, and biochemical characterization of enzymes; T.F., P.B., and C.F. solved the crystal structures; M.V.R and R.V. characterize reaction products; T.F, X.F.L., M.F.L, F.S, and L.M. performed MD simulations, docking, and computational analysis. All authors contributed to the experimental design, interpretation of results, and writing the paper.

Competing Interests

The authors declare no competing interests.

References

1. Hultin, P.G. Bioactive C-glycosides from bacterial secondary metabolism. *Curr Top Med Chem* 5, 1299-331 (2005).

2. Zhang, Y.Q., Zhang, M., Wang, Z.L., Qiao, X. & Ye, M. Advances in plant-derived C-glycosides: Phytochemistry, bioactivities, and biotechnological production. *Biotechnol Adv* 60, 108030 (2022).

3. Yang, Y. & Yu, B. Recent Advances in the Chemical Synthesis of C-Glycosides. *Chem Rev* 117, 12281-12356 (2017).

4. Bililign, T., Griffith, B.R. & Thorson, J.S. Structure, activity, synthesis and biosynthesis of aryl-C-glycosides. *Nat Prod Rep* 22, 742-760 (2005).

5. Kytidou, K., Artola, M., Overkleeft, H.S. & Aerts, J.M.F.G. Plant Glycosides and Glycosidases: A Treasure-Trove for Therapeutics. *Front Plant Sci* 11 (2020).

6. Braune, A., Engst, W. & Blaut, M. Identification and functional expression of genes encoding flavonoid O- and C-glycosidases in intestinal bacteria. *Environ Microbiol* 18, 2117-2129 (2016).

7. Nakamura, K., Zhu, S., Komatsu, K., Hattori, M. & Iwashima, M. Deglycosylation of the Isoflavone C-Glucoside Puerarin by a Combination of Two Recombinant Bacterial Enzymes and 3Oxo-Glucose. *Appl Environ Microbiol* **86** (2020).

8. Kim, E.M., Seo, J.H., Baek, K. & Kim, B.G. Characterization of two-step deglycosylation via oxidation by glycoside oxidoreductase and defining their subfamily. *Sci Rep* **5** (2015).

9. Mori, T. et al. C-Glycoside metabolism in the gut and in nature: Identification, characterization, structural analyses and distribution of C-C bond-cleaving enzymes. *Nat Commun* **12** (2021).

10. Kumano, T. et al. FAD-dependent C-glycoside-metabolizing enzymes in microorganisms: Screening, characterization, and crystal structure analysis. *Proc Nat Acad Sci USA* **118** (2021).

11. Sanugul, K., Akao, T., Nakamura, N. & Hattori, M. Two proteins, Mn2+, and low molecular cofactor are required for C-glucosyl-cleavage of mangiferin. *Biol Pharm Bull* **28**, 2035-2039 (2005).

12. Sutzl, L., Foley, G., Gillam, E.M.J., Boden, M. & Haltrich, D. The GMC superfamily of oxidoreductases revisited: analysis and evolution of fungal GMC oxidoreductases. *Biotechnol Biofuels* **12** (2019).

13. Kostelac, A. et al. Biochemical Characterization of Pyranose Oxidase from *Streptomyces canus*-Towards a Better Understanding of Pyranose Oxidase Homologues in Bacteria. *Int J Mol Sci* **23** (2022).

14. Abrera, A.T., Sutzl, L. & Haltrich, D. Pyranose oxidase: A versatile sugar oxidoreductase for bioelectrochemical applications. *Bioelectrochem* **132** (2020).

15. Wongnate, T. & Chaiyen, P. The substrate oxidation mechanism of pyranose 2-oxidase and other related enzymes in the glucose-methanol-choline superfamily. *FEBS J* **280**, 3009-27 (2013).

16. Sutzl, L. et al. Multiplicity of enzymatic functions in the CAZy AA3 family. *Appl Microbiol Biotechnol* **102**, 2477-2492 (2018).

17. Hallberg, B.M., Leitner, C., Haltrich, D. & Divne, C. Crystal structure of the 270 kDa homotetrameric lignin-degrading enzyme pyranose 2-oxidase. *J Mol Biol* **341**, 781-96 (2004).

18. Bannwarth, M., Bastian, S., Heckmann-Pohl, D., Giffhorn, F. & Schulz, G.E. Crystal structure of pyranose 2-oxidase from the white-rot fungus *Peniophora* sp. *Biochemistry* **43**, 11683-90 (2004).

19. Hassan, N. et al. Crystal structures of *Phanerochaete chrysosporium* pyranose 2-oxidase suggest that the N-terminus acts as a propeptide that assists in homotetramer assembly. *FEBS Open Bio* **3**, 496-504 (2013).

20. Mendes, S. et al. Characterization of a bacterial pyranose 2-oxidase from *Arthrobacter siccitolerans*. *J. Mol. Cat. B: Enzymatic* **133**, 34-43 (2016).

21. Busse, H.J. Review of the taxonomy of the genus *Arthrobacter*, emendation of the genus *Arthrobacter sensu lato*, proposal to reclassify selected species of the genus *Arthrobacter* in the novel genera *Glutamicibacter* gen. nov., *Paeniglutamicibacter* gen. nov., *Pseudoglutamicibacter* gen. nov., *Paenarthrobacter* gen. nov and *Pseudarthrobacter* gen. nov., and emended description of *Arthrobacter roseus*. *Int J Syst Evol Microbiol* **66**, 9-37 (2016).

22. Prongjit, M., Sucharitakul, J., Wongnate, T., Haltrich, D. & Chaiyen, P. Kinetic mechanism of pyranose 2-oxidase from *Trametes multicolor*. *Biochemistry* **48**, 4170-80 (2009).

23. Sucharitakul, J., Wongnate, T. & Chaiyen, P. Kinetic isotope effects on the noncovalent flavin mutant protein of pyranose 2-oxidase reveal insights into the flavin reduction mechanism. *Biochemistry* **49**, 3753-65 (2010).

24. Fossati, P., Prencipe, L. & Berti, G. Use of 3,5-dichloro-2-hydroxybenzenesulfonic acid/4-aminophenazone chromogenic system in direct enzymic assay of uric acid in serum and urine. *Clin Chem* **26**, 227-31 (1980).

25. Herzog, P.L. et al. Versatile Oxidase and Dehydrogenase Activities of Bacterial Pyranose 2-Oxidase Facilitate Redox Cycling with Manganese Peroxidase In Vitro. *Appl Environ Microbiol* **85** (2019).

26. Hashimoto, K., Madej, T., Bryant, S.H. & Panchenko, A.R. Functional states of homooligomers: insights from the evolution of glycosyltransferases. *J Mol Biol* **399**, 196-206 (2010).

27. Hashimoto, K. & Panchenko, A.R. Mechanisms of protein oligomerization, the critical role of insertions and deletions in maintaining different oligomeric states. *Proc Natl Acad Sci U S A* **107**, 20352-7 (2010).

28. Kumari, N. & Yadav, S. Modulation of protein oligomerization: An overview. *Prog Biophys Mol Biol* **149**, 99-113 (2019).

29. Caldinelli, L., Pedotti, M., Motteran, L., Molla, G. & Pollegioni, L. FAD binding in glycine oxidase from Bacillus subtilis. *Biochimie* **91**, 1499-508 (2009).

30. Fraaije, M.W., van Den Heuvel, R.H., van Berkel, W.J. & Mattevi, A. Structural analysis of flavinylation in vanillyl-alcohol oxidase. *J Biol Chem* **275**, 38654-8 (2000).

31. Villegas, J.M. et al. FAD binding properties of a cytosolic version of Escherichia coli NADH dehydrogenase-2. *Biochim Biophys Acta* **1844**, 576-84 (2014).

32. Wongnate, T., Surawatanawong, P., Chuaboon, L., Lawan, N. & Chaiyen, P. The Mechanism of Sugar C-H Bond Oxidation by a Flavoprotein Oxidase Occurs by a Hydride Transfer Before Proton Abstraction. *Chem Eur J* **25**, 4460-4471 (2019).

33. Wongnate, T., Sucharitakul, J. & Chaiyen, P. Identification of a catalytic base for sugar oxidation in the pyranose 2-oxidase reaction. *Chembiochem* **12**, 2577-86 (2011).

34. Kujawa, M. et al. Structural basis for substrate binding and regioselective oxidation of monosaccharides at C3 by pyranose 2-oxidase. *J Biol Chem* **281**, 35104-15 (2006).

35. Tan, T.C., Haltrich, D. & Divne, C. Regioselective control of beta-d-glucose oxidation by pyranose 2-oxidase is intimately coupled to conformational degeneracy. *J Mol Biol* **409**, 588-600 (2011).

36. Spadiut, O., Tan, T.C., Pisanelli, I., Haltrich, D. & Divne, C. Importance of the gating segment in the substrate-recognition loop of pyranose 2-oxidase. *FEBS J* **277**, 2892-909 (2010).

37. Tan, T.C. et al. H-bonding and positive charge at the N5/O4 locus are critical for covalent flavin attachment in trametes pyranose 2-oxidase. *J Mol Biol* **402**, 578-94 (2010).

38. Kabsch, W. Xds. *Acta Crystallogr. D* **66**, 125-132 (2010).

39. Evans, P.R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D* **67**, 282-92 (2011).

40. Evans, P.R. & Murshudov, G.N. How good are my data and what is the resolution? *Acta Crystallogr D* **69**, 1204-14 (2013).

41. Vonrhein, C. et al. Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr D* **67**, 293-302 (2011).

42. Kantardjieff, K.A. & Rupp, B. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Prot Sci* **12**, 1865-1871 (2003).

43. Matthews, B.W. Solvent Content of Protein Crystals. *J. Mol. Biol.* **33**, 491-497 (1968).

44. Oeffner, R.D., Bunkoczi, G., McCoy, A.J. & Read, R.J. Improved estimates of coordinate error for molecular replacement. *Acta Crystallogr D* **69**, 2209-15 (2013).

45. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60**, 2126-32 (2004).

46. Blanc, E. et al. Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr D* **60**, 2210-21 (2004).

47. Roversi, P., Blanc, E., Vonrhein, C., Evans, G. & Bricogne, G. Modelling prior distributions of atoms for macromolecular refinement and completion. *Acta Crystallogr D* **56**, 1316-23 (2000).

48. Borges, P.T. et al. Unveiling molecular details behind improved activity at neutral to alkaline pH of an engineered DyP-type peroxidase. *Comput Struct Biotechnol J* **20**, 3899-3910 (2022).

49. Stein, A. & Kortemme, T. Improvements to Robotics-Inspired Conformational Sampling in Rosetta. *Plos One* **8**, e63090 (2013).

50. Kocourek, J., Ticha, M. & Kostir, J. The use of diphenylamine-alanine-phosphoric acid reagent in the detection and differentiation of monosaccharides and their derivatives on paper chromatograms. *J Chromatogr* **24**, 117-24 (1966).

51. Krieger, E., Nielsen, J.E., Spronk, C.A.E.M. & Vriend, G. Fast empirical pK(a) prediction by Ewald summation. *J Mol Graph Model* **25**, 481-486 (2006).

52. Krieger, E. & Vriend, G. New ways to boost molecular dynamics simulations. *J Comput Chem* **36**, 996-1007 (2015).

53. Hornak, V. et al. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins* **65**, 712-725 (2006).

54. Mark, P. & Nilsson, L. Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *J. Phys. Chem. B* **105**, 9954-9960 (2001).

55. Miao, Y., Feher, V.A. & McCammon, J.A. Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. *J Chem Theory Comput* **11**, 3584-3595 (2015).

56. D.A. Case, H.M.A., K. Belfon, I.Y. Ben-Shalom, J.T. Berryman, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, G.A. Cisneros, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, and P.A. Kollman AMBER 2020. (University of California, San Francisco., 2020).

57. Wang, J.N. et al. Gaussian accelerated molecular dynamics: Principles and applications. *Wiley Interdisciplinary Reviews-Computational Molecular Science* **11** (2021).

58. Canutescu, A.A. & Dunbrack, R.L. Cyclic coordinate descent: A robotics algorithm for protein loop closure. *Prot Sci* **12**, 963-972 (2003).

59. Trott, O. & Olson, A.J. Software News and Update AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J Comput Chem* **31**, 455-461 (2010).

60. Land, H. & Humble, M.S. YASARA: A Tool to Obtain Structural Guidance in Biocatalytic Investigations. *Methods Mol Biol* **1685**, 43-67 (2018).

61. Krieger, E., Darden, T., Nabuurs, S.B., Finkelstein, A. & Vriend, G. Making optimal use of empirical energy functions: Force-field parameterization in crystal space. *Proteins-Structure Function and Bioinformatics* **57**, 678-683 (2004).

Methods

Bacterial strains, plasmids, and cultivation media. *Escherichia coli* strains, plasmids, and primers used in this work are summarized in **Table S1**. *E.coli* strain DH5α (Novagen, Darmstadt, Germany) was used to propagate and amplify plasmid constructs. *E. coli* Rosetta pLysS (DE3, Novagen, Darmstadt, Germany) was used to express the wild-type *PsG3Ox* and its variants cloned in the pET-15b plasmid (Novagen, Darmstadt, Germany). KRX (Promega, Wiscosin, USA) was used as expression strain in the SSM library screenings. Luria Bertani medium (LB) was used for cell cultivation, supplemented with 100 μg ml^{-1} of ampicillin (NZYTech, Lisbon, Portugal) and, in the case of Rosetta pLysS, also with 20 μg ml^{-1} of chloramphenicol (NZYTech, Lisbon, Portugal).

Construction of variants by site-directed mutagenesis. The Quick-change mutagenesis protocol (Stratagene, California, USA) was used in the construction of single variants, as well as to delete two regions of the enzyme, the insertion-1 loop, between residues 73 and 93, and the substrate-binding loop between residues 345 and 359, in variant Δloop (345-359) (**Table S1**). The plasmid pSM-1, carrying the wild-type *PsG3Ox* gene, was used as a DNA template with appropriate primers. PCRs were performed in a thermal cycler (MyCycler^{TM} thermocycler, Biorad) in 50 μL reaction volume containing 100 ng of DNA template, 1 μM of primers (forward and reverse), and 200 μM of dNTPs (NZYTech, Lisbon, Portugal). 1 U of NZYProof polymerase (NZYTech, Lisbon, Portugal) was used to amplify the DNA, except for pAT27 and pAT28 where 1 U of Q5 High-Fidelity DNA polymerase (New England BioLabs, Massachusetts, USA) was used. For the single and double mutants, after an initial denaturation of 4 min at 95 °C, 25 cycles of 1 min at 95 °C, 1.5 min at 72 °C, and 10 min at 72 °C were performed, followed by a final elongation of 10 min at 72 °C. Amplification of truncated variants was performed after an initial denaturation of 30 sec at 98 °C, 35 cycles of 10 sec at 98 °C, 30 sec at 72 °C and 4 min at 72 °C, followed by a final elongation of 2 min at 72 °C. For all the PCR products, the DNA template was digested with 10 U of *Dpn*I (ThermoFisher, Massachusetts, USA) at 37 °C for 6 h, followed by purification using Illustra GFX PCR DNA kit (GE Healthcare, Illinois, USA). In the truncated variants, an overnight ligation with T4 ligase (ThermoFisher,

Massachusetts, USA) was performed at room temperature, followed by purification with the abovementioned kit. The PCR products were transformed in *E. coli* strains using electroporation, and the presence of the desired mutation(s) or deletions was confirmed by DNA sequencing.

Production and purification of *PsG3Ox*. The recombinant strains *E. coli* Rosetta pLysS carrying the genes coding for wild-type *PsG3Ox* and variants were grown in 2.5 L LB media supplemented with 100 µg ml^{-1} of ampicillin and 20 µg ml^{-1} of chloramphenicol in Corning® 5L Baffled PETG Erlenmeyer flasks. The cultures were incubated at 37 °C, 100 rpm (Innova 44 incubator shaker, New Brunswick Scientific). Cultures were induced with 100 µM of isopropyl β-D-1-thiogalactopyranoside (IPTG) at an $OD_{600\ nm}$ = 0.8, the temperature was lowered to 25 °C, and cells were collected by centrifugation (4420 × g, 10 min, 4°C) after 16 h of cultivation. The purification of wild-type *PsG3Ox* and variants was performed using a Histrap HP column (Cytiva, Massachusetts, USA) as previously described (Mendes et al., 2016). Enzyme preparations for crystallographic trials were purified using a 1 ml-Resource-Q column (Cytiva, Massachusetts, USA) using 20 mM Tris-HCl pH 7.6 as running buffer and a gradient of 0-500 mM of NaCl for elution. Before the enzyme crystallization, the His(6x)-tag was cut from the enzyme using the thrombin cleavage kit (Abcam, Cambridge, UK) following the manufacturer protocol at 20ºC for 16 h. Afterward, the preparation was loaded on the Histrap HP column, and the flowthrough (containing the untagged enzyme) was collected. The total protein concentration was determined by Bradford assay using bovine serum albumin as standard. Abs_{450nm} (M^{-1} cm^{-1}) of purified preparations was measured to assess the functional fraction of enzyme preparations (*e.g.*, for kinetic measurements).

Crystallization and cryoprotection. Crystallization conditions were screened with a nanodrop crystallization robot (Cartesian, Genomic Solutions) using the sitting drop vapour diffusion method with round-bottom Greiner 96-well CrystalQuickTM plates (Greiner Bio-One, Kremsmünster, Austria). The Structure Screen I and II (Molecular Dimensions) led to the formation of *PsG3Ox* crystals within seven days at 20 ºC using 2 M ammonium sulfate and 0.1 M Tris-HCl, pH 8.5 in drops of 0.1 µl protein solution (18 mg ml^{-1}) plus 0.1 µl reservoir solution. Following this crystallization hit, microliter-scale crystals optimization proceeded using the hanging drop vapour diffusion method in XRL 24-well crystallization plates with 500 µL of the reservoir solution (Molecular Dimensions, Newmarket, UK). Several conditions were tested: ammonium sulfate concentration ranging from 0.5 to 2M, Tris-HCl, pH 7.0-8.5, and different ratios (1:2, 1:1, and 2:1) of protein: reservoir solution volumes. Yellow and round crystals appeared after 7-10 days, reaching dimensions of 100 µm in their three dimensions when using 2 M ammonium sulfate and 0.1 M Tris-HCl, pH 8.5, and 1 µL of protein (18 mg ml^{-1}) and 2 µL of reservoir solution at 20 °C. Crystals of the *PsG3Ox*-substrate complex were obtained by soaking *PsG3Ox* crystals in 1) the reservoir solution containing 2M D-Glc for 30 min and 2) the reservoir solution containing 1 mM Mang for 1 min. Crystals were cryo-protected by plunging in a reservoir solution supplemented with 20% (*v/v*) glycerol before flash-cooling in liquid nitrogen.

Data collection and processing. Diffraction data were measured in ID23-2 and ID30A-1 beamlines at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) and in XALOC beamline at ALBA

(Barcelona, Spain) for *PsG3Ox*, *PsG3Ox*-Glc and *PsG3Ox*-Mang complex crystals, respectively. Diffraction images of *PsG3Ox* were obtained with a DECTRIS PILATUS3 X 2M detector, using 0.8731 Å radiation wavelength, crystal-to-detector distance 232 mm, and oscillations width 0.20° in a total of 360° rotation. The diffraction data of the *PsG3Ox*-Glc complex were obtained with a PILATUS3 2M detector, using 0.9654 Å radiation wavelength, crystal-to-detector distance of 238 mm, and oscillations width of 0.20° in a total of 180° rotation. The diffraction data of the *PsG3Ox*-Mang complex were obtained with a DECTRIS PILATUS 6M detector, radiation wavelength 0.97926 Å, crystal-to-detector distance of 553.95 mm, and oscillations width of 0.20° in a total of 180° rotation. Data were indexed and integrated with XDS[38] in a space group determined with POINTLESS[39], and the data scaled with AIMLESS[39,40]. These programs were used within the autoPROC data processing pipeline [41]. Data collection details and processing statistics are listed in **Table S2**.

Structure determination, refinement, and analysis. The three crystals belong to the same space group and show similar cell dimensions. The distribution of their Matthews coefficient[42,43] indicated a high probability of a single molecule in their asymmetric units. The phase problem of *PsG3Ox* was solved by molecular replacement using MORDA that selected the coordinates of *P. chrysosporium* POx (*PcP2Ox*, PDB 4MIF) and *Alkalihalobacillus halodurans* ATP phosphoribosyltransferase regulatory subunit structure (PDB 3OD1) as search models, which led to a 99% probability solution. PHASER, within the PHENIX suite, was used to localize the two *PsG3Ox*-substrates structures using native *PsG3Ox* structure as a search model, which led to TFZ values of 35 and 42, indicating successful structures solutions[44]. Automated model rebuilding and completion were performed with PHENIX.AUTOBUILD followed by manual model building, performed with COOT[45], and iterative refinement cycles, using PHENIX.REFINE.

Structure refinement included atomic coordinates, isotropic atomic displacement parameters (*a.d.p.s*), and domains of translation, libration, and screw refinement of anisotropic *a.d.p.s* (TLS), previously defined with TLSMD server (http://skuld.bmsc.washington.edu/~tlsmd). Approximately 1.5% of reflections were randomly excluded from monitoring the refinement strategy. Solvent water molecules were automatically assigned from σA difference maps peaks neighboring hydrogen bonding acceptors/donors within 2.45-3.40 Å distances. Other solvent molecules were identified through a comparison of their shapes against electron density blobs, as well as by comparing their refined *a.d.p.s* with those of neighboring atoms. Some atoms were modeled with partial occupancies when hinted by the difference in Fourier maps and neighboring *a.d.p.s* values. As some regions of the *PsG3Ox* and *PsG3Ox*-substrate complex structures were not visible in the electron density maps, a BUSTER protocol was applied to search for missing atoms[46,47]. The stereochemistry of the refined structures was analyzed with MOLPROBITY. Three-dimensional superposition of polypeptide chains was performed with MODELLER. Figures of structural models were prepared with PyMOL. Refinement statistics are presented in **Table 1**. Structure factors and associated structure coordinates of *PsG3Ox*, *PsG3Ox*-Glc, and *PsG3Ox*-Mang complex were deposited in the Protein Data Bank (www.rcsb.org) with PDB codes 7QF8, 7QFD, and 7QVA, respectively.

Loop modeling. Rosetta loop modeling was used to build the non-visible loops in the *Ps*G3Ox (202-204 and 309-318), *Ps*G3Ox-Glc (77-89, 201-206, 308-320, and 345-359), and *Ps*G3Ox-Mang (74-90, 201-206, 309-324 and 351-359) complex structures using a methodology previously described[48]; a full-atom refinement step was performed with the next-generation kinematic (NGK) closure robotics-inspired conformational sampling protocol[49]. The crystal structures of *Ps*G3Ox, *Ps*G3Ox-Glc, and *Ps*G3Ox-Mang were kept intact except for the loop regions that were created, with repacking of the side chains within 10Å of the remodeled region. A total of 500 loops were built, and to find the best loop candidate; each model was scored by its Rosetta energy score and contacts with the substrate molecules.

Apparent steady-state kinetics. Apparent steady-state kinetics measurements were performed at 37 ºC in 100 mM sodium phosphate buffer, pH 7.5, and reactions were started with the addition of enzyme. The kinetic parameters for D-glucose (D-Glc, PanReac Applichem, Darmstadt, Germany), D-galactose (PanReac Applichem, Darmstadt, Germany), D-ribose (VWR, Pennsylvania, USA), D-xylose (Sigma Aldrich, Missouri, USA), and L-arabinose (Sigma Aldrich, Missouri, USA) were measured using coupling assay containing 0.1 mM 4-Aminoantipyrine (AAP, Acros organics, Geel, Belgium), 1 mM 3,5-dichloro-2-hydroxybenzenesulfonic acid sodium salt (DCHBS, Alfa Aesar, Massachusetts, USA), 8 U ml^{-1} Horseradish peroxidase (HRP, PanReac Applichem, Darmstadt, Germany) and different concentrations of substrate. Enzymatic activity was monitored using a Synergy2 microplate reader (BioTek, Vermont, USA) following the formation of N-(4-antipyryl)-3-chloro-5-sulfonate-p-benzoquinone-monoimine (a pink chromogen) at 515 nm (ε_{515}= 26,000 M^{-1} cm^{-1}). The kinetic parameters for molecular oxygen were measured in an Oxygraph system (Hansatech instruments, Pentney, UK) to follow oxygen consumption in reactions containing 1 M D-Glc as an electron donor and different oxygen concentrations pre-set by bubbling O_2 or N_2 gas. The oxidation of the glycosides mangiferin (Sigma Aldrich, Missouri, USA), rutin (Acros Organics, Geel, Belgium), and carminic acid (Sigma Aldrich, Missouri, USA) were followed by oxygen consumption in the Oxygraph apparatus in reactions containing 0 - 2 mM of Mang, 0 - 0.5 mM of rutin or 0 − 1 mM of carminic acid. Specific activity was calculated considering the preparation's functional (FAD-loaded) enzyme ratio. Apparent steady-state kinetic parameters (k_{cat} and K_m) were determined by fitting data directly into the Michaelis-Menten equation using Origin-Lab software. For inhibition assays, the steady-state kinetics for D-Glc were performed as described below in the presence of 0 - 0.5 mM rutin or 0 - 0.2 mM carminic acid. The data was represented using a Lineweaver Burk plot and the inhibition constants were estimated based on a secondary plot of the slopes against inhibitor concentration.

Identification of mangiferin oxidation product. Oxidation of Mang was performed under aerobic conditions at 25ºC, pH 7.5 in 30 mL of Milli-Q water containing 20 mg of Mang, 1 U ml^{-1} of Catalase (Sigma Aldrich, Missouri, USA), and 1 U ml^{-1} of *Ps*G3Ox. To estimate the time needed to have a high yield of oxidized Mang, a time-course of the reaction was performed in a thin layer chromatography (TLC) on silica gel 60 F254 sheet (Merck, Darmstadt, Germany) using a mixture of butanol, acetic acid, and water in the proportions 4:1:2.2 (v/v) as mobile phase. The TLC revealed a diphenylamine-aniline-phosphoric acid reagent[50], a system used to distinguish sugars. For the NMR characterization, the reaction occurred for 30 min, and then the enzymes were removed by ultrafiltration using a vivaspin20 of 30 kDa cutoff

(Cytiva, Massachusetts, USA). The water in the mixture was evaporated under low pressure on a rotary vacuum evaporator, and the resulting sediment was resuspended in ~ 600 of dimethyl sulfoxide-d6 (Merck, Darmstadt, Germany). The Mang and the reaction product, both in DMSO-d6, were analyzed through ^{1}H, ^{13}C APT, COSY, and HMQC NMR in a Bruker Avance II+400. ^{1}H NMR spectra were obtained at 400 MHz and ^{13}C at 100.61 MHz.

Molecular dynamics simulations. Molecular dynamics simulations (MD) of *Ps*G3Ox were carried out to further explore the conformational preferences of the enzyme at different stages of the catalytic cycle. First, MD simulations of 400 ns production runs were performed of four systems built as follows: i) Model I: taken from the *Ps*G3Ox crystal structure (without substrates), which shows closed substrate and insertion-1 loops; ii) Model II: obtained from the same crystal structure but with the substrate loop sampled with the Yasara Sample Loop function and D-Glc docked (see below) in a binding mode compatible with C2 oxidation; this model presents a semi-open substrate loop and a closed insertion-1 loop; iii) Model III: prepared from the *Ps*G3Ox-Mang crystal structure, with Mang removed, the missing parts of the loops were built with the Yasara Build Loop function, and D-Glc docked in a C2 binding mode; this model shows open substrate and insertion-1 loops; iv) Model IV: derived from the *Ps*G3Ox-Mang crystal structure (in a binding mode compatible with C3 oxidation) by building the missing parts of the open substrate and insertion-1 loops. Missing hydrogen atoms were added, and the protonation state of the titratable residues was assigned with the Yasara hydrogen bond networks optimization and pKa prediction tools at pH 7[51]. The systems were solvated with a solvation box and neutralized with NaCl. The conventional MDs, cMDs, were set up and run using Yasara[52]; the AMBER14 force field[53] and TIP3P water model[54] were used. A two-step equilibration was carried out for 400 ps: first, the system's temperature was increased in 10 steps from 30K up to 300K, followed by a second step with constant temperature and box dimensions. The bonds and angles involving a hydrogen atom were fixed. A restrain on all non-hydrogen atoms of the complexed protein-ligand was applied so that the equilibration would mainly affect the system's solvent. Bonded and non-bonded forces were updated every 2 and 5 fs, respectively. The protein-ligand complex was then released, and a production step of 400 ns was carried out. Gaussian accelerated MDs (GaMDs)[55] were then run for Model I - III to access effective longer simulation times and wider conformational explorations, especially of the substrate loop. The starting structures for these GaMD simulations were the 100 ns structure from the corresponding classical MDs. An extra model (Model III*) was built from Model III by removing the bound D-Glc substrate to see if loop transitions could be observed. The GaMD simulations were run with the AMBER 20 program[56]. The GaMD protocol consisted of an initial equilibration stage where the potential boost was applied, boost parameters were updated, and production runs were updated with fixed boost parameters. A dual boost on dihedral and total potential energy was applied (igamd = 3). Two simulations of 600 ns were run for each Model I, II, and III*, whereas the simulation for Model III was stopped after 200 ns as D-Glc was observed to left the active-site. GaMD inputs were generated following the recommendations from the developers[57]. The convergence of all simulations was assessed by calculating the RMSD values of the

protein C_{alpha} atoms. The distances between FAD^{N5} and A352, as well as G84, were used as an indicator of the conformational state of the substrate and insertion-1 loops, respectively.

Protein-ligand docking. Protein-ligand docking calculations were carried out on different protein structures from the cMD trajectories using AutoDock VINA[58-60] D-Glc was docked in *Ps*G3Ox's active site among 100 frames from semi-open (Model II) and open (Model III) substrate's loop MDs, respectively. On the other hand, Mang was docked among 100 frames from the open substrate loop MD (Model IV). The rest of the substrate's loop conformations, particularly those with the closed loop, did not allow proper positioning of the substrate in the active site due to steric clashes, and they were discarded. D-Glc was docked using the YAMBER force field[61] on a 20 × 20 × 20 Å cuboid docking cell centered on FAD^{N5}, while Mang was docked on a 34 × 34 × 34 Å cuboid docking cell centered on the same atom. A total of 16 ligand conformations were generated per frame, potentially yielding 1600 ligand-enzyme combinations per ligand. Distances between the glycoside's D-Glc group and FAD^{N5} (N5-HC2; N5-HC3), as well as between H440 (NE2-HO2; NE2-HO3), were measured to filter structures in potential catalytically relevant C2 (NE2-HO2 and N5-HC2 < 4 Å) and C3 (NE2-HO2 and N5-HC2 < 4 Å) binding modes that could go into an optimization protocol, with a minimization step of the whole complex. The initial (pre-optimization) and final distances between FAD^{N5}/H440 and the D-Glc group of each glycoside, as well as the total energies of the entire system and the ligand binding energies, were obtained for further analysis. The minimum distances between the ligand and residues Q297, Q340, R94, T129, K55, were measured. The number of events (frequency) for each measured distance and energy, depending on the ligand-binding mode and substrate loop conformation, were plotted in different histograms. Rutin and carminic acid were docked following the same protocol as for Mang.

Table

Table 1. Apparent steady-state kinetic parameters of wild-type and variants *Ps*G3Ox for different substrates. The catalytic parameters for D-Glc and other monosaccharides were estimated using the HRP-AAP/DCHBS coupled assay; for molecular oxygen, the reactions were followed in an Oxygraph in the presence of 1 M D-Glc. The reactions with Mang were monitored by oxygen consumption in an Oxygraph. All reactions were performed in 100 mM sodium phosphate buffer at pH 7.5 and 37 ºC. The kinetic parameters were determined by fitting the data directly on the Michaelis-Menten equation using OriginLab.

Enzyme	Substrate	k_{cat} (s^{-1})	K_m (M)	k_{cat}/K_m (M^{-1}s^{-1})
Wild-type	D-Glucose	0.19 ± 0.03	0.46 ± 0.13	0.45 ± 0.09
Wild-type	D-Xylose	0.13 ± 0.02	1.00 ± 0.20	0.13 ± 0.01
Wild-type	D-Galactose	0.02 ± 0.00	0.58 ± 0.16	0.03 ± 0.01
Wild-type	L-Arabinose	0.02 ± 0.00	0.57 ± 0.04	0.03 ± 0.00
Wild-type	D-Ribose	0.01 ± 0.00	0.19 ± 0.08	0.06 ± 0.01
Wild-type	Dioxygen[a]	0.14 ± 0.04	(5.83 ± 2.10) ×10^{-6}	(19.87 ± 4.82) ×10^3
Wild-type	Mangiferin	8.13 ± 1.67	(0.49 ± 0.10) ×10^{-3}	(19.22 ± 2.71) ×10^3
Wild-type	Rutin	nd	-	-
Wild-type	Carminic acid	nd	-	-
Δloop (345-359)	D-Glucose	0.06 ± 0.01	0.36 ± 0.13	0.19 ± 0.06
Δloop (345-359)	Mangiferin	1.4 ± 0.4	(0.08 ± 0.02) ×10^{-3}	(23.6 ± 0.3) ×10^3
Δinsert1 (73-93)[b]	D-Glucose	-	-	(0.19 ± 0.01) ×10^{-2}
Δinsert1 (73-93)[b]	Mangiferin	[c]		

[a] D-Glc was used as an electron donor; nd – not detected

[b] Assays performed after *in vitro* flavinylation of the purified preparation; k_{cat}/K_m was obtained from the first-order approximation of the Michaelis-Menten equation ([S] « K_m)

[c] Residual activity was detected using 2.5 mM of Mang (V_{max} = 13.4 ± 1.7 nmol min^{-1} mg^{-1})
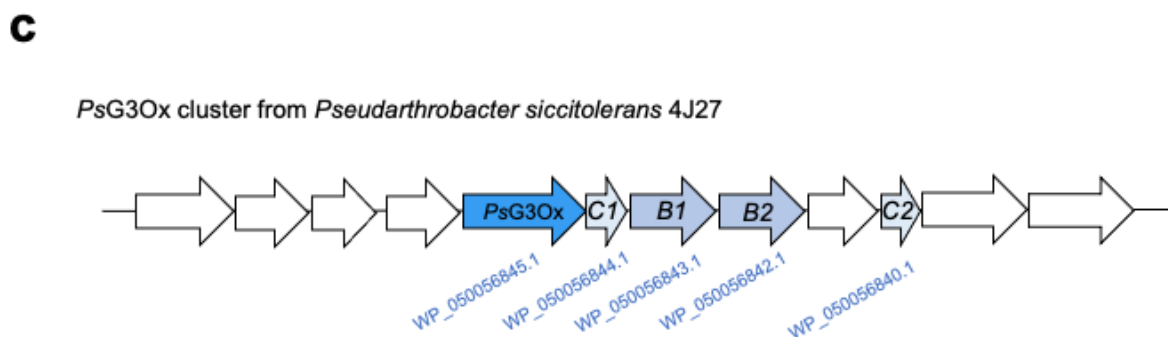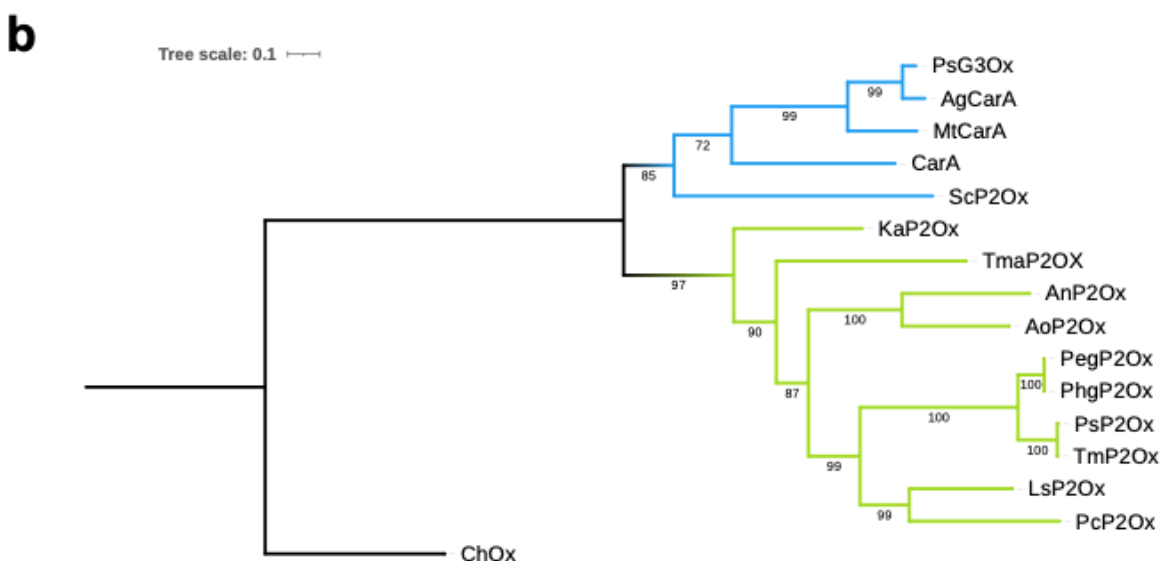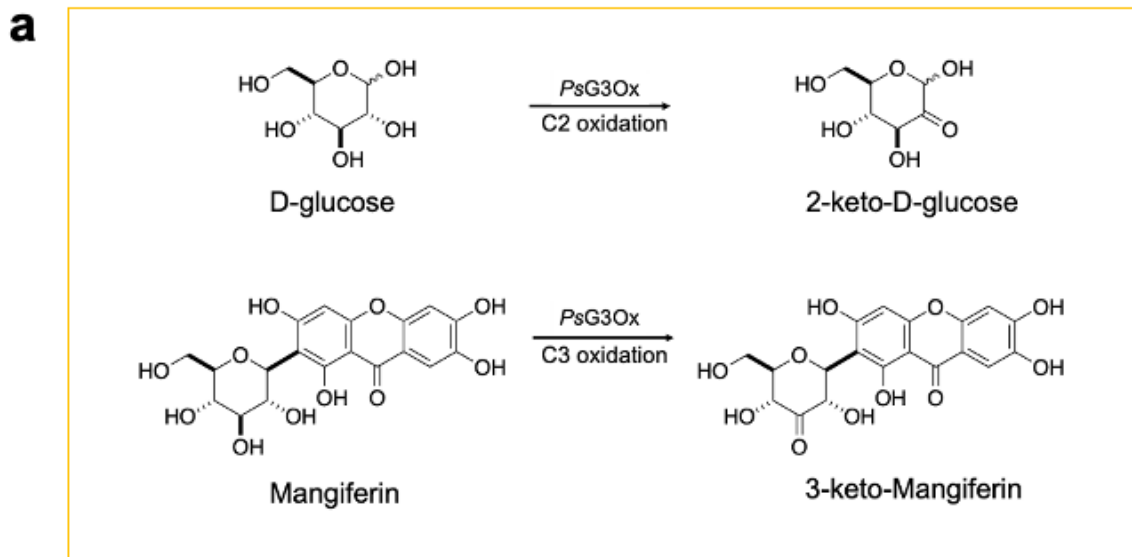
Figures

Figure 1

Reaction scheme of *Ps*G3Ox, phylogenetic analysis of P2Oxs/G3Oxs and *Ps*G3Ox *C*-deglycosilation gene cluster. **(a)** The reaction catalyzed by *Ps*G3Ox to convert D-Glc to 2-keto-D-Glc [20] and the conversion of Mang to 3-keto-Mang. **(b)** Maximum likelihood phylogenetic relationship between characterized P2Oxs; cholesterol oxidase sequence, ChOx, from *Streptomyces sp.* was used as outgroup. The blue cluster represent the bacterial enzymes that shows higher specificity for glycosides and the green cluster group

the dimeric bacterial enzyme (*Ka*P2Ox) and the tetrameric fungal enzymes that display higher specificity towards D-Glc. (c) The gene coding for *Ps*G3Ox is part of a putative *C*-deglycosylation catabolic pathway. Four proteins with similarity to known C-deglycosylates (CGDs) are encoded in close vicinity to the *psg3ox* gene in *P. siccitolerans* 4J27 genome: *C*1 and *C*2 display 37 %, and 33 % identity compared to CarC and *B1* and *B2* which displayed 49 % and 38 % identity compared to CarB present from *Microbacterium sp.* 5-2b [9].
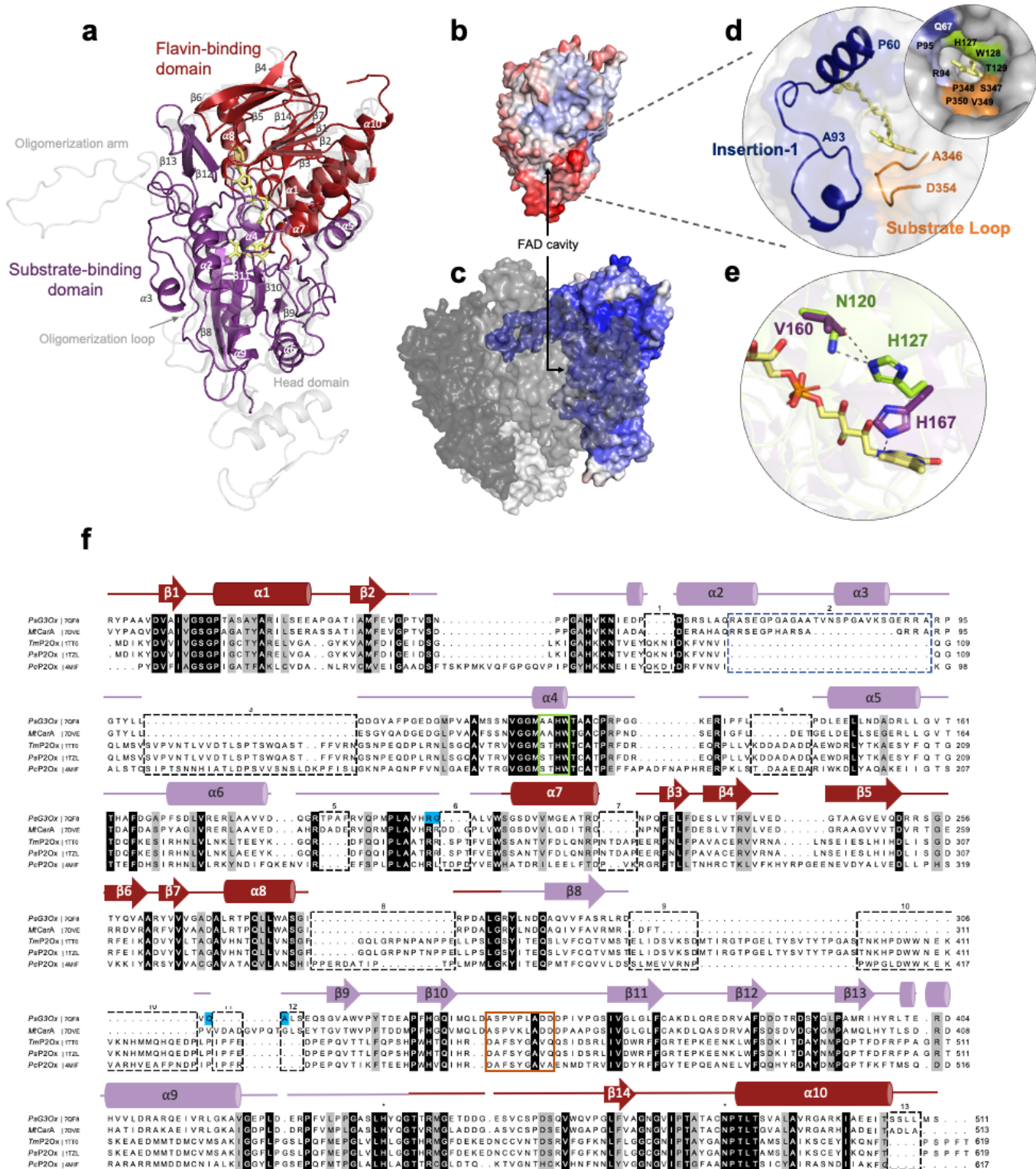
Figure 2

The structural fold of bacterial *Ps*G3Ox. (a) The overall structure of *Ps*G3Ox displays the flavin- and substrate-binding domains highlighted in dark red and purple, respectively. The monomer of *Tm*P2Ox (PDB 1TT0, light grey) is superimposed. The solvent-accessible surface of *Ps*G3Ox **(b)** and of one subunit of *Tm*P2Ox **(c)** are represented according to the *a.d.p* values, blue (6 Å^2) to red (107 Å^2). **(d)** Structural elements that surround the FAD cavity in *Ps*G3Ox and delimiting amino acid residues (defined using a 1.4-Å rolling probe). **(e)** Comparison of *Ps*G3Ox (residues in green) and *Tm*P2Ox (PDB 2IGK, residues in purple) flavinylation site. The H bonds are shown as black dashed lines. In all structures, the FAD is shown as sticks in yellow color. **(f)** Amino acid sequence alignment of bacterial and fungal POxs based on 3D superpositions of the crystal structure. The α-helices or β-chains are numbered and colored as in **(a)**. Catalytic residues are highlighted with *. The flavinylation motif and the substrate loop in fungal P2Oxs are marked with light green and orange boxes, respectively. The insertions and deletions regions are highlighted with dashed boxes. The cyan-marked residues correspond to the non-visible regions. Strictly conserved amino acids are represented on black background, whereas dark grey represents the most conserved residues among the selected sequences.
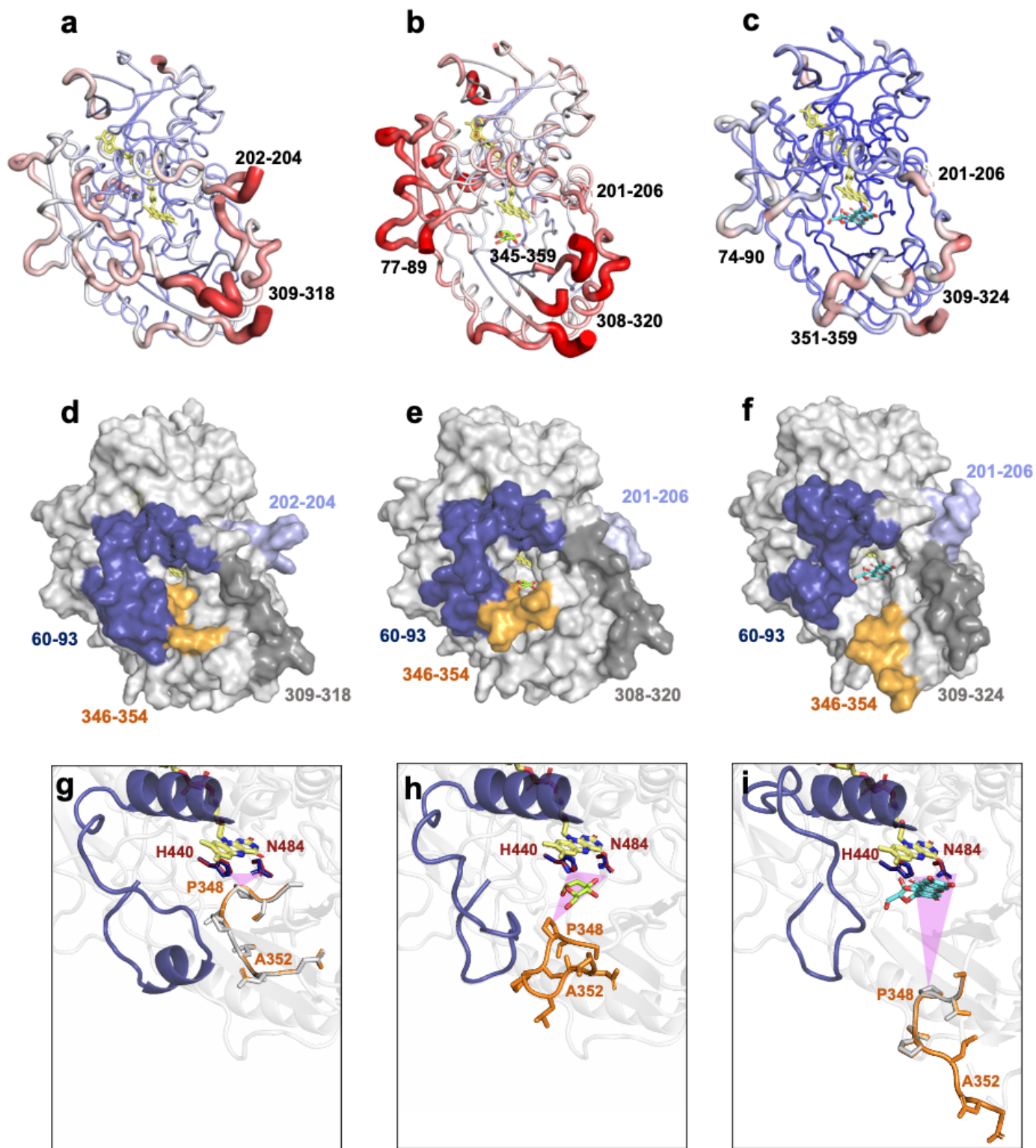
Figure 3

Conformational changes of *Ps*G3Ox upon substrate binding. X-ray structure of **(a)** substrate-free *Ps*G3Ox, **(b)** *Ps*G3Ox-Glc, and **(c)** *Ps*G3Ox-Mang complexes with thickness proportional to *a.d.p.* values, color-coded from blue (6 Å²) to red (107 Å²). Regions without electron density are highlighted near the structures. Structural models of **(d)** substrate-free *Ps*G3Ox, **(e)** *Ps*G3Ox-Glc, and **(f)** *Ps*G3Ox-Mang with the non-visible regions in the crystal structure modeled by Rosetta. The insertion-1 and substrate loop

(including the modeled segments) are colored dark blue and orange, respectively. Cartoon representation of the active site highlighting the insertion-1 and the conformation of the substrate loop (residues 346-354) in **(g)** substrate-free *Ps*G3Ox, **(h)** *Ps*G3Ox-Glc, and **(i)** *Ps*G3Ox-Mang complexes. The catalytic residues are shown as sticks colored in dark red or blue for the crystal structures or models. The residues in the substrate loop are shown as sticks in grey color and orange for crystal structures and models. The purple triangles represent the interatomic distances between the catalytic pair and the residue P348 of the substrate loop. In all structures, the FAD and the substrates D-Glc and Mang are shown as sticks and colored yellow, green, and cyan, respectively.
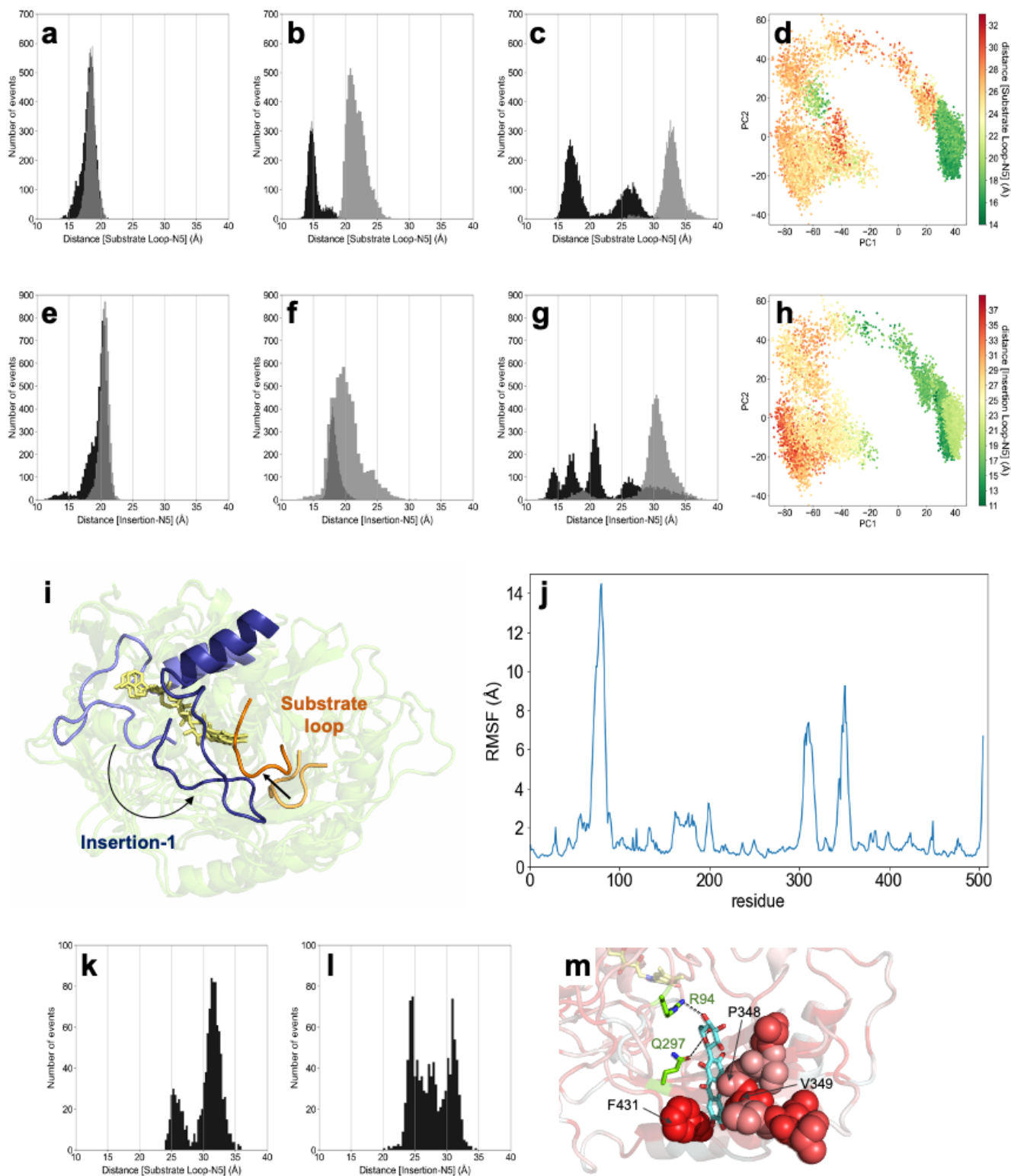
Figure 4

Conformational transitions of loops close to *Ps*G3Ox active site. Histograms of the substrate-loop distance to FAD (A352-FAD^{N5}) in GaMD simulations run for **(a)** Model I, which has no substrate and starts with a closed loops conformation, **(b)** Model II, which contains D-Glc and starts at semi-open conformation; and for **(c)** Model III*, that starts at open conformation and has no substrate. **(d)** Projection of the Model III* first trajectory onto the two first principal components obtained by PCA analysis, colored

accordingly to A352-FAD^{N5} distance, two main areas of structures are populated along PC1, which roughly correspond to the closed (green) and open (yellow and red) conformations of substrate loop. Histograms of the insertion-1 distance to FAD (G84-FAD^{N5}) in GaMD simulations run for **(e)** Model I, **(f)** Model II, and **(g)** Model III*. **(h)** Projection of the Model III* first trajectory onto the two first principal components obtained by PCA analysis, colored accordingly to G84-FAD^{N5} distance, two main areas of structures are populated along PC1, which roughly correspond to the closed (green) and open (yellow and red) conformations of the insertion 1. In all histograms of GaMD simulations, the black and grey colors represent two distinct simulations of 600 ns. **(i)** Cartoon representation of the first and last (600 ns) frames of the GaMD simulation of Model III*. The transition from open to closed state of substrate loop and insertion-1 are represented by a black arrow. FAD is colored yellow. **(j)** root-mean-square fluctuation (RMSF) of Model III* GaMD. Histograms of the **(k)** substrate-loop and **(i)** insertion-1 distances to FAD^{N5} in cMD simulation for Model IV that contains Mang and starts with open loops conformations. **(m)** Visual representation of the final frame (400 ns) of the corresponding simulation, where residues are colored according to hydrophobicity (white to red).
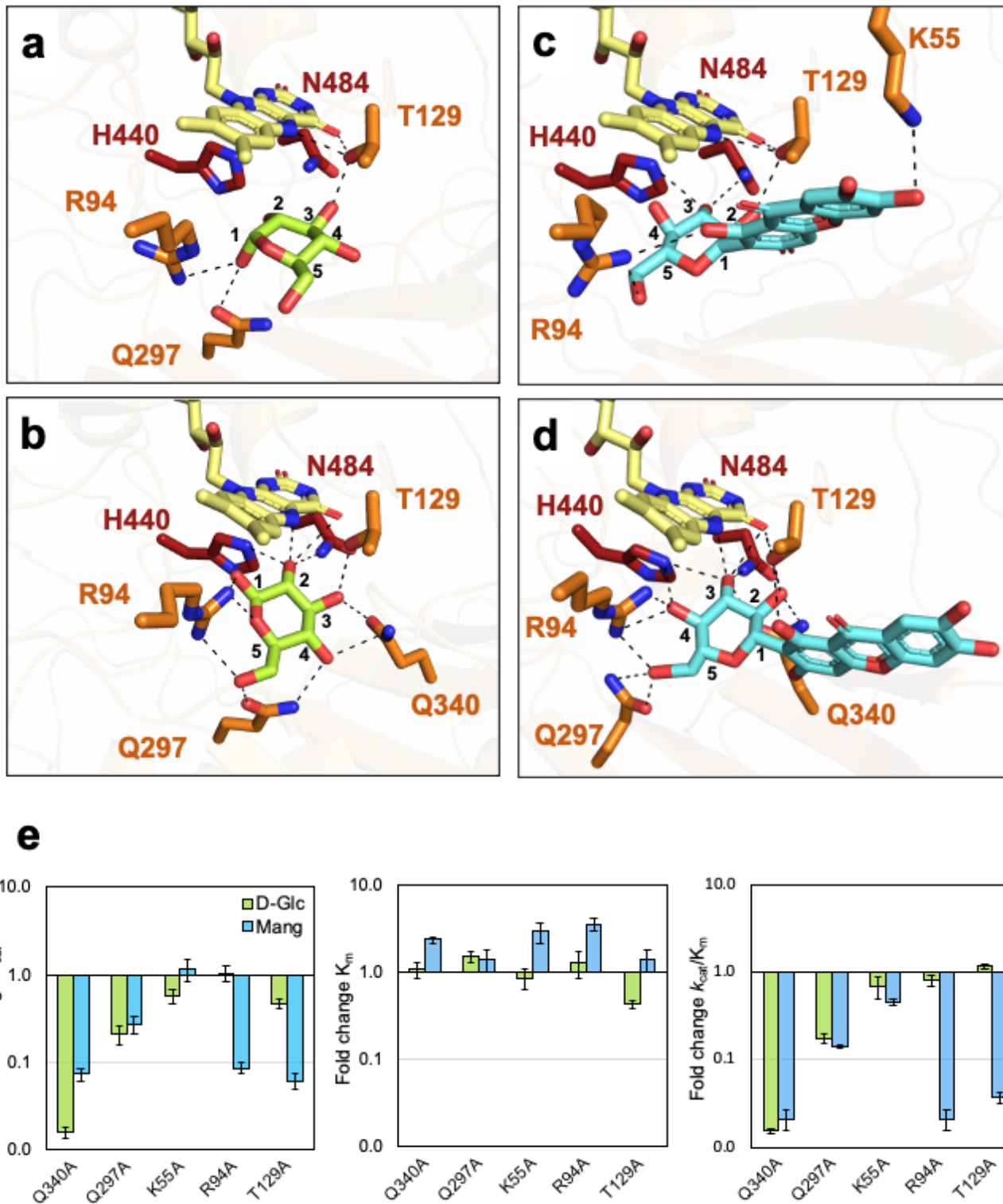
Figure 5

Interactions of *Ps*G3Ox with substrates. Binding of D-Glc in **(a)** *Ps*G3Ox-Glc crystal structure and **(b)** docking model. Binding of Mang in **(c)** *Ps*G3O*x*-Mang crystal structure and in **(d)** docking model. The catalytic residues (H440 and N484) are shown as sticks colored in dark-red and the non-catalytic interacting residues are colored in orange. The FAD and the substrates D-Glc and Mang are shown as sticks and colored in yellow, green, and cyan, respectively. The hydrogen bonds are shown as black

dashed lines. **(e)**Fold-change of the catalytic parameters compared with the wild-type *Ps*G3Ox for the alanine mutants at the non-catalytic interacting residues. The catalytic parameters for D-Glc were estimated using the HRP-AAP/DCHBS coupled assay, whereas reactions with Mang were monitored by oxygen consumption in an Oxygraph. All reactions were performed in 100 mM sodium phosphate buffer at pH 7.5 at 37 ºC. The kinetic parameters were determined by fitting the data directly on the Michaelis-Menten equation using OriginLab. The catalytic parameters are displayed in Table S6.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SIPsG3OxFeb13.docx