Directed divergent evolution of a thermostable D-tagatose epimerase towards improved activity for two hexose substrates

Andreas Bosshart[a], Chee Seng Hee[b], Matthias Bechtold[a], Tilman Schirmer[b] and Sven Panke*[a]

Abstract: We exploit the functional promiscuity of an engineered thermostable variant of a promiscuous D-tagatose epimerase (DTE Var8) to morph it into two efficient catalysts for the C3 epimerization of D-fructose to D-psicose and of L-sorbose to L-tagatose. Iterative single-site randomization and screening of 48 residues in the first and second shell around the substrate binding site of Var8 yielded 8-sites mutant IDF8 with an 9-fold improved k_{cat} for the epimerization of D-fructose, and the 6-sites mutant ILS6 with an 14-fold improved epimerization of L-sorbose compared to Var8. Structure analysis of IDF8 revealed a charged patch at the entrance of its active site that supposedly facilitates the entry of the polar substrate, whereas the improvement in variant ILS6 is thought to relate to subtle changes in the hydration of the bound substrate. The structures will inform future engineering efforts of these and other isomerizing enzymes for the production of rare.

Introduction

In the last decades, directed evolution of proteins has become a powerful tool to tailor enzymes for their application in industrial biocatalysis[1]. Directed evolution, relying on iterative cycles of mutagenesis and screening of the resulting libraries for variants that exhibit the desired traits[2], has been shown to allow improving almost any property that is of importance for an industrially useful biocatalyst, including thermostability[3], enantioselectivity[4] or catalytic rate[5].

The enzyme D-tagatose epimerase (DTE) catalyzes the interconversion of C3 ketohexose epimers[6] (**Figure 1**). It constitutes the central enzyme for biocatalytic access to rare monosaccharides, which have recently attracted great interest as low calorie sweeteners, chiral building blocks or as active pharmaceutical ingredients[7]. By combining DTE with maximally two additional isomerases, the whole set of 24 hexoses can be generated in a short cascade reaction from only 4 starting materials that are cheaply available (D-glucose, D-fructose, D-galactose, L-sorbose)[8]. D-Tagatose epimerases from various organisms are known to date [9], however none of them exhibits catalytic rates on the respective substrates that would make them attractive for application in an industrial context.

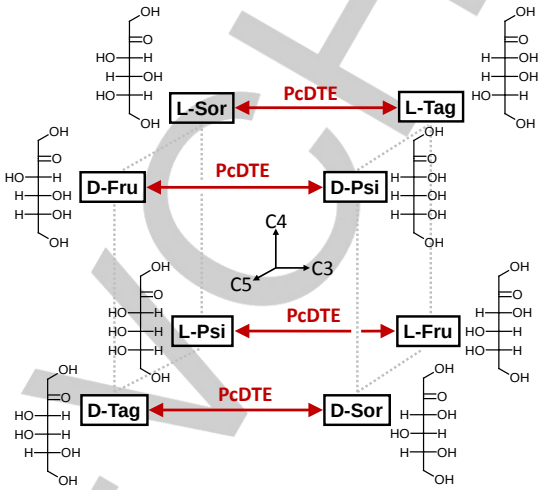


Figure 1. Schematic representation of all 8 ketohexoses in a sugar cube . Each of The X-axis relates C3-epimers (catalysed by the enzyme D-tagatose epimerase from *P. cichorii*, PcDTE), the Y-axis relates C4-epimers, and the Z-axis relates C5-epimers.

On the other hand, no active DTE homolog from thermophilic origin is known that could serve as an optimal starting point for improving catalytic efficiency. There is a growing body of evidence that enzymes from thermophiles are more suitable to serve as starting points for directed evolution than enzymes from mesophiles[10]. The major reason for this is that most mutations, and mutations altering protein function in particular, are destabilizing, increasing the chance that the mutation results in a protein that is not folded (correctly) and thus not functional anymore[10c]. Proteins from thermophilic organisms on the other hand have to encode enzymes that are functional at high temperatures, thus exhibiting a larger free energy difference (ΔG_u) between folded (native) and unfolded state than mesophilic enzymes[10d) 11]. This free folding energy can be considered as a 'stability reservoir' that can be consumed by accumulating mutations that encode for new protein functions but are destabilizing in terms of free energy of folding. This larger stability reservoir of thermostable enzymes therefore allows exploring a larger fraction of protein sequence space, which is often referred to as 'evolvability'[12], before the limit of folding stability is reached[10a].

These reasons clearly suggest proceeding in two steps for DTE. First, a mesostable enzyme that catalyzes the desired reaction should be (thermo-) stabilized and in a second step this thermostable enzyme variant should be evolved towards the desired (novel) functions.

Following this strategy we recently described the thermostabilization of a dimeric DTE from the mesophilic *Pseudomonas cichorii* (PcDTE) by systematically optimizing the dimeric interface interactions, generating a variant termed PcDTE Var8[3b]. This enzyme shows promiscuous enzymatic activity for C3 epimerization of all 4 ketohexose pairs and was

[a] A. Bosshart, Dr. M. Bechtold, Prof. S. Panke
Department of Biosystems Science and Engineering
ETH Zurich
Mattenstrasse 26, 4058 Basel, Switzerland
E-mail: sven.panke@bsse.ethz.ch
[b] Dr. C.S. Hee, Prof. T. Schirmer
Focal Area of Structural Biology and Biophysics
Biozentrum, University of Basel
CH-4056 Basel, Switzerland

Supporting information for this article is given via a link at the end of the document

therefore an interesting starting point for further catalytic rate optimization. We decided to divergently evolve this thermostable Var8 for improved production of two rare ketohexoses, specifically D-psicose from D-fructose and L-tagatose from L-sorbose. Both products are rare hexoses that are of considerable interest as low-calorie sweetener or chiral building blocks[13]. Var8 shows modest catalytic activity towards the epimerization of D-fructose to D-psicose (k_{cat}: 12.1 s^{-1} at 30°C) and only poor activity towards the epimerization of L-sorbose to L-tagatose (k_{cat}: 0.24 s^{-1} at 30°C). Regarding a suitable engineering strategy for these objectives, it has been shown previously that a limited number of amino acid substitutions in vicinity to the active site are sufficient to change the substrate specificity and catalytic activity of a promiscuous enzyme[14], potentially limiting the required effort to obtain improved biocatalysts substantially. Accordingly, we reasoned that a directed evolution strategy based on iterative saturation mutagenesis (ISM)[15] of residues that surround the active site of Var8 would be a promising method to improve catalytic efficiency towards D-fructose or L-sorbose. Besides the obvious benefit of superior biocatalysts for the production of rare sugars, the central position of DTE in sugar interconversions made us investigate the development of catalytic parameters for other ketohexose substrates as well.

In this work we describe the successful divergent evolution of thermostable Var8 towards the efficient epimerization of D-fructose/D-psicose (IDF variants) and L-sorbose/L-tagatose (ILS variants) by targeting residues that are in the first and second sphere around the active site. We characterized all variants of the two divergent evolutionary trajectories in terms of substrate specificity and catalytic activity towards six distinct ketohexoses. Furthermore, the crystal structures of Var8 in its substrate-free form as well as the crystal structure of the final variants IDF8 in complex with D-fructose and ILS6 in complex with L-sorbose give insight into the molecular basis of the altered enzymatic specificities.

Results and Discussion

Establishment of a high-throughput screening protocol

The screening effort that comes with large libraries is arguably still the main limiting factor in directed evolution. The strategy to consider only the 45 residues that are in proximity to the active site already reduces the workload significantly but still requires the screening of several thousands of clones. Therefore the establishment of a microtiter plate-based screening procedure was necessary as an HPLC-based assay described earlier[3b] seemed infeasible. A galactitol dehydrogenase (RsGD) from *Rhodobacter sphaeroides* was previously described for the NADH-dependent oxidation of galactitol to L-tagatose. L-Tagatose, the C3 epimerization product of L-sorbose, can hence be detected using RsGD in the reverse direction by following the oxidation of NADH at 340 nm. Indeed, RsGD reliably detected L-tagatose in the presence of L-sorbose (**Figure S3B**), a prerequisite for the screening assay.

A dehydrogenase that specifically reduces D-psicose in presence of D-fructose was less readily available. *Klebsiella pneumonia* strain 3321 was reported to contain a ribitol dehydrogenase that can specifically reduce D-psicose without reducing D-fructose[16]. The gene sequence coding for this protein was not specified, but the genomes of several *K. pneumoniae* strains have been sequenced to date. Based on sequence homology data, a ribitol 2-dehydrogenase (KpRD) was identified in the genome of *K. pneumoniae* MGH 48, codon optimized for expression in *E. coli* (see **Figure S1** for amino acid sequence), chemically synthesized, and expressed in *E. coli*.

The purified KpRD indeed showed high specificity for the reduction of D-psicose under concomitant oxidation of NADH and very low activity towards the reduction of D-fructose, making it a suitable enzyme for the determination of D-psicose in our screening assay (**Figure S3A**).

In summary, we could establish a microtiter plate-based screening system for the reliable detection of D-psicose in presence of D-fructose by KpRD and for L-tagatose in presence of L-sorbose by RsGD. These UV-VIS assays allowed the rapid screening of over 6,000 clones from IDF libraries (improved for D-fructose) and over 5,000 clones from ILS libraries (improved for L-sorbose), a number that would be highly laborious to meet by HPLC-based screening.

Selecting sites for saturation mutagenesis of Var8

Based on the available crystal structure of PcDTE WT co-crystallized with D-fructose (PDB 2QUN[9d]) we selected the 22 amino acid residues that are located within 5Å of the C3-atom of D-fructose and categorized them as 1st-sphere residues, whereas the 28 amino acid residues located between 5Å and 10Å were categorized as 2nd-sphere residues (**Figure 2**). From the residues belonging to the 1st sphere we excluded the 5 residues that either make up the catalytic duo (E152 and E246), bind the metal ion Mn(II) (D185), or stabilize the cis-enediol transition state (H188 and R217)[9d] (**Figure 2** and **Figure S5**). To reduce the screening effort further we applied a "short-cut" version of the established ISM method[15]. Instead of screening each of the remaining 17 sites separately first and then re-diversifying each beneficial site once another mutation has been fixed, we fixed the first beneficial site that we encountered in the first round and then immediately used this variant to diversify the next site (**Figure S5**).

Evolving PcDTE towards improved D-fructose epimerization

Upon screening the first shell of residues, only one mutation (S37N) was found that improved conversion of D-fructose to D-psicose by 1.4-fold compared to Var8. This mutation constituted the first variant and was termed IDF1 (**Figure S2**).

A)

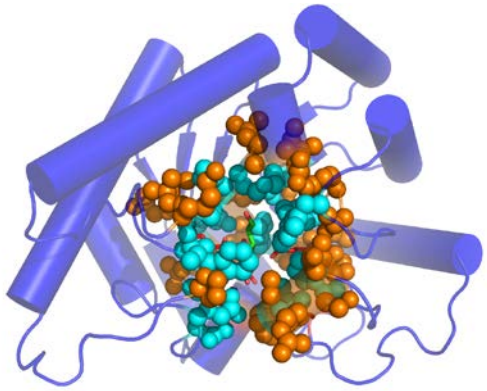


Figure 2. Residues targeted during divergent directed evolution are depicted in PcDTE WT (PDB 2QUN) with bound D-fructose (green sticks): residues that are located within 5Å of the C3 of D-fructose (1st sphere residues) are colored cyan, and residues that are located between 5Å and 10Å (2nd sphere) are colored in orange. Catalytic residues (E152 and E248) as well as highly conserved residues (D185, H188 and R217) that were excluded from screening are shown as red sticks.

When screening the second-shell residues of variant IDF1, mutation H209V was obtained in protein IDF2 that yielded a 1.6-fold higher activity compared to IDF1. Mutation G39E (IDF3) was found that increased activity 2.7-fold (IDF3), and in round 4 the silent mutation T9T (IDF4, codon ACC to ACT) that increased protein expression by 1.3-fold. In the next rounds the mutations A258D (IDF5, 1.6-fold improvement), T109N (IDF6, 1.4-fold improvement), L212I (IDF7, 1.3-fold improvement), and S256G (IDF8, 1.4-fold improvement, final variant) were fixed. The total improvement in catalysis from Var8 to IDF8 as determined from lysate was 14.4-fold, comprised of the 8.6-fold catalytic improvement (**Figure 3A**) and a (calculated) 1.7-fold improvement in protein expression yield.

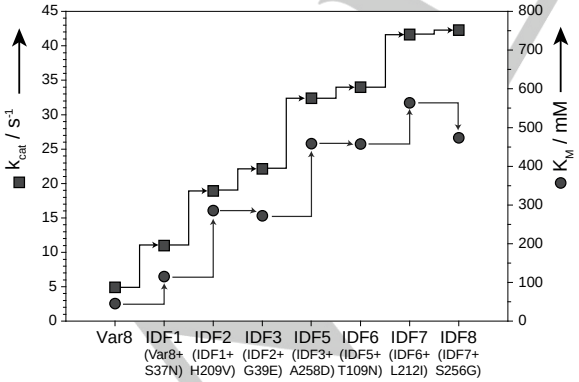
Next, all IDF variants were overexpressed, purified, and k_{cat} and K_m values were determined for each variant with the substrates D-fructose and D-psicose using an HPLC-based detection assay. For the starting template and the final variant IDF8, these values were also determined for the substrates D-tagatose, D-sorbose, L-sorbose and L-fructose (**Figure 3A**, **Figure 4A**, and **Table S4**). We observed a steady increase in k_{cat} from Var8 (4.9 s^{-1}) to IDF8 (42.3 s^{-1}) for the substrate D-fructose but also a concomitant increase in k_{cat} for substrate D-psicose from Var8 (2.8 s^{-1}) to IDF8 (57.1 s^{-1}). Therefore, the net catalytic efficiency (k_{cat}/K_m) remained nearly unchanged as the increase in k_{cat} was accompanied by a concurrent increase in K_m for both D-fructose and D-psicose.

For substrates D-tagatose and D-sorbose, which are supposed to be the original substrate pair of PcDTE[17] no net decrease in k_{cat} was observed for the IDF evolution trajectory (**Figure 4A**), however the catalytic efficiency (k_{cat}/K_m) for these two epimers was significantly reduced when going from Var8 to IDF8 (**Figure 4C**). For substrates L-sorbose and L-fructose net catalytic rate was increased moderately whereas catalytic efficiency remained largely unchanged (**Figure 4A and C**).

Evolving PcDTE towards improved L-sorbose epimerization

Screening the first shell for improved L-sorbose to L-tagatose conversion also yielded only one single mutation (Q183H) that was termed variant ILS1 (**Figure 3B**). This mutation, however, increased specific catalytic activity of L-sorbose to L-tagatose by 8.2-fold as determined from lysate (not corrected for protein expression). We started to screen the second shell residues for improved catalytic activity based on variant ILS1 and found mutation V153A (ILS2) that increased catalytic activity 1.4-fold. Further rounds of directed evolution resulted in fixation of mutation T9S (ILS3) that brought an improvement of 1.2-fold, mutation G39S (ILS4) that improved activity by 1.3-fold and mutation T109N (ILS5) that resulted in 1.3-fold higher activity compared to ILS4, too. The final mutation M245I made up variant ILS6 and improved activity by another 1.3-fold. In total, a 30-fold activity improvement determined from cell lysate was obtained compared to Var8. **Figure 5C** shows the accumulated mutations around the active site and their localization in the molecule.

A)



B)

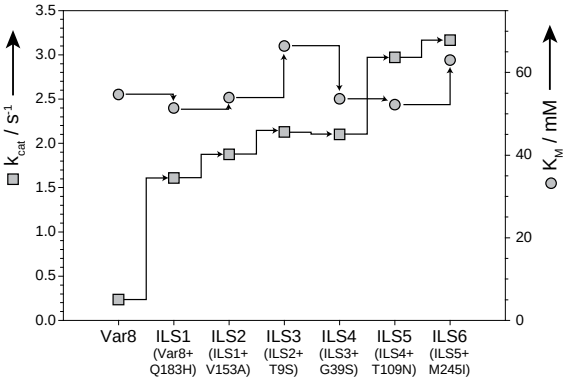


Figure 3. Development of enzyme kinetic parameters k_{cat} and K_m for each improved variant IDF and ILS, determined at 25°C. A) k_{cat} (dark grey square) and K_m values (dark grey circle) for Var8 and IDF variants with D-fructose as substrate. B) k_{cat} (light grey square) and K_m (light grey circle) values for Var8 and ILS variants with L-sorbose as substrate.

Enzyme kinetic parameters were determined from purified protein for each variant in the evolutionary trajectory for substrate L-sorbose and for ILS1 and ILS6 for substrates L-fructose, D-fructose, D-psicose, D-tagatose and D-sorbose. The most dramatic improvement in kinetic parameters was found with mutation Q183N (ILS1) and resulted in a 6.9-fold higher k_{cat} value, from a k_{cat} of 0.24 s^{-1} (Var8) to a k_{cat} of 1.61 s^{-1} (ILS1) (**Figure 3B**). Mutation G39S forming ILS4 resulted in no improvement in terms of k_{cat} which indicates that this mutation was found due to an increased expression level of soluble protein. Finally, ILS6 had a 13.5-fold improvement in terms of k_{cat} compared to the starting variant Var8 and hence a roughly two-fold higher level in soluble protein expression (see above). Interestingly, catalytic improvement for ILS6 was not accompanied by a loss in thermostability compared to starting variant Var8 (ΔT_{50}^{20} = +0.2°C), in contrast to IDF8 that lost considerable thermostability (ΔT_{50}^{20} = -11.8°C; **Figure S5**).

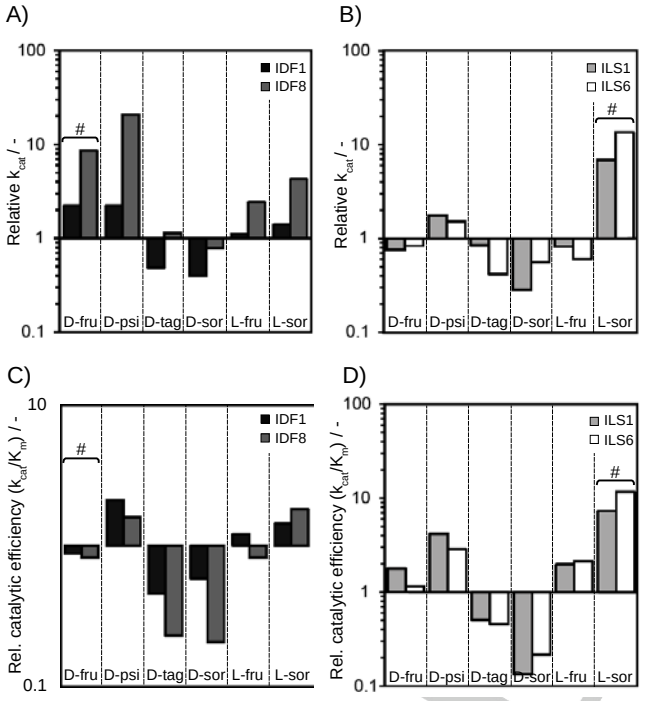


Figure 4. Development of k_{cat} and catalytic efficiencies (k_{cat}/K_m) for the first and the final variant discovered during the screening relative to the values of k_{cat} and k_{cat}/K_m of parent Var8 for the same substrate. The substrate for which the variants were evolved is marked with a hash (#). A) Relative k_{cat} for IDF1 (black bar) and IDF8 (dark grey bar) for 6 different substrates. B) Relative k_{cat} for ILS1 (light grey bar) and ILS6 (white bar). C) Relative catalytic efficiencies for IDF1 (black bar) and IDF8 (dark grey bar). D) Relative catalytic efficiencies for ILS1 (light grey bar) and ILS6 (white bar). Detailed kinetic parameters are listed in Table S4.

The mutation leading to ILS1 effectively reverses the preference of the enzyme for the C4-hydroxy group of the hexose substrate from the S- to the R-configuration and that of the C5-hydroxy group from R to S. This single mutation changes the catalytic efficiency between substrates D- (3R,4S,5R) and L-sorbose (3S,4R,5S) by 56-fold (**Figure 4D**). As a result, in the final variant ILS6, net catalytic rates as well as catalytic efficiency for the native substrates D-sorbose (3R,4S,5R) and D-tagatose (3S,4S,5R) were both reduced, whereas the k_{cat} for D-fructose (3S,4R,5R), D-psicose (3R,4R,5R) and L-fructose (3R,4S,5S) was unchanged but catalytic efficiency for D-psicose and L-fructose was moderately increased (**Figure 4B and D**).

Structural basis of the change in substrate specificity and catalytic activity

To investigate the structural changes in PcDTE upon mutagenesis, we determined the high resolution crystal structure of the starting variant Var8 in the substrate-free form, and of IDF8 and ILS6 in complex with their respective substrates. The structures were solved at resolutions of 1.8 Å (Var8), 2.1 Å (ILS6) and 1.9 Å (IDF8) (**Table 1**), giving precise insight into the detailed geometry of the active site and the interactions between PcDTE variants and their substrates.

Structural comparison between the wildtype (WT, PDB 2QUN) and the 3 variants (Var8, IDF8, and ILS6) showed very little difference in tertiary structure (**Figure S6**). Pairwise alignment of the dimeric structures yielded low root-mean square deviation (rmsd) values ranging from 0.43 Å (IDF8/ILS6) to 0.69 Å (Var8/IDF8).

In the IDF8 and ILS6 structures, the electron density indicated that the ketohexose ligands were present in both C3 epimeric forms. Apparently, the added substrates had effectively turned over by the enzymes in the crystals, reaching the thermodynamic equilibrium of the respective epimerization reaction.

Structural features of Var8

The structure of Var8 at 1.8 Å resolution shows that the active site geometry is virtually identical to the substrate-bound (PDB 2QUN) or substrate free form (PDB 2QUL) of the WT enzyme[9d] (**Figure S7**). Although Var8 was crystallized without substrate, the active sites showed the presence of ligands that could be accounted for as glycerol which we used for cryoprotection. Two of its hydroxyl groups coordinate to the bound manganese ion (**Figure 5B**), thus mimicking the interactions previously observed for D-fructose or D-tagatose bound to PcDTE[9d].

The Var8 structure confirmed our previous speculations on the stabilizing effect of F157Y, T194N and A215Q in that these mutants show additional H-bonds at the dimer interface[3b] (**Figure 5A**). Unexpectedly, a metal ion is bound to H116 (which is a serine in WT) and its symmetry mate (**Figure S8**). Next, mutation G260C appears to increase the complementarity of the dimer interface, thus strengthening the subunit-subunit interaction via van der Waals contacts[18]. The contributions of mutations K122V and K251T to the Var8 stability are likely due to the reduction of dimer interface entropy by substituting larger side chains with smaller ones. Finally, the stabilizing effect of the M265L mutation is difficult to rationalize based on the structure.

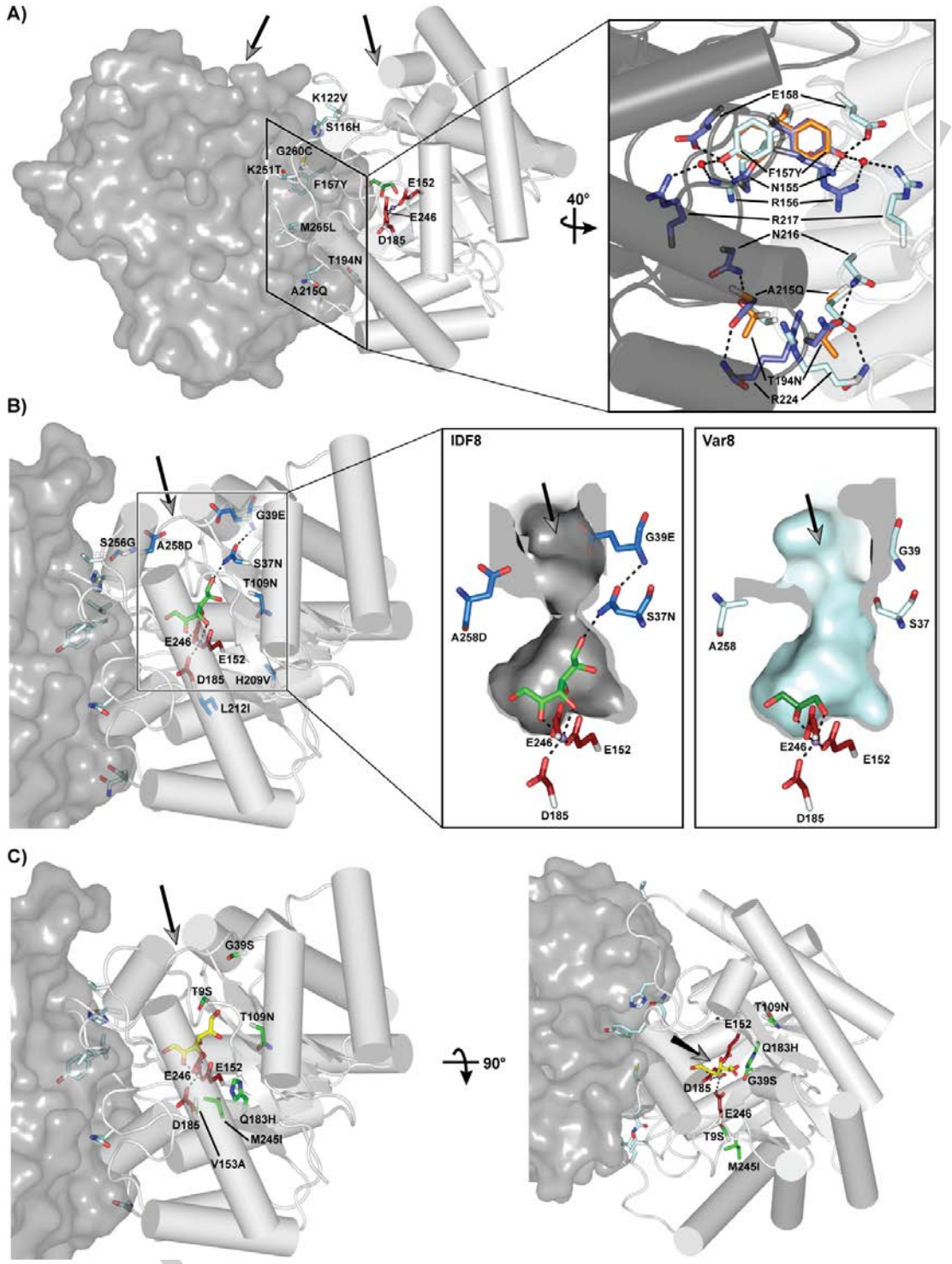


Figure 5. Structural details of PcDTE variants. **A)** Mutations of thermostable Var8 (PDB 4Q7I) at the dimeric interface are depicted in chain A, with inter-chain hydrogen bond interactions established by mutations F157Y and T194N shown in detail (right panel). The active site tunnel of each subunit is indicated by an arrow and the catalytic residues E152 and E246 as well as metal-coordinating residue D185 are depicted as red sticks. Glycerol in the active site is shown as green sticks and dark-blue sticks indicate residues from chain B, whereas light-cyan sticks indicate residues from chain A and orange sticks indicate WT residues (right panel). **B)** Activity-enhancing mutations around the active site of IDF8 (PDB 4PFH) are shown as blue sticks on chain B, chain A is shown in surface representation, mutations that derived from the precursor variant Var8 are depicted as light-cyan sticks and substrate D-fructose is shown as green sticks. The active site cavity of IDF8 with the three mutations S37N, G39E and A258D at the entrance are depicted and reveal a tightly constricted active site entrance compared to Var8 (far right). **C)** Mutations that accumulated during screening for activity on L-sorbose are depicted as green sticks in ILS6, chain C (PDB 4PGL) whereas chain D is shown in surface representation and mutations derived from Var8 are depicted as light-cyan sticks. Substrate L-sorbose is shown as yellow sticks and the entrance of the active site is indicated by arrows.

Structural features of IDF8

In the IDF8 structure, one of the 6xHis tags (residues H293 – H298) of a neighbouring dimer protrudes into the cleft next to the active site of chain A (**Figure S10**), forming hydrogen bonds to H116 and R257. Since these interactions are crystallographic artifacts, only chain B is considered in the following analysis.

Crystals of variant IDF8 had been soaked in D-fructose for 10 min prior to freezing in order to determine the structure in the substrate-bound state. The active site of IDF8 exhibits additional electron density consistent with the presence of both D-fructose and D-psicose epimers, which were modelled in the electron densities with occupancies of 0.7 and 0.3, respectively, according to the reported thermodynamic equilibrium of the epimerization reaction[13] (**Figure 6B**).

Hydroxyl groups 1 to 3 of both epimers are engaged in the same interactions. However, due to the distinct C3 chirality, C1 to C3 do not superimpose. This probably explains the discontinuity in the electron density at that part of the sugar. The electron density is also rather weak at the substrate C4–C6 moiety, which appears to be due to the lack of interactions with surrounding residues as observed in other PcDTE structures in their substrate-bound form[9d) 19].

The geometry of the IDF8 active site is highly similar to that of WT and Var8, which is illustrated by the nearly perfect superposition of catalytic residues E152 and E246 as well as the metal-coordinating residue D185 (**Figure S7**). The most prominent changes of IDF8 are due to mutation S37N, G39E, and A258D. The substitutions of small amino acid residues by larger ones constrict the active site entrance tunnel (**Figure** 5**B** and **Figure S9**). The constriction of the tunnel might shield the active site from bulk solvent and in this way lead to an increase in catalytic efficiency, as proposed for Kemp eliminase[20] and haloalkane dehalogenase[21]. Additionally, mutations G39E and A258D increase the charge at the tunnel opening.

Asparagine 37 of the S37N mutant (that constitutes IDF1) not only constricts the substrate channel but also forms a hydrogen bond via its amide group to the O6 of the bound substrate D-fructose. Thus, the mutation may have improved the correct positioning of the substrate in the active site (**Figure 6A**), since, in WT, the O6 group interacts only indirectly via an ordered water molecule with S37 (**Figure 6B**). Mutation G39E in IDF3 further extends this hydrogen bond network by interacting with N37 via an ordered water molecule and via its backbone amide (**Figure 5B**).

Mutation H209V that yielded IDF2 breaks a hydrogen bond network, extending from the manganese ion to the surface of the enzyme (**Figure 6A/B**). This leads to a slight repositioning of H211, the manganese ion, and of the ordered water that contacts O3 and O5 of substrate. The effects of the remaining mutations (T109N, L212I and S256G) are difficult to predict from the crystal structure.

Structural features of ILS6

Pairwise alignment with WT as well as with Var8 revealed a highly similar tertiary structure of ILS6 (**Figure S6 and Table S5**). The active site geometry, as judged from the superposition of catalytic residues E152 and E246 together with metal-binding residue D185, is very similar as well (**Figure S7**).

Crystals of ILS6 were soaked in L-sorbose for 3 min before they were subjected to flash freezing in order to obtain the structure in its substrate-bound state. As in the case of IDF8, the electron density indicated the presence of both C3 epimers. The L-sorbose and the L-tagatose epimers were modeled with an occupancy of 0.7 and 0.3, respectively (**Figure 6D**).

Mutation Q183H, first introduced in ILS1, resulted in the highest increase in k_{cat} for the of L-sorbose/L-tagatose epimerization reaction. When comparing the substrate binding modes of L-sorbose to ILS1 with that of D-fructose to the wild-type enzyme, distinct hydration patterns at the distal part of the saccharides are evident (**Figure 6C**). Whereas in the wild-type enzyme a water molecule is linking the sugar OH5 hydroxyl with the Q183 side-chain, this interaction is lost upon the Q183H mutation in ILS1. Whether this changed hydration pattern results in an improved positioning of L-sorbose (which is the C5 epimer of D-fructose) with respect to the catalytic center in ILS6 *versus* WT is hard to predict. The structural consequences of the other mutations all appear to be very minor as are their effects on k_{cat} and Km (**Figure 3B**) with the exception of mutation T109N that improves k_{cat} by a factor of about 1.5 (ILS5). This mutation affects a second shell residue that is completely buried. Replacement of the wild-type threonine side-chain by the longer asparagine side-chain is accommodated by a slight displacement of main-chain segment 106 to 108. This structural change is propagated towards segment 67 to 69 that forms part of the substrate entry channel. Probably, this small but significant change (0.3 Å) is not responsible alone for the altered enzyme characteristics. Surely the dynamic behavior of the aforementioned segments will be important for substrate entry and product release and may well be affected by the mutation at this buried site.

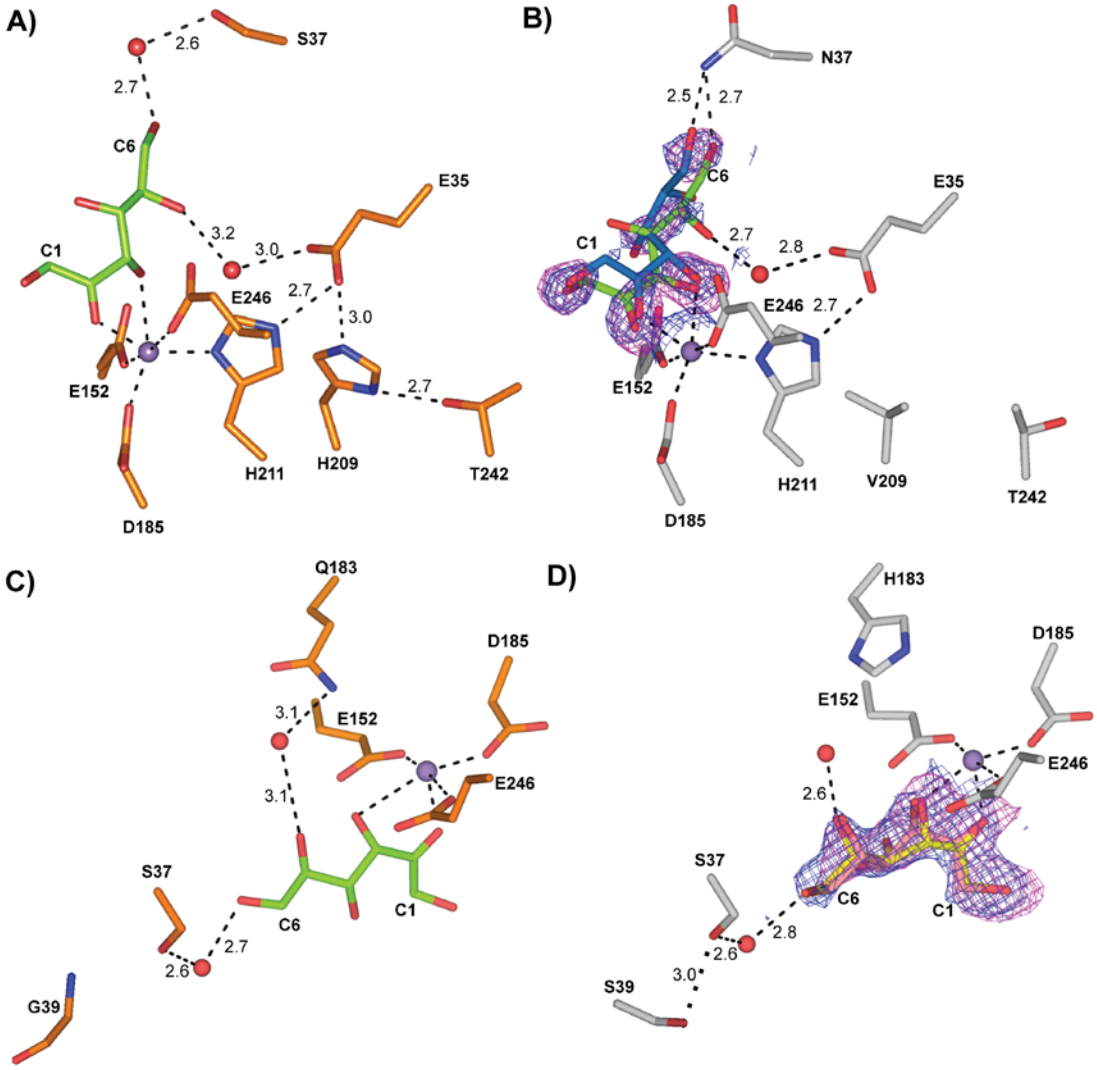


Figure 6. Comparisons of interactions between WT and improved variants. Interactions are shown as black dashed lines and distances indicated in Å. Manganese ions are displayed as violet spheres and water molecules as red spheres. $2F_o$-F_C and F_o-F_c omit maps are contoured at 1σ (blue mesh) and at 3σ (pink mesh), respectively. **A)** Interaction network in vicinity of WT active site (PDB 2QUN) with the substrate D-fructose represented as green sticks. **B)** Altered interaction network in IDF8 (PDB 4PFH) compared to WT (A) due to the mutation H209V that interrupts the interaction network between T242 and the substrate. D-fructose is shown as green sticks and C3-epimer D-psicose as blue sticks. **C)** Interaction network in vicinity of WT active site (PDB 2QUN) with D-fructose shown as green sticks. **D)** Changes in interactions in the active site of ILS6 (PDB 4PGL) compared to WT (C) as a result of the mutations G39S and Q183H. ILS6 substrates L-sorbose and L-tagatose are depicted as yellow and pink sticks, respectively.

pH-Rate profiles for WT, Var8, IDF8 and ILS6

To elucidate whether a change in pK_a of catalytically relevant residues might have had an impact on the increased catalytic activity during directed evolution for IDF8 or ILS6, we recorded pH-rate profiles for the four PcDTE variants. The epimerization reaction between two ketohexoses at the C3-position by D-tagatose epimerase is suggested to be catalyzed by two glutamate residues (E154 and E246) that are both in their deprotonated state and can act as catalytic bases. Mechanistically, the two glutamates have been suggested to proceed by abstracting a proton from the C3, forming a cis-enediolate intermediate, followed by re-protonation of the C3 by the other glutamate residue[9d]. Other ionizable residues contribute as well to the binding and positioning of the substrate, suggesting that pH might significantly affect the catalytic rate.

By fitting the standard Henderson-Hasselbach equation (Eq. 1, see methods for details) to the normalized activities, an apparent $pK_{a,app}$ of 6.1 +/- 0.05 could be determined for WT, whereas the apparent $pK_{a,app}$ of Var8 was significantly lower (4.5 +/- 0.02) and ILS6 showed a $pK_{a,app}$ of 5.25 +/- 0.06 (**Figure 7**). Surprisingly, the pH-rate curve of IDF8 could not be described adequately by a simple Henderson-Hasselbach equation,

suggesting cooperativity between several different ionizable residues within the active site of IDF8. Therefore a modified Hill-equation which accommodates for cooperativity was employed for fitting[22] (Equation 2). To assess which residues could elicit this effect we determined pK_a values of all ionizable groups of WT and IDF8 by means of the web-based prediction software H++[23]. Predictions were made with explicit incorporation of the manganese ion in the active site as it significantly influences the pK_a values of the adjacent residues. We found three residues which showed a significant difference in pK_a for WT and IDF8 and which are close to the active site, namely E35, E39 and H209 (**Table S6**).

Hence, two different explanations are conceivable for the cooperative pH-rate profile of IDF8. First, the mutation H209V disrupts a hydrogen-bond network between T242, H209, H211 and E35 (**Figure 6A/B**) and thus changes the predicted pK_a of E35 significantly (pK_a < 0 for WT, pK_a = 5.6 for IDF8). Protonation of E35 in IDF8 would supposedly severely disturb the non-covalent interactions with H211 and thus with the catalytic Mn(II) ion, leading to the break-up of these noncovalent interactions in a domino-effect and abolish the catalytic activity. Second, the protonation of E39 (pK_a = 5.9 for IDF8) could lead to a loss of the extensive hydrogen bond network at the entrance of the substrate channel in a cooperative manner, leading to an impaired entrance of the substrate.

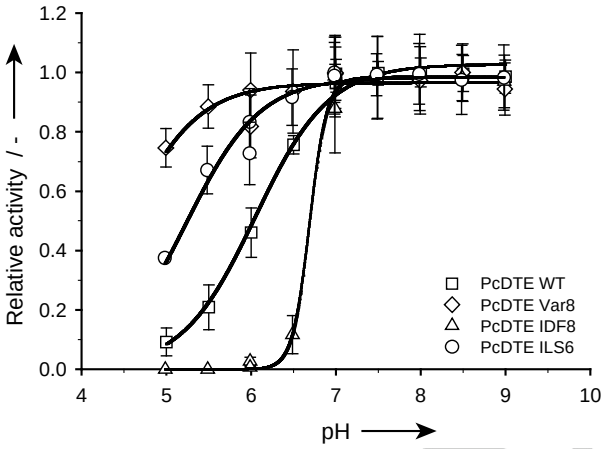


Figure 7. The pH-rate profiles of WT (squares), thermostabilized Var8 (diamonds), and catalytically improved variants IDF8 (triangles) and ILS6 (circles) are shown. Activities at the respective pH were determined with D-fructose (WT, Var8, IDF8) or L-sorbose (ILS6) as substrate and normalized to the maximum activity of the respective variant. Different pH's were set with actetate buffer (pH 5 – pH 6), phosphate buffer (pH 6 – pH 7) and Tris buffer (pH 7 – pH 9). Error bars indicate standard deviations from mean from triplicate measurements (Var8, ILS6, IDF8) and duplicate measurements (WT). Solid lines indicate the best fit of the experimental values to a pH-rate equation (see methods for details).

Conclusions

Screening of the thermostable Var8 towards catalytic efficiency for two different substrates revealed two interesting evolutionary trajectories for each branch. In both cases we were able to find only one single beneficial mutation that is directly adjacent to the active site (so called first-sphere residues), mutation S37N in IDF1 and Q183H for ILS6. IDF1 had only a moderate increase in k_{cat} for D-fructose whereas ILS1 showed a much more pronounced increase in catalytic activity as well as catalytic efficiency for L-sorbose. We suppose that mutation Q183H 'switches' the substrate preference from hexoses with a 4S,5R configuration to those with a 4R,5S configuration based on the catalytic efficiencies of the ILS1 mutant for 6 different substrates (**Figure 4**). Building on this initial improvement, catalytic activity was then further improved by 5 additional mutations, leading finally to ILS6 with an improvement in k_{cat} of over 13-fold compared to Var8 for L-sorbose. These results underline the importance of second-shell residues on catalytic activity. Tokuriki et al. showed for example that mutations that were randomly introduced by error-prone PCR (epPCR) into a phosphotriesterase to change it to an arylesterase were stepwise accumulated in a radial fashion from the center out to the periphery, with only 4 of the total of 18 mutations directly adjacent to the substrate binding site[24]. We observed a similar behavior for the evolution of ILS variants insofar as an initial big improvement in catalytic rate was followed by less pronounced improvements from more distant mutations in the second sphere.

In contrast to the evolution of the ILS variants of PcDTE, where a strong initial improvement was followed by minor incremental improvements, the IDF variants showed a smooth stepwise increase in k_{cat} over the whole trajectory (**Figure 3**). Second, the increase in catalytic activity (k_{cat}) for IDF variants was accompanied by a loss of affinity (K_m) for D-fructose during the evolutionary trajectory whereas the ILS branch showed no significant change in affinity to L-sorbose despite the increase in catalytic activity. It is difficult to explain this finding as the substrate concentration used for screening for both branches was the same (200 mM). Due to the increase in K_m for IDF variants no net increase in catalytic efficiency (k_{cat}/K_m) could be observed during the directed evolution that finally resulted in variant IDF8. This does however not impair the use of IDF8 as a promising catalyst for the production of D-psicose, as the substrate concentration that is applied in industrial settings and in particular in sugar-processing plants is usually high (> 1 M). Therefore the impact of K_m is often considered to be of subordinate interest in industrial biocatalysis[25].

As expected, the overall structural changes for IDF8 as well as for ILS6 from the WT are very small. Upon detailed inspection of the structures, the most prominent change for IDF8 is the accumulation of negatively charged residues at the entrance of the substrate channel (**Figure 5B**). It has been shown previously that residues lining the substrate channel or its mouth have a high probability of returning catalytic improvement[20-21]. The two mutations that encode negatively charged residues (G39E and A258D) lead to a conformational change of the two positively charged residues K70 and R257 at the entrance of the substrate tunnel. Together, this highly charged patch at the mouth of the entrance might facilitate the entry of the polar sugar substrate, although the entry tunnel of IDF8 is supposed to be also significantly constricted by these mutations compared to the WT enzyme.

Finally we conclude that our strategy of improving catalytic activity of enzymes by screening residues around the active site, going radially from the center to the periphery, is an efficient method for directed evolution of enzymes where screening is the limiting step, which is still the case for most industrially relevant biocatalysts[1a]. In this specific case, it delivered improvements in k_{cat} of a factor 8.6 for the conversion of D-fructose (IDF8) and of 13.5 for L-sorbose (ILS6). As perhaps expected, the phenotype of the variants cannot easily be correlated with the respective structures of the substrate complexes. A full understanding of the effects would require extensive molecular dynamics/quantum mechanics calculations and would be very sensitive to the proper estimation of the on- and off-rates through the tight access channel and the structure of the transition state. The structures of the optimized variants, however, can now immediately be utilized for the identification of new randomization sites for further enzyme improvement.

Experimental Section

If not stated otherwise all chemicals were purchased from Sigma Aldrich (Buchs, Switzerland). NADH was purchased from GERBU (Heidelberg, Germany), D-fructose, D-sorbose and L-fructose were purchased from Carbosynth (Berkshire, UK), D-tagatose and L-sorbose were obtained from Sigma Aldrich (Buchs, Switzerland). D-psicose was produced in-house by epimerization of D-fructose to D-psicose with D-tagatose epimerase and separation of D-psicose from D-fructose by continuous chromatography[26]. Restriction enzymes and polymerases were obtained from New England Biolabs (Ipswich, MA, USA) and oligonucleotides from Microsynth (Balgach, Switzerland).

Molecular Biology: General molecular biology was performed according to standard protocols[27]. All PCRs were generated using Phusion High-Fidelity Polymerase (NEB). Primers used for cloning are listed in Table S1 and all plasmids used and generated in this work are listed in Table S2. All general cloning work was done in *E. coli* Top10 cells (Invitrogen).

Cloning of KpRD from *K. pneumoniae* and RsGD from *R. sphaeroides*: The genetic sequence of ribitol dehydrogenase (KpRD) from *Klepsiella pneumoniae* was retrieved from the NCBI database (GenBank Nr. ESM54230) based on a report of Takeshita et al.[16]. The gene was codon-optimized for expression in *E. coli* and synthesized by Geneart (Regensburg, Germany). Plasmid pAB139 was constructed as follows (see Table S2 for details): the origin of replication pBR322 was amplified from pRK793 using primer pBR322_for-AscI and pBR322_rev-FseI and used to exchange the origin of replication of pSEVA131 via unique restriction sites AscI and FseI, resulting in plasmid pAB73. The tetR-Ptet-PT7 cassette was amplified from plasmid pKTS using primers Ptet-SEVA_for and Ptet-SEVA_rev and inserted into pAB73 via restriction sites SpeI and HindIII, resulting in construct pAB92. The bla-resistance gene of pAB92 was exchanged against the kamR resistance gene from plasmid pSEVA231 using unique restriction sites SwaI and FseI, giving vector pAB228. The *kpRD* gene was amplified from the vector pMA-T-KpRD using primers KpRD_N6H_f (introducing a 6xHis tag) and KpRD_s_rev and inserted into pAB92 using restriction enzymes HindIII and EcoRI, resulting in construct pAB138.
The galactitol dehydrogenase (RsGD)[28] was isolated from the genomic DNA of R. sphaeroides (DSM No. 8371). The gene was amplified directly from whole R. sphaeroides cells using primers RsGDH_NheI_for and RsGDH_s_rev and inserted via restriction sites NheI and EcoRI into pAB139, giving construct pAB140.

Expression and purification of KpRD and RsGD: *E. coli* BL21(DE3) cells were transformed with plasmid pAB139 or pAB140. Cells containing the respective plasmid were pre-cultured in 5 mL of Luria Bertani broth supplemented with 50 µg mL^{-1} kanamycin, then 200 µL of this preculture was used to inoculate 50 mL of M9-medium containing 0.4 % D-glucose and 50 µg mL^{-1} kanamycin, which after re-incubation served to inoculate a 1 L fed-batch reactor. Protein production was induced with 0.2 mM IPTG when an OD600 of 50 was reached and protein was synthesized for 6 h at 37°C. Cells were harvested by centrifugation (20 min at 6,000 rcf) and stored at -80°C until further use. To obtain purified RsGD, 10 g of wet cell pellet from the RsGD cultivation was resuspended in 15 mL^{-1} of RsGD-lysis buffer (50 mM Tris, pH 6.8; 100 mM NaCl, 20 mM imidazol, 0.2 mg mL^{-1} lysozyme) and incubated 20 min at room temperature before the cell suspension was frozen at -80°C for 20 min. The cell suspension was then thawed at room temperature, $MnCl_2$ to a final concentration of 1 mM and a spatula-tip of DNase was added and the suspension was sonicated for 10 min in an ultrasonication waterbath. Cell debris was removed by centrifugation (20 min at 48'384 rcf and 4°C) and the cleared lysate was loaded on 2 mL of Ni-Sepharose 6 Fast Flow (GE Healthcare) in a gravity-flow column. The column was extensively washed with lysis buffer before protein was eluted with elution buffer (50 mM Tris pH 6.8, 200 mM imidazol, 100 mM NaCl). The main fractions containing RsGD were analyzed by SDS-PAGE, pure fractions (> 95 %) were pooled and dialyzed twice against 1 L of buffer A (20 mM Tris, pH 6.8; 1 mM $MnCl_2$; 10% sucrose) before it was dialyzed once against 200 mL of buffer B (20 mM Tris, pH 6.8; 1 mM $MnCl_2$; 30% sucrose). The protein solution was sterile filtered through an 0.2 µm filter and stored at 4°C. KpRD was purified in the same way as RsGD, except for the following modifications: The wet pellet was resuspended in 15 mL KpRD-lysis buffer (50 mM Tris, pH 8.0; 100 mM NaCl, 20 mM imidazol, 0.2 mg mL^{-1} lysozyme) and the IMAC-purified KpRD was dialyzed twice against 2 L of buffer C (20 mM Tris, pH 8.0) before being aliquoted and stored at -80°C. Protein concentrations were determined spectrophotometrically at 280 nm wavelength for both RsGD (MW=27.5 kDa, ε =21.1 x 10^3 M^{-1} cm^{-1}) and KpRD (MW=27.4 kDa, ε=35.1 x 10^3 M^{-1} cm^{-1}).

D-Psicose quantification assay using KpRD: For the qualitative determination of conversion of D-psicose from D-fructose, a screening assay was developed based on the reduction of D-psicose by the enzyme KpRD and coenzyme NADH. The screening assay was performed in a two-step fashion. First, epimerization of D-fructose to D-psicose by a PcDTE enzyme variant was done for a certain time period (see below for details). Second, NADH and KpRD was added which converted D-psicose to allitol. The concomitant oxidation of NADH can be followed spectrophotometrically at 340 nm. PcDTE was not stopped before the KpRD was added for reasons of simplification, therefore this assay allows only for a qualitative comparison of PcDTE activity, which is however sufficient for screening purposes. As D-fructose is also to a small extent a substrate for KpRD, a calibration curve for the oxidation of different starting concentrations of D-psicose in presence of D-fructose was recorded by following the rate of NADH consumption at 340 nm in a Perkin Elmer Wallac 1420 Victor plate reader (Perkin Elmer, MA, USA). Six different calibration samples were prepared in 50 mM Tris buffer (pH 8.0) with the following concentrations (in mM) of D-fructose: 100; 99; 98; 95; 90; 80. The difference to a total hexose concentration of 100 mM was made up by D-psicose. Next, an aliquot of 200 µL of each calibration sample was supplied with 1 mM NADH (Gerbu Biotechnik GmbH, Heidelberg, Germany) and 25 µg of KpRD in a 96 well microplate (Greiner Bio-One, Germany), and the rates of NADH consumption were recorded at 30°C. This rate was a linear function of the D-psicose concentration at least in the range of 0 to 10 % (Figure S4A), and this

part was then used as calibration curve to determine the amount of D-psicose in the screening assay.

L-Tagatose quantification assay using RsGD: Detection of epimerization of L-tagatose from L-sorbose was done similarly as outlined above for detection of D-psicose. Epimerization of L-tagatose from L-sorbose was detected using RsGD that preferentially reduces L-tagatose to galactitol with concomitant oxidiation of NADH. A calibration curve for L-tagatose in presence of L-sorbose was generated in a similar fashion as for D-psicose and KpRD. In short, 200 µL of L-tagatose calibration sample in 50 mM Tris buffer (pH 8.0) was supplied with 1 mM NADH, 1 mM $MgCl_2$ and 4.8 µg of RsGD in a 96 well microplate and NADH decline rates were recorded at 30°C. The linear part of the slope was used as calibration curve (see Figure S4B).

Cloning of Var8 and library generation: The gene of thermostabilized D-tagatose epimerase from P. cichorii (Var8)[3b]was amplified using primers DTEci_HindIII_f and DTEci-ss_EcoRI_r and inserted into pAB92, giving plasmid pAB174 that served as template for mutant library generation. Saturation mutagenesis libraries on pAB174 were generated as described previously[3b]. Primers (listed in Table S7) with NNK-degenerated codons[29] were used to randomize selected sites according to the QuikChange protocol (Stratagene) using Phusion High-Fidelity Polymerase (NEB). The product was digested directly in the polymerase buffer with 10 U of DpnI for at least 2 h at 37°C in order to remove the template before 5 µL were used to transform 70 µL of chemo-competent *E. coli* Top10 cells.

Expression and screening of saturation-mutagenesis libraries of Var8: Clones from saturation mutagenesis libraries were expressed as described previously[3b]. For details please see the supporting information.

Expression and purification of improved IDF and ILS variants: Variants IDF1 – IDF8 and ILS1 – ILS6 were expressed using the T7 expression system, utilizing the Ptet-PT7 fusion promoter (see Table S2). Plasmids encoding these variants were isolated from *E. coli* Top10 cells and used for transformation of chemo-competent *E. coli* BL21 (DE3) cells. All variants were expressed in ZYM-5052 autoinduction medium[30]. A volume of 1 mL of overnight culture of the respective variant was used to inoculate 250 mL of ZYM-5052 medium supplemented with 100 µg mL^{-1} ampicillin in a 1 L Erlenmeyer flask and incubated for 16 h at 30°C and 220 rpm. Cells were harvested by centrifugation (6,238 rcf, 20 min, 4°C) and stored as a pellet at -20°C until further use. Cells were resuspended in 10 mL lysis buffer (50 mM Tris (pH 8.0), 0.2 mg mL^{-1} lysozyme) and incubated for 30 min at room temperature. The cells suspension was then lysed by one freeze/thaw cycle (20 min at -80°C, thaw at room temperature) before $MnCl_2$ was added to a final concentration of 1 mM and DNase (5 µL of a 5 mg mL^{-1} solution) was added to reduce viscosity. The cell lysate was heat-treated for 10 min at 70°C in a water bath before cell debris and denatured host proteins were removed by centrifugation (20 min at 48'384 rcf). Cleared lysate was applied to 2 mL of Ni-Sepharose 6 Fast Flow (GE Healthcare) in a gravity-flow column. The column was extensively washed with wash buffer (50 mM Tris (pH 8.0), 100 mM NaCl, 30 mM imidazole) before protein was eluted with elution buffer (50 mM Tris pH 8.0, 100 mM NaCl, 200 mM imidazole). Main fractions containing the PcDTE variants were pooled and dialyzed against buffer D (10 mM Tris (pH 8.0), 1 mM $MnCl_2$) and then 2 times against 10 mM Tris (pH 8.0). Dialyzed proteins were then aliquoted and stored at -80°C. All variants showed >95% protein purity as judged by SDS-PAGE. Enzyme concentration was determined spectrophotometrically (see Table S8 for extinction coefficients and molecular weights of all variants).

Enzyme kinetic measurements: Enzyme kinetic constants k_{cat} and K_m were determined from progress curves with 6 different ketohexose substrates (D-fructose, D-psicose, D-tagatose, D-sorbose, L-sorbose, L-fructose) epimerizing to the respective stereoisomer. In detail, 200 µL of a solution containing 50 mM sodium-phosphate buffer (pH 7.0) and different substrate concentrations (4.4 mM, 22 mM, 110 mM, 550 mM, 1.1 M, 2.2 M) was mixed with 20 µL of a solution containing different amounts of purified enzyme (76 µg to 184 µg) in a U-shaped 96-well plate (NUNC) on a BioShake iQ shaker (Q.Instruments, Jena, Germany) at 25 °C and 1'200 rpm. Reaction was stopped by adding 20 µL of the reaction mix to 145 µL of 0.1 M HCl, followed by the addition of 135 µL of 0.1 M NaOH after 5 min. Conversion of the substrate to the respective epimer was determined by HPLC (see above). Kinetic parameters K_m and k_{cat} were obtained by fitting initial velocities to the Michaelis-Menten kinetic model using SigmaPlot 12.2 (Systat Software Inc., CA, USA).

pH-Dependent activity profile: Initial catalytic rates were determined with D-fructose as substrate for WT, Var8 and IDF8 and L-sorbose as substrate for ILS6 at a final concentration of 90 mM at pH 5.0 - 9.0, initial rates for ILS6 were determined with 90 mM L-sorbose as substrate. 80 mM acetate buffer (pH 5.0 – 6.0), 80 mM phosphate buffer (pH 6.0 – 7.0) and 80 mM Tris buffer (pH 7.0 – 9.0) was used to set pH. Enzyme stocks were stored in 10 mM phosphate buffer, pH 7.0 supplemented with 1 mM $MnCl_2$. The reactions were performed at 25°C in a 96well plate as described above for the enzyme kinetic measurements. Reactions were stopped at 4 different time-points by adding 20 µL of a reaction mix to 145 µL of 0.1 M HCl. An aliquot of 135 µL of 0.1 M NaOH was added after 5 min and conversion to the respective product was determined by HPLC as decribed above. pH rate data were fitted for variants WT, Var8 and ILS6 using equation 1[31].

$$V_H = (V_{max} \cdot 10^{(-pK_{a,app})})/(10^{(-pH)}+10^{(-pK_{a,app})}) \quad \text{Equation 1}$$

with V_H as the pH-dependent reaction rate, V_{max} as the maximum reaction rate and $pK_{a,app}$ as the apparent pK_a value of the acidic groups. For IDF8 a modified Hill equation[22] was used for fitting:

$$V_H = (V_{max} \cdot 10^{(-pK_{a,app})})/(10^{(-pH \cdot n)}+10^{(-pK_{a,app} \cdot n)}) \quad \text{Equation 2}$$

with V_H as the pH-dependent reaction rate, n as the Hill coefficient, V_{max} as the maximum reaction rate and $pK_{a,app}$ as the apparent pK_a value of the acidic groups.

Protein production and purification for crystallization: The three variants Var8, IDF8 and ILS6 were expressed and purified in a first step as described above. For more details please refer to the supporting information.

Crystallization of Var8, IDF8 and ILS6: Crystals were obtained by the sitting-drop vapor diffusion method (IDF8 and ILS6) in Intelli-Plates 96 (Hampton Research, CA, USA), MRC2 crystallization plates (Molecular Dimensions, Suffolk, UK) or by the hanging-drop vapor diffusion method (Var8) in ComboPlates on siliconized cover slides (Jena Bioscience, Jena, Germany). All crystallization trials were performed at 20°C. Initial crystallization hits were obtained using commercial screens JBScreen Classic 5 (Jena Bioscience, Jena, Germany) for Var8, NeXtal PEGs-Suite (Quiagen, VA, USA) for IDF8 and PEGRx HT (Hampton Research, CA, USA) for ILS6. Crystallization conditions were optimized by varying the precipitant concentration and the pH around the initial hit conditions. The optimized crystallization conditions were as follows: 0.1 M Tris-HCl (pH 7.7), 18% (w/v) PEG 8K, 0.2 M $LiSO_4$ for Var8; 0.1 M Na-HEPES (pH 7.5), 25% (w/v) PEG 6K for IDF8; 0.1 M MES (pH 5.7), 10% (w/v) PEG 4K for ILS6. Crystals for Var8 were cryoprotected by dipping into mother

liquor complemented with 20% glycerol (v/v) and flash-frozen in liquid nitrogen. Crystals for IDF8 were soaked for 10 min in mother liquor complemented with 0.8 M D-fructose and crystals for ILS6 were soaked for 3 min in mother liquor complemented with 0.6 M L-sorbose. Soaked crystals were then flash-frozen without addition of any further cryoprotectant.

Table 1. Data collection and structure refinement statistics.

Crystal structure	PcDTE Var8 (PDB: 4Q7I)	PcDTE IDF8 (PDB: 4PFH)	PcDTE ILS6 (PDB: 4PGL)
Data collection			
Space group	C 1 2 1	P 1 2_1 1	P 1 2_1 1
a, b, c (Å)	110.5, 47.5, 124.7	57.5, 86.7, 61.8	102.8, 47.4, 126.4
α, β, γ (°)	90, 103.7, 90	90, 90.1, 90	90, 102.5, 90
Resolution (Å) *	53-1.80 (1.85-1.80)	56-1.90 (1.94-1.90)	36-2.10 (2.15-2.10)
Unique reflections	58426 (5789)	46594 (2988)	69460 (4447)
Multiplicity	3.1 (2.9)	5.4 (5.5)	6.2 (5.8)
R_{merge} (%)	5 (33.6)	11 (53.5)	9.9 (42.5)
I/σ(I)	11.3 (3.4)	9.6 (4.6)	12.1 (3.7)
Completeness (%)	99.5 (99.6)	99.2 (71.0)	99.1 (98.4)
Refinement			
R_{work} / R_{free} (%)	17.6 / 20.8	14.4 / 17.7	15.2 / 18.1
RMSD			
Bond lengths (Å)	0.013	0.013	0.011
Bond angles (°)	1.5	1.5	1.4
No. of atoms			
Protein	4697	4814	9384
Ligand	29	48	156
Metals	4	3	5
Water	460	420	381
Average B-factor (Å^2)			
Protein (main chain)	6.2	6.6	13.3
Protein (side chain)	8.2	9.7	17.5
Metals	23.1	15.7	22.3
Water	35.9	26.9	25.4
Other ligands	36.0	26.6	38.9
Ramachandran statistics (%)			
Favored regions	98.3	97.8	97.8
Allowed regions	1.6	2.2	2.2

* Highest resolution shell is shown in parenthesis.

Co-ordinates and structure factors have been deposited in the Protein Data Bank under the accession code 4Q7I (PcDTE Var8), 4PFH (PcDTE IDF8), and 4PGL (PcDTE ILS6).

Structural analysis: The substrate entry tunnel was analyzed by the PyMOL plugin Caver 3.0[32], potential hydrogen bonds and salt bridges were predicted by PyMOL or the PDBePISA webserver (http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver). All figures were prepared with PyMOL.

Keywords: rare monosaccharides • iterative saturation mutagenesis • D-tagatose epimerase • X-ray crystallography • enzyme engineering

Acknowledgements

The authors want to thank the beamline staff from the Swiss Light Source (SLS) for their support during data collection. A.B. is indebted to the Swiss National Science foundation for funding (grant 200021-121918).

[1] a) U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robins, Nature 2012, 485, 185-194; b) S. Luetz, L. Giver, J. Lalonde, Biotechnol. Bioeng. 2008, 101, 647-653.

[2] C. A. Tracewell, F. H. Arnold, Curr. Opin. Chem. Biol. 2009, 13, 3-9.

[3] a) Y. Gumulya, M. T. Reetz, ChemBioChem 2011, 12, 2502-2510; b) A. Bosshart, S. Panke, M. Bechtold, Angew. Chem., Int. Ed. 2013, 52, 9673-9676.

[4] M. T. Reetz, S. Prasad, J. D. Carballeira, Y. Gumulya, M. Bocola, J. Am. Chem. Soc. 2010, 132, 9144-9152.

[5] C. K. Savile, J. M. Janey, E. C. Mundorff, J. C. Moore, S. Tam, W. R. Jarvis, J. C. Colbeck, A. Krebber, F. J. Fleitz, J. Brands, P. N. Devine, G. W. Huisman, G. J. Hughes, Science 2010, 329, 305-309.

[6] S. Ito, Appl. Microbiol. Biotechnol. 2009, 84, 1053-1060.

[7] K. Beerens, T. Desmet, W. Soetaert, J. Ind. Microbiol. Biotechnol. 2012, 39, 823-834.

[8] T. B. Granstrom, G. Takata, M. Tokuda, K. Izumori, J. Biosci. Bioeng. 2004, 97, 89-94.

[9] a) M. Jia, W. M. Mu, F. F. Chu, X. M. Zhang, B. Jiang, L. L. Zhou, T. Zhang, Appl. Microbiol. Biotechnol. 2014, 98, 717-725; b) W. L. Zhang, D. Fang, Q. C. Xing, L. Zhou, B. Jiang, W. M. Mu, PLoS One 2013, 8; c) L. Zhang, W. Mu, B. Jiang, T. Zhang, Biotechnol. Lett. 2009, 31, 857-862; d) H. Yoshida, M. Yamada, T. Nishitani, G. Takada, K. Izumori, S. Kamitori, J. Mol. Biol. 2007, 374, 443-453.

[10] a) W. Besenmatter, P. Kast, D. Hilvert, Proteins: Struct., Funct., Bioinf. 2007, 66, 500-506; b) J. D. Bloom, S. T. Labthavikul, C. R. Otey, F. H. Arnold, Proc. Natl. Acad. Sci. U. S. A. 2006, 103, 5869-5874; c) J. D. Bloom, F. H. Arnold, Proc. Natl. Acad. Sci. U. S. A. 2009, 106, 9995-10000; d) N. Tokuriki, D. S. Tawfik, Curr. Opin. Struct. Biol. 2009, 19, 596-604.

[11] K. M. Polizzi, A. S. Bommarius, J. M. Broering, J. F. Chaparro-Riggers, Curr. Opin. Chem. Biol. 2007, 11, 220-225.

[12] N. Tokuriki, D. S. Tawfik, Science 2009, 324, 203-207.

[13] W. M. Mu, W. L. Zhang, Y. H. Feng, B. Jiang, L. Zhou, Appl. Microbiol. Biotechnol. 2012, 94, 1461-1467.

[14] a) Y. Yoshikuni, T. E. Ferrin, J. D. Keasling, Nature 2006, 440, 1078-1082; b) A. Aharoni, L. Gaidukov, O. Khersonsky, S. M. Gould, C. Roodveldt, D. S. Tawfik, Nat. Genet. 2005, 37, 73-76.

[15] M. T. Reetz, J. D. Carballeira, Nat. Protoc. 2007, 2, 891-903.

[16] K. Takeshita, Y. Ishida, G. Takada, K. Izumori, J. Biosci. Bioeng. 2000, 90, 545-548.

[17] Y. Ishida, T. Kamiya, H. Itoh, Y. Kimura, K. Izumori, J. Ferment. Bioeng. 1997, 83, 529-534.

[18] S. J. Hubbard, P. Argos, Protein Sci. 1994, 3, 2194-2206.

[19] K. Kim, H. J. Kim, D. K. Oh, S. S. Cha, S. Rhee, J. Mol. Biol. 2006, 361, 920-931.

[20] R. Blomberg, H. Kries, D. M. Pinkas, P. R. E. Mittl, M. G. Grutter, H. K. Privett, S. L. Mayo, D. Hilvert, Nature 2013, 503, 418-+.

[21] M. Pavlova, M. Klvana, Z. Prokop, R. Chaloupkova, P. Banas, M. Otyepka, R. C. Wade, M. Tsuda, Y. Nagata, J. Damborsky, Nat. Chem. Biol. 2009, 5, 727-733.

[22] J. L. Markley, Acc. Chem. Res. 1975, 8, 70-80.

[23] R. Anandakrishnan, B. Aguilar, A. V. Onufriev, Nucleic Acids Res. 2012, 40, W537-W541.

[24] N. Tokuriki, C. J. Jackson, L. Afriat-Jurnou, K. T. Wyganowski, R. Tang, D. S. Tawfik, Nat. Commun. 2012, 3, 1257.

[25] R. J. Fox, M. D. Clay, Trends Biotechnol. 2009, 27, 189-189.

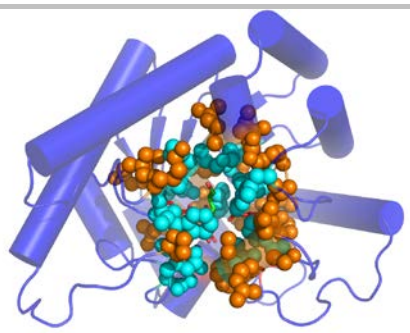
[26] N. Wagner, M. Fuereder, A. Bosshart, S. Panke, M. Bechtold, Org. Process Res. Dev. 2012, 16, 323-330.

[27] J. Sambrook, D. W. Russell, Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory Press, 2001.

[28] P. Kornberger, J. Gajdzik, H. Natter, G. Wenz, F. Giffhorn, G. W. Kohring, R. Hempelmann, Langmuir 2009, 25, 12380-12386.

[29] M. T. Reetz, D. Kahakeaw, R. Lohmer, ChemBioChem 2008, 9, 1797-1804.
[30] F. W. Studier, Protein Expression Purif. 2005, 41, 207-234.
[31] O. Khersonsky, D. Rothlisberger, O. Dym, S. Albeck, C. J. Jackson, D. Baker, D. S. Tawfik, J. Mol. Biol. 2010, 396, 1025-1042.
[32] J. Damborsky, J. Brezovsky, Curr. Opin. Chem. Biol. 2009, 13, 26-34.

Divergent evolution: Saturation mutagenesis of residues around the active site of a thermostable D-tagatose epimerase Var8 and screening for epimerization of two different hexose substrates resulted in two variants with much improved specific activities. Crystal structure analysis of the two final variants as well as of the parent Var8 revealed different mutation trajectories for the increase in activity.

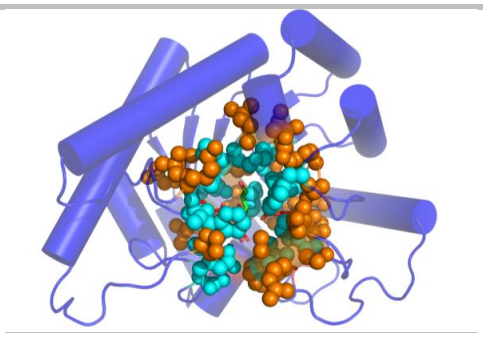


*A. Bosshart, C. S. Hee, M. Bechtold, T. Schirmer and S. Panke**

Page No. – Page No.

Directed divergent evolution of a thermostable D-tagatose epimerase towards improved activity for two different hexose substrates

Divergent evolution: Saturation mutagenesis of residues around the active site of a thermostable D-tagatose epimerase Var8 and screening for epimerization of two different hexose substrates resulted in two variants with much improved specific activities. Crystal structure analysis of the two final variants as well as of the parent Var8 revealed different mutation trajectories for the increase in activity.



*A. Bosshart, C. S. Hee, M. Bechtold, T. Schirmer and S. Panke**

Page No. – Page No.

Directed divergent evolution of a thermostable D-tagatose epimerase towards improved activity for two different hexose substrates

Supporting Information

Click here to download Supporting Information: SuppInfo_ChemBioChem_V5.docx