Characterization of Sorbitol Dehydrogenase SmoS from *Sinorhizobium meliloti* 1021

MacLean G. Kohlmeier, Ben A. Bailey-Elkin, Brian L. Mark, and Ivan J. Oresnik*

Department of Microbiology, University of Manitoba, Winnipeg, MB, Canada

* Corresponding author

E-mail: Ivan.Oresnik@umanitoba.ca

Abstract

Sinorhizobium meliloti 1021 is a Gram-negative alphaproteobacterium with a robust capacity for carbohydrate metabolism. The enzymes that facilitate these reactions assist in the survival of the bacterium across a range of environmental niches, and they may also be suitable for use in industrial processes. SmoS is a dehydrogenase that catalyzes the oxidation of the commonly occurring sugar alcohols sorbitol and galactitol into fructose and tagatose respectively using NAD^+ as a cofactor. The main objective of this study is to evaluate SmoS using biochemical techniques. The nucleotide sequence was codon optimized for heterologous expression in *E. coli* BL21 (DE3) GOLD cells, the protein was subsequently overexpressed and purified. Size exclusion chromatography and X-ray diffraction experiments suggest that SmoS is a tetrameric peptide. SmoS was crystallized to 2.1 Å in the absence of substrate and 2.0 Å in complex with sorbitol. SmoS was characterized kinetically and shown to have a preference for sorbitol despite a higher affinity for galactitol. Computational ligand docking experiments suggest that galactitol oxidation proceeds slowly because tagatose binds the protein in a more energetically favorable complex than fructose, and is retained in the active site for a longer time frame following oxidation which reduces the rate of the reaction. These results supplement the inventory of biomolecules with the potential for industrial applications and enhance our understanding of metabolism in the model organism *S. meliloti*.

Introduction

Sugar alcohols, also called polyols, are carbohydrate compounds that can be formed by the reduction of an aldo or keto sugar. The first polyols were identified from honeydew, a substance secreted by aphids as they feed on plant sap [1]. The most commonly encountered sugar alcohols in nature are sorbitol, mannitol, and galactitol (also known as dulcitol or melampyrite) [2]. These linear, six carbon polyols were named for the higher plants from which they originated; sorbitol from *Sorbus aucuparia*, mannitol from *Fraxinus ornis* or manna ash, and galactitol from *Melampyrum nemorosum* [1].

Sugar alcohols and their derivatives have a variety of applications. Sorbitol is commonly included in food products for sweetness, texture, and preservation, and can be present in pharmaceuticals [3, 4]. D-tagatose, a product of galactitol oxidation, is classified as a rare sugar and is being considered as a treatment for diabetes due to its insulin independent metabolism in humans and potential to lower blood glucose levels [5-7]. The concentrations of sugar alcohols in plant tissue are typically too low for chemical extractions to generate sufficient yields, therefor polyols are often synthesized for commercial use via catalytic hydrogenation of more readily available sugars [4]. However, biological enzymes can serve as biocatalysts for the generation of sugar alcohols and related molecules at an industrial scale. Some advantages to biocatalysts include high product selectivity and low environmental or physiological toxicity [8]. As an example, galactitol dehydrogenase has been immobilized on gold electrodes for use in electrochemical reactors with the goal of generating precursor molecules for pharmaceuticals via reactions that regenerate reduced cofactors [9-11].

66 Enzymes of microbial origin are ideal with respect to industrial applications as
67 they can be produced in an easy, cost effective, and consistent manner [12]. Carbohydrate
68 metabolism in plant associated soil bacteria has been studied in great detail due to the
69 involvement of carbon utilization in symbiotic establishment and efficiency [13, 14].
70 Transport genes responsible for the uptake of sorbitol, mannitol, and galactitol are
71 induced in the rhizosphere [15]. In bacteria, the initial step of sugar alcohol metabolism is
72 often oxidation into a keto sugar, followed by phosphorylation [16]. The root-nodulating
73 bacterium *Sinorhizobium meliloti* has been shown to produce a D-sorbitol specific
74 dehydrogenase (SDH), which uses NAD^+ as a cofactor [17]. A mutant lacking fructose
75 kinase activity was unable to grow using sorbitol as a sole carbon source, suggesting that
76 fructose is the product of sorbitol oxidation in *S. meliloti* [18]. A mutation to a gene
77 annotated as a putative sorbitol dehydrogenase *smoS* resulted in a strain with the inability
78 to grow on several sugar alcohols, including sorbitol [19], suggesting that *smoS* encodes
79 the SDH protein. *smoS* was first identified as encoding a SDH in *Rhodobacter*
80 *sphaeroides,* in which it was described as one gene in a novel polyol metabolic operon, as
81 well as a member of the short-chain dehydrogenase/reductase (SDR) family [20]. SDR
82 proteins are typically about 250 amino acids in length and despite having low sequence
83 identity at 20-30%, members of this family share a similar overall three-dimensional
84 structure [21]. Currently there are over 230,000 members of the SDR family in the
85 UniProt database, and a recently devised nomenclature system based on Hidden Markov
86 Models placed SmoS within the SDR196C subfamily [22, 23]. *Rs*SDH is dependent on
87 NAD^+ as a cofactor and has activity on sorbitol and galactitol (Fig. 1) [20]. Structural
88 studies on the protein in the absence of bound substrate were some of the first structures

of a bacterial SDH in the SDR family [24]. The purpose of this study is to characterize SmoS from *S. meliloti* with respect to its structure, as well as kinetic and physical properties.

Fig 1. Enzymatic reactions catalyzed by SmoS. Sorbitol or galactitol are oxidized at carbon 2, using NAD^+ as a cofactor, producing fructose or tagatose respectively, as well as NADH. Both sugar alcohols are viable substrates for SmoS due to the identical orientation of hydroxyl groups about carbons 1, 2, and 3.

Results

Structural characterization of *Sm*SmoS

Size exclusion chromatography of purified SmoS showed two distinct peaks at elution volumes of 49.22 mL and 55.35 mL (Fig. 2A), with the most prominent peak at ~55 mL likely representative of a tetrameric complex of SmoS. To corroborate these results, the column fractions were separated by nondenaturing polyacrylamide gel electrophoresis and stained with Coomassie Brilliant Blue. Fractions 2-4 showed two distinct bands, while fractions 5-9 contain a single band, which mimics the migration distance of the lower band from fractions 2-4 (Fig. 2B). Both protein bands are capable of sorbitol oxidation when the gel is stained for dehydrogenase activity (Fig. 2B), and resolve to a molecular weight of 27 kDa when SDS is included in the gel matrix (Fig. 2B), suggesting that both bands observed are due to the presence of SmoS.

Fig 2. Size exclusion chromatography of purified SmoS from a Superdex 75 gel filtration column and analysis by polyacrylamide gel electrophoresis (PAGE). (A) UV trace of elutions from the S75 column displaying two peaks at approximately 49 mL and 55 mL. (B) Elutions separated by nondenaturing PAGE stained with coomassie blue (top), elutions separated by nondenaturing PAGE stained for sorbitol dehydrogenase activity (middle), and elutions separated denaturing PAGE and stained with coomassie blue (bottom).

To further characterize SmoS, the enzyme was crystallized and determined to a resolution of 2.1Å (Fig. 3; Table 1). Consistent with this observation, SmoS crystallized as a tetramer, with four copies in the asymmetric unit arranged as a dimer of dimers, similarly to a previously determined structure of a *Bradyrhizobium japonicum* D-sorbitol dehydrogenase (Fig. 3A) [25]. These results are consistent with SmoS being present in two distinct conformations in solution, with the majority being tetrameric.

Table 1. Crystallographic and refinement statistics for SmoS and SmoS-sbt structures

Crystal	SmoS	SmoS-sbt
X-ray source	Rigaku MicroMax-007HF	Rigaku MicroMax-007HF
Crystal geometry		
Space group	P 1 2_1 1	P 1 2_1 1
Unit cell (Å)	*a*=83.30 *b*=88.30 *c*=87.32; α=90.00° β=117.39° γ=90.00°	*a*=83.30 *b*=88.30 *c*=87.32; α=90.00° β=117.39° γ=90.00°
Crystallographic data		
Wavelength (Å)	1.5419	1.5419
Resolution range (Å)	43.70-2.10 (9.62-2.10)*	39.60-2.0 (10.0-2.0)
Total observations	223880 (15515)	224791 (13489)
Unique reflections	65492 (4602)	74395 (4478)
Multiplicity	3.4 (3.4)	3.0 (3.0)
Completeness (%)	99.8 (99.9)	98.0 (99.7)
R_{merge}	0.117 (0.586)	0.139 (0.522)
CC1/2	0.99 (0.71)	0.98 (0.82)
I/σI	8.7 (2.2)	5.4 (2.0)
Wilson B-factor (Å^2)	19.11	20.80
Refinement statistics		
Reflections in test set	3314	3581
Protein atoms	7608	7588
Solvent molecules	737	914
R_{work}/R_{free}	0.192 / 0.252	0.194 / 0.249
RMSDs		
Bond lengths/angles (Å/°)	0.0081 / 1.464	0.0146 / 1.927
Ramachandran plot		
Favored/allowed (%)	97.44 / 2.17	97.15 / 2.26
Average B factor (Å^2)		
Macromolecules	25.09	20.64
Solvent	29.46	27.66

*Values in parentheses refer to the highest resolution shell

The SmoS monomer adopts a structural fold similar to other previously determined Zn-independent SDR enzymes, comprising of an NAD-binding Rossman fold centralized around a core 7-stranded parallel β-sheet, and an extended α-helical clamp-like lobe formed by helices α7 and α8 involved in substrate binding [24] (Fig. 3B). A DALI search [26] to identify structural homologues of SmoS identified a previously determined *Sm*SmoS structure (deposited by the New York Structural Genomic Consortium), and a *R. sphaeroides* sorbitol dehydrogenase (81% sequence amino acid identity), which aligned to *Sm*SmoS with an RMSD of 0.7 Å over 256 C_α atoms, and adopted a nearly identical structural fold [24].

Fig 3. Crystal structure of SmoS from *S. meliloti* 1021. (A) Cartoon representation of the SmoS quaternary structure. SmoS forms a homotetramer; the individual monomers are colored magenta, green, blue, and yellow. (B) Cartoon representation of the SmoS monomer (grey). Secondary structure elements are labeled numerically (α, α-helix; β, β strand; π, π helix; η, 3_{10} helix).

In an attempt to uncover the residues involved in substrate binding, SmoS was also crystallized in the presence of sorbitol and a structure determined to 2.0 Å (Fig.3; Table 1). Consistent with other described Zn-independent SDR enzymes, conserved active site residues Tyr153, Lys157, Ser140 and Asn111 form the active site (Fig. 4A). Residue Asn111 resides on a π-bulge motif formed by an atypical backbone hydrogen bond disrupting helix α4. This deformation allows the backbone carbonyl group of Asn111 to form a hydrogen bond with a water molecule likely to be involved in the

formation of a proton relay system similar to what has been described for the *Comamonas testosterone* hydroxysteroid dehydrogenase [27, 28]. Clear electron density representing sorbitol was visible near the active site of each of the four monomers in the asymmetric unit, with sorbitol coordinated near the active site through a hydrogen-bonding network mediated by SmoS residues Gln141, Glu147, Gly184 and His190 (Fig. 4A and B). A comparison of the apo and sorbitol-bound forms of SmoS reveals a slight change in the position of the clamp domain formed by helices α7 and α8, which moves inward during sorbitol binding and allows for the satisfaction of a hydrogen bond between His190 and sorbitol OH1 (Fig. 4C). Interestingly, while clear density for sorbitol was observed in all *Sm*SmoS monomers, the substrate does not appear to be positioned appropriately within in the active site to permit NAD^+-mediated oxidation at C2. In order for the reaction mechanism to proceed as described, the sorbitol C2 hydroxyl group would need to be positioned within hydrogen bonding distance from Tyr153, to allow for Tyr153-mediated proton abstraction and subsequent oxidation of C2 *via* the nicotinamide moiety of NAD^+. In the SmoS-sbt structure, the C2 hydroxyl group is situated ~5.9 Å away from Tyr153, and points away from the active site residue in an arrangement that would not permit the conversion of sorbitol to fructose.

Fig 4. Crystal structure of the SmoS-sbt complex. (A) Close up on the active site of sorbitol-bound SmoS. Catalytic residues are shown as blue sticks, and residues involved in the coordination of sorbitol are shown as a cyan sticks. Sorbitol is shown as yellow sticks, surrounded by an mF_o-DF_c omit map generated using phenix.polder ([29]; green mesh) contoured to 3.0σ. (B) Two-dimensional representation of the H-bonding network

174 observed in the SmoS-sbt complex. Carbon atoms are black, oxygen atoms are red,
175 nitrogen atoms are blue, H-bonds are shown as green dashed lines with corresponding
176 bond lengths (Å). Figure was generated using LigPlot [30]. (C) Superposition of apo
177 SmoS (grey), and SmoS-sbt (blue) depicted in ribbon diagrams with the movement of
178 helix α7 indicated by arrows.
179

180 SmoS has a high pH optimum and a preference for sorbitol

181 It has been reported that functionally related enzymes to SmoS have optimum
182 activity at alkaline pH levels [31, 32]. To investigate the pH preference of *S. meliloti*
183 SmoS, sorbitol dehydrogenase assays were conducted across a pH gradient facilitated by
184 several solutions of differed buffering capacities. 1 µg of SmoS was added to the assay
185 mixture along with 10 mM sorbitol and 1.5 mM NAD^+, the buffers included MES,
186 MOPS, TRIS, and CAPS, each at a concentration of 20 mM, which allowed for a pH
187 gradient spanning pH 5.5-12.5. An optimum enzyme activity of 57.8 mM/min/mg was
188 observed at pH 11; activities recorded across the gradient are reported relative to this
189 value (Fig. 5). Fifty percent of this activity was found at pH 9.5. All subsequent activity
190 assays were conducted in a solution buffered with 20 mM CAPS pH 11. This result is
191 consistent with observations made in *R. sphaeroides* [33].
192

193 **Fig 5. Effect of pH on SmoS dehydrogenase activity.** Reactions were carried out
194 with10 mM sorbitol using 200 mM MES, MOPS, TRIS, or CAPS buffers over their
195 appropriate pH ranges. Activity at the optimal pH was defined as 100%.
196

Despite the previous inability to detect galactitol dehydrogenase activity [17], recent work has shown that *S. meliloti* is capable of galactitol oxidation and that SmoS is responsible for this activity [34]. The ability of SmoS to oxidize sorbitol and galactitol is likely due to the stereochemistry of the functional groups about carbon's 1, 2, and 3, which are identical for both substrates (Fig. 1). To determine the substrate preference of the enzyme, reaction rates were determined by measurement of NADH accumulation over time in a spectrophotometer at 340 nm. Saturation curves for sorbitol and galactitol dehydrogenase activities were generated along with double reciprocal plots facilitating the determination of Michaelis-Menten reaction constants (Fig. 6). It was determined that SmoS has a K_M of 2.5 mM for sorbitol, and a K_M of 1.2 mM for galactitol (Table 2), however, the maximum velocity (V_{max}) of the sorbitol oxidation reaction was calculated to be 50.8 mM/min, while galactitol oxidation proceeded at only 6.4 mM/min (Table 2). Despite a higher affinity, the low reaction velocity of galactitol oxidation greatly reduces the overall reaction efficiency (k_{cat}/K_M). We note that L-iditol shares hydroxyl group orientation about carbons 1, 2, and 3, with sorbitol and galactiol, however this substrate was not tested due to lack of availability [24].

Fig 6. Kinetic characteristics of analysis of *Sm*SmoS. Analysis of sorbitol (A) and galactitol (B) oxidation by using Michaelis-Menten and Lineweaver-Burk plots.

Table 2. Kinetic properties of *Sm*SmoS

Substrate	K_M (mM)	V_{max} (mM/min)	k_{cat} (s^{-1})	V_{max}/K_M (min^{-1})	k_{cat}/K_M ($mM^{-1}s^{-1}$)
Sorbitol	2.5	50.8	25625.6	20.6	10419.7
Galactitol	1.2	6.4	3227.1	5.2	2638.9

SmoS-tagatose complex is predicted to be in a lower energy state than SmoS-fructose complex

Kinetic analysis revealed that galactitol turnover is much less efficient than sorbitol oxidation (Fig. 6, Table 2). This observation was particularly interesting due to the K_M value of galactitol oxidation, which suggested that the enzyme's affinity for galactitol was higher than for sorbitol (Table 2). This led to the hypothesis that tagatose is a poor leaving group in comparison to fructose and the inability of tagatose to quickly leave the active site results in low reaction turnover. This hypothesis is supported by our inability to detect fructose in the active site of SmoS structures determined from crystals grown in the presence of a large concentration (20%) of fructose. To test this hypothesis, computational ligand docking analysis was conducted using the Rosetta Ligand Docking Protocol on the ROSIE server [35-38]. D-fructose and D-tagatose model files were submitted to the ligand dock protocol along with apo SmoS monomer structure, and the outputs were analyzed for indications of the energy state of the complexes. The server generated 200 docking predictions for each SmoS-ligand complex, which were organized via their interface delta scores. The interface delta score represents the total energy of the

237 complex in isolation subtracted from the total energy of the complex with the substrate
238 bound [39]. The ten models with the lowest interface delta score from each complex were
239 selected. The scores from the SmoS-fru model complexes were consistently higher than
240 the scores reported for the SmoS-tag complexes, suggesting that the SmoS-tag complex is
241 in a lower energy state with higher stability than the SmoS-fru complex (Fig. 7A). The
242 data from each SmoS-ligand complex were analyzed for significance via a students *t* test,
243 revealing a *P* value of 1.3×10^{-6}. The entire process from submission to the server through
244 data collection and analysis was repeated independently to evaluate reproducibility; the
245 SmoS-tag complexes were consistently in a lower energy state than the SmoS-fructose
246 complexes. The *P* value for the second trial was 2.4×10^{-6}. An examination of the
247 hydrogen bonding interactions that mediate binding reveals that the SmoS-tag complex
248 forms an additional hydrogen bond that is not present in the SmoS-fru complex, which
249 further stabilizes the tagatose bound structure (Fig. 7B and C). These data suggest that the
250 SmoS-tag complex is a lower energy and more stable complex than the SmoS-fru
251 complex, and that the predicted interface energies from the SmoS-fru complexes and the
252 SmoS-tag complexes are statistically different. They also support the hypothesis that
253 tagatose is a poor leaving group in comparison with fructose and are consistent with
254 observations of the kinetic properties of the enzyme.
255
256 **Fig 7. SmoS-fructose and SmoS-tagatose binding complexes predicted by the Ligand**
257 **docking protocol housed on the ROSIE server.** (A) The distribution of the top ten
258 interface delta scores displayed as box and whisker plots. The tips of the whiskers
259 represent the maximum and minimum values, the horizontal lines represent the first,

second, and third quartiles, and the orange dots represent the averages of the data sets. P value of 1.3x10-6. The lowest energy docking prediction for the SmoS-fructose complex (B) and the SmoS-tagatose complex (C). Fructose is shown in magenta and tagatose in blue.

Discussion

S. meliloti SmoS appears to be most similar to the sorbitol dehydrogenase from *R. sphaeroides*, these enzymes share kinetic characteristics [33], operon structures [20], pH preferences [33], and overall quaternary structure [24]. SmoS can be classified within a group of "high-alkaline enzymes," which are enzymes with a pH optimum from pH 10-11. These enzymes are useful in industry due to their high durability [40]. Similar to *Rs*SmoS, *Sm*SmoS was found to have a higher affinity toward galactitol compared to sorbitol, but turned over sorbitol at a faster rate [33].

Most of the crystal structures of SmoS related enzymes have reported tetrameric structures found in the crystal packing [24, 25, 41], however reports differ on the structure of the enzyme in solution. *R. sphaeroides* SmoS has been reported as dimeric in solution, on the basis of gel filtration chromatography as well as sucrose gradient centrifugation experiments [33]. However the enzyme was later predicted to function as a tetramer based on predicted surface area exposure [24], and these results were supported by size exclusion chromatography and light scattering experiments [41]. *Bj*SDH had been proposed to exist as a trimer in solution [42] but researchers later suggested that a tetramer was more likely [25]. A galactitol dehydrogenase from *Rhizobium leguminosarum* 3841 has also been reported to be tetrameric in solution [31]. The data

283 presented clearly shows that SmoS from *S. meliloti* is present as a tetramer in solution but
284 with a small subset seemingly present as a hexamer or an octamer made up of a dimer of
285 tetramers (Fig. 2). Of note, it appears that both the tetrameric as well as the higher
286 oligomeric forms show sorbitol dehydrogenase activity (Fig. 2). Tetrameric
287 configurations are reported most often and likely represent the majority of SDR protein
288 structures in solution [43].

289 The SmoS-sbt structure shows that the hydroxyl group bonded to C1 of sorbitol
290 associating with catalytic residue Tyr153, and that the structure has a subtle difference
291 from the apo structure in that residues His190 and Trp191 in alpha helix 7 are contorted
292 slightly to accommodate the presence of the substrate (Fig. 4C). As well, residues
293 Asn111, Ser140, Tyr153, and Lys157, which have been proposed to be involved in
294 electron transfer, are too distant from the substrate for catalysis (Fig. 4A).

295 If the positioning of Tyr153 were correct, it would imply that sorbitol should be
296 oxidized to glucose. Based on the available genetic and physiological data it is clear that
297 both sorbitol and galactitol catabolism mediated by SmoS generate fructose and tagatose
298 via an enzymatic reaction in which the hydroxyl group on C2 of the substrate is oxidized
299 forming a planar carbonyl carbon [18, 34]. We also note that enzymes catalyzing the
300 oxidation of sorbitol into glucose are known as sorbitol oxidase (SOX) proteins [44, 45].
301 These enzymes are dissimilar to SDH enzymes of the SDR family [46, 47].

302 This anomaly could be due to the absence of NAD^+ in the binding pocket. NAD^+ was
303 left out of the crystallization solution because its presence would result in an enzymatic
304 reaction, which would prevent the capture of a substrate-bound complex. However, SDR
305 reactions proceed with the coenzyme binding first and leaving last [48], which may help

306 to explain not only why sorbitol is found in an atypical position, but also why fructose
307 was not found in the active site of the fructose grown crystal structures despite its
308 presence at high concentrations. In addition, modeling of NAD^+ and sorbitol into the *R.*
309 *sphaeroides* predicted direct contact and a sandwiching of the C2 carbon of sorbitol
310 between the active site tyrosine, and the nicotinamide ring. Taken together these may
311 explain the observed structure.

312 Thermal stability of an enzyme can affect its ability to be exploited in industrial
313 processes [8]. It has been proposed that the increased thermal stability of SDH is due to
314 the abundance of proline residues and the proline to glycine ratio in its primary amino
315 acid sequence [25]. Proline is a rigid residue with low configurational entropy due to its
316 pyrrolidine ring hindrance, there are several studies that suggest protein thermostability
317 can be influenced by proline content [49-51]. *Rs*SDH contains 6 proline residues and a
318 Pro/Gly ratio of 0.22, while *Bj*SDH has 13 prolines with a ratio of 0.86. The melting
319 temperatures were found to be 62°C and 47°C respectively (25, 39). The SmoS from *S.*
320 *meliloti* has 5 proline residues and the Pro/Gly ratio is 0.2, additionally the position of the
321 residues appears to be conserved, indicating that it's thermostability is likely more similar
322 to *Rs*SDH (Fig. 8).
323
324 **Fig 8. Comparison of the position and distribution of proline residues.** *S. meliloti*
325 SmoS (green), *R. sphaeroides* SmoS (blue; PDB ID: 1K2W), and *B. japonicum* SDH
326 (pink; PDB ID: 5JO9), proline residues are shown in orange.
327

The structure and characterization of *S. meliloti* SmoS provides a high quality structure with sorbitol within the active site. In addition, the characterization and determination of its affinities for its substrates provides insight into why the growth rate of the organism on what should be two equivalent carbon substrates shows great differences. This information is invaluable for higher order resolution of metabolism in *S. meliloti*.

Experimental procedures

Bacterial strains and culture conditions

E. coli BL21 (DE3) GOLD cells were grown on Luria Bertani (LB) medium [52] at 37°C; when necessary, kanamycin was added to a final concentration of 10 µg/mL in liquid media.

Overexpression and purification of SmoS

S. meliloti smoS is a 774 bp gene with a GC content of 64.5%, the overall GC content of *E. coli* K-12 is 50.8% [53]. To accommodate this disparity, the *smoS* nucleotide sequence was codon optimized for expression in *E. coli*. Translation of *smoS* is predicted to generate a 257 amino acid sequence with a molecular weight of 27.2 kDa [54]. *smoS* was cloned into overexpression vector pET-28a as a *Bam*HI-*Hind*III fragment (GenScript, Piscataway, NJ, USA) and this construct was transformed into competent *E. coli* BL21 (DE3) GOLD cells.

Cultures were grown in 1 L volumes of LB medium at 37°C to an OD_{600} of ~0.6. Induction with 1 mM isopropyl-β-D-galactopyranoside (IPTG) preceded growth overnight, shaking, at 16°C. Cells were pelleted by centrifugation at 10000 rpm for 10 min and stored at -80°C. Pellets were resuspended in 30 mL cold lysis buffer consisting of 50 mM Tris pH 8.0, 300 mM NaCl, 2 mM dithiothreitol (DTT), 10 mM imidazole, and lysed by French Press. Cell debris were removed from extracts by centrifugation at 12000 rpm for 1 hour at 4°C. The cell free lysate was applied to a nickel nitrilotriacetic acid (Ni-NTA) column, which was washed with 10 column volumes of lysis buffer and followed by a second wash with 10 column volumes of lysis buffer with 25 mM imidazole. Final elution was prompted by washing with 3 column volumes of buffer with 500 mM imidazole. Eluted protein was dialyzed against 20 mM HEPES pH 7.5, 150 mM NaCl, 10% (v/v) glycerol, and further purified by gel filtration through a Superdex 75 gel filtration column.

SmoS crystallization

Purified SmoS was concentrated to 10 mg/mL and screened by sitting drop vapour diffusion using a Gryphon (Art Robbins Instruments, Sunnyvale, CA, USA) robotic drop setter. Screening was performed using 600 nL drops containing SmoS and crystallization solution at a 1:1 ratio, equilibrated against 50 µL of reservoir solution. Initial crystallization hits were identified in 100 mM HEPES pH 7.4, 50 mM sodium acetate and 20% PEG 3000, and further optimized by hanging-drop vapour diffusion using 48-well VDX plates. Crystals of apo-SmoS were grown in 100 mM HEPES pH 7.4, 50 mM sodium acetate and 18% PEG 3000 and crystals of the SmoS-sorbitol complex

were grown under the same conditions supplemented with 20% sorbitol, galactitol, tagatose, or fructose. Crystallization with galactitol and tagatose was not pursued due to poor solubility or lack of availability of these respective substrates. Crystals in which sorbitol or fructose were included in the reservoir solution were morphologically indistinguishable from the native crystals.

X-ray data collection and structure solution

X-ray data for individual SmoS crystals were collected at 100K on a Rigaku MicroMax 007-HF equipped with a RAXIS IV++ detector. X-ray diffraction images were integrated and scaled using XDS [55], and merged using Aimless [56]. Initial phase estimates for apo-SmoS were determined by molecular replacement within Phaser using the deposited structure of *S. meliloti* SmoS (PDB ID: 4E6P) as a search model, and phase estimates for the SmoS-sbt complex were determined using the refined apo-SmoS structure. Structure refinement and model building were performed using REFMAC [57] and Coot [58], respectively within the ccp4i2 software package. All structure figures were generated using PyMOL [59]. The coordinates and structure factors for the apo SmoS and SmoS-sbt structures have been deposited to the Protein Data Bank under PDB ID 6PEI and 6PEJ, respectively.

Enzyme assays

Spectrophotometric dehydrogenase assays were conducted by measuring the reduction of NAD^+ at 340 nm for 60 seconds. Reaction buffer consisted of 20 mM CAPS pH 11, 1.5 mM NAD^+, and increasing concentrations of sorbitol or galactitol, in a total

395 volume of 1 mL. 1µg SmoS was added per reaction. The optimum pH for enzyme
396 activity was determined by measuring dehydrogenase activity using 200 mM MES,
397 MOPS, TRIS, or CAPS to buffer the reaction mixtures over their appropriate pH ranges.
398 All pH-profiling reactions were initiated with 10 mM sorbitol. Additionally, native gel
399 dehydrogenase assays were performed as previously described [60]. Following elution
400 from the S75 column, fractions were separated by nondenaturing polyacrylamide gel
401 electrophoresis; subsequently the gels were stained for dehydrogenase activity with an
402 assay reagent containing Tris pH 8.0, phenazine methosulfate, nitroblue tetrazolium,
403 NAD^+, and sorbitol.
404

405 ## Ligand docking analysis

406 D-fructose and D-tagatose model files, in SDF file format, were submitted to the
407 Ligand Docking Protocol on the ROSIE server, found at http://rosie.rosettacommons.org,
408 along with the apo SmoS monomer structure in PDB file format. The ligand SDF files
409 were downloaded from Research Collaboratory for Structural Bioinformatics Protein
410 Data Bank (RCSB PDB) at https://www.rcsb.org. These ligand models were manipulated
411 to within 5 Å of the SmoS substrate binding pocket using PyMOL [59] prior to
412 submission to add coordinate data to the files. The top ten predicted models with the
413 lowest interface delta scores were collected and the distribution of these data sets was
414 compared with box and whisker plots. A paired *t* test was performed on the score arrays,
415 a *P* value of less that 0.01 was considered significant.
416
417

References

1. Bieleski RL. Sugar Alcohols. In: Loewus FA, Tanner W, editors. Plant Carbohydrates I: Intracellular Carbohydrates. Berlin, Heidelberg: Springer Berlin Heidelberg; 1982. p. 158-92.
2. Williamson JD, Jennings DB, Guo W-W, Pharr DM, Ehrenshaft M. Sugar Alcohols, Salt Stress, and Fungal Resistance: Polyols—Multifunctional Plant Protection? J Am Soc Hortic Sci. 2002;127(4):467-73.
3. Silveira M, Jonas R. The biotechnological production of sorbitol. Appl Microbiol Biot. 2002;59(4):400-8. doi: 10.1007/s00253-002-1046-0.
4. Rapaille A, Goosens J, Heume M. SUGAR ALCOHOLS. In: Caballero B, editor. Encyclopedia of Food Sciences and Nutrition (Second Edition). Oxford: Academic Press; 2003. p. 5665-71.
5. Ensor M, Banfield AB, Smith RR, Williams J, Lodder RA. Safety and efficacy of D-tagatose in glycemic control in subjects with type 2 diabetes. J Endocrinol Diabetes Obes. 2015;3(1):1065. Epub 12/31. PubMed PMID: 27054147.
6. Espinosa I, Fogelfeld L. Tagatose: from a sweetener to a new diabetic medication? Expert Opin Investig Drugs. 2010;19(2):285-94. doi: 10.1517/13543780903501521.
7. Lu Y, Levin GV, Donner TW. Tagatose, a new antidiabetic and obesity control drug. Diabetes Obes Metab. 2008;10(2):109-34. doi: doi:10.1111/j.1463-1326.2007.00799.x.
8. Chapman J, Ismail AE, Dinu CZ. Industrial applications of enzymes: Recent advances, techniques, and outlooks. Catalysts. 2018;8(6):238. PubMed PMID: doi:10.3390/catal8060238.
9. Gajdzik J, Lenz J, Natter H, Kohring G-W, Giffhorn F, Wenz G, et al. Directed immobilisation of modificated galactitol-dehydrogenase on gold electrodes for electrochemical cofactor regeneration. ECS Trans. 2010;25(28):13-20. doi: 10.1149/1.3309673.
10. Gajdzik J, Lenz J, Natter H, Walcarius A, Kohring GW, Giffhorn F, et al. Electrochemical screening of redox mediators for electrochemical regeneration of NADH. J Electrochem Soc. 2011;159(2):F10-F6. doi: 10.1149/2.056202jes.
11. Kornberger P, Gajdzik J, Natter H, Wenz G, Giffhorn F, Kohring GW, et al. Modification of galactitol dehydrogenase from *Rhodobacter sphaeroides* D for immobilization on polycrystalline gold surfaces. Langmuir. 2009;25(20):12380-6. doi: 10.1021/la9010168.
12. Raveendran S, Parameswaran B, Ummalyma SB, Abraham A, Mathew AK, Madhavan A, et al. Applications of microbial enzymes in food industry. Food Technol Biotechnol. 2018;56(1):16-30. Epub 2018/05/26. doi: 10.17113/ftb.56.01.18.5491. PubMed PMID: 29795993; PubMed Central PMCID: PMCPMC5956270.
13. Geddes BA, Oresnik IJ. Physiology, genetics, and biochemistry of carbon metabolism in the alphaproteobacterium *Sinorhizobium meliloti*. Can J Microbiol. 2014;60(8):491-507. doi: 10.1139/cjm-2014-0306.

14. Udvardi M, Poole PS. Transport and Metabolism in Legume-Rhizobia Symbioses. Annu Rev Plant Biol. 2013;64(1):781-805. doi: doi:10.1146/annurev-arplant-050312-120235. PubMed PMID: 23451778.
15. Ramachandran VK, East AK, Karunakaran R, Downie JA, Poole PS. Adaptation of *Rhizobium leguminosarum* to pea, alfalfa and sugar beet rhizospheres investigated by comparative transcriptomics. Genome Biology. 2011;12(10):R106. doi: 10.1186/gb-2011-12-10-r106.
16. Mortlock RP, editor. Microorganisms as Model Systems for Studying Evolution. New York: Plenum Press; 1984.
17. Martinez De Drets G, Arias A. Metabolism of some polyols by *Rhizobium meliloti*. J Bacteriol. 1970;103(1):97-103. PubMed PMID: 5423374.
18. Gardiol A, Arias A, Cerveñansky C, Gaggero C, Martínez-Drets G. Biochemical characterization of a fructokinase mutant of *Rhizobium meliloti*. J Bacteriol. 1980;144(1):12-6.
19. Jacob AI, Adham SAI, Capstick DS, Clark SRD, Spence T, Charles TC. Mutational analysis of the *Sinorhizobium meliloti* short-chain dehydrogenase/reductase family reveals substantial contribution to symbiosis and catabolic diversity. Mol Plant Microbe In. 2008;21(7):979-87. doi: 10.1094/mpmi.
20. Stein MA, Schäfer A, Giffhorn F. Cloning, nucleotide sequence, and overexpression of *smoS*, a component of a novel operon encoding an ABC transporter and polyol dehydrogenases of *Rhodobacter sphaeroides* Si4. J Bacteriol. 1997;179(20):6335-40.
21. Persson B, Kallberg Y. Classification and nomenclature of the superfamily of short-chain dehydrogenases/reductases (SDRs). Chemico-Biological Interactions. 2013;202(1):111-5. doi: https://doi.org/10.1016/j.cbi.2012.11.009.
22. Persson B, Kallberg Y, Bray JE, Bruford E, Dellaporta SL, Favia AD, et al. The SDR (short-chain dehydrogenase/reductase and related enzymes) nomenclature initiative. Chemico-Biological Interactions. 2009;178(1):94-8. doi: https://doi.org/10.1016/j.cbi.2008.10.040.
23. Sola-Carvajal A, García-García MI, García-Carmona F, Sánchez-Ferrer Á. Insights into the evolution of sorbitol metabolism: phylogenetic analysis of SDR196C family. BMC Evolutionary Biology. 2012;12(1):147. doi: 10.1186/1471-2148-12-147.
24. Philippsen A, Schirmer T, Stein MA, Giffhorn F, Stetefeld J. Structure of zinc-independent sorbitol dehydrogenase from *Rhodobacter sphaeroides* at 2.4 A resolution. Acta Crystallographica Section D. 2005;61(4):374-9. doi: doi:10.1107/S0907444904034390.
25. Fredslund F, Otten H, Gemperlein S, Poulsen J-CN, Carius Y, Kohring G-W, et al. Structural characterization of the thermostable *Bradyrhizobium japonicum* D-sorbitol dehydrogenase. Acta Crystallographica Section F. 2016;72(11):846-52. doi: doi:10.1107/S2053230X16016927.
26. Holm L, Laakso LM. Dali server update. Nucleic Acids Res. 2016;44(W1):W351-5. Epub 2016/05/01. doi: 10.1093/nar/gkw357. PubMed PMID: 27131377; PubMed Central PMCID: PMCPMC4987910.
27. Filling C, Berndt KD, Benach J, Knapp S, Prozorovski T, Nordling E, et al. Critical residues for structure and catalysis in short-chain

dehydrogenases/reductases. J Biol Chem. 2002;277(28):25677-84. Epub 2002/04/27. doi: 10.1074/jbc.M202160200. PubMed PMID: 11976334.

28. Jornvall H, Persson M, Jeffery J. Alcohol and polyol dehydrogenases are both divided into two protein types, and structural properties cross-relate the different enzyme activities within each type. Proc Natl Acad Sci U S A. 1981;78(7):4226-30. Epub 1981/07/01. PubMed PMID: 7027257; PubMed Central PMCID: PMCPMC319762.

29. Liebschner D, Afonine PV, Moriarty NW, Poon BK, Sobolev OV, Terwilliger TC, et al. Polder maps: improving OMIT maps by excluding bulk solvent. Acta Crystallogr D Struct Biol. 2017;73(Pt 2):148-57. doi: 10.1107/S2059798316018210. PubMed PMID: 28177311; PubMed Central PMCID: PMCPMC5297918.

30. Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein engineering. 1995;8(2):127-34. Epub 1995/02/01. PubMed PMID: 7630882.

31. Jagtap SS, Singh R, Kang YC, Zhao H, Lee J-K. Cloning and characterization of a galactitol 2-dehydrogenase from *Rhizobium legumenosarum* and its application in D-tagatose production. Enzyme and Microbial Technology. 2014;58-59:44-51. doi: https://doi.org/10.1016/j.enzmictec.2014.02.012.

32. Lee J-K, Koo B-S, Kim S-Y. Cloning and characterization of the *xyl1* gene, encoding an NADH-preferring xylose reductase from *Candida parapsilosis*, and Its functional expression in *Candida tropicalis*. Appl Environ Microb. 2003;69(10):6179-88. doi: 10.1128/aem.69.10.6179-6188.2003.

33. Schauder S, Schneider K-H, Giffhorn F. Polyol metabolism of *Rhodobacter sphaeroides*: biochemical characterization of a short-chain sorbitol dehydrogenase. Microbiology. 1995;141(8):1857-63. doi: doi:10.1099/13500872-141-8-1857.

34. Kohlmeier MG, White CE, Fowler JE, Finan TM, Oresnik IJ. Galactitol catabolism in *Sinorhizobium meliloti* is dependent on a chromosomally encoded sorbitol dehydrogenase and a pSymB-encoded operon necessary for tagatose catabolism. Molecular Genetics and Genomics. 2019. doi: 10.1007/s00438-019-01545-z.

35. Combs SA, DeLuca SL, DeLuca SH, Lemmon GH, Nannemann DP, Nguyen ED, et al. Small-molecule ligand docking into comparative models with Rosetta. Nature Protocols. 2013;8:1277. doi: 10.1038/nprot.2013.074
https://www.nature.com/articles/nprot.2013.074 - supplementary-information.

36. DeLuca S, Khar K, Meiler J. Fully Flexible Docking of Medium Sized Ligand Libraries with RosettaLigand. PloS One. 2015;10(7):e0132508. doi: 10.1371/journal.pone.0132508.

37. Kothiwale S, Mendenhall JL, Meiler J. BCL::Conf: small molecule conformational sampling using a knowledge based rotamer library. Journal of Cheminformatics. 2015;7(1):47. doi: 10.1186/s13321-015-0095-1.

38. Lyskov S, Chou F-C, Conchúir SÓ, Der BS, Drew K, Kuroda D, et al. Serverification of Molecular Modeling Applications: The Rosetta Online Server That Includes Everyone (ROSIE). PloS One. 2013;8(5):e63906. doi: 10.1371/journal.pone.0063906.

39. Kaufmann KW, Meiler J. Using RosettaLigand for small molecule docking into comparative models. PloS One. 2012;7(12):e50769-e. doi: 10.1371/journal.pone.0050769. PubMed PMID: 23239984.

40. Fujinami S, Fujisawa M. Industrial applications of alkaliphiles and their enzymes--past, present and future. Environmental technology. 2010;31(8-9):845-56. Epub 2010/07/29. doi: 10.1080/09593331003762807. PubMed PMID: 20662376.

41. Carius Y, Christian H, Faust A, Zander U, Klink BU, Kornberger P, et al. Structural insight into substrate differentiation of the sugar-metabolizing enzyme galactitol dehydrogenase from *Rhodobacter sphaeroides D*. The Journal of Biological Chemistry. 2010;285(26):20006-14. doi: 10.1074/jbc.M110.113738. PubMed PMID: PMC2888412.

42. Gauer S, Wang Z, Otten H, Etienne M, Bjerrum MJ, Lo Leggio L, et al. An L-glucitol oxidizing dehydrogenase from *Bradyrhizobium japonicum* USDA 110 for production of D-sorbose with enzymatic or electrochemical cofactor regeneration. Appl Microbiol Biot. 2014;98(7):3023-32. doi: 10.1007/s00253-013-5180-7.

43. Zhu L, Wang S, Tian W, Zhang Y, Song Y, Zhang J, et al. Stabilization of Multimeric Proteins via Intersubunit Cyclization. Appl Environ Microb. 2017;83(18):e01239-17. doi: 10.1128/aem.01239-17.

44. Hiraga K, Kitazawa M, Kaneko N, Oda K. Isolation and Some Properties of Sorbitol Oxidase from *Streptomyces* sp. H-7775. Bioscience, Biotechnology, and Biochemistry. 1997;61(10):1699-704. doi: 10.1271/bbb.61.1699.

45. Yamaki S. A sorbitol oxidase that converts sorbitol to glucose in apple leaf1. Plant and Cell Physiology. 1980;21(4):591-9. doi: 10.1093/oxfordjournals.pcp.a076034.

46. Forneris F, Heuts DPHM, Delvecchio M, Rovida S, Fraaije MW, Mattevi A. Structural Analysis of the Catalytic Mechanism and Stereoselectivity in Streptomyces coelicolor Alditol Oxidase. Biochemistry. 2008;47(3):978-85. doi: 10.1021/bi701886t.

47. Heuts DPHM, van Hellemond EW, Janssen DB, Fraaije MW. Discovery, Characterization, and Kinetic Analysis of an Alditol Oxidase from *Streptomyces coelicolor*. J Biol Chem. 2007;282(28):20283-91. doi: 10.1074/jbc.M610849200.

48. Kavanagh K, Jörnvall H, Persson B, Oppermann U. The SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes2008. 3895-906 p.

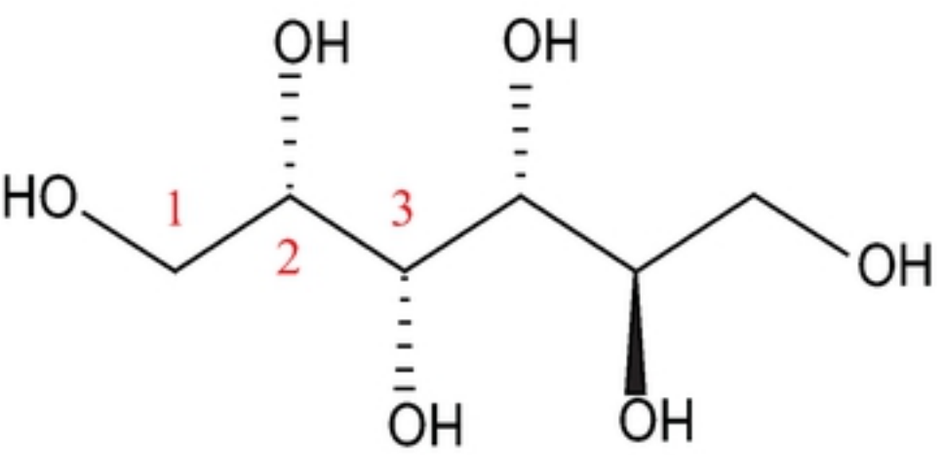
49. Matthews BW, Nicholson H, Becktel WJ. Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. Proc Natl Acad Sci. 1987;84(19):6663-7. doi: 10.1073/pnas.84.19.6663.

50. Suzuki Y. A General Principle of Increasing Protein Thermostability. Proceedings of the Japan Academy, Series B. 1989;65(6):146-8. doi: 10.2183/pjab.65.146.

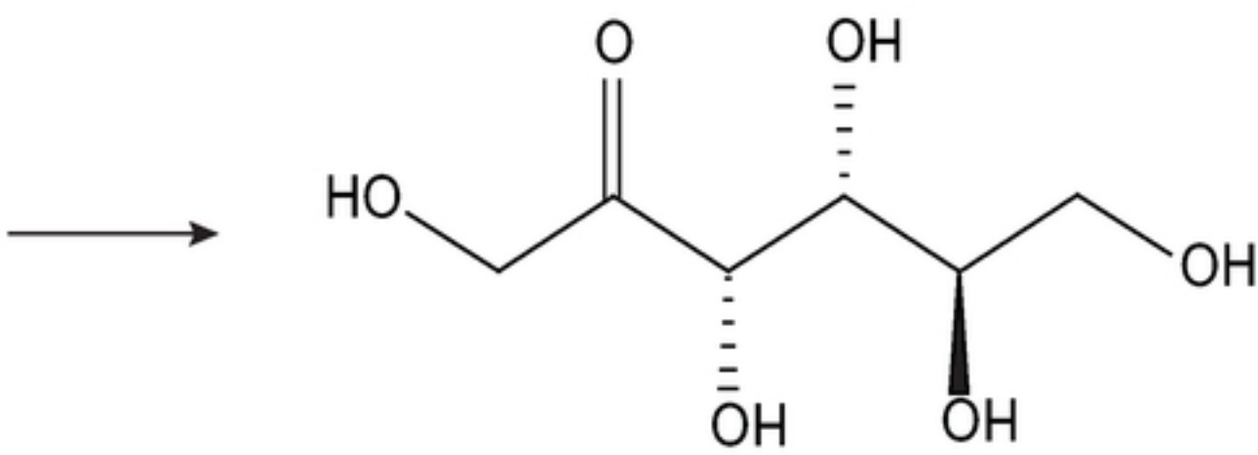
51. Suzuki Y, Oishi K, Nakano H, Nagayama T. A strong correlation between the increase in number of proline residues and the rise in thermostability of five *Bacillus* oligo-1,6-glucosidases. Appl Microbiol Biot. 1987;26(6):546-51. doi: 10.1007/bf00253030.

52. Cold Spring Harbor Protocols. LB (Luria-Bertani) liquid medium. Cold Spring Harb Protoc. 2006;2006(1):pdb.rec8141. doi: 10.1101/pdb.rec8141.

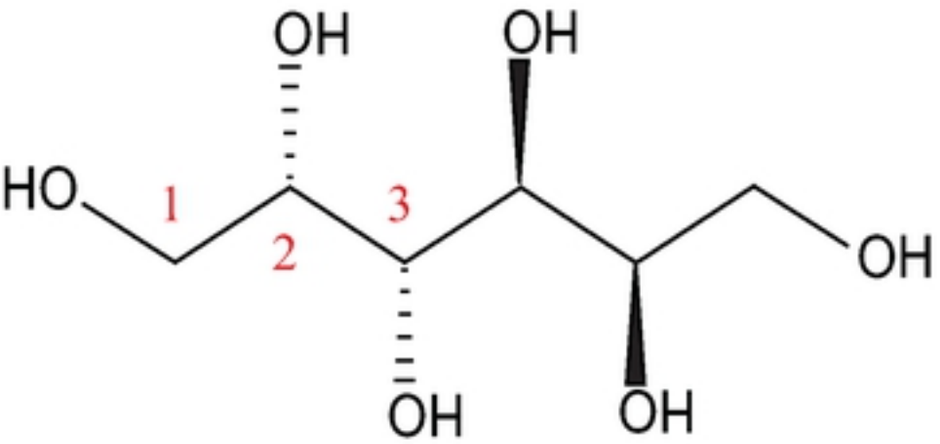
598 53. Riley M, Abe T, Arnaud MB, Berlyn MKB, Blattner FR, Chaudhuri RR, et al.
599 *Escherichia coli* K-12: a cooperatively developed annotation snapshot—2005.
600 Nucleic Acids Res. 2006;34(1):1-9. doi: 10.1093/nar/gkj405. PubMed PMID:
601 PMC1325200.
602 54. Gasteiger E. HC, Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A.
603 Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM,
604 editor. The Proteomics Protocols Handbook: Humana Press; 2005. p. 571-607.
605 55. Kabsch W. XDS. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 2):125-32.
606 Epub 2010/02/04. doi: 10.1107/s0907444909047337. PubMed PMID: 20124692;
607 PubMed Central PMCID: PMCPMC2815665.
608 56. Evans PR. An introduction to data reduction: space-group determination, scaling
609 and intensity statistics. Acta Crystallogr D Biol Crystallogr. 2011;67(Pt 4):282-
610 92. Epub 2011/04/05. doi: 10.1107/s090744491003982x. PubMed PMID:
611 21460446; PubMed Central PMCID: PMCPMC3069743.
612 57. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures
613 by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr.
614 1997;53(Pt 3):240-55. Epub 1997/05/01. doi: 10.1107/s0907444996012255.
615 PubMed PMID: 15299926.
616 58. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot.
617 Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 4):486-501. Epub 2010/04/13.
618 doi: 10.1107/s0907444910007493. PubMed PMID: 20383002; PubMed Central
619 PMCID: PMCPMC2852313.
620 59. Schrodinger L. The PyMOL Molecular Graphics System, Version 1.8. 2015.
621 60. Pickering BS, Oresnik IJ. Formate-dependent autotrophic growth in
622 *Sinorhizobium meliloti*. J Bacteriol. 2008;190(19):6409-18. doi:
623 10.1128/jb.00757-08.
624



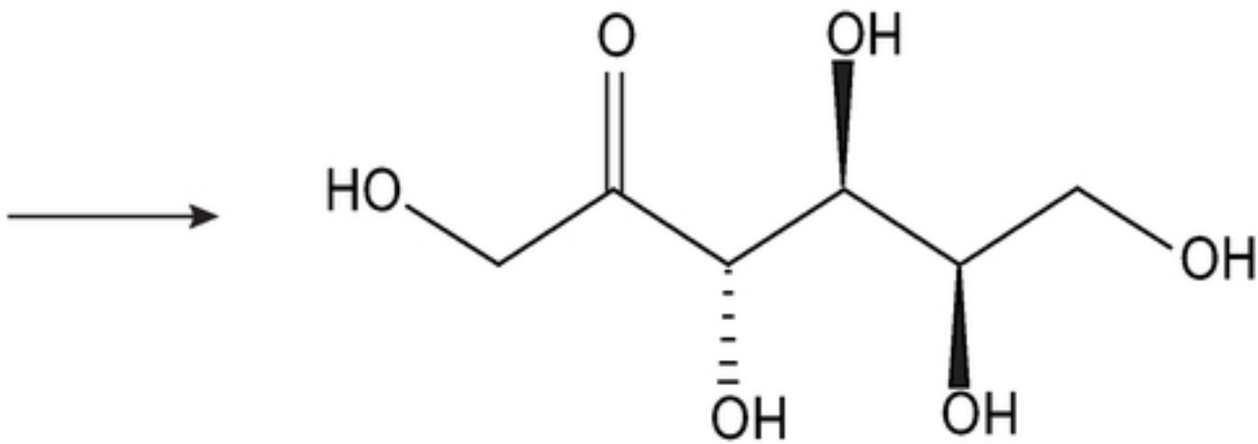
sorbitol



fructose



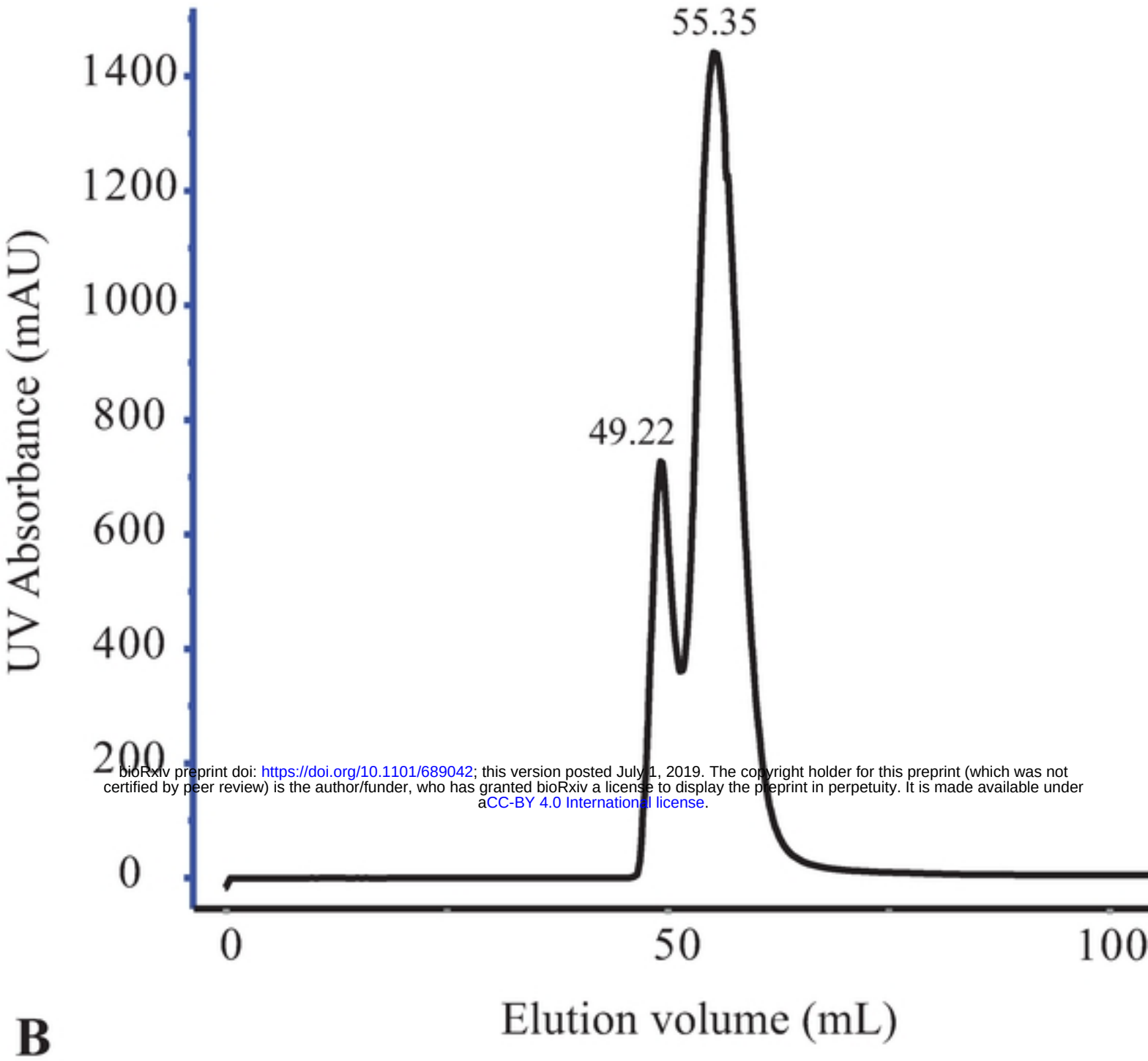
galactitol



tagatose

Figure 1

A



B

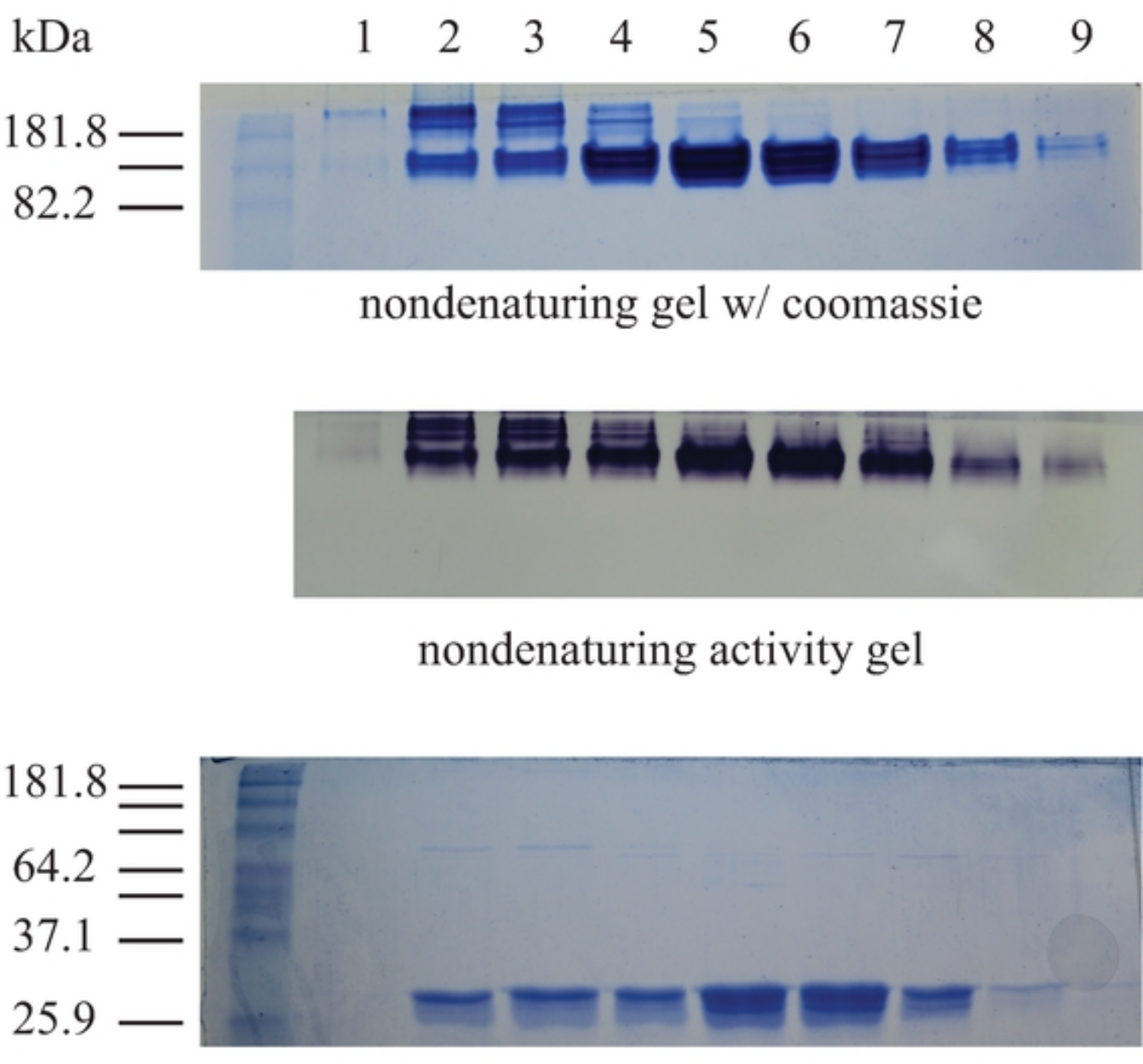


Figure 2

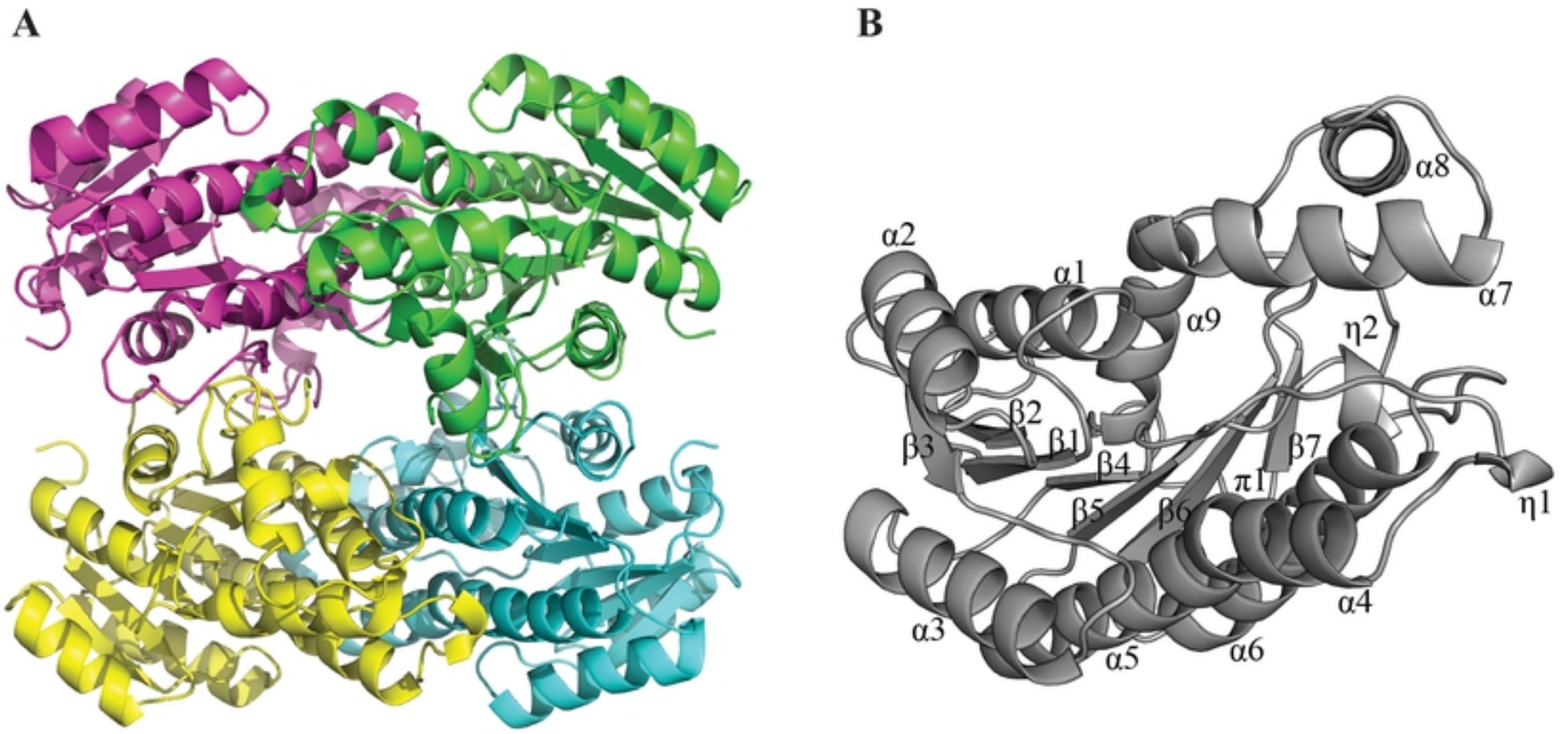


Figure 3

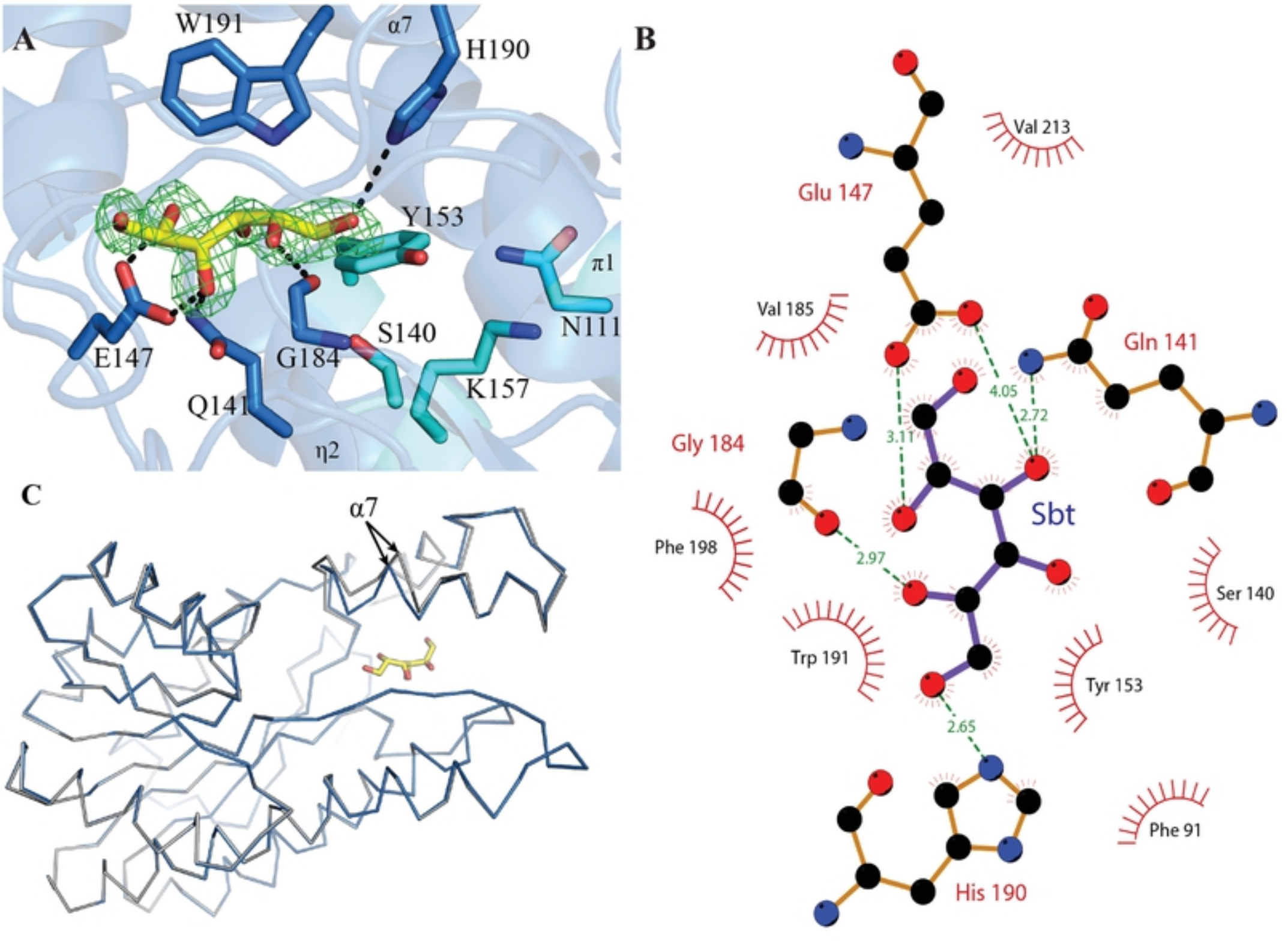


Figure 4

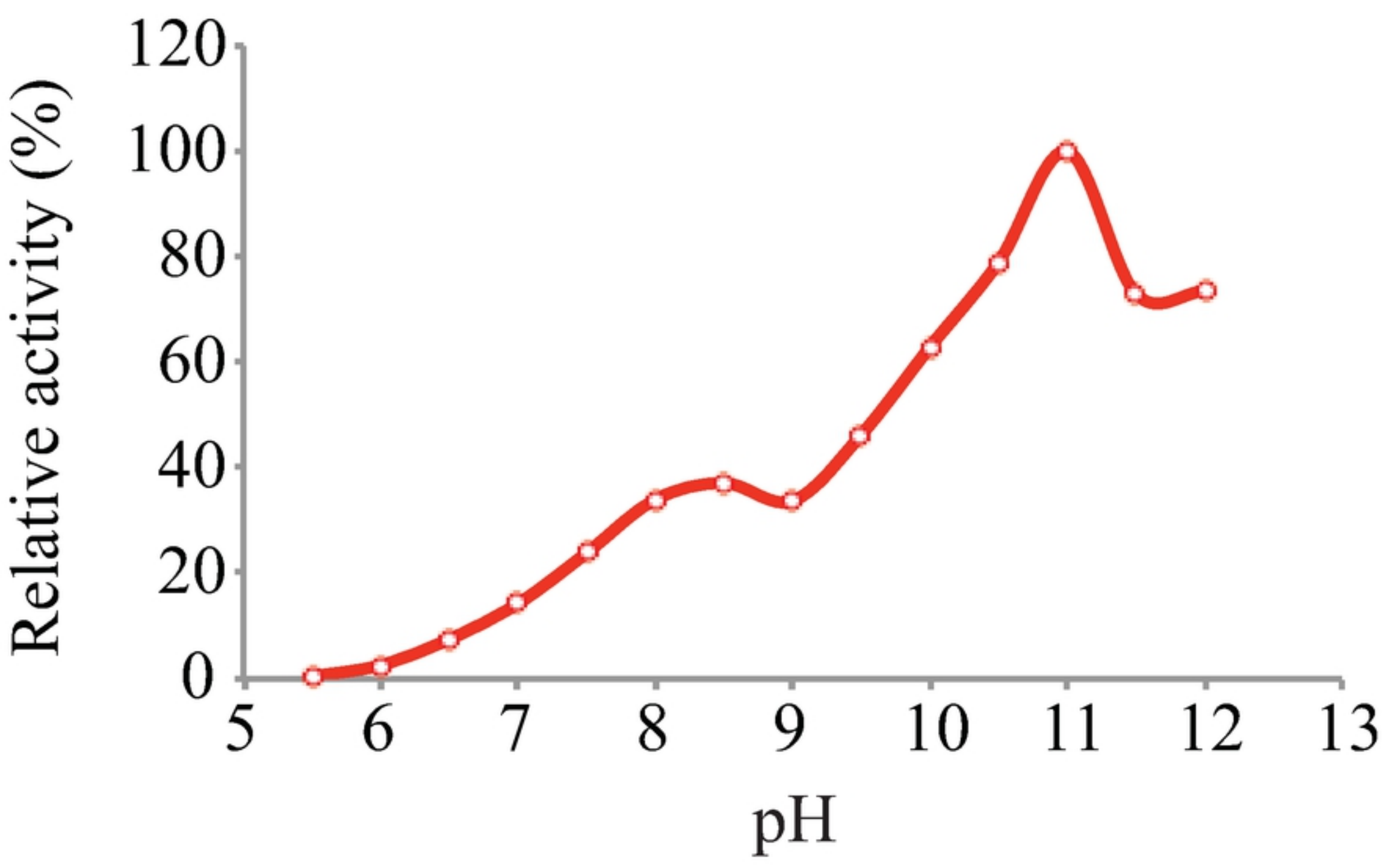


Figure 5

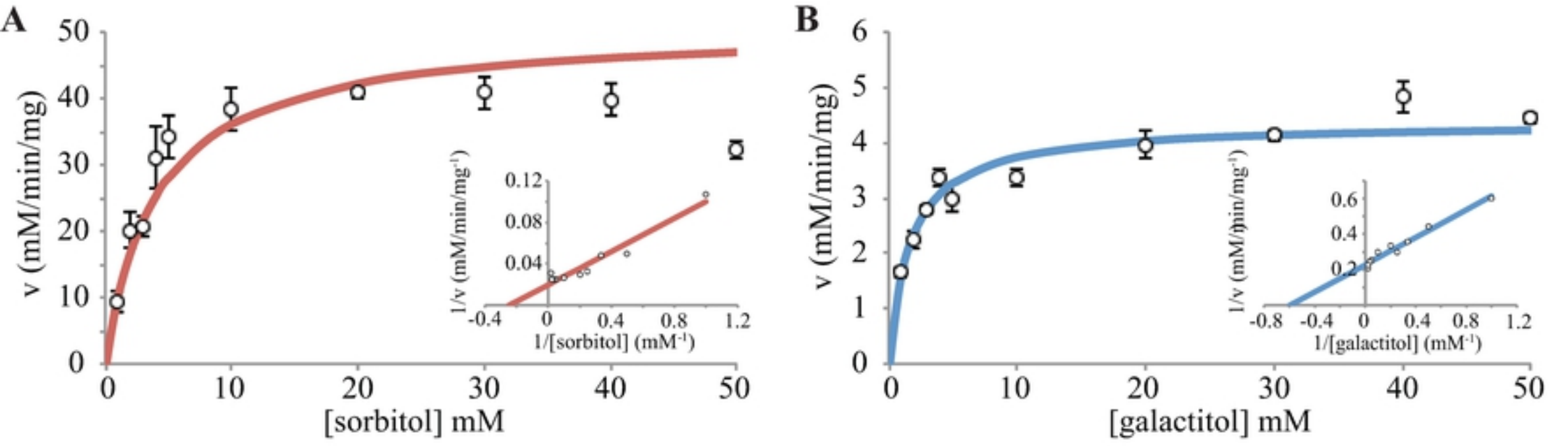
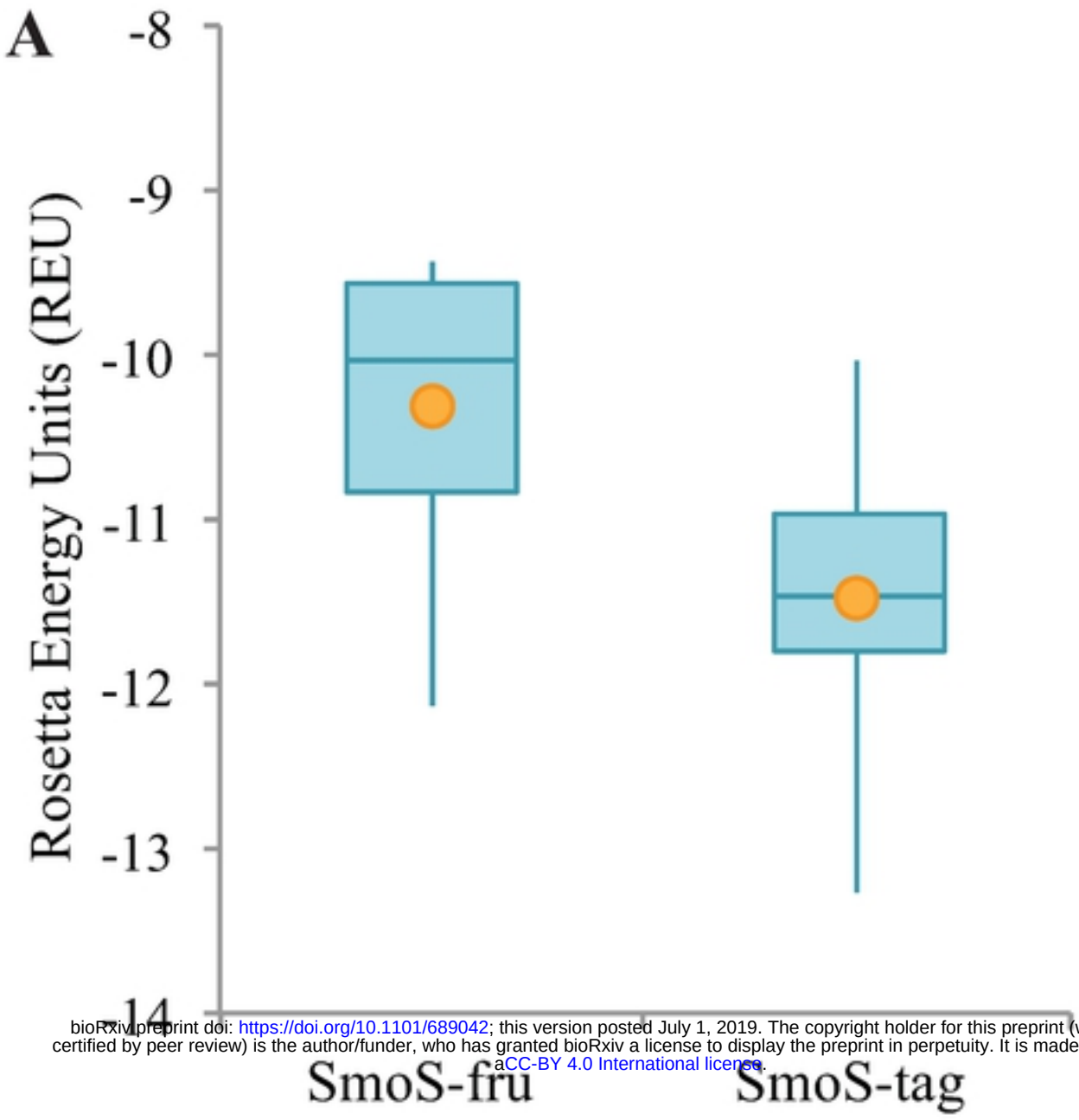
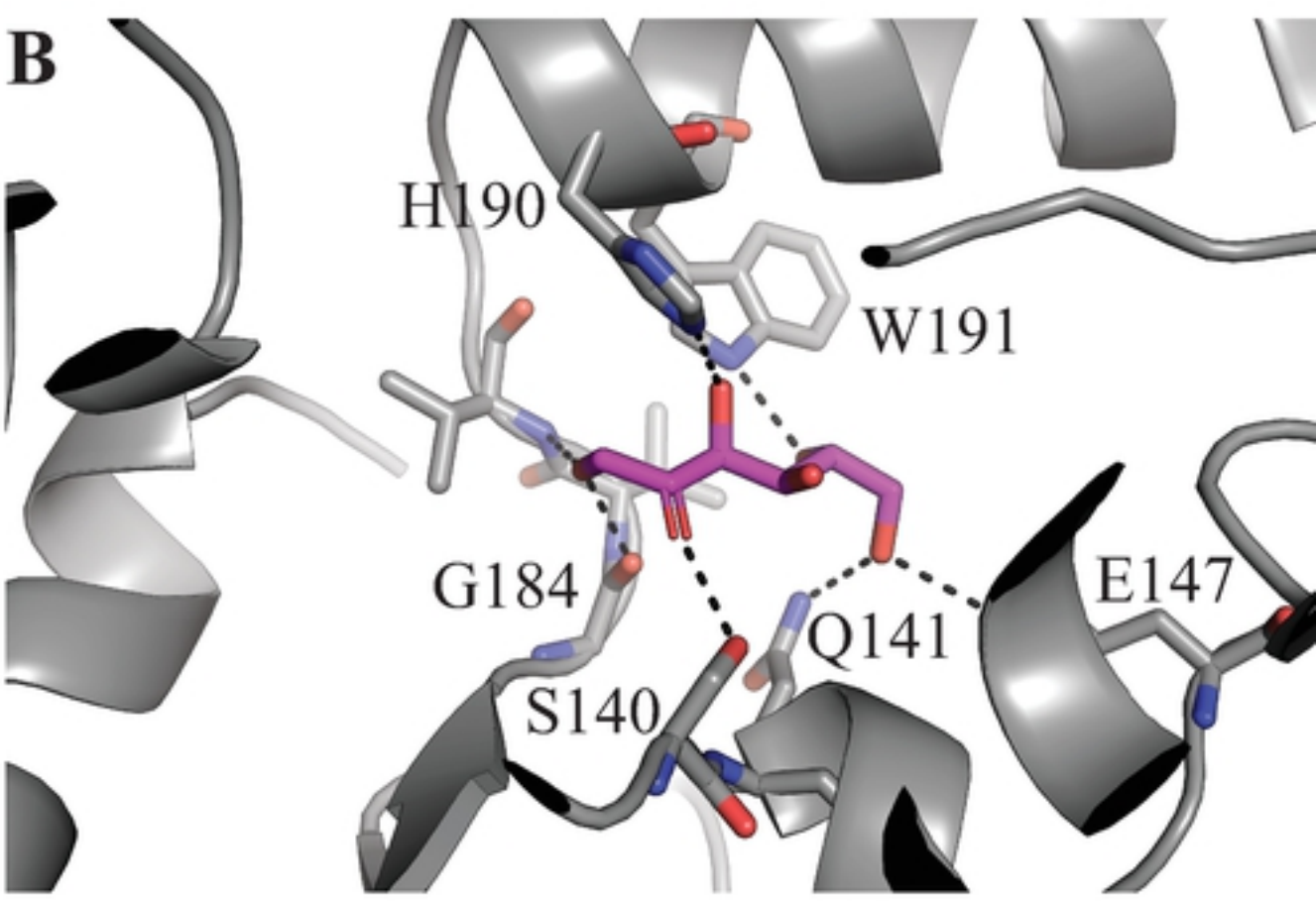

A
v (mM/min/mg)
[sorbitol] mM
1/v (mM/min/mg⁻¹)
1/[sorbitol] (mM⁻¹)
B
v (mM/min/mg)
[galactitol] mM
1/v (mM/min/mg⁻¹)
1/[galactitol] (mM⁻¹)

Figure 6





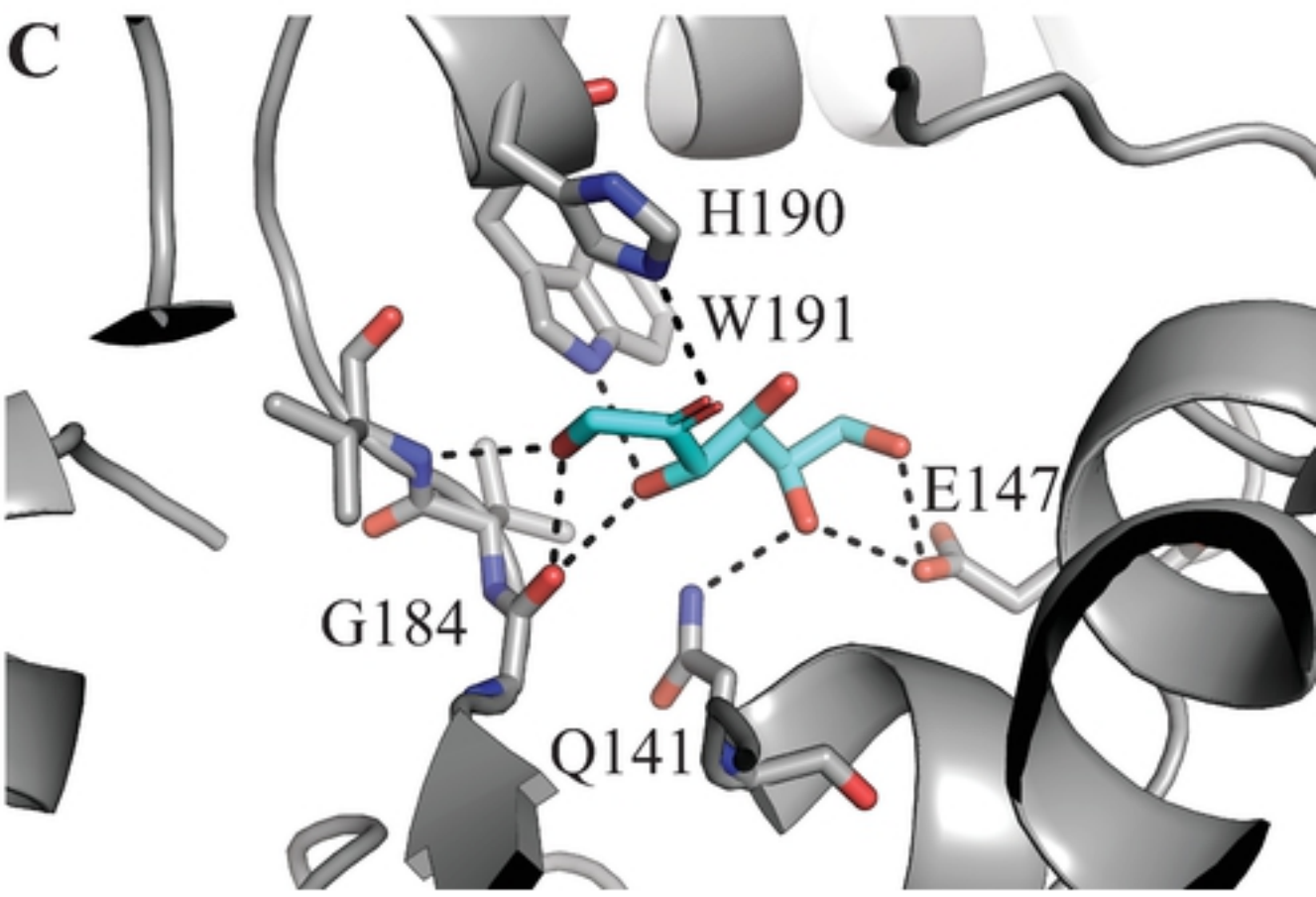
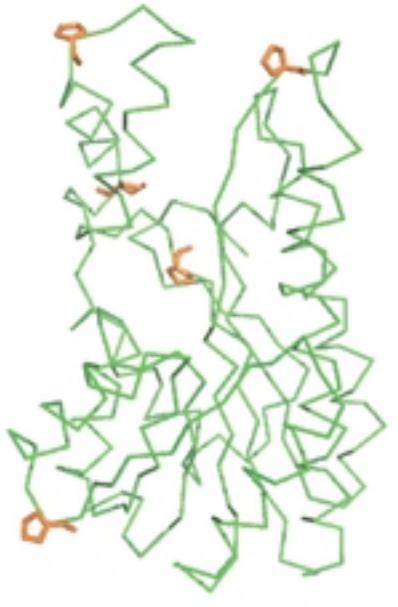
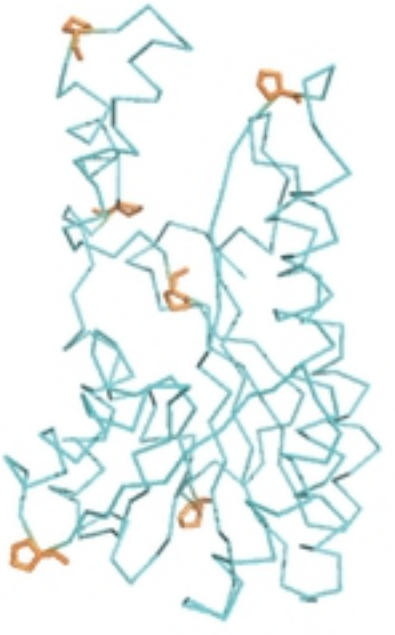


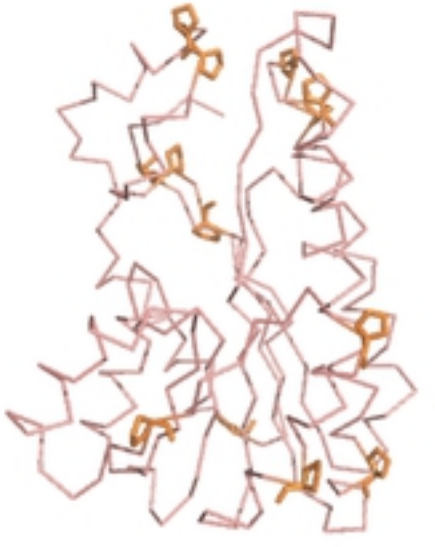
Figure 7



*Sm*SmoS



*Rs*SmoS



*Bj*SDH

Figure 8