PCP
PLANT & CELL PHYSIOLOGY

Regular Paper

Structural Determinants of Sugar Alcohol Biosynthesis in Plants: The Crystal Structures of Mannose-6-Phosphate and Aldose-6-Phosphate Reductases

Romina I. Minen[1,2,†], Jaina A. Bhayani[2,†], Matías D. Hartman[1,2,†], Antonela E. Cereijo[1,2], Yuanzhang Zheng[2], Miguel A. Ballicora[2], Alberto A. Iglesias[1], Dali Liu[2,*] and Carlos M. Figueroa[1,*]

[1]UNL, CONICET, FBCB, Instituto de Agrobiotecnología del Litoral, UNL, CONICET, FBCB, Santa Fe 3000, Argentina
[2]Department of Chemistry and Biochemistry, Loyola University Chicago, Chicago, IL 60660, USA
[†]These authors contributed equally.
*Corresponding authors: Dali Liu, E-mail, dliu@luc.edu; Carlos M. Figueroa, E-mail, carfigue@fbcb.unl.edu.ar
(Received 13 September 2021; Accepted 3 March 2022)

Sugar alcohols are major photosynthetic products in plant species from the Apiaceae and Plantaginaceae families. Mannose-6-phosphate reductase (Man6PRase) and aldose-6-phosphate reductase (Ald6PRase) are key enzymes for synthesizing mannitol and glucitol in celery (*Apium graveolens*) and peach (*Prunus persica*), respectively. In this work, we report the first crystal structures of dimeric plant aldo/keto reductases (AKRs), celery Man6PRase (solved in the presence of mannonic acid and $NADP^+$) and peach Ald6PRase (obtained in the apo form). Both structures displayed the typical TIM barrel folding commonly observed in proteins from the AKR superfamily. Analysis of the Man6PRase holo form showed that residues putatively involved in the catalytic mechanism are located close to the nicotinamide ring of $NADP^+$, where the hydride transfer to the sugar phosphate should take place. Additionally, we found that Lys^{48} is important for the binding of the sugar phosphate. Interestingly, the Man6PRase K48A mutant had a lower catalytic efficiency with mannose-6-phosphate but a higher catalytic efficiency with mannose than the wild type. Overall, our work sheds light on the structure–function relationships of important enzymes to synthesize sugar alcohols in plants.

Keywords: Aldo/keto reductase superfamily • Aldose-6-phosphate reductase • *Apium graveolens* • Mannose-6-phosphate reductase • *Prunus persica* • Sugar alcohols

Introduction

Certain plant species use a significant portion of photosynthetically fixed carbon to synthesize sugar alcohols (Loescher and Everard 2000). The synthesis of acyclic polyols is widely distributed in angiosperms, but the type of sugar alcohol produced depends on each organism. For example, glucitol (Gol, also known as sorbitol) is a major photosynthetic product in species from both the Rosaceae (which includes apple, peach and pear, among others) and Plantaginaceae (such as *Plantago* and *Veronica* species) families (Webb and Burley 1962). Similarly, mannitol (Mol) is a significant photosynthate in species from the Apiaceae (such as celery, carrot and parsley) and Oleaceae (including olive and *Jasminum* species) families (Loescher et al. 1992). Carbon partitioning shows an extra layer of complexity in these species, since they synthesize three primary photosynthates: starch, sucrose and the sugar alcohol (Figueroa et al. 2019). In Gol- and Mol-producing species, the synthesis of the sugar alcohol mainly occurs in source tissues (i.e. mature leaves). In contrast, degradation takes place in sink tissues (such as fruits, roots and non-photosynthetic leaves). In addition to their role as primary photosynthates, sugar alcohols play a central function against different abiotic stress conditions (Figueroa et al. 2019).

The enzymes responsible for Mol and Gol syntheses are mannose-6-phosphate reductase (Man6PRase, EC 1.1.1.224) and aldose-6-phosphate reductase (Ald6PRase, EC 1.1.1.200), respectively (Loescher and Everard 2000). Both enzymes belong to the aldo/keto reductase (AKR) superfamily (Hyndman et al. 2003, Kratzer et al. 2006). AKRs are NAD(P)H-dependent enzymes that reduce a wide variety of aldehydes and ketones, monosaccharides, ketosteroids and prostaglandins (Hyndman et al. 2003). AKRs are found in all kingdoms of life, and the AKR superfamily comprises >190 members, further classified into 16 families according to their amino acid sequence similarity (Penning 2015). AKRs characterized thus far are ~320 amino acids long and display the characteristic $(\beta/\alpha)_8$ barrel fold of the triose-phosphate isomerase, commonly named the TIM barrel (Hyndman et al. 2003). Most AKRs are monomeric, but the enzymes included in the AKR2, AKR6 and AKR7 families may exist in oligomeric forms. The AKR2 subfamily contains dimeric enzymes such as Man6PRase from celery and Ald6PRases from peach, apple and rice (Loescher et al. 1992, Figueroa and Iglesias 2010, Yadav and Prasad 2014, Hartman et al. 2017).

Plant Cell Physiol. 63(5): 658–670 (2022) doi:https://doi.org/10.1093/pcp/pcac029, Advance Access publication on 3 March 2022, available online at https://academic.oup.com/pcp

Man6PRases (AKR2A2) from celery, *Striga hermonthica* and *Orobanche ramosa* (Loescher et al. 1992, Everard et al. 1997, Robert et al. 1999) have been purified and kinetically characterized in the direction of mannose-6-phosphate (Man6P) reduction in all cases. It has been reported that celery Man6PRase has high specificity for both Man6P and NADPH, as no activity was detected when Man6P was replaced with Glc6P, Fru6P or Man1P or when NADPH was replaced with NADH (Loescher et al. 1992). The recombinant Man6PRase from celery leaves was kinetically and structurally indistinguishable from the source-purified enzyme (Everard et al. 1997). Ald6PRase (AKR2A1) has been purified from numerous plant sources, such as loquat leaves (Hirai 1981), apple seedlings (Kanayama and Yamaki 1993) and apple leaves (Negm and Loescher 1981, Zhou et al. 2003). In these studies, enzyme activity was determined in both directions of the reaction, Glc6P reduction and Gol6P oxidation, the former being the physiologically relevant reaction (Kanayama and Yamaki 1993). Recombinant Ald6PRases from apple and peach have been previously purified and characterized (Figueroa and Iglesias 2010, Hartman et al. 2017).

Most plant AKRs previously studied belong to the AKR4C subfamily (Simpson et al. 2009), while all structures of plant AKRs solved so far are monomeric (Bomati et al. 2005, Olsen et al. 2008, Simpson et al. 2009, Giuseppe et al. 2016, Songsiriritthigul et al. 2020). Hence, there is a lack of structural information regarding multimeric plant AKRs involved in synthesizing sugar alcohols. In this work, we report the crystal structures of celery Man6PRase (*Agr*Man6PRase) in complex with $NADP^+$ and mannonic acid and the apo form of peach Ald6PRase (*Ppe*Ald6PRase). To the best of our knowledge, these are the first solved structures of plant enzymes belonging to the AKR2 family. Using structural data, we defined the residues putatively involved in Man6P binding to *Agr*Man6PRase and determined that Lys^{48} plays a key role in the selectivity of the sugar phosphate. Our results support the structural divergence of dimeric members of the AKR2 family and a common catalytic mechanism for all AKRs.

Results

Production of *Agr*Man6PRase and *Ppe*Ald6PRase

The coding sequence for *Agr*Man6PRase was amplified using total RNA extracted from celery leaves and comprises an open reading frame of 930 bp (**Supplementary Text S1**). This sequence encodes a putative protein of 309 amino acids (**Supplementary Text S2**), with an estimated mass of 35.2 kDa, similar to that described for Man6PRase purified from celery leaves (Loescher et al. 1992, Everard et al. 1997) and other reductases from the plant AKR2 family (Figueroa and Iglesias 2010, Hartman et al. 2017). The cloned sequence differs in five bases from the one reported by Everard et al. (1997), resulting in the replacement of four amino acids, located at positions 130, 134, 199 and 272 (**Supplementary Fig. S1**). More recently, Khalil et al. (2017) reported a different sequence for Man6PRase from celery, which differs from our sequence in two bases, but only one amino acid at position 134 (**Supplementary Fig. S1**). We further tested the relevance of these changes for the catalytic activity and quaternary structure of *Agr*Man6PRase (see below).

Table 1 Kinetic characterization of the *Agr*Man6PRase and K48A mutant (reactions were carried out at pH 7.5 in the presence of 3 mM DTT).

Enzyme	Substrate	K_m (mM)	V_{max} (U mg^{-1})
Man6PRase	Man6P	4.4 ± 0.1	3.5 ± 0.1
	$NADPH_{(Man6P)}$	0.026 ± 0.002	
	Man	145 ± 68	0.17 ± 0.03
	$NADPH_{(Man)}$	0.14 ± 0.01	
	Glc6P	12.3 ± 1.8	0.073 ± 0.009
	$NADPH_{(Glc6P)}$	0.020 ± 0.001	
Man6PRase-K48A	Man6P	16.3 ± 2.5	0.68 ± 0.18
	$NADPH_{(Man6P)}$	0.029 ± 0.002	
	Man	147 ± 15	0.76 ± 0.03
	$NADPH_{(Man)}$	0.042 ± 0.003	
	Glc6P	14.9 ± 2.5	0.017 ± 0.001
	$NADPH_{(Glc6P)}$	0.058 ± 0.006	
Ald6PRase[a]	Glc6P	11 ± 1	3.1 ± 0.1
	$NADPH_{(Glc6P)}$	0.031 ± 0.001	

[a]Data obtained from Hartman et al. (2017).

The sequence encoding *Agr*Man6PRase was subcloned into the expression vector and the recombinant enzyme purified by immobilized metal affinity chromatography (IMAC). Analysis of the purified protein by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) showed a major band of ~37 kDa (**Fig. 1A**), while the native protein eluted from the size-exclusion column as a 65-kDa protein (**Fig. 1B**), indicating that recombinant *Agr*Man6PRase is a homodimer. Next, we determined the kinetic parameters of *Agr*Man6PRase in the physiological direction of the reaction (reduction of Man6P), obtaining a V_{max} of 3.5 U mg^{-1} and K_m values of 4.4 and 0.026 mM for Man6P and NADPH, respectively (**Table 1**). Recombinant *Ppe*Ald6PRase was produced and purified as previously described (Hartman et al. 2017). The kinetic parameters of this enzyme are presented in **Table 1**.

Structural analysis of *Agr*Man6PRase and *Ppe*Ald6PRase

The crystal structure of substrate-bound *Agr*Man6PRase was determined at a resolution of 1.72 Å, while *Ppe*Ald6PRase crystals diffracted with a resolution of 1.61 Å (see crystallographic parameters in **Table 2**). Both proteins crystallized in the monoclinic space group C121. The crystallographic asymmetric units contained four polypeptide chains (two dimeric molecules) for the holo form of *Agr*Man6PRase and two polypeptide chains (one dimer) for the apo form of *Ppe*Ald6PRase.

Overall structure. Both *Agr*Man6PRase and *Ppe*Ald6PRase are dimers, and each subunit contains an active site. These enzymes share the canonical TIM barrel fold, comprised

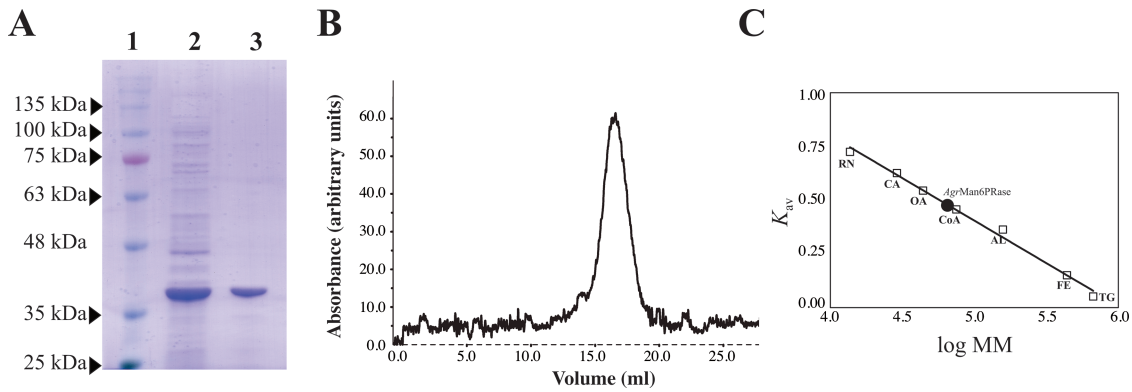

Fig. 1 Analysis of *Agr*Man6PRase by SDS–PAGE and gel filtration chromatography. (A) Reducing SDS–PAGE. Lane 1, molecular mass markers; lane 2, soluble fraction of recombinant cells expressing *Agr*Man6PRase (crude extract); lane 3, *Agr*Man6PRase purified by IMAC. (B) The elution profile of *Agr*Man6PRase from the Superdex 200 10/300 column. (C) Gel filtration chromatography. A plot of K_{av} versus log (molecular mass). Standard proteins (□): RN, ribonuclease (13.7 kDa); CA, carbonic anhydrase (29 kDa), OA, ovalbumin (44 kDa); CoA, conalbumin (75 kDa); AL, aldolase (158 kDa); FE, ferritin (440 kDa) and TG, thyroglobulin (669 kDa); (●) *Agr*Man6PRase.

by an alternating arrangement of β-strands and α-helices (**Fig. 2A, B**). Eight β-strands (β3 to β10) form the central region, and nine α-helices are located in the outer area (according to *Agr*Man6PRase, α1 to α6, α8, α10 and α11; **Supplementary Fig. S2**). One α-helix (according to *Agr*Man6PRase, α9) is located between the central and the outer regions. Two short, antiparallel β-strands (β1 and β2) are located at the N-terminus and form a hairpin covering the bottom of the barrel, whereas the C-terminus has a short α-helix (according to *Agr*Man6PRase, α12). As shown in **Supplementary Fig. S2**, three long loops are found in both enzymes: loop A (between β6 and α4, *Agr*Man6PRase residues His^{107} to Ser^{137}), loop B (between β9 and α8, *Agr*Man6PRase residues Thr^{209} to Ser^{225}) and loop C (between α11 and α12, *Agr*Man6PRase residues Thr^{289} to Pro^{299}). *Agr*Man6PRase was crystallized with substrates; thus, loop B has an ordered α-helix (α7), absent in the apo form of *Ppe*Ald6PRase (**Fig. 2** and **Supplementary Fig. S3**). The structures of *Agr*Man6PRase and *Ppe*Ald6PRase are substantially similar [root-mean-square deviation (RMSD) of 0.742 Å over 287 C_α atoms], and the loops involved in substrate binding account for the major structural differences observed between both enzymes (**Supplementary Fig. S3**).

*Agr*Man6PRase and *Ppe*A6PRase dimerize through an interface that includes residues from loop A, helices α5 and α6, the loop located between β8 and α6, and the C-terminal loop from both subunits (**Fig. 3** and **Supplementary Fig. S4**). Specifically, residues located in loop A and the C-terminus from one subunit interact with residues located in helices α5 and α6 and the loop β8-α6 located in the opposite subunit (**Supplementary Figs. S5, S6**). This subunit arrangement orientates the active sites of the dimers in an antiparallel conformation (**Fig. 2**). We found 18 and 17 hydrogen bonds between both subunits of *Agr*Man6PRase and *Ppe*A6PRase, respectively, mainly located in loop A, the α5 helix and the C-terminal loop (**Supplementary Figs. S5, S6**). Furthermore, 136 and 158 contacts were found between both subunits in *Agr*Man6PRase and *Ppe*Ald6PRase,

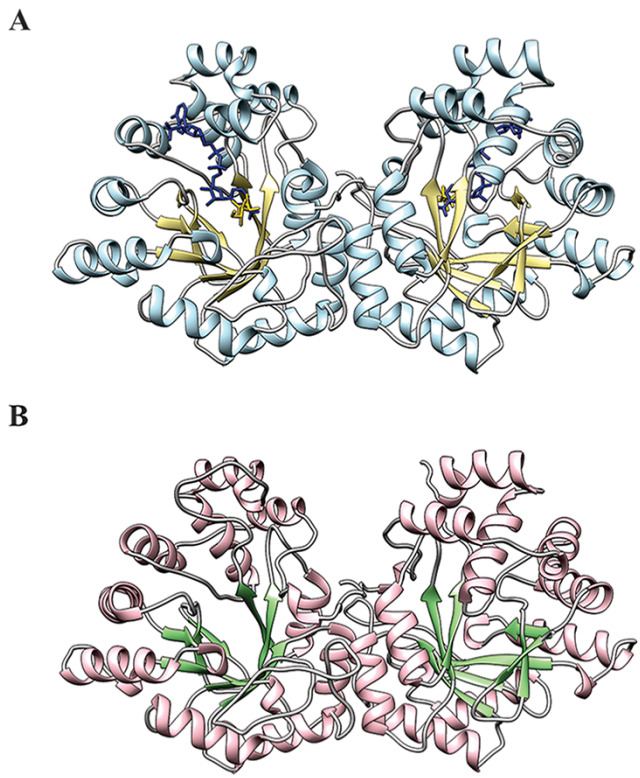

Fig. 2 Crystal structures of *Agr*Man6PRase and *Ppe*Ald6PRase. (A) The holo form of *Agr*Man6PRase (α-helices and β-strands are colored in light blue and yellow, respectively) in complex with $NADP^+$ (dark blue) and mannonic acid (yellow). (B) The apo-form of *Ppe*Ald6PRase (α-helices and β-strands are colored in pink and green, respectively).

respectively, including amino acids located in loop A, helices α5 and α6, and the C-terminal loop (**Fig. 3** and **Supplementary Figs. S4–S6**). In *Agr*Man6PRase, we identified three amino acids (Tyr^{165}, Arg^{168} and Ser^{172}, located within the α5 helix) that generate >10 contacts with four or more amino acids in the opposite subunit, most of them located in loop A (**Supplementary Table S1**). We observed a similar pattern in *Ppe*Ald6PRase,

Table 2 Data collection and refinement statistics for the crystal structures of *Agr*Man6PRase and *Ppe*Ald6PRase.

Structure	*Agr*Man6PRase	*Ppe*Ald6PRase
Data processing		
Space group	C121	C121
Cell dimension		
α, β, γ (deg)	152.2,61.8,148.1	178.4,50.0,111.8
a, b, c (Å)	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Processed resolution (Å)	38.01–1.73	31.6–1.61
Wilson B-factor (Å)	18.20	12.36
R_{merge} [a] (%)	8.1(112.0) [b]	5.4 (37.1)
R_{pim} [c] (%)	5.5 (43.9)	2.3 (15.9)
I/σ (I)	15.0 (2.2)	28.9 (5.6)
CC $^1/_2$ [d] (%)	99.7 (71.3)	99.8 (95.3)
Completeness (%)	95.8 (98.2)	99.0 (98.5)
Multiplicity	6.0 (6.0)	7.2 (7.1)
No. of reflections	791,677	624,800
No. of unique reflections	132,086	86,724
Refinement		
R_{work} [e]/R_{free} [f] (%)	18.08/21.05	15.67/17.94
No. of atoms		
Protein	9,906	5,735
Ligand	244	N.A.
Water	1,668	859
Average B-factor (Å^2)		
Protein	24.75	19.20
Ligand	21.78	31.19
RMSD		
Bond lengths (Å)	0.005	0.007
Bond angles (deg)	0.97	0.88
Ramachandran plot (%)		
Favored	97.56	96.87
Allowed	2.44	3.13
Outliers	0.00	0.00

[a] $R_{merge} = \Sigma |Iobs - Iavg| / \Sigma Iavg$.
[b] The values for the highest-resolution bin are in parentheses.
[c] Precision indicating merging *R*.
[d] Pearson correlation coefficient of two 'half' data sets.
[e] $R_{work} = \Sigma |Fobs - Fcalc| / \Sigma Fobs$.
[f] Five percent of the reflection data was selected at random as a test set, and only these data were used to calculate R_{free}.

where Phe[166], Arg[169] and Ser[173] generate multiple contacts in the opposite subunit. As observed in *Agr*Man6PRase, most of the contacts are between the α5 helix from one subunit and loop A from the opposite subunit (**Supplementary Table S1**).

As previously mentioned, the sequence coding for *Agr*Man6PRase cloned in this study showed five point mutations compared with that reported by Everard et al. (1997). Two of these changes produced point mutations in amino acids located in loop A (L130W and T134A; **Supplementary Figs. S1, S7A–B**), which seem to be important for dimerization (**Fig. 3** and **Supplementary Fig. S5**). Thus, we synthesized the sequence reported by Everard et al. (1997) and expressed the recombinant enzyme (from now on, the quadruple mutant, **Supplementary Fig. S7C**), using the same system developed for expression of the 'wild-type' *Agr*Man6PRase (**Fig. 1**). Interestingly, the quadruple mutant was a monomer (**Supplementary Fig. S7D**) with negligible enzymatic activity. Although Leu[130] does not interact with any residue in the opposite subunit, the L130W mutation could hinder the interaction of the vicinal Val[129] with Lys[196] (**Supplementary Fig. S7A**); on the other hand, the T134A mutation would prevent the occurrence of five interactions between Thr[134] and Leu[171] and Tyr[201] (**Supplementary Fig. S7B**).

Cofactor binding to AgrMan6PRase and PpeAld6PRase. Based on the observed electron density in the active site (**Fig. 4D**), $NADP^+$ was built in the *Agr*Man6PRase structure and refined. The interactions between cofactor $NADP^+$ and protein are quite extensive (**Supplementary Fig. S8**). A hydrogen-bonding network including 20 hydrogen bonds are formed between the cofactor and 17 enzyme residues, namely Val[18], Trp[19], Arg[20], Asp[42], Ser[160], Asn[161], Glu[182], Thr[209], Leu[211], Gly[213], Lys[261], Ser[262], Ser[263], Glu[270] and Asn[271]. Two Pi–Pi stacking interactions include one between His[208] and nicotinamide ring as well as another between Arg[267] and adenosine ring (**Fig. 4A**). Various Van der Waals interactions also contribute to the binding of $NADP^+$. All the residues mentioned above are located at <3.5 Å from the cofactor.

It is worth mentioning that most of the amino acids that interact with $NADP^+$ in *Agr*Man6PRase are well conserved in *Ppe*Ald6PRase (**Supplementary Fig. S2**); moreover, we observed only minor structural differences between the key amino acids that interact with $NADP^+$ when both structures were superimposed (**Fig. 4B**). $NADP^+$ sits at the bottom of the cofactor-binding site; in the holo form of *Agr*Man6PRase, loop B closes the $NADP^+$-binding cavity and interacts with amino acids located at the N-terminal loop β3-α1 (**Fig. 4C**). Specifically, Ala[216], Asn[217] and Arg[220] from loop B interact with Trp[19] and Arg[20] from loop β3-α1 in a fashion resembling a locked seat belt. Cofactor release would require a conformational change to open-loop B by unlocking the seat belt (i.e. breaking such interactions).

As shown in **Supplementary Fig. S3**, loop B comprises a region with high RMSD values between *Agr*Man6PRase and *Ppe*Ald6PRase, despite they share 71% identity in such domain. *Agr*Man6PRase was co-crystallized in presence of $NADP^+$, which induced a conformational change in loop B, thus producing the observed differences between both structures (**Fig. 4B** and **Supplementary Fig. S3**).

Man6P binding in AgrMan6PRase and conservation of the catalytic site in both AgrMan6PRase and PpeAld6PRase. The *Agr*Man6PRase crystal was produced in the presence of $NADP^+$ and Man6P; however, the final model was obtained in complex with the cofactor and mannonic acid (**Fig. 2A**). This result suggests the occurrence of chemical modifications on Man6P during the crystallization process; hence, we performed different experiments to test the stability of substrates in conditions resembling those used for crystallization of *Agr*Man6PRase. Data presented in **Supplementary Fig. S9** indicate that the mannonic acid molecule present

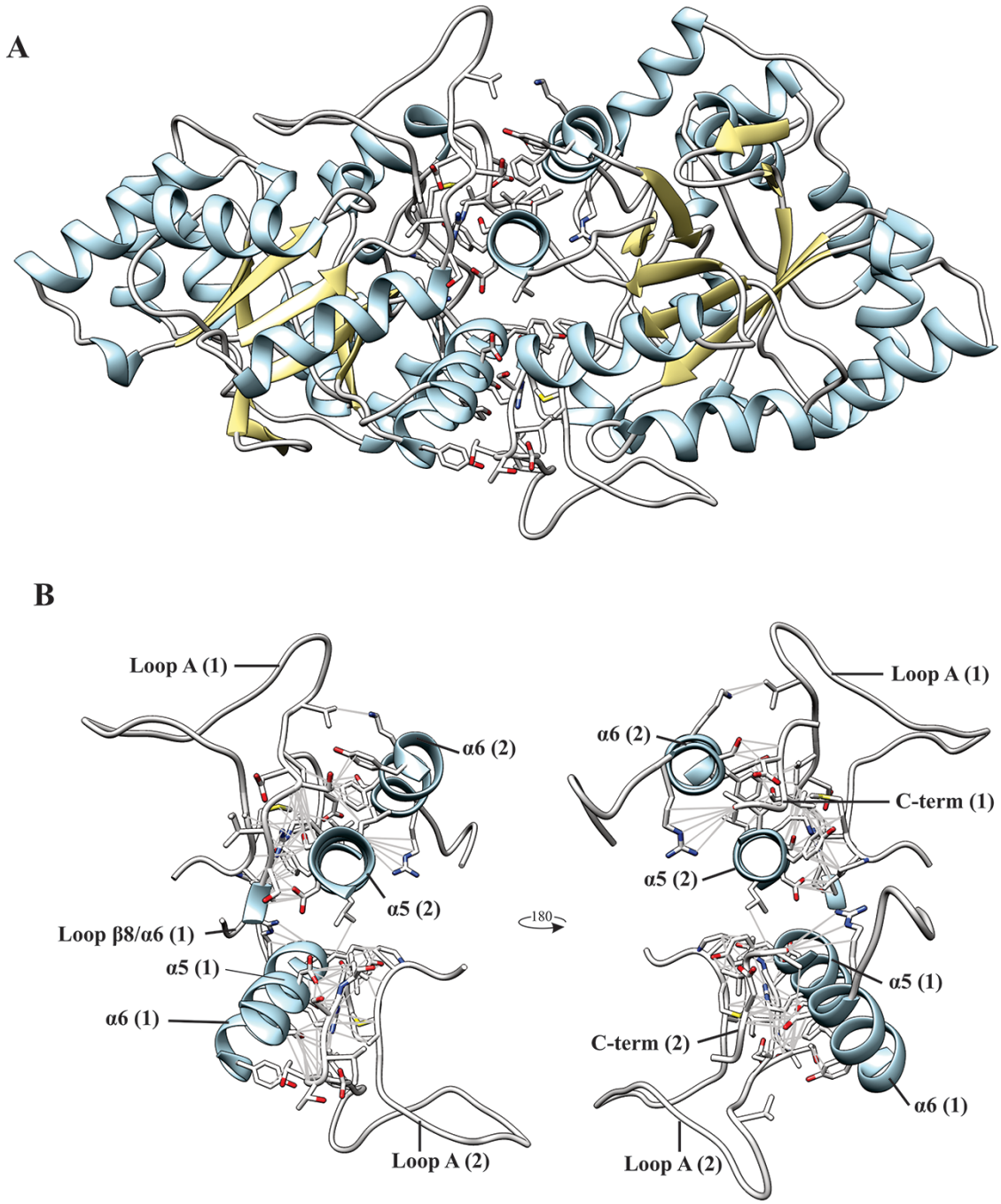

Fig. 3 Interactions (contacts) found in the dimer interface for *Agr*Man6PRase. (A) *Agr*Man6PRase dimer. (B) Loops and α-helices involved in the interaction extracted from A (left) and rotated 180° (right). The interactions between residues from different subunits are shown in gray lines. A detailed visualization of the dimer interface is presented in **Supplementary Fig. S5**.

in the crystal structure of *Agr*Man6PRase might be produced from Man6P due to its sequential dephosphorylation and oxidation (see **Supplementary Text S5** for further details).

Based on the observed electron density in the active site (**Fig. 4D**), we modeled a mannonic acid molecule within it; the refinement result is consistent with the interpretation that a mannonic acid was likely co-purified with the enzyme and has been co-crystallized with it. Mannonic acid locates in the vicinity of the $NADP^+$ cofactor mimicking the substrate position. Active site residues that directly interact with mannonic acid include Asp^{46}, Tyr^{47}, His^{107} and Asn^{297}, which form hydrogen bonds with mannonic acid, and Trp^{19} and Trp^{78}, which provide Van der Waals contacts as the boundaries

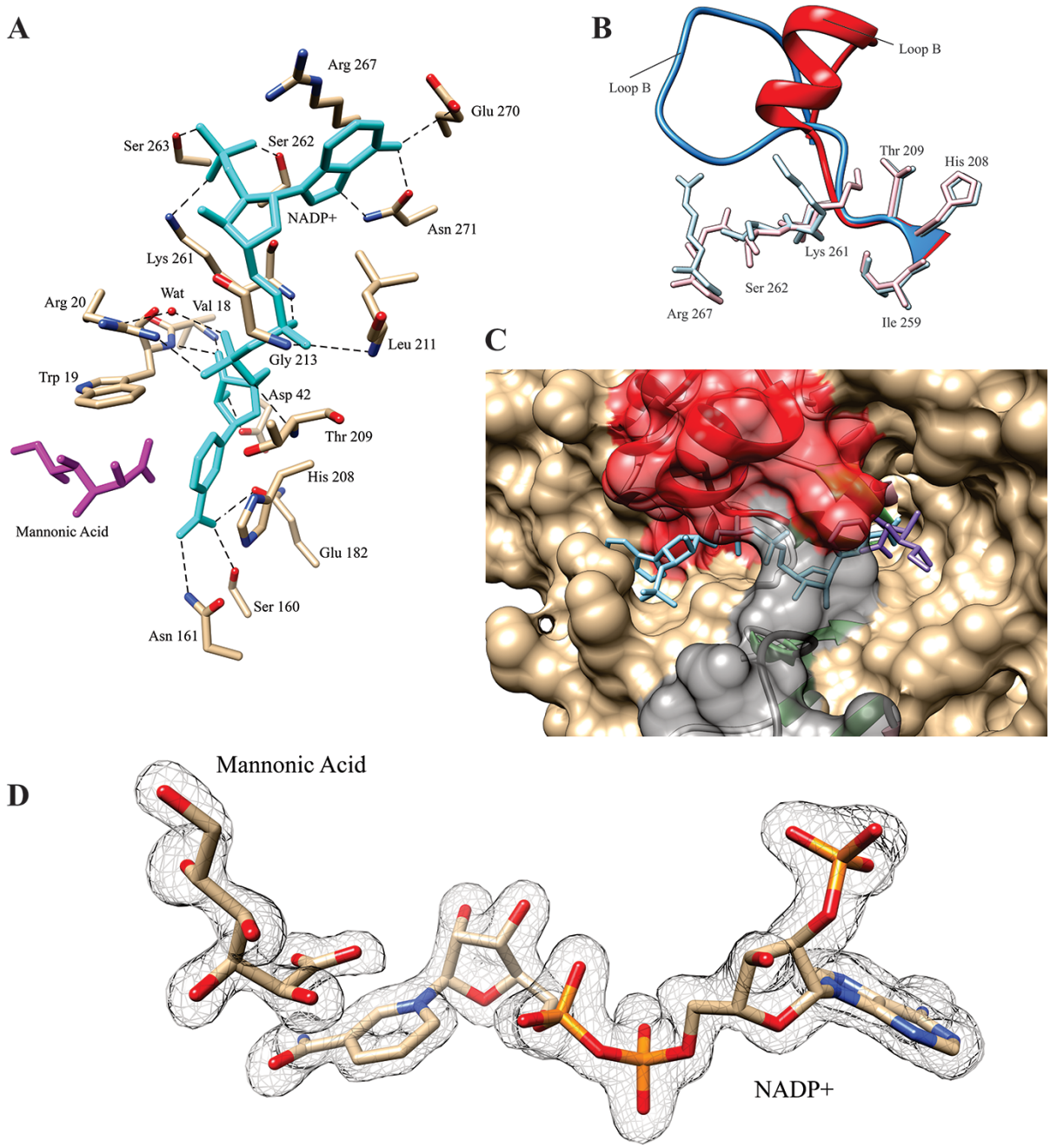

Fig. 4 Cofactor binding in *Agr*Man6PRase and *Ppe*Ald6PRase. (A) Residues involved in cofactor binding are shown as sticks and colored by atom type. The cofactor and mannonic acid are presented as sticks and colored in cyan and magenta, respectively. A water molecule (Wat) is shown as a red ball. The hydrogen bonds formed between the protein and the cofactor with distances ranging from 2.7 to 3.4 Å are shown as dashed lines. (B) Conservation of key $NADP^+$ interactors between *Agr*Man6PRase (light blue) and *Ppe*Ald6PRase (pink). Loop B adopts a helix conformation in the holo form of *Agr*Man6PRase (red), absent in the apo form of *Ppe*Ald6PRase (blue). (C) Surface representation of *Agr*Man6PRase in the surroundings of the substrates' binding pocket. Both loop B (red) and loop $\beta3$-$\alpha1$ (gray) close the structure on top of the substrates. $NADP^+$ and mannonic acid are colored in light blue and violet, respectively. (D) Detail of the electron density map of $NADP^+$. The omit map (Fo-Fc) is shown as the gray mesh around the ligand atoms at 3.2 sigma level.

of the binding cavity (**Fig. 5A**). These residues are located at <3.5 Å from the cofactor. Amino acids from the catalytic tetrad (Asp^{42}, Tyr^{47}, Lys^{76} and His^{107}) are close to both mannonic acid and $NADP^+$ (**Fig. 5B**); comparative analysis made between *Agr*Man6PRase and *Ppe*Ald6PRase revealed high structural conservation within the catalytic tetrad (**Fig. 5C**), even though *Ppe*Ald6PRase was crystallized in the absence of substrates.

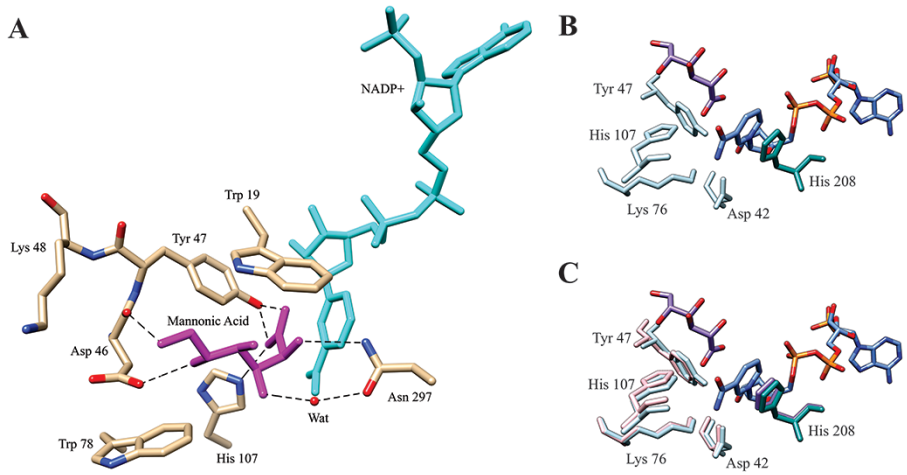

Fig. 5 Binding of mannonic acid to *Agr*Man6PRase and active site of *Agr*Man6PRase and *Ppe*Ald6PRase. (A) Binding residues are shown as sticks and colored by atom type. The cofactor and mannonic acid are presented as sticks and colored in cyan and magenta, respectively. A water molecule (Wat) is shown as a red ball. The hydrogen bonds formed with the cofactor with distances ranging from 2.6 to 3.2 Å are shown as dashed lines. (B) Catalytic tetrad (light blue residues) in *Agr*Man6PRase. (C) Conservation of the catalytic tetrad between *Agr*Man6PRase (light blue residues) and *Ppe*Ald6PRase (pink residues).

Importance of Lys^{48} for the selectivity of Man6P in *Agr*Man6PRase

To determine the residues that would be involved in Man6P binding, we first selected the amino acids located within 6 Å from mannonic acid: Trp^{19}, Ala^{44}, Asp^{46}, Tyr^{47}, Lys^{48}, Lys^{76}, Trp^{78}, His^{107}, Phe^{108}, Gln^{182}, Arg^{220}, Phe^{221}, Asn^{297}, Pro^{299} and Trp^{303} (**Fig. 6A**). Afterward, we did a multiple sequence alignment containing *Agr*Man6PRase and AKRs using different sugars as substrates, namely peach and apple Ald6PRases (which reduce Glc6P), *Candida tenuis* xylose reductase and human aldose reductase (which use non-phosphorylated substrates; **Supplementary Fig. S10**). The alignment showed that residues Trp^{19}, Ala^{44}, Tyr^{47}, Lys^{76}, Trp^{78}, His^{107}, Phe^{108}, Gln^{182} and Pro^{299} are conserved in all enzymes, while Asp^{46}, Arg^{220} and Asn^{297} are only present in *Agr*Man6PRase. Lys^{48}, Phe^{221} and Trp^{303} are conserved among the reductases using sugar phosphates (i.e. *Agr*Man6PRase and both Ald6PRases; **Supplementary Fig. S10**). Then, we performed a detailed analysis using the crystal structure of *Agr*Man6PRase and determined that Lys^{48} is the only amino acid located near the C6-OH group from mannonic acid, where the phosphate group from Man6P would potentially be located. To further investigate the possible interplay between Lys^{48} and the phosphate moiety of Man6P, we modeled a mannonic acid 6-phosphate molecule within the *Agr*Man6PRase structure (**Fig. 6A**). As shown in **Fig. 6A** and **Supplementary Fig. S11**, Lys^{48} interacts with the phosphate group of mannonic acid 6-phosphate.

We then produced the *Agr*Man6PRase K48A mutant to test the hypothesis that Lys^{48} is important for the binding of the sugar phosphate. The purified K48A mutant displayed a similar molecular size to the wild-type enzyme when analyzed both by SDS–PAGE (~35 kDa) and by size-exclusion chromatography (64 kDa), suggesting the mutation did not alter the homodimeric quaternary structure observed for the wild-type enzyme (**Supplementary Fig. S12**). We characterized both the wild-type and the K48A mutant enzymes with Man6P, Man and Glc6P in the presence of $NADP^+$ as a cofactor (**Table 1**), to evaluate the substrate specificity. Results presented in **Table 1** were further processed to estimate catalytic efficiency values (**Fig. 6B-C**). The wild-type enzyme displayed similar catalytic efficiencies for Man and Glc6P, although these values are 3 orders of magnitude lower than the one calculated for Man6P (**Fig. 6B**) as a consequence of decreased V_{max} and increased K_m values for Man and Glc6P compared to Man6P (**Table 1**). Similar results were observed for the catalytic efficiencies of NADPH in the presence of Man6P, Man and Glc6P (**Fig. 6C** and **Table 1**). Interestingly, the K48A mutant showed similar kinetic parameters for Man and Glc6P compared to the wild-type enzyme; however, the use of Man6P was dramatically impaired in the K48A mutant: the K_m increased 4-fold and the V_{max} decreased 5-fold (**Table 1**), thus reducing 20-fold the catalytic efficiency (**Fig. 6B**). Conversely, the K_m for NADPH with Man decreased 3-fold and the V_{max} increased 4-fold in the K48A mutant compared to the wild-type enzyme, thus producing a 15-fold higher catalytic efficiency for NADPH with Man in the K48A mutant than in the wild-type enzyme (**Table 1** and **Fig. 6C**).

Discussion

Members of the AKR superfamily are found in all living organisms, from prokaryotes to animals (Hyndman et al. 2003). This superfamily includes aldose reductases, aldehyde reductases, hydroxyl-steroid dehydrogenases and dihydrodiol dehydrogenases (Jez et al. 1997), which reduce a plethora of substrates

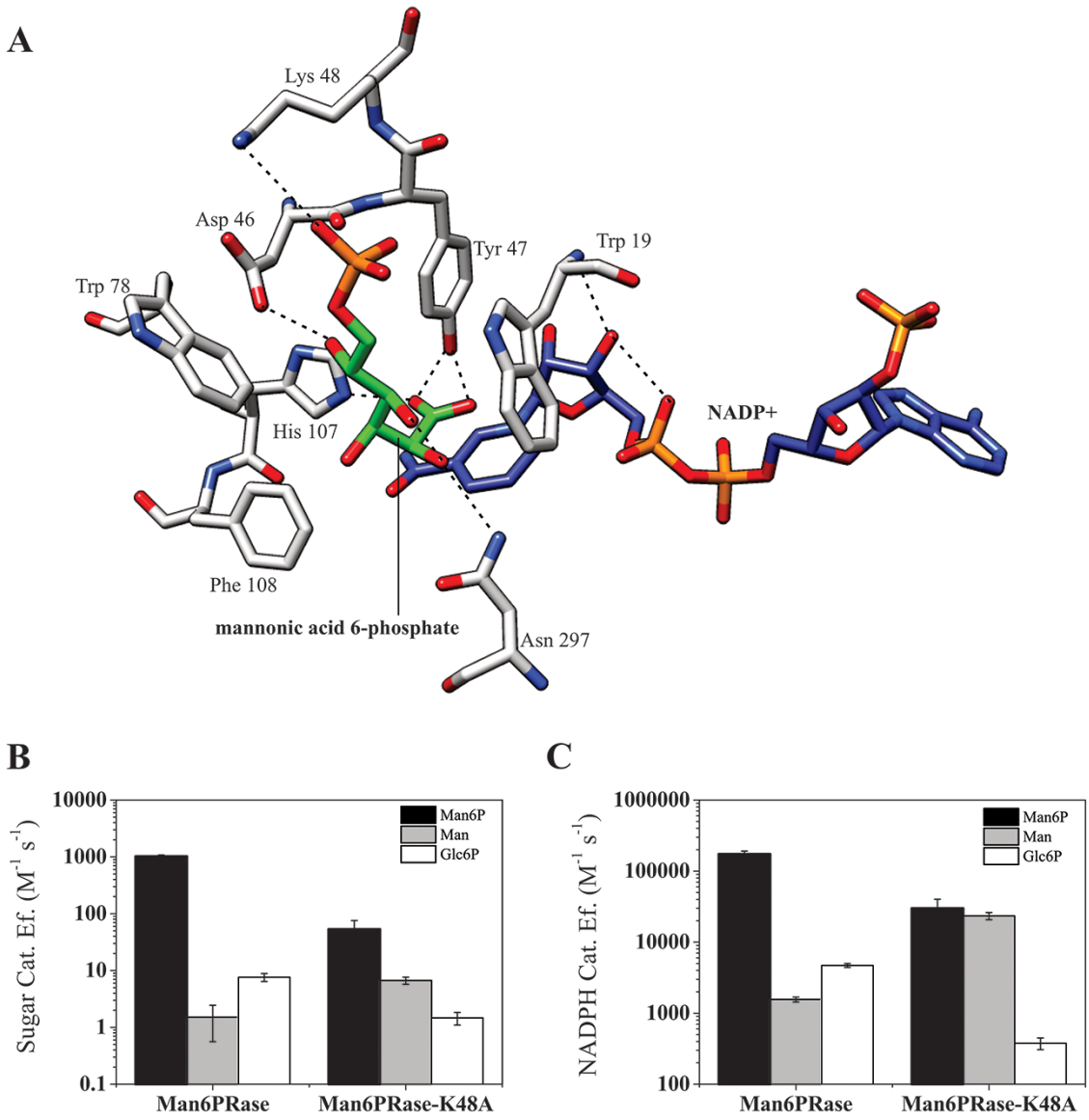

Fig. 6 Lys[48] is important for Man6P selectivity in *Agr*Man6PRase. (A) Hydrogen bonds (dashed lines) between *Agr*Man6PRase and the modeled mannonic acid 6-phosphate molecule (green). Amino acids were colored in by atom type, and $NADP^+$ was colored in blue. (B) Comparison of the catalytic efficiencies of *Agr*Man6PRase and the *Agr*Man6PRase K48A mutant for Man6P, Man and Glc6P. (C) Comparison of the catalytic efficiencies of *Agr*Man6PRase and the *Agr*Man6PRase K48A mutant for $NADP^+$ with Man6P, Man and Glc6P. To calculate catalytic efficiencies, we used the theoretical molecular mass of the recombinant *Agr*Man6PRase, including the His-tag (77.2 kDa). Error bars indicate the propagated error from data presented in **Table 1**.

such as sugar aldehydes, keto-steroids, keto-prostaglandins, vitamin A aldehydes, quinones and by-products of lipid peroxidation (Penning 2015). Most AKRs are monomeric; however, members of the AKR2, AKR6 and AKR7 families can adopt multimeric structures (Hyndman et al. 2003). Sengupta et al. (2015) reported the existence of 35 different plant AKRs; from these, 28 proteins were included in the AKR superfamily database (https://hosting.med.upenn.edu/akr/; last updated 8 May 2019). Conversely, the latter contains eight sequences not included in the work of Sengupta et al. (2015), four of them already presented in published papers (Bashir et al. 2006, de Sousa et al. 2009, Jirschitzka et al. 2012, Pan et al. 2019). According to the AKR superfamily database, these enzymes belong to families 2, 4 and 6, and they function in the detoxification of reactive aldehydes, production of osmolytes and secondary metabolites, and transport across membranes (Sengupta et al. 2015). Based on the AKR superfamily database, the AKR2 family has 22 members, including celery Man6PRase (AKR2A1), apple and rice Ald6PRases (AKR2A2 and AKR2A3, respectively), 10 xylose reductases from yeast (AKR2B1–9) and 10 proteins (AKR2C1, AKR2D1 and AKR2E1–8) with diverse activities and from different non-plant organisms.

In this work, we cloned the gene coding for *Agr*Man6PRase and characterized the recombinant protein. Our results were similar to those reported for the enzyme purified from celery leaves and for other plant AKRs, as shown in **Supplementary Table S2**. We determined that recombinant *Agr*Man6PRase was a dimer, which is in good agreement with data previously reported for the same enzyme (Loescher et al. 1992, Everard et al. 1997), as well as with results obtained by our group for recombinant Ald6PRases from apple (Figueroa and Iglesias 2010) and peach (Hartman et al. 2017).

We obtained crystals of *Agr*Man6PRase (in the presence of $NADP^+$ and mannonic acid) and *Ppe*Ald6PRase (in the apo form) and elucidated their structures. As far as we know, these are the first structures of dimeric plant AKR proteins to be solved. The nature of specific residues located in particular structural elements would be important for AKR2 dimerization. The dimer is formed through the specific interactions of amino acids located in loop A and the C-term from one subunit and amino acids positioned in helices α5, α6 and the loop β8/α6 from the opposite subunit, and vice versa. Strikingly, we found three amino acids in the α5 helix that generates >10 contacts with multiple residues located in loop A from the opposite subunit. A closer analysis of the α5 helix indicates that these amino acids are highly conserved between dimeric AKRs but are not present in monomeric plant AKRs (**Supplementary Fig. S13**). Furthermore, loop A is a highly divergent region in the multiple alignment, with an insertion-deletion of nine amino acids, being those from dimeric proteins the larger ones (**Supplementary Fig. S13**). These data suggest that dimeric AKRs evolved to promote the interaction between the α5 helix and loop A from the opposite subunit, by introducing amino acids that are not present in monomeric AKRs (**Supplementary Fig. S13**). The dimeric structure of plant Man6PRase and Ald6PRase (both from the AKR2 family) seems crucial for their enzymatic activity. The transition from an active dimer to an inactive monomer by introducing a His-tag at the C-terminus was already reported for apple Ald6PRase (Figueroa and Iglesias 2010). In this manuscript, we showed that the quadruple mutant of *Agr*Man6PRase is a monomer with negligible enzymatic activity. Whether the loss of activity is a consequence of the dimer-to-monomer transition or due to the accumulation of point mutations remains to be elucidated.

Both *Agr*Man6PRase and *Ppe*Ald6PRase structures are very similar, with the most significant variations in the regions comprising the loops that determine cofactor binding. The central role of such loops was demonstrated when the loop elements from a donor enzyme were grafted into an acceptor enzyme, switching the substrate specificity in the acceptor enzyme toward that of the donor enzyme (Campbell et al. 2013). These results agree with already published data, which mention that a small portion of loop B undergoes a conformational change when the cofactor binds, thus locking NADPH within the active site; this is the rate-limiting step of the reaction mechanism (Grimshaw et al. 1990, Borhani et al. 1992, Kubiseski et al. 1992).

The crystal structures of *Agr*Man6PRase and *Ppe*Ald6PRase displayed a TIM barrel folding, which is conserved in all proteins from the AKR superfamily. They also showed large loops at the bottom of the barrel, which are putatively involved in determining the specificity for the sugar phosphate (Harrison et al. 1994, Kavanagh et al. 2002). Both structures also showed a cofactor-binding domain, which is well conserved in other family members (Harrison et al. 1994, Kavanagh et al. 2002, Songsiriritthigul et al. 2020). Enzymes from the AKR superfamily display a bi-bi-ordered kinetic mechanism, in which the cofactor binds first and exits last (Grimshaw et al. 1995, Cooper et al. 2007). Comparative structural studies between the apo and holo forms of other aldose reductases established that NADPH not only provides the hydride for substrate reduction but also plays a key role as a structural component. Thus, NADPH is important for the organization of the active site residues and the binding pocket of the second substrate by producing a catalytically competent form of the enzyme (Sanli and Blaber 2001, Sanli et al. 2003).

The spatial distribution of the residues putatively involved in the catalytic mechanism (Asp^{42}, Tyr^{47}, Lys^{76} and His^{107} in *Agr*Man6PRase) was well conserved in the structures of *Agr*Man6PRase and *Ppe*Ald6PRase. Previous studies (Bohren et al. 1994, Barski et al. 1995, Schlegel et al. 1998) suggested that the hydroxyl group of Tyr^{47} is the most likely candidate to provide acid catalytic assistance for the reduction of the carbonyl group of the substrate. Tyr mutants were inactive, whereas mutation of the His residue generated partially active enzymes, as described below (Bohren et al. 1994). Also, studies performed with xylose reductases from *C. tenuis* and *Debaryomyces nepalensis* demonstrated that the amino group of the side chain from Lys^{76} facilitates the reduction of the carbonyl group by the cofactor providing electrostatic stabilization of the transition state (Kratzer and Nidetzky 2005, Kratzer et al. 2006, Paidimuddala et al. 2018). Regarding the function of Asp^{42}, studies on mutant enzymes at this position are consistent with the structural indication that the Asp carboxylate side chain forms a salt bridge with Lys^{76}. Such interaction stabilizes Lys^{76} and correctly positions its protonated side chain (Kavanagh et al. 2002, Kratzer and Nidetzky 2005, Kratzer et al. 2006, Paidimuddala et al. 2018).

Several previous studies have used site-directed mutagenesis to analyze the involvement of His^{107} in the active site (Bohren et al. 1994, Barski et al. 1995, Schlegel et al. 1998, Kratzer et al. 2006). The low but measurable activities of mutants at this position substituted with Asn, Gln, Glu or Ala suggested that His^{107} plays a 'helper' function in the catalytic mechanism of AKRs. Trying to reconcile the evidence obtained by previous studies in a unified catalytic mechanism (which would be consistent with the conservation pattern of this residue), it was proposed that His^{107} plays a preponderant role in positioning the carbonyl group of the substrate for catalysis (Kratzer et al. 2006). Additionally, we observed that Trp^{19} is in the vicinity of the catalytic site. It has been suggested that the side chain of Trp^{19} is necessary for optimal catalytic efficiency, although this residue is not

part of the catalytic tetrad. Indeed, it has been shown that Trp[19] is involved in the selective binding of the aldose (Kratzer et al. 2006). The proposed function of Trp[19] may be of general importance because aldehyde-preferring enzymes from AKR families 1–5 typically have a Trp residue in a position homologous to Trp[19]. In contrast, AKRs that prefer ketones often present Tyr or Phe at the equivalent position to that of Trp[19] (Kratzer et al. 2006).

Although the crystallization trials for *Agr*Man6PRase were carried out in the presence of $NADP^+$ and Man6P, we found mannonic acid instead in the refined structure. We propose that, during crystallization, the Man6P added to the medium lost the phosphate group, and the resulting Man was oxidized to mannonic acid. Among the residues interacting with the phosphate group of Man6P in the active site, most are conserved in different reductases; however, Lys[48] is only present in reductases whose substrates are sugar phosphates. Plant AKRs usually have a Lys at this position, and, in some cases, this residue is replaced by Arg or Gln. In non-plant homologs, such as mammalian AKR1s, the analogous position is occupied by Gln, Asn or Gly (**Supplementary Fig. S13**). To the best of our knowledge, this is the first study showing the importance of the residue located at this position (regardless of the amino acid that is found in the investigated sequence). Moreover, this Lys[48] is near the position where the phosphate group from Man6P should be located. We speculated that this residue could interact with the phosphate group from Man6P, which was supported by modeling mannonic acid 6-phosphate into the *Agr*Man6PRase structure. In comparison with the wild-type enzyme, the *Agr*Man6PRase K48A mutant showed a lower catalytic efficiency with Man6P and a higher catalytic efficiency with Man. The ratio $\frac{\left(\frac{k_{cat}}{K_m}\right)_{Man6P}}{\left(\frac{k_{cat}}{K_m}\right)_{Man}}$ in the wild-type enzyme is 677, whereas in the K48A mutant is 8.1. These results strongly suggest that Lys[48] is important for binding the phosphate group from Man6P, probably by stabilizing its negative charge (Ballicora et al. 2005). Consequently, the evolutionary role of this residue seems to be to increase the specificity for the sugar phosphate compared to the sugar.

To conclude, we showed that AKRs involved in sugar alcohol synthesis in plants share similar structural properties to their counterparts in other kingdoms (Kavanagh et al. 2002). We also proved that the quaternary structure of proteins from the ARK2 family is important for the enzymatic activity, as we previously showed for the apple Ald6PRase (Figueroa and Iglesias 2010). Additionally, we demonstrated that Lys[48] is relevant for the selectivity of the sugar phosphate. This information would be relevant to develop enzymes with novel catalytic properties, which could be further used in different industrial processes (Hartman et al. 2019). Overall, our work lays the ground for further studies on enzymes regulating the synthesis of sugar alcohols in plants, which should help to determine structure-to-function relationships on this important group of enzymes.

Materials and Methods

Plant material, bacterial strains and reagents

Celery (*Apium graveolens*) plants were grown under a 16-h light/8-h dark diel cycle, with a light intensity of 100 μmol m^{-2} s^{-1}, at 23°C and 70% relative humidity. Fully developed leaves from 6-week-old plants were harvested and rapidly frozen with liquid nitrogen and stored at −80°C until use. *Escherichia coli* TOP10 (Invitrogen, Waltham, MA, USA) was used for cloning procedures and plasmid maintenance. Protein expression was carried out using *E. coli* BL21 Star (DE3) (Invitrogen). The substrates used to determine enzyme activity were from Sigma-Aldrich. Crystallization screen solutions and other supplies were purchased from Hampton Research. All other reagents were of the highest purity available.

Cloning of the sequence coding for *Agr*Man6PRase from celery leaves

Total RNA from celery leaves was isolated using TransZol (TransGen Biotech, China). The complementary DNA (cDNA) was synthesized at 42°C for 1 h in a 15-μl reaction mixture containing 1,600 ng of RNA, 1 mM dNTPs, 80 pmol of oligo(dT) primer and 40 U of M-MLV reverse transcriptase (Promega). To amplify the gene coding for *Agr*Man6PRase, we designed the primers *Agr*Man6PRase-Fw (CATATGGCAATAACTCTTAACAGCGG, the *Nde*I site is underlined) and *Agr*Man6PRase-Rv (GGATCCTCAAGCATAAACATCTATTCC, the *BamH*I site is underlined), based on the published sequence coding for celery Man6PRase (NCBI nucleotide ID U83687.1 and protein IDAAB97617.1). The coding sequence was amplified by polymerase chain reaction (PCR) with 2 μl of cDNA solution, 2 mM dNTPs, 100 pmol of each primer and 2.5 U of Pfu DNA polymerase (Productors Bio-L'ogicos, Buenos Aires, Argentina), using the following conditions: 5 min at 95°C, 30 cycles of 1 min at 95°C, 1 min at 50°C and 2 min at 68°C, followed by 10 min at 68°C. The amplified DNA was cloned into the pGEM-T Easy vector (Promega), and the resulting construct was used to transform *E. coli* TOP10 cells. The identity of the cloned sequence was determined by automated sequencing (Macrogen, South Korea) of at least two clones from two independent PCR events. The sequence encoding *Agr*Man6PRase was then subcloned into the pET28c vector (Novagen) between the *Nde*I and *BamH*I sites to obtain the recombinant protein fused to a His_6-tag at the N-terminus.

We used the QuikChange Site-Directed Mutagenesis Kit (Agilent) to introduce the K48A mutation in the [pET28c/*AgrMan6PRase*] construct with the following primers (mutated codons are underlined): GCTGCTGA CTACGCGAATGAGTTAGAAG (forward) and CTTCTAACTCATTCGCGTAG TCAGCAGC (reverse). Insertion of the mutation was confirmed by automated DNA sequencing (Macrogen).

DNA synthesis of the sequence reported by Everard et al. (1997) and subcloning into the pET28b vector between the *Nde*I and *Sac*I sites was performed by BioBasic (Canada). The DNA sequence was codon optimized for expression in *E. coli* (**Supplementary Text S3**). The resulting protein sequence is presented in **Supplementary Text S4**.

The construct [pET19b/*PpeAld6PRase*] used in this work was previously reported by our group (Hartman et al. 2017).

Protein expression and purification

The constructs [pET28c/*AgrMan6PRase*] and [pET28c/*AgrMan6PRase-K48A*] were used to transform *E. coli* BL21 Star (DE3) cells. Expression of the recombinant proteins was performed by inoculating 1 l of lysogeny broth medium (supplemented with 50 μg ml^{-1} kanamycin) with a 1/100 dilution of an overnight culture. Cells were grown at 37°C and 180 rpm in an orbital shaker until the OD_{600} reached ∼0.6, induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside at 25°C overnight, harvested at 5,000×*g* at room temperature for 15 min and kept at −20°C until use.

The cell paste was resuspended with 25 ml of Buffer A [25 mM Tris–HCl pH 8.0, 300 mM NaCl, 5% (v/v) glycerol, 10 mM imidazole] and disrupted by sonication. The resulting suspension was centrifuged twice at 20,000×*g* at 4°C for 10 min. The supernatant was loaded onto a 1 ml HisTrap column connected to an ÄKTA Explorer 100 purification system (GE Healthcare, chicago, IL, USA), previously equilibrated with Buffer A. The column was washed with 10 ml of Buffer A and the recombinant protein was eluted with a linear gradient of imidazole
(10–300 mM, 50 ml). The fractions containing the enzyme of interest were collected and concentrated.

Expression and purification of *Ppe*Ald6PRase were performed as previously described (Hartman et al. 2017).

Protein methods

Protein electrophoresis was performed under denaturing conditions (SDS–PAGE), as described by Laemmli (1970). Protein concentration was determined following the procedure described by Bradford (1976), using bovine serum albumin as the standard.

Native molecular mass determination

The native molecular mass of *Agr*Man6PRase was determined by using a Superdex 200 10/300 column (GE Healthcare) equilibrated with Buffer G (50 mM HEPES–NaOH pH 8.0; 100 mM NaCl; 0.1 mM EDTA). A calibration curve was constructed by plotting the K_{av} values versus the log of the molecular mass of standard proteins (ribonuclease, 13.7 kDa; carbonic anhydrase, 29 kDa; ovalbumin, 43 kDa; conalbumin, 75 kDa; aldolase, 158 kDa; ferritin, 440 kDa; and thyroglobulin, 669 kDa). The K_{av} was calculated as $(V_e - V_0)/(V_t - V_0)$, where V_e is the elution volume of the protein, V_0 is the elution volume of Dextran Blue and V_t is the total volume of the column.

Enzyme activity assay and determination of kinetic constants

The activity of *Agr*Man6PRase was determined by following NADPH oxidation at 340 nm. The standard assay mixture contained (unless otherwise specified) 50 mM Tris–HCl pH 7.5, 3 mM DTT, 0.3 mM NADPH, 30 mM Man6P and enzyme in an appropriate dilution. Reactions were carried out for 10 min in a final volume of 50 µl at 25°C in a 384-microplate reader (Multiskan GO, Thermo Scientific). One unit of enzyme activity (U) is defined as the amount of enzyme catalyzing the oxidation of 1 µmol NADPH in 1 min under the specified assay conditions.

Kinetic constants were determined by measuring enzyme activity at different concentrations of one substrate while keeping a fixed and saturating concentration of the other substrate. Data of the initial velocity (v) were plotted versus the substrate concentration and fitted to the Michaelis–Menten equation, $v = V_{max}[S]/(K_m + [S])$, where K_m is the concentration of the substrate (S) producing 50% of the maximal velocity (V_{max}). Kinetic constants and their corresponding standard errors were calculated by the computer program Origin 8.1 (OriginLab Corporation, Northampton, MA, USA), using the mean of at least two independent sets of data. The catalytic efficiency was calculated as the ratio k_{cat}/K_m, whereas k_{cat} was calculated using the V_{max} and the theoretical molecular mass of Man6PRase (77.2 kDa).

The biochemical characterization of the recombinant *Ppe*Ald6PRase was reported by our group in a previous publication (Hartman et al. 2017).

Crystallization and data collection

To carry out the crystallization assays, *Agr*Man6PRase and *Ppe*Ald6PRase were first purified by IMAC and then by size-exclusion chromatography. For this purpose, a HiLoad 16/600 Superdex 200-pg column (GE Healthcare) previously equilibrated with Buffer G was used. Finally, the fractions containing the enzyme of interest were pooled and concentrated.

Crystals were first obtained using the sitting-drop vapor diffusion method at 20°C. *Agr*Man6PRase drops were prepared with 6 µl of 10 mg ml^{-1} enzyme and 6 µl of reservoir solution, containing 0.2 M ammonium formate (pH 6.6) and 20% (w/v) PEG 3500. The initial crystallization condition was optimized by adding the metabolites Man6P and $NADP^+$. The optimal condition for obtaining *Ppe*Ald6PRase crystals was 0.1 M HEPES (pH 7.5) and 20% (w/v) PEG 8000 for drops consisting of 6 µl of 14 mg ml^{-1} protein and 6 µl of reservoir solution. Crystals with a good morphology and large size were transferred to a cryo-condition, which contained 25% (v/v) glycerol in addition to the components of the reservoir solution, before being frozen in liquid nitrogen.

X-ray diffraction data sets were collected at the SBC19-ID beamline at the Advanced Photon Source (Argonne National Laboratory, Chicago, IL, USA). The wavelength used in the monochromatic data collection was 1.008 Å. All the collected data sets were indexed and integrated and scaled using HKL2000 (Otwinowski and Minor 1997). The molecular replacement was conducted using Phaser (McCoy et al. 2007). For the Man6PRase structure, a known structure model (PDB ID 2IKI) of a human aldose reductase was used as the starting model for molecular replacement. For the Ald6PRase structure, a known structural model (PDB ID 1HQT) of an aldehyde reductase from *Sus scrofa* was used as the starting search model. The refinement and model building were done using Phenix (Adams et al. 2010) and Coot (Emsley et al. 2010) programs.

Modeling of mannonic acid-6-phosphate into the *Agr*Man6PRase structure

Modeling of the ligand mannonic acid-6-phosphate was performed as previously indicated (Bhayani et al. 2019). The oxygen of the hydroxyl group in position 6 of the mannonic acid already present in the crystal structure (PDB ID 7S5F) was manually replaced by a phosphate group. The coordinates of the phosphorus and oxygens were later optimized for angles and bond lengths. This optimization was performed with Gnumeric 1.12 using the non-linear regression 'Solver' utility. No further optimization of the protein structure was performed since no clash was observed using the program Chimera 1.13.1 (Pettersen et al. 2004).

Supplementary Data

Supplementary data are available at *PCP* online.

Data Availability

The nucleotide sequence of celery Man6PRase reported in this article has been submitted to the NCBI (nucleotide ID OK043677). The crystal structures of celery Man6PRase and peach Ald6PRase have been submitted to the PDB (structure IDs 7S5F and 7S5I, respectively).

Funding

Agencia Nacional de Promoción de la Investigación, el Desarrollo Tecnológico y la Innovación (PICT-2017-1515, PICT-2018-00929, PICT-2018-00865); Universidad Nacional del Litoral (CAI+D 2020); Consejo Nacional de Investigaciones Científicas y Técnicas (PUE-2016-0040); National Science Foundation (MCB 1616851); Max Planck Society (Partner Group for Plant Biochemistry).

Acknowledgements

R.I.M. was a fellow from Agencia I+D+i and CONICET. M.D.H. is a fellow from Agencia I+D+i. M.D.H. and A.E.C. were Fulbright fellows. A.A.I. and C.M.F. are researchers from CONICET.

Author Contributions

Conceptualization: M.A.B., A.A.I., D.L. and C.M.F.; investigation: R.I.M., J.A.B., M.D.H., A.E.C. and Y.Z.; formal analysis: all authors; writing—original draft: R.I.M., M.D.H., D.L. and C.M.F.; writing—review and editing: all authors.

Disclosures

The authors have no conflicts of interest to declare.

References

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66: 213–221.

Ballicora, M.A., Dubay, J.R., Devillers, C.H. and Preiss, J. (2005) Resurrecting the ancestral enzymatic role of a modulatory subunit. *J. Biol. Chem.* 280: 10189–10195.

Barski, O.A., Gabbay, K.H., Grimshaw, C.E. and Bohren, K.M. (1995) Mechanism of human aldehyde reductase: characterization of the active site pocket. *Biochemistry* 34: 11264–11275.

Bashir, K., Inoue, H., Nagasaka, S., Takahashi, M., Nakanishi, H., Mori, S., et al. (2006) Cloning and characterization of deoxymugineic acid synthase genes from graminaceous plants. *J. Biol. Chem.* 281: 32395–32402.

Bhayani, J.A., Hill, B.L., Sharma, A., Iglesias, A.A., Olsen, K.W. and Ballicora, M.A. (2019) Mapping of a regulatory site of the *Escherichia coli* ADP-glucose pyrophosphorylase. *Front. Mol. Biosci.* 6: 89.

Bohren, K.M., Grimshaw, C.E., Lai, C.-J., Harrison, D.H., Ringe, D., Petsko, G.A., et al. (1994) Tyrosine-48 is the proton donor and histidine-110 directs substrate stereochemical selectivity in the reduction reaction of human aldose reductase: enzyme kinetics and crystal structure of the Y48H mutant enzyme. *Biochemistry* 33: 2021–2032.

Bomati, E.K., Austin, M.B., Bowman, M.E., Dixon, R.A. and Noel, J.P. (2005) Structural elucidation of chalcone reductase and implications for deoxychalcone biosynthesis. *J. Biol. Chem.* 280: 30496–30503.

Borhani, D.W., Harter, T.M. and Petrash, J.M. (1992) The crystal structure of the aldose reductase·NADPH binary complex. *J. Biol. Chem.* 267: 24841–24847.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein using the principle of protein dye binding. *Anal. Biochem.* 72: 248–254.

Campbell, E., Chuang, S. and Banta, S. (2013) Modular exchange of substrate-binding loops alters both substrate and cofactor specificity in a member of the aldo-keto reductase superfamily. *Protein Eng. Des. Sel.* 26: 181–186.

Cooper, W.C., Jin, Y. and Penning, T.M. (2007) Elucidation of a complete kinetic mechanism for a mammalian hydroxysteroid dehydrogenase (HSD) and identification of all enzyme forms on the reaction coordinate: the example of rat liver 3α-HSD (AKR1C9). *J. Biol. Chem.* 282: 33484–33493.

de Sousa, S.M., Rosselli, L.K., Kiyota, E., da Silva, J.C., Souza, G.H., Peroni, L.A., et al. (2009) Structural and kinetic characterization of a maize aldose reductase. *Plant Physiol. Biochem.* 47: 98–104.

Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66: 486–501.

Everard, J.D., Cantini, C., Grumet, R., Plummer, J. and Loescher, W.H. (1997) Molecular cloning of mannose-6-phosphate reductase and its developmental expression in celery. *Plant Physiol.* 113: 1427–1435.

Figueroa, C.M. and Iglesias, A.A. (2010) Aldose-6-phosphate reductase from apple leaves: importance of the quaternary structure for enzyme activity. *Biochimie* 92: 81–88.

Figueroa, C.M., Minen, R.I., Podestá, F.E. and Iglesias, A.A. (2019) Carbon metabolic pathways and relationships with plant stress. *In* Handbook of Plant and Crop Stress, 4th edn. Edited by Pessarakli, M. pp. 389–404. CRC Press, Taylor & Francis Group, Boca Raton.

Giuseppe, P.O., Santos, M.L., Sousa, S.M., Koch, K.E., Yunes, J.A., Aparicio, R., et al. (2016) A comparative structural analysis reveals distinctive features of co-factor binding and substrate specificity in plant aldo-keto reductases. *Biochem. Biophys. Res. Commun.* 474: 696–701.

Grimshaw, C.E., Bohren, K.M., Lai, C.J. and Gabbay, K.H. (1995) Human aldose reductase: subtle effects revealed by rapid kinetic studies of the C298A mutant enzyme. *Biochemistry* 34: 14366–14373.

Grimshaw, C.E., Putney, C.G. and Shahbaz, M. (1990) Spectroscopic and kinetic characterization of nonenzymic and aldose reductase mediated covalent NADP-glycolaldehyde adduct formation. *Biochemistry* 29: 9936–9946.

Harrison, D.H., Bohren, K.M., Ringed, D., Petsko, G.A. and Gabbay, K.H. (1994) An anion binding site in human aldose reductase: mechanistic implications for the binding of citrate, cacodylate, and glucose 6-phosphate. *Biochemistry* 33: 2011–2020.

Hartman, M.D., Figueroa, C.M., Arias, D.G. and Iglesias, A.A. (2017) Inhibition of recombinant aldose-6-phosphate reductase from peach leaves by hexose-phosphates, inorganic phosphate and oxidants. *Plant Cell. Physiol.* 58: 145–155.

Hartman, M.D., Minen, R.I., Iglesias, A.A. and Figueroa, C.M. (2019) Cofactor specificity switch on peach glucitol dehydrogenase. *Biochemistry* 58: 1287–1294.

Hirai, M. (1981) Purification and characteristics of sorbitol-6-phosphate dehydrogenase from loquat leaves. *Plant Physiol.* 67: 221–224.

Hyndman, D., Bauman, D.R., Heredia, V.V. and Penning, T.M. (2003) The aldo-keto reductase superfamily homepage. *Chem. Biol. Interact.* 143–144: 621–631.

Jez, J.M., Bennett, M.J., Schlegel, B.P., Lewis, M. and Penning, T.M. (1997) Comparative anatomy of the aldo-keto reductase superfamily. *Biochem. J.* 326: 625–636.

Jirschitzka, J., Schmidt, G.W., Reichelt, M., Schneider, B., Gershenzon, J. and D'Auria, J.C. (2012) Plant tropane alkaloid biosynthesis evolved independently in the Solanaceae and Erythroxylaceae. *Proc. Natl. Acad. Sci. U.S.A.* 109: 10304–10309.

Kanayama, Y. and Yamaki, S. (1993) Purification and properties of NAD-dependent sorbitol dehydrogenase from apple fruit. *Plant Cell Physiol.* 34: 819–823.

Kavanagh, K.L., Klimacek, M., Nidetzky, B. and Wilson, D.K. (2002) The structure of apo and holo forms of xylose reductase, a dimeric aldo-keto reductase from *Candida tenuis*. *Biochemistry* 41: 8785–8795.

Khalil, S.R.M., Ibrahim, A.S., Hussien, B.A., Hussien, E.A. and Tawfik, M.S. (2017) Cloning of a functional mannose-6-phosphate reductase (M6PR) gene homolog from Egyptian celery plants (*Apium graveolens*): overexpression in non-mannitol producing plants resulted in mannitol accumulation in transgenic individuals. 3 *Biotech* 7: 341.

Kratzer, R. and Nidetzky, B. (2005) Electrostatic stabilization in a preorganized polar active site: the catalytic role of Lys-80 in *Candida tenuis* xylose reductase (AKR2B5) probed by site-directed mutagenesis and functional complementation studies. *Biochem. J.* 389: 507–515.

Kratzer, R., Wilson, D.K. and Nidetzky, B. (2006) Catalytic mechanism and substrate selectivity of aldo-keto reductases: insights from structure-function studies of *Candida tenuis* xylose reductase. *IUBMB Life* 58: 499–507.

Kubiseski, T.J., Hyndman, D.J., Morjana, N.A. and Flynns, T.G. (1992) Pig muscle aldose reductase. *Biochemistry* 267: 6510–6517.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227: 680–685.

Loescher, W. and Everard, J. (2000) Regulation of sugar alcohol biosynthesis. *Photosynth. Physiol. Metab.* 9: 275–299.

Loescher, W.H., Tyson, R.H., Everard, J.D., Redgwell, R.J. and Bieleski, R.L. (1992) Mannitol synthesis in higher plants. *Plant Physiol.* 98: 1396–1402.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* 40: 658–674.

Negm, F.B. and Loescher, W.H. (1981) Characterization and partial purification of aldose-6-phosphate reductase (alditol-6-phosphate: NADP1-oxidoreductase) from apple leaves. *Plant Physiol.* 67: 139–142.

Olsen, J.G., Pedersen, L., Christensen, C.L., Olsen, O. and Henriksen, A. (2008) Barley aldose reductase: structure, cofactor binding, and substrate recognition in the aldo/keto reductase 4C family. *Proteins Struct. Funct. Genet.* 71: 1572–1581.

Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Meth. Enzymol.* 276: 307–326.

Paidimuddala, B., Mohapatra, S.B., Gummadi, S.N. and Manoj, N. (2018) Crystal structure of yeast xylose reductase in complex with a novel NADP-DTT adduct provides insights into substrate recognition and catalysis. *FEBS J.* 285: 4445–4464.

Pan, L., Yu, Q., Han, H., Mao, L., Nyporko, A., Fan, L., et al. (2019) Aldo-keto reductase metabolizes glyphosate and confers glyphosate resistance in *Echinochloa colona*. *Plant Physiol.* 181: 1519–1534.

Penning, T.M. (2015) The aldo-keto reductases (AKRs): overview. *Chem. Biol. Interact.* 234: 236–246.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., et al. (2004) UCSF Chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25: 1605–1612.

Robert, S., Simier, P. and Fer, A. (1999) Purification and characterization of mannose 6-phosphate reductase, a potential target for the control of *Striga hermonthica* and *Orobanche ramosa*. *Aust. J. Plant Physiol.* 26: 233–237.

Sanli, G. and Blaber, M. (2001) Structural assembly of the active site in an aldo-keto reductase by NADPH cofactor. *J. Mol. Biol.* 309: 1209–1218.

Sanli, G., Dudley, J.I. and Blaber, M. (2003) Structural biology of the aldo-keto reductase family of enzymes: catalysis and cofactor binding. *Cell Biochem. Biophys.* 38: 79–101.

Schlegel, B.P., Ratnam, K. and Penning, T.M. (1998) Retention of NADPH-linked quinone reductase activity in an aldo-keto reductase following mutation of the catalytic tyrosine. *Biochemistry* 37: 11003–11011.

Sengupta, D., Naik, D. and Reddy, A.R. (2015) Plant aldo-keto reductases (AKRs) as multi-tasking soldiers involved in diverse plant metabolic processes and stress defense: a structure-function update. *J. Plant Physiol.* 179: 40–55.

Simpson, P.J., Tantitadapitak, C., Reed, A.M., Mather, O.C., Bunce, C.M., White, S.A., et al. (2009) Characterization of two novel aldo-keto reductases from Arabidopsis: expression patterns, broad substrate specificity, and an open active-site structure suggest a role in toxicant metabolism following stress. *J. Mol. Biol.* 392: 465–480.

Songsiriritthigul, C., Narawongsanont, R., Tantitadapitak, C., Guan, H.H. and Chen, C.J. (2020) Structure-function study of AKR4C14, an aldo-keto reductase from Thai jasmine rice (*Oryza sativa* L. ssp. *indica* cv. KDML105). *Acta Crystallogr. D Struct. Biol.* 76: 472–483.

Webb, K.L. and Burley, J.A.W. (1962) Sorbitol translocation in apple. *Science* 137: 766.

Yadav, R. and Prasad, R. (2014) Identification and functional characterization of sorbitol-6-phosphate dehydrogenase protein from rice and structural elucidation by *in silico* approach. *Planta* 240: 223–238.

Zhou, R., Sicher, R.C., Cheng, L. and Quebedeaux, B. (2003) Regulation of apple leaf aldose-6-phosphate reductase activity by inorganic phosphate and divalent cations. *Funct. Plant Biol.* 30: 1037–1043.