ACS Catalysis

pubs.acs.org/acscatalysis

Research Article

Structurally Informed Mutagenesis of a Stereochemically Promiscuous Aldolase Produces Mutants That Catalyze the Diastereoselective Syntheses of All Four Stereoisomers of 3-Deoxy-hexulosonic Acid

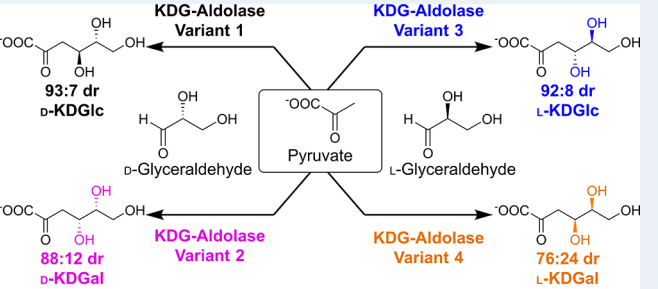
Sylvain F. Royer, Xuan Gao, Robin R. Groleau, Marc W. van der Kamp, Steven D. Bull, Michael J. Danson, and Susan J. Crennell*

Cite This: *ACS Catal.* 2022, 12, 11444−11455

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: A 2-keto-3-deoxygluconate aldolase from the hyperthermophile *Sulfolobus solfataricus* catalyzes the nonstereoselective aldol reaction of pyruvate and D-glyceraldehyde to produce 2-keto-3-deoxygluconate (D-KDGlc) and 2-keto-3-deoxy-D-galactonate (D-KDGal). Previous investigations into curing the stereochemical promiscuity of this hyperstable aldolase used high-resolution structures of the aldolase bound to D-KDGlc or D-KDGal to identify critical amino acids involved in substrate binding for mutation. This structure-guided approach enabled mutant variants to be created that could stereoselectively catalyze the aldol reaction of pyruvate and natural D-glyceraldehyde to selectively afford D-KDGlc or D-KDGal. Here we describe the creation of two further mutants of this *Sulfolobus* aldolase that can be used to catalyze aldol reactions between pyruvate and non-natural L-glyceraldehyde to enable the diastereoselective synthesis of L-KDGlc and L-KDGal. High-resolution crystal structures of all four variant aldolases have been determined (both unliganded and liganded), including Variant 1 with D-KDGlc, Variant 2 with pyruvate, Variant 3 with L-KDGlc, and Variant 4 with L-KDGal. These structures have enabled us to rationalize the observed changes in diastereoselectivities in these variant-catalyzed aldol reactions at a molecular level. Interestingly, the active site of Variant 4 was found to be sufficiently flexible to enable catalytically important amino acids to be replaced while still retaining sufficient enzymic activity to enable production of L-KDGal.



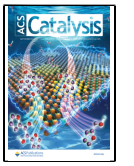
Wild-type KDG-aldolase catalysed aldol reactions of D-glyceraldehyde and L-glyceraldehyde produce ~50:50 mixtures of aldol diastereomers

KEYWORDS: *enzyme engineering, stereospecificity, aldolase, crystal structures, Sulfolobus solfataricus, carbon−carbon bond formation*

INTRODUCTION

Aldolases are a class of carbon−carbon bond-forming enzymes that catalyze the reversible aldol reaction between a carbonyl donor and an aldehyde acceptor to afford chiral aldol products containing up to two stereocentres.[1,2] They have been widely used as biocatalysts for the stereoselective transformation of protecting-group-free substrates into sugar derivatives, enzyme inhibitors, metabolites, heterocycles, and drug precursors, all under environmentally friendly aqueous conditions.[3−6] Aldolases are normally specific for their natural nucleophilic donors, which can significantly restrict the range of aldol products that can be produced.[7] However, many aldolases display greater promiscuity toward their electrophilic aldehyde substrates, which allows significant structural diversity to be incorporated into their aldol products.[8] Most aldolases produce aldol products as single stereoisomers containing stereocenters with a defined configuration, meaning that there is often no way of accessing stereoisomeric aldol products that are potentially desirable.[9] This has led to molecular biology and protein engineering techniques being developed to improve the specificity and/or stereoselectivity profiles of aldolases toward their donor and acceptor substrates.[10] Site-directed mutagenesis, gene shuffling, and directed evolution approaches have been used to evolve the reactivity profiles of aldolases toward non-natural donors and aldehyde electrophiles, with the aim of using mutant biocatalysts to produce chiral aldol products from non-native substrates.[11] Similar success has also been achieved in improving the stereoselectivities of aldolase-catalyzed reactions using mutant aldolases that exhibit improved or inverted enantioselectivities/diastereoselectivities in their aldol

Received: July 8, 2022
Revised: July 28, 2022
Published: September 6, 2022



reactions.[12−14] However, many of these approaches require expensive and time-consuming screening of large libraries of mutant enzymes to achieve the desired changes in their reactivity profiles. Alternative approaches to modify specificity profiles have also been explored, based on structurally informed site-directed mutagenesis of key amino acid residues in the aldolase active site.[15] This approach requires an intimate knowledge of the catalytic mechanism of the aldolase and an understanding of how donor/acceptor substrates (or aldol products) bind to the aldolase active site. This structural information can then be used to target key amino acid "hot spots" for replacement to produce mutants with new or improved selectivity profiles.

Pyruvate (and other 2-oxo acid) aldolases are an attractive class of carbon−carbon bond-forming enzymes that catalyze the aldol reaction of inexpensive pyruvate with a range of aldehydes to afford synthetically versatile chiral γ-hydroxy-α-keto acids that contain four contiguous carbon atoms at different oxidation states.[16] Screening of aldolase collections has identified aldolases with complementary stereoselectivities for the same aldol reaction, thus providing biocatalysts for the stereodivergent syntheses of enantiomeric aldol products.[17] Extensive protein engineering studies have been carried out to widen the substrate specificity profiles of pyruvate aldolases, with the aim of improving/inverting their stereoselectivities toward their aldehyde substrates.[16f,18] An alternative approach to inverting the stereochemical preference of a stereoselective aldolase is to mutate an aldolase that catalyzes aldol reactions with poor levels of stereocontrol to produce variants that produce stereoisomeric aldol products with better levels of stereocontrol. This stereodivergent approach is a potentially efficient way of generating complementary mutant aldolase activities because the wild-type enzyme already displays good catalytic activity for production of each aldol stereoisomer. The challenge in this approach is to increase the catalytic activity of the aldolase for formation of the desired aldol stereoisomer and/or decrease the rate of formation of the unwanted aldol stereoisomer(s).

Sulfolobus solfataricus is a hyperthermophilic archaeon that grows optimally at 80 °C and metabolizes sugars via a nonphosphorylative variant of the Entner−Doudoroff pathway, with a series of stereochemically promiscuous enzymes used to metabolize the four most-common sugars in Nature, namely, D-glucose, D-galactose, D-ribose, and L-arabinose.[19,20] The aldolase from this pathway [SsKDG-aldolase][21] catalyzes retro-aldol cleavage of both 2-keto-3-deoxy-D-gluconate and 2-keto-3-deoxy-D-galactonate into pyruvate and glyceraldehyde, as well as catalyzing cleavage of 2-keto-3-deoxy-D-xylonate and 2-keto-3-deoxy-L-arabinonate to pyruvate and glycolaldehyde.[22] This aldolase is thermally stable, accepts a wide range of non-natural aldehyde substrates (e.g., the enantiomers of glyceraldehyde, threose, and erythrose), is tolerant to a range of cosolvents, and its gene has been efficiently expressed in *Escherichia coli*.[23] However, its aldol reactions often proceed with poor levels of diastereocontrol, with wild-type SsKDG-aldolase catalyzing the aldol reaction of pyruvate and its natural D-glyceraldehyde substrate to afford a mixture of 2-keto-3-deoxy-D-gluconate (D-KDGlc) and 2-keto-3-deoxy-D-galactonate (D-KDGal) in a 55:45 diastereomeric ratio (dr) (Figure 1a).[24] Conversely, the broad specificity profile of this aldolase means that it can also be used to catalyze aldol reaction of pyruvate with non-natural L-glyceraldehyde, affording a complementary 53:47 mixture of 2-keto-3-deoxy-L-gluconate (L-KDGlc) and 2-keto-3-deoxy-L-galactonate (L-KDGal) (Figure 1b).[24]

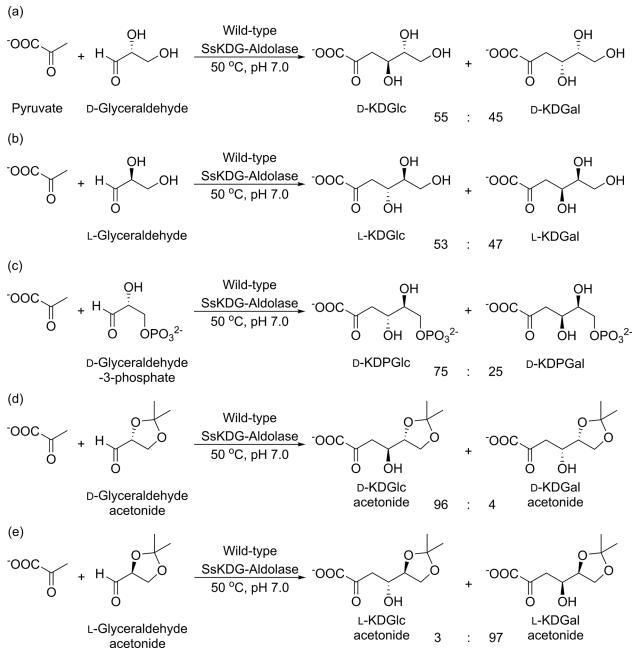


Figure 1. Wild-type SsKDG-aldolase catalyzed reactions of pyruvate with (a) D-glyceraldehyde; (b) L-glyceraldehyde; (c) D-glyceraldehyde-3-phosphate; (d) D-glyceraldehyde acetonide; and (e) L-glyceraldehyde acetonide. All aldol products are shown in their open-chain forms for clarity.

Although demonstrating good catalytic activity toward a range of aldehyde substrates, the lack of stereocontrol in these SsKDG-aldolase-catalyzed reactions is potentially problematic from a biocatalytic perspective because these aldol reactions produce mixtures of water-soluble diastereomeric aldol products that are difficult to separate. However, access to X-ray crystal structures of the wild-type aldolase bound to substrates and aldol products[25] provided us with the opportunity to develop strategies to "cure" the aldolase's inherent stereochemical promiscuity toward its natural D-glyceraldehyde substrate. Initial attempts to improve diastereocontrol in these aldolase reactions used a substrate-based strategy inspired by Nature, where charged phosphorylated substrates are often used to control binding affinities and selectivities in enzyme-catalyzed reactions. Therefore, it was found that the SsKDG-aldolase catalyzes the reaction of D-glyceraldehyde-3-phosphate with pyruvate[26,27] to afford a mixture of D-KDPGlc and D-KDPGal in an improved 3:1 dr (Figure 1c) (S. F. Royer, unpublished data). Subsequent evaluation of aldolase crystal structures containing D-KDGlc and D-KDGal led us to consider using a more rigid cyclic D-glyceraldehyde acetonide substrate, which resulted in a highly diastereoselective aldol reaction that produced D-KDGlc and D-KDGal in 96:4 dr (Figure 1d).[24] Interestingly, use of L-glyceraldehyde acetonide as a substrate resulted in an aldol reaction with opposing diastereoselectivity, affording a mixture of L-KDGlc and L-KDGal in 3:97 dr, thus indicating that the facial selectivity of both these aldol reactions occurs under enzymatic control (Figure 1e).[24,28]

We then employed X-ray crystal structures of the wild-type SsKDG-aldolase bound to D-KDGlc and D-KDGal (Figure 2) to identify a series of active-site amino acids for mutation that enabled stereocomplementary mutants to be identified, which were then used to selectively transform natural D-glyceraldehyde into D-KDGlc and D-KDGal, respectively. Therefore, a double

https://doi.org/10.1021/acscatal.2c03285
ACS Catal. 2022, 12, 11444−11455

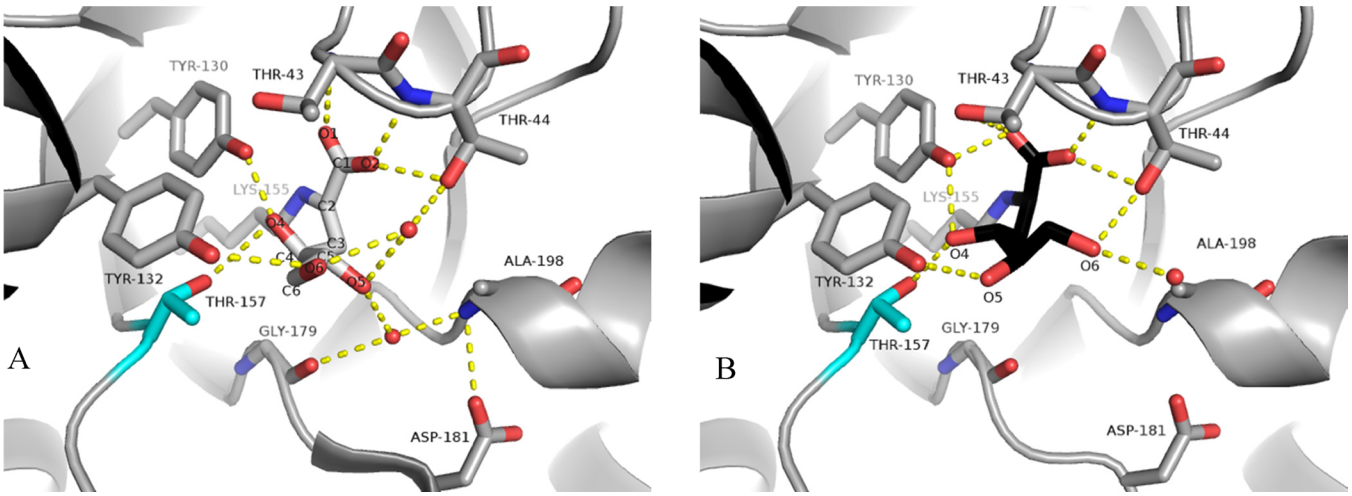


Figure 2. Active site of wild-type SsKDG-aldolase in complex with (A) D-KDGlc (PDB code: 1W3N, white bonds); and (B) D-KDGal (PDB code: 1W3T, black bonds) bound in the active site. The side chains of amino acids that interact with both aldol products are shown with hydrogen bonds indicated as yellow dotted lines. Bridging water molecules involved in key hydrogen bonding bridging are shown as red spheres.

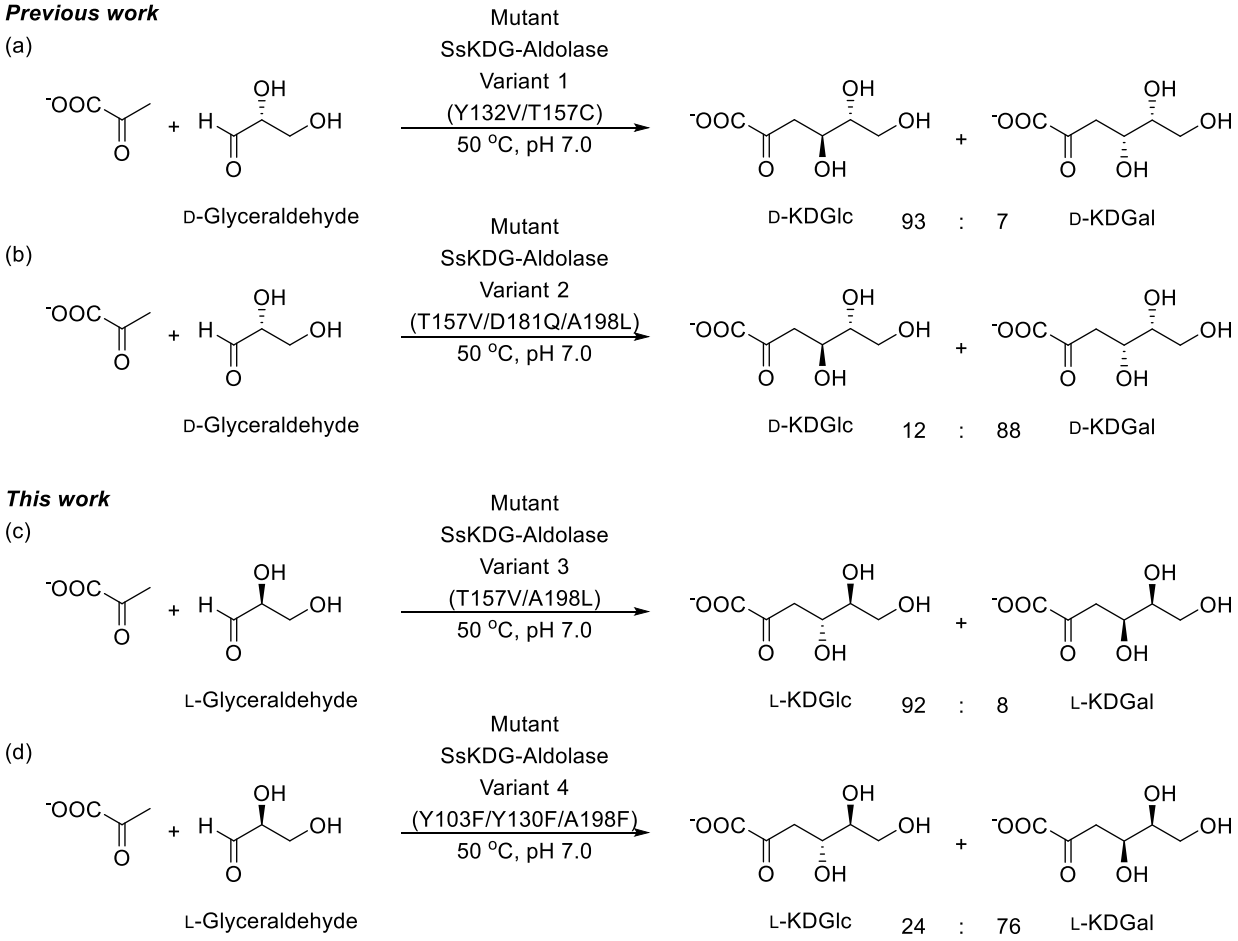


Figure 3. (a) Mutant SsKDG-aldolase Variant 1 for the stereoselective synthesis of D-KDGlc. (b) Mutant SsKDG-aldolase Variant 2 for the stereoselective synthesis of D-KDGal. (c) Mutant SsKDG-aldolase Variant 3 for the stereoselective synthesis of L-KDGlc. (d) Mutant SsKDG-aldolase Variant 4 for the stereoselective synthesis of L-KDGal. All aldol products are drawn in their open-chain forms for clarity.

mutant (Y132V/T157C; Variant 1) was identified that increased the proportion of D-KDGlc (over D-KDGal) from 55:45 to 93:7 dr (Figure 3a), while a triple mutant (T157V/D181Q/A198L; Variant 2) was created that increased the proportion of D-KDGal (over D-KDGlc) from 45:55 to 88:12 dr (Figure 3b).[29]

In this current study, we now report X-ray crystal structures of Variants 1 and 2 and Variant 1 bound to D-KDGlc that have enabled us to rationalize changes in the diastereoselectivity

https://doi.org/10.1021/acscatal.2c03285
ACS Catal. 2022, 12, 11444−11455

profiles of the aldol reactions of both these mutants at a molecular level. Furthermore, we also now report construction of two new SsKDG-aldolase variants that can be used to selectively transform non-natural L-glyceraldehyde into L-KDGlc or L-KDGal with good levels of diastereocontrol. This was achieved using structurally guided mutation approaches to create a double mutant (T157V/A198L; Variant 3) that increases the proportion of L-KDGlc produced from 53:47 to 92:8 dr (Figure 3c), while creation of a triple mutant (Y103F/Y130F/A198F; Variant 4) increased the amount of L-KDGal epimer produced from 43:57 to 76:24 dr (Figure 3d). Once again, these improvements in diastereocontrol have been rationalized at a molecular level by examining changes in the X-ray crystal structures of these new mutants complexed to L-KDGlc (Variant 3) and L-KDGal (Variant 4) in comparison with wild-type SsKDG-aldolase structures.

METHODS AND MATERIALS

Bacterial Strains and Plasmids. *Escherichia coli* BL21 (DE3) competent cells were obtained from Stratagene, U.K., and used in conjunction with the pET-3a expression vector as described previously.[23]

Creation and Purification of Site-Directed Mutants. Creation of Variant 1 (Y132V/T157C) and Variant 2 (T157V/D181Q/A198L) by site-directed mutagenesis has been reported previously.[29] In the current study, two further mutant aldolases [Variant 3 (T157V/A198L) and Variant 4 (Y103F/Y130F/A198F)] were similarly created using a QuikChange II PCR kit obtained from Stratagene. Plasmids carrying the mutant genes were transformed into and expressed in *E. coli* BL21 (DE3), with the variant aldolases then purified from cell extracts[29] using a combination of heat treatment, anion exchange chromatography on a HiTrap Q HP column from GE Healthcare, and gel filtration on a Superdex 200 10/300GL (GE Healthcare).

Enzyme Assay. Kinetic analyses of the new variant aldolases (Variants 3 and 4) were carried out using a modification of the thiobarbituric acid assay.[21] Assays were carried out at 70 °C in 50 mM sodium phosphate buffer, pH 6.0, using various concentrations of pyruvate and D- and L-glyceraldehyde. Aldol reactions were started by the addition of enzyme and stopped at known time intervals by the addition of 12% (w/v) trichloroacetic acid, with the presence of any 3-deoxy-hexulosonic acid products then detected using the thiobarbituric acid assay.

Determination of the Diastereoselectivities of Aldolase-Catalyzed Reactions. Biotransformation reactions between 600 mM pyruvate and 100 mM either D-glyceraldehyde or L-glyceraldehyde were carried out at 50 °C in water using the four variant aldolases (10 μg). The diastereomeric ratios of the aldol products produced in these aldol reactions were then determined by high-pressure liquid chromatography (HPLC) analysis using an Agilent 1200 machine fitted with a BioRad Aminex HPX-87H column (300 mm × 7.8 mm) running at 0.4 mL/min in 8 mM H_2SO_4 at 60 °C, with peak detection carried out using a refractive index detector.

Biotransformations. The variant aldolases 1, 2, 3, and 4 were used to prepare D-KDGlc, D-KDGal, L-KDGlc, and L-KDGal, respectively, using our previously published methods.[23,29] Enantiopure D- or L-glyceraldehyde (150 mg, 1.67 mmol) and sodium pyruvate (1.00 g, 9.00 mmol) were dissolved in 100 mL of water containing 10 mg of the appropriate aldolase variant. The reaction was heated to 50 °C in a shaking incubator. When consumption of glyceraldehyde was sufficient (monitored by HPLC, see the conditions above), the reactions were lyophilized to produce 0.7−1.0 g of crude product. Portions of the crude product (100 mg) were then purified by semipreparative HPLC using a Bio-Rad Aminex HPX-87H column to remove excess pyruvate, with yields of up to 20 mg (per portion) of each purified 3-deoxy-hexulosonic acid obtained in >95:5 dr. Alternatively, purification and separation of diastereomers could be carried out by anion exchange chromatography (0−0.6 M aqueous formic acid, Dowex 1X8-formate resin). See the Supporting Information for representative characterization data.

Crystallization and X-ray Data Collection and Processing. Variants were crystallized by the hanging-drop vapor diffusion method at 18 °C. Substrate complexes were obtained by soaking crystals in 30 mM solutions of the 3-deoxy-hexulosonic acid. Variant 2 was cocrystallized with pyruvate by adding 100 mM pyruvate to the crystallization conditions. Glycerol (15−20%, v/v) was added to the crystals as a cryoprotectant prior to data collection. X-ray diffraction data were collected at the Diamond Light Source (experiments mx1226-10, mx1226-11, and mx1226-23) or in-house using Cu Kα radiation from a Rigaku MicroMax 007HF generator using a Saturn 944+ detector. Data were processed initially with DIALS, but if this was unsuccessful, they were then processed with MOSFLM, or if necessary with HKL2000 as a final choice, to obtain data with the lowest R_{merge} and the highest completeness. The wild-type aldolase structure (1W37) was used to solve the variant structures using BALBES,[30] with the models subsequently refined using Refmac[31] or Phenix.[32]

RESULTS AND DISCUSSION

Characterization of SsKDG-Aldolase and Its Variants. The approach used to construct Variant 1 (Y132V/T157C) and Variant 2 (T157V/D181Q/A198L) aldolases that exhibit selectivity for formation of D-KDGlc and D-KDGal, respectively, has been described previously.[29] Briefly, X-ray crystallographic structures of the wild-type aldolase bound to D-KDGlc and D-KDGal were used to identify specific active site amino acids responsible for producing stabilizing interactions to each of the aldol diasteroisomers (Figure 2). These residues were then targeted for mutation with the aim of disrupting stabilizing active site binding interactions that were responsible for binding D-KDGlc (or D-KDGal). This approach resulted in rapid identification of a Variant 1 mutant that favored formation of D-KDGlc in 93:7 dr and a Variant 2 mutant that produced D-KGal in 88:12 *dr*. A similar strategy (outlined in detail below) has been used in this study to create Variant 3 (T157V/A198L) and Variant 4 (Y103F/Y130F/A198F) mutants that exhibit complementary diastereoselectivity profiles for formation of non-natural L-KDGlc and L-KDGal, respectively. The kinetic parameters of the wild-type and variant aldolases used in this study were determined using a modified thiobarbituric acid assay, as described in the Methods and Materials section. Each variant was assayed with each enantiomer of glyceraldehyde, with the dr of their aldol products then determined by HPLC analysis (HPLC traces are shown in Figure S1).

The enzyme kinetic and dr data obtained for the variants are summarized in Table 1. It should be noted that the K_M values are operational values; that is, they are the substrate concentrations giving half V_{max} in the presence of near-saturating (≥90%) concentrations of the second substrate. The precise constants comprising these K_M values will depend on the mechanism of the catalyzed reaction (random rapid equilibrium or compulsory order).

https://doi.org/10.1021/acscatal.2c03285
ACS Catal. 2022, 12, 11444−11455

Table 1. Kinetic Parameters of SsKDG-Aldolase and Its Variants[a]

	K_M (mM)	V_{max} (μmol/min/mg)	dr of aldol products KDGlc:KDGal
wild-type			
pyruvate	1.1 (±0.1)	8.3 (±0.1)	55:45
D-glyceraldehyde	5.3 (±0.3)	8.9 (±0.3)	
pyruvate	0.5 (±0.1)	5.4 (±0.1)	53:47
L-glyceraldehyde	3.8 (±0.1)	5.6 (±0.04)	
Variant 1			
pyruvate	4.1	0.13	93:7
D-glyceraldehyde	8.0	0.11	
Variant 2			
pyruvate	39.6	0.12	12:88
D-glyceraldehyde	9.1	0.13	
Variant 3			
pyruvate	6.0 (±0.2)	5.7 (±0.1)	92:8
L-glyceraldehyde	37.6 (±0.2)	6.1 (±0.1)	
pyruvate	2.9 (±0.1)	1.7 (±0.02)	28:72
D-glyceraldehyde	19.5 (±0.2)	1.8(±0.01)	
Variant 4			
pyruvate	1.1 (±0.15)	0.06 (±0.01)	24:76
L-glyceraldehyde	6.3 (±0.18)	0.06 (±0.01)	
pyruvate	0.3 (±0.04)	0.09 (±0.01)	75:25
D-glyceraldehyde	6.9 (±0.2)	0.11 (±0.01)	

[a]Data for the wild-type enzyme and for Variants 1 and 2 are taken from Royer et al.[29] and are included for comparison with the data generated in the current paper for Variants 3 and 4. The errors given are for two repeats of each experiment (n = 3).

The synthetic utility of the SsKDG-aldolase variants as biocatalysts has been demonstrated by using them to produce their respective aldol products for use as docking substrates to carry out aldolase-ligand crystallization studies.

Crystallization of SsKDG-Aldolase Variants. Crystals of the various SsKDG-aldolase variants were obtained using 11−15% (v/v) PEG4K, 4−12% (v/v) isopropanol, and 0.1 M HEPES, pH 5.9−6.8. Data were collected for each of the four uncomplexed variants, with subsequent data sets then collected for each variant following their soaking in a 30 mM concentration of their preferred 3-deoxy-hexulosonic acid ligand (major aldol product), with soaking times ranging from 1 min to 24 h. Unfortunately, crystals of a complex of Variant 2 docked with D-KDGal could not be obtained, even after soaking was carried out for 24 h, although its cocrystallization with pyruvate was successful. Data collection and processing statistics for Variants 1−4 are presented in Table S1.

Structural Analysis of the Wild-Type SsKDG-Aldolase Bound to D-KDGlc and D-KDGal. The previously reported X-ray crystal structure of the wild-type SsKDG aldolase−D-KDGlc complex (PDB code 1W3N) revealed that the 4-OH group of D-KDGlc forms hydrogen bonds with the phenol and hydroxyl groups of the Y130 and T157 residues, with its 6-OH group interacting with the phenol group of Y132 (Figure 2A). The crystal structure of the wild-type enzyme complexed with D-KDGal (PDB code 1W3T) revealed that D-KDGal makes strong hydrogen bonds between the 4-OH and the T157 hydroxyl group, with its 5-OH group interacting with the Y132 phenol group and its 6-OH hydroxyl group interacting with the T44 hydroxyl group (Figure 2B). With this structural information in hand, it was proposed that the D-KDGlc selectivity observed for Variant 1 (Y132V/T157C) was due to elimination of stabilizing contributions from Y132 and T157 in the aldol transition state that favors D-KDGal formation in the wild-type enzyme.

Similar arguments were used to explain the improved D-KDGal selectivity of Variant 2 (T157V/D181Q/A198L), which was suggested to be due to disruption of stabilizing hydrogen bond interactions as well as the introduction of a hydrophobic region in the binding site that disfavors the transition state leading to D-KDGlc formation in the wild type. Consequently, we decided to seek further structural evidence for these mechanistic explanations by acquiring the crystal structures of Variant 1 complexed to D-KDGlc and Variant 2 complexed to D-KDGal, whose structures would then be compared to those of the wild-type aldolase bound to D-KDGlc and D-KDGal.

Structural Analysis of the D-KDGlc Selective Variant 1 (Y132V/T157C). As previously reported,[29] the effect of introducing the Y132V and T157C mutations into Variant 1 is to increase selectivity for D-KDGlc production from 55:45 to 93:7 dr (Table 1). The overall structure of this variant was found to be very similar to the wild-type enzyme (0.3 Å RMSD over all 293 Cα positions for Variant 1 and 0.4 Å for the Variant 1−D-KDGlc complex obtained at a lower resolution). This indicates that the mutations in Variant 1 do not cause major structural disruption, with the only significant movements (more than 3× RMSD) occurring at the mutation sites. Comparison of the unsoaked structure (2.0 Å resolution) of Variant 1 with the wild-type aldolase revealed that the only major side-chain movement was of Y130, which moves 1.5 Å away from K155. This is potentially significant because the T157C mutation of Variant 1 means that Y130 is the only remaining amino acid capable of forming stabilizing interactions with the 4-OH groups of both D-KDGlc and D-KDGal. An unexpected consequence of the mutations in Variant 1 was the formation of a strong interaction between the Nζ amino group of K155 (the Schiff-base forming lysine residue) to the introduced Sγ thiol group of C157 (Figure S2). This potentially strong interaction needs to be disrupted for catalysis to occur, which, combined with displacement of Y130 (that also plays a role in the catalytic mechanism),[25,33] potentially explains the lower V_{max} value of this variant (Table 1). The presence of this K155−C157 interaction may also explain the difficulties experienced in obtaining a complexed structure, with a 24-hour soak required to obtain a low-resolution complex (3.15 Å) of Variant 1 with D-KDGlc, with strong electron density for bound D-KDGlc only seen in the active site of one of the two aldolase molecules present in the crystal unit cell.

As seen in the other variants, examination of the structure of Variant 1 complexed with D-KDGlc revealed that interactions between the active site residues and the acid region of D-KDGlc were essentially unchanged when compared to the wild-type structures (Figures 4 and S3, upper part of the bound substrate). A relatively short distance (2.35 Å) between the 4-OH of bound D-KDGlc and the Sγ thiol of C157 was observed in this variant, with the corresponding distance in a KDGal complex predicted to be even shorter (compared to the short distance between O4 of D-KDGal and T157 in the wild-type enzyme). This agrees with our previous suggestion[29] that the sterically demanding Cys residue in Variant 1 should disfavor D-KDGal formation. Further evidence that steric interactions with C157 may be responsible for the change in specificity is that the corresponding T157S variant has specificity almost unchanged from WT, while the T157C single mutant produces 75% dr favoring D-KDGlc.[29] The movement of Y130 away from K155 removes any interactions with the 4-OH group of D-KDGlc in Variant 1, with Y130 now

https://doi.org/10.1021/acscatal.2c03285
ACS Catal. 2022, 12, 11444−11455

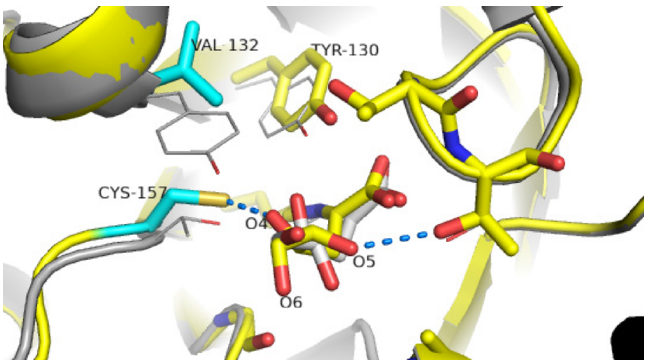


Figure 4. Active site of Variant 1 in complex with D-KDGlc superimposed onto the corresponding D-KDGlc complex of the wild-type enzyme. Variant 1 amino acids are shown as yellow sticks, with mutations shown in cyan. Conformations of D-KDGlc in wild-type (gray) and Variant 1 (yellow) complexes are shown, together with stabilizing hydrogen bond interactions (blue dotted lines) between O4 and O5 of D-KDGlc and residues in the active site.

only interacting with the O1 carbonyl of D-KDGlc, which means that it is unlikely to affect the diastereoselectivity of the Variant 1 aldol reaction.

Potential stabilizing interactions between the 5-OH and 6-OH groups of KDGlc and the phenol group of Y132 in the wild-type enzyme are removed through introduction of the Y132V mutation into Variant 1, which results in the C4−C5 bond of D-KDGlc rotating to enable a new hydrogen bond between the 5-OH group and the T44 hydroxyl residue. A hydrogen bond between the 6-OH and T44 hydroxyl group already exists in the wild-type enzyme−D-KDGal complex, so this interaction does not contribute additional stabilization to the transition state leading to D-KDGal formation. Therefore, destabilizing steric interactions between the C157 residue and the D-KDGal 4-OH hydroxyl group and removal of stabilizing hydrogen bonding interactions between the Y132 phenol group and the D-KDGal 4-OH and 5-OH groups, coupled with a new stabilizing interaction between the D-KDGlc 5-OH and T44 hydroxyl group, result in Variant 1 favoring the selective formation of D-KDGlc in 93:7 dr.

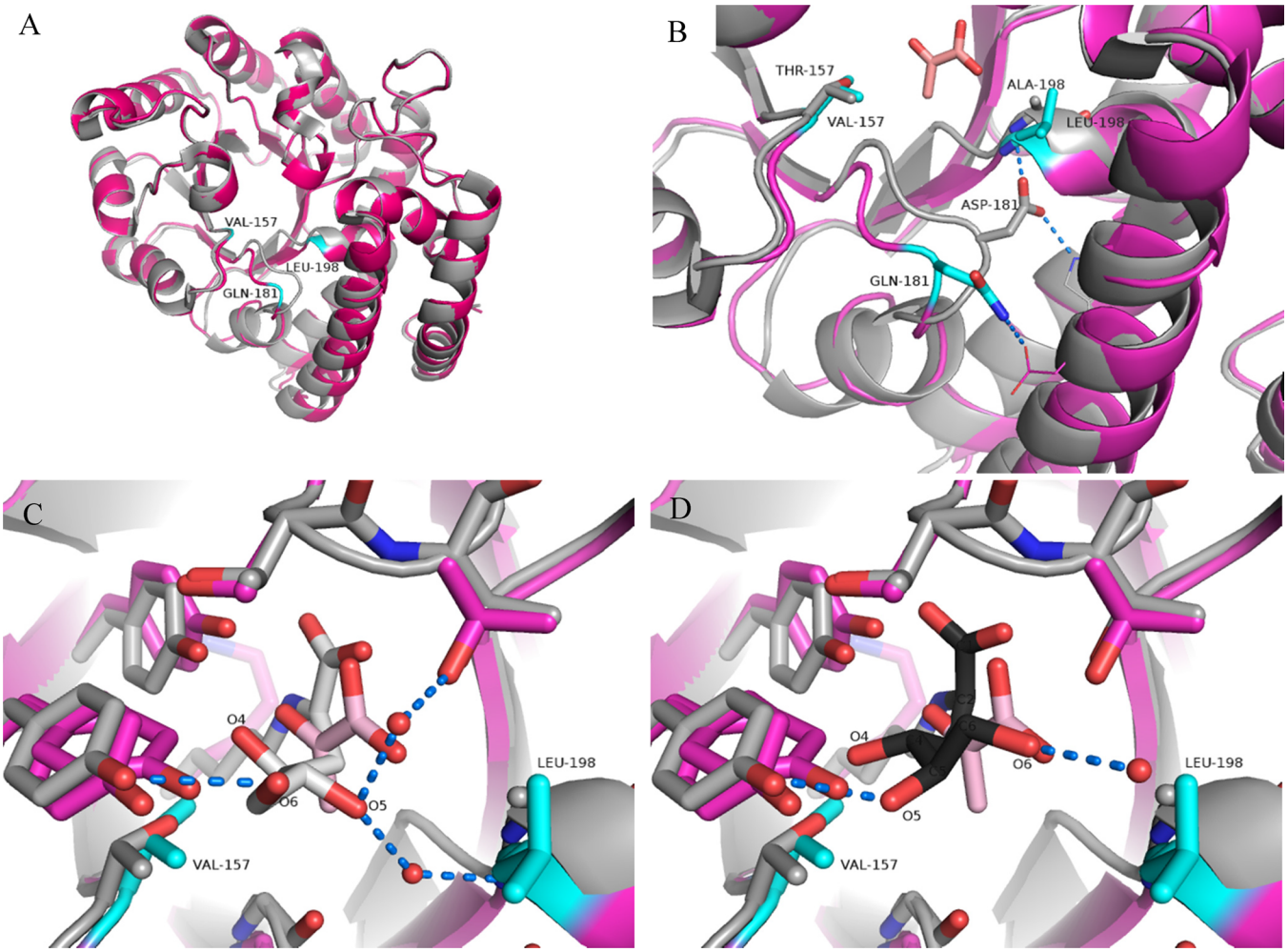


Figure 5. (A) Superimposition of the Variant 2 (magenta) and wild-type (gray, PDB code: 1W37) SsKDG-aldolase structures with mutation sites labeled and colored in cyan. Major movements occur in the lower middle portion of figure (A), which are magnified in figure (B). Pyruvate, bound in the variant structure, is shown in pink. Hydrogen bonds are shown by blue dotted lines. The active site of Variant 2 and the wild-type enzyme structure (gray) shown in complex with (C) D-KDGlc (white, PDB code: 1W3N) or (D) D-KDGal (black, PDB code: 1W3T), both of which show displacement of pyruvate within the Variant 2 active site. Water molecules involved in hydrogen bonds in the wild-type enzyme are shown as red spheres. Hydrogen bonds in figures (C and D) for the 5-OH and 6-OH groups of KDGlc and KDGal are those present in the corresponding wild-type SsKDG-aldolase structures.

https://doi.org/10.1021/acscatal.2c03285
ACS Catal. 2022, 12, 11444−11455

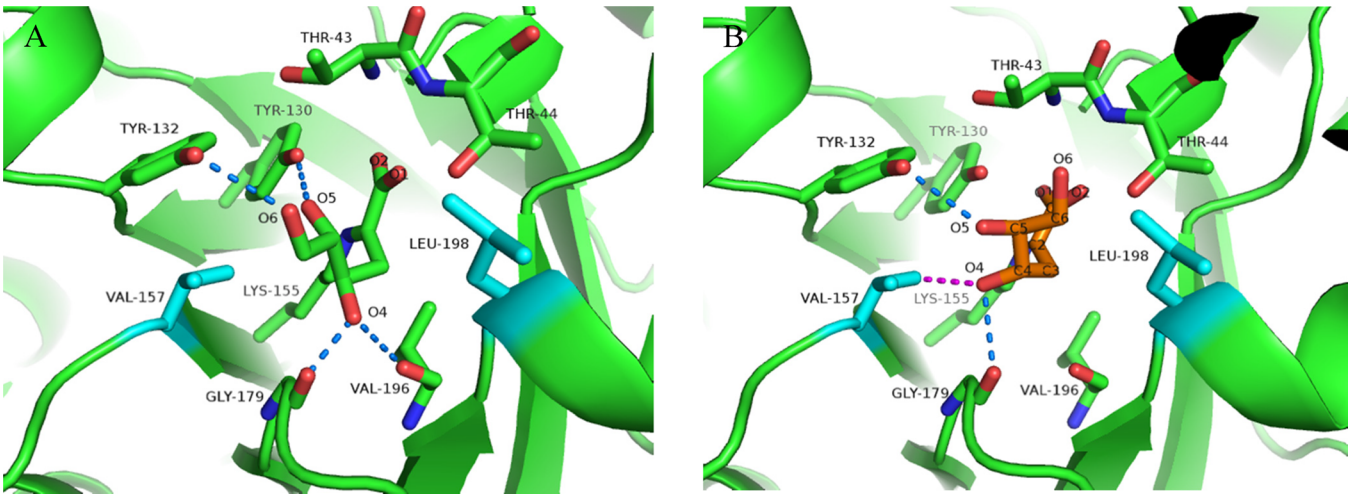


Figure 6. Active site structure of SsKDG-aldolase Variant 3 shown in green, with mutated residues shown in cyan. (A) L-KDGlc is shown having formed a Schiff-base link to Lys155, with key hydrogen bonds shown by blue dotted lines. (B) Superimposition of the conformation of L-KDGal (orange) from the Variant 4−L-KDGal complex (see Figure 7) onto the Variant 3 active site reveals that the 4-OH group of L-KDGal is likely to make unfavorable hydrophobic interactions with Val157 (magenta dotted line).

Structural Analysis of the D-KDGal Selective Variant 2 (T157V/D181Q/A198L). As previously reported,[29] the effect of introducing T157V, D181Q, and A198L mutations into Variant 2 results in a mutant that catalyzes selective formation of D-KDGal in 88:12 dr, which is a considerable increase over the 45% selectivity for D-KDGal produced by the wild-type aldolase. The kinetic parameters of Variant 2 are similar to those of Variant 1, (i.e., a similarly reduced k_{cat}) with the exception of its K_M for pyruvate, which was 10-fold greater than Variant 1 and 40-fold greater than that for the wild-type enzyme (Table 1).

X-ray crystallographic data for both the unsoaked (1.57 Å) and the pyruvate-complex structures (2.17 Å) were collected at a higher resolution than was achieved for Variant 1; however, these were in a space group $P2_1$ not seen in the other variants. Density for the Schiff base lysine (K155) could not be observed, even when the variant was cocrystallized with pyruvate, and no complexes of Variant 2 with D-KGal were obtained from multiple attempts.

The four molecules of the tetramer within the Variant 2 structure are very similar, having an average RMSD within the uncomplexed structure of only 0.2 Å (and between this structure and the pyruvate complex); however, the 0.7 Å difference when compared to the wild-type aldolase was larger than those for the other variants. Looking at the superimposed structures (Figure 5A), most of the wild-type and Variant 2 residues align to a much greater degree than might be expected from these RMSD values. However, significant displacement of the main chain backbone between residues 180 and 191 is seen, which is presumably caused by the D181Q mutation (Figure 5B). We had previously suggested[29] that inclusion of a longer glutamine residue for D181 in the loop underneath L198 (see the orientation shown in Figure 5B) would restrict the conformational freedom of this leucine residue, thus making this region of the active site more hydrophobic. This structural change was predicted to promote preferential binding of D-KDGal over D-KDGlc by Variant 2 because it would disrupt water-mediated hydrogen bonds present in the active site that we predicted would be more important for producing D-KDGlc over D-KDGal in the wild-type enzyme. The crystal structure of Variant 2 revealed that the primary amide group of the Q181 residue no longer makes interactions with the amide group of L198 (as seen in the wild-type enzyme), with Q181 now pointing in the opposite direction to form a hydrogen bond with the acid group of D230. This removes any conformational restraints on the L198 loop region of the active site (Figure 5B), which can relax away from the catalytic center, resulting in the floor of the active site being lowered below the pyruvate binding site. This relatively large conformational change may explain this variant's significantly greater K_M for pyruvate.

Comparison of the structure of Variant 2 (cocrystallized with pyruvate) with the wild-type enzyme complexed with D-KDGlc and D-KDGal (Figure 2) reveals that pyruvate is displaced from its active position, making no hydrogen bonds with T43, (Figures 5C and S4) while electron density for the Schiff base K155 is not visible in the active site. Nevertheless, our original hypothesis that was proposed to explain the D-KDGal selectivity of Variant 2 appears to still be valid. Replacement of T157 by a hydrophobic valine group will remove stabilizing hydrogen bond interactions to the 4-OH groups of both D-KDGlc and D-KDGal. However, introduction of the Leu198 residue will create a new hydrophobic region in the active site, which is likely to displace proximal active site water molecules. These bridging water molecules provide greater stabilizing hydrogen bond interactions with both the 5-OH and 6-OH groups of bound D-KDGlc than with the single 6-OH group of bound D-KDGal (whose 5-OH is also stabilized by interactions with T192) in wild-type enzyme complexes. Therefore, their absence in Variant 2 is likely to disfavor D-KDGlc formation, thus resulting in preferential D-KDGal in 88:12 dr (Figure 5).

Structural Analysis of the L-KDGlc Selective Variant 3 (T157V/A198L). Identification of an SsKDG-aldolase variant that could selectively transform non-natural L-glyceraldehyde into L-KDGlc began by screening the selectivity profile of the library of T157 saturation mutants generated previously to identify Variants 1 and 2.[29] This identified that T157V (used in Variant 2 to produce D-KDGal selectivity) was a suitable starting point for generating L-KDGlc selectivity, with inclusion of a second substitution (A198L) to create a hydrophobic pocket to disrupt the proximal hydrogen bonding water network (as used in the design of Variant 2) affording Variant 3 that produced L-KDGlc in 92:8 dr. The kinetic parameters of this variant were then determined (Table 1), with its V_{max} found to be very similar

https://doi.org/10.1021/acscatal.2c03285
ACS Catal. 2022, 12, 11444−11455

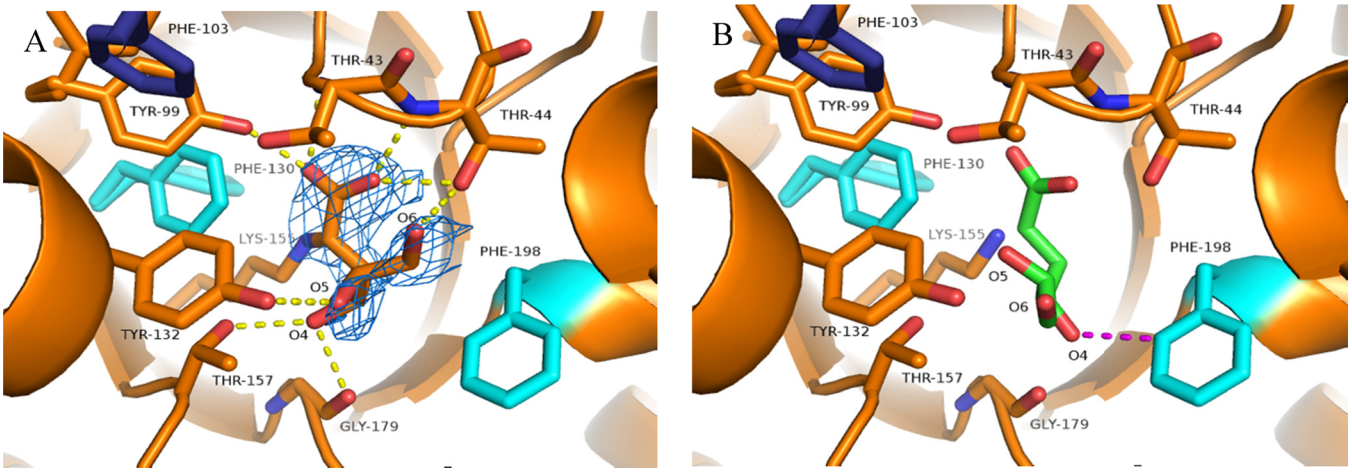


Figure 7. (A) Structure of SsKDG-aldolase Variant 4 with ligand-interacting active-site residues shown in orange. Mutated amino acids in the same monomer are shown in cyan, with mutated amino acids from a different aldolase monomer shown in dark blue. 2FoFc density (contoured at 1σ) shown around the bound L-KDGal residue. Stabilizing hydrogen bond interactions are indicated by yellow dotted lines. (B) Variant 4 structure with L-KDGlc conformation present in the Variant 3−L-KDGlc complex superimposed. Close proximity of the destabilizing 4-OH group of the bound L-KDGlc to Phe198 is shown by the magenta dotted line.

to that of the wild-type aldolase; however, the K_M values of Variant 3 for both aldol substrates were found to be increased. Crystallization of Variant 3 with L-KDGlc was carried out under standard conditions; however, its higher activity than those of Variants 1 and 2 meant that effective trapping of a ligand complex required shorter soaking times. For instance, a 4-min soak of Variant 3 with L-KDGlc resulted in a covalent Schiff-base lysine complex of pyruvate (formed from retro-aldol cleavage of L-KDGlc; data not shown). However, data acquired after a 2 min soak revealed greater continuous electron density within the active site, which enabled us to determine the structure of the L-KDGlc complex (Figure S5).

The structure of Variant 3 is essentially unchanged from the wild-type enzyme (RMSD of both soaked and unsoaked structures compared to the wild-type structure 1W37 is 0.3 Å, as are the RMSD between molecules within these structures). Unfortunately, structures of the wild-type SsKDG-aldolase complexed with L-KDGlc or L-KDGal were not available for comparison with the Variant 3−L-KDGlc structure. However, comparison with the wild-type enzyme complexed with D-KDGal (which has the same (*R*)-configuration as L-KDGlc at O4) (Figure 2) revealed that the 4-OH of L-KDGlc in Variant 3 had rotated away from V157 to make a new hydrogen bond interaction with the main chain peptide bonds of G179 and V196 at the bottom of the active-site cleft (Figure 6). This results in the 5-OH and 6-OH groups of L-KDGlc now making hydrogen bonds with the phenol groups of Y130 and Y132, respectively. This conformational realignment means that L-KDGlc presents a nonpolar face to the introduced V157 and L198 residues, which both combine to form a hydrophobic sandwich on either side of the active site.

Comparison of the structures of the Variant 3−L-KDGlc complex to the Variant 4−L-KDGal complex (vide infra) revealed that the 4-OH group of L-KDGal would be predicted to make unfavorable hydrophobic interactions with Val157 in Variant 3 (Figure 6B).

Rotation of the 4-OH group of L-KDGal to relieve this unfavorable interaction (to occupy a position similar to that observed for the 4-OH group of docked L-KDGlc in Variant 3) would result in its 5-OH group being brought into close proximity to V157, which is equally unfavorable. Therefore, the ability of L-KDGlc to present a hydrophobic face to the introduced hydrophobic amino acids (V157 and L198), while still maintaining multiple stabilizing hydrogen bonding interactions between its 5-OH and 6-OH hydroxyl groups, most likely explains the selectivity of Variant 3 for L-KDGlc formation in 92:8 dr. The presence of the two hydrophobic amino acids (V157 and L198) is likely to also restrict the overall conformational mobility of the aldehyde substrates within the active site of Variant 3, which may explain its higher K_M for L-glyceraldehyde.

Structural Analysis of the L-KDGal Selective Variant 4 (Y103F/Y130F/A198F). Producing a variant with a significant preference for L-KDGal over L-KDGlc proved challenging, with most structure-guided variants explored proving nonselective for formation of L-KDGal or displaying precipitous losses in catalytic activity. Saturation mutagenesis at T157 did not reveal a lead variant with L-KDGal selectivity. The other residue that interacts with O4 is Y130; however, this residue is responsible for donating a proton to O4 in the wild-type mechanism.[25,33] Nonetheless, it was decided to explore the activity of a Y103F variant, in the hope that the Y99 side chain could mimic the role of the Y130 OH residue of the wild-type enzyme. Further specificity was achieved through replacement of the A198 residue with a hydrophobic phenylalanine residue to disrupt the previously discussed hydrogen bonding water network that favored D-KDGlc formation. This triple mutant Variant 4 (Y103F/Y130F/A198F) was shown to selectively transform L-glyceraldehyde into L-KDGal in 76:24 dr (cf 47:53 dr in the wild-type aldolase).

The crystal structure of the wild-type enzyme−D-KDGal complex reveals that the Y130 residue (together with T157) interacts with the 4-OH group in the wild-type SsKDG−aldolase complexes (Figure 2); however, Y130 is also intimately involved in the catalytic mechanism of the aldolase (see Figure 8). Consequently, removal of the stabilizing Y130 phenol group significantly affected the catalytic activity of Variant 4, resulting in a marked reduction in its k_{cat} value (see Table 1). However, Variant 4 still retained sufficient catalytic activity to be useful as a biocatalyst for the synthesis of synthetically useful amounts of L-KDGal. The main chain structure of the active site of Variant 4 remains unchanged on binding to L-KDGal, as shown by an

https://doi.org/10.1021/acscatal.2c03285
ACS Catal. 2022, 12, 11444−11455

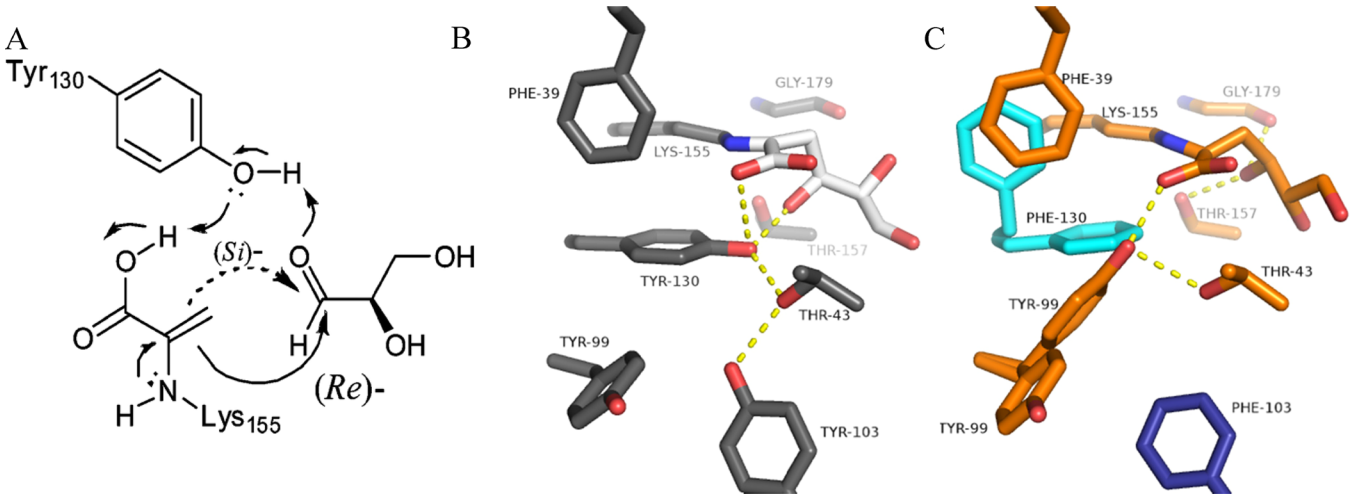


Figure 8. (A) Y130 residue of the wild-type aldolase is intimately involved in facilitating reaction of the K155-derived enamine derivative of pyruvate at the *Re*- or *Si*-face of D-glyceraldehyde to afford D-KDGlc or D-KDGal.[25,33] (B) Positions of amino acids around Y130 in the wild-type SsKDG-aldolase structure shown in gray, with bound D-KDGlc shown in white. Hydrogen bonds made by the phenol group of Y130 are shown as yellow dotted lines. (C) Positions of equivalent amino acids in Variant 4 shown in orange, with mutated residues in the same monomer shown in cyan and residues from a neighboring aldolase monomer shown in dark blue. Hydrogen bonds made by the phenol group of the replacement Y99 and the 4-OH group of L-KDGal are shown as dotted yellow lines.

RMSD of only 0.3 Å between Variant 4 and the wild-type structure and between the structures of soaked and unsoaked Variant 4. Comparison of the structure of Variant 4 bound to L-KDGal to the wild-type structure bound to D-KDGlc (Figure 2A, same (*R*)-4-OH configuration) revealed that the 4-OH group of L-KDGal in Variant 4 rotates away from F130 to make a new hydrogen bond with the backbone amide of G179 while still maintaining an interaction with the T157 hydroxyl group (Figure 7A). This enables the 5-OH and 6-OH groups of Variant 4-bound L-KDGal to make hydrogen bonds directly to both the phenol group of Y132 and the hydroxyl group of T44, which results in L-KDGal presenting a hydrophobic face toward the aryl ring of the introduced F198 residue.

Superimposition of the structure of L-KDGlc−Variant 3 onto L-KDGal−Variant 4 reveals that the 4-OH hydroxyl group of L-KDGlc likely occupies a position close to F198 when docked into Variant 4 (Figure 7B), and so its binding energy would be predicted to be significantly higher than the binding energy required to accommodate L-KDGal. Rotation of the C3−C4 bond of L-KDGlc to allow its 4-OH group to adopt a similar conformation to the 4-OH group of bound L-KDGal (i.e., to allow hydrogen bonding interactions with T157 and G179) would result in the L-KDGlc 5-OH group being brought into unfavorably close proximity to the hydrophobic F198 residue.

Therefore, removal of the 4-OH-Y130 hydrogen bond from Variant 4 that preferentially favors KDGlc formation in the wild-type enzyme, combined with the presence of a new hydrophobic pocket formed by the A198F mutation, results in Variant 4 promoting selective formation of L-KDGal in 76:24 dr. While the proposed structural interpretation can be used to explain the observed shift in L-KDGal selectivity from 55:45 in the wild-type to 24:76 in Variant 4, it is recognized that this difference in diastereoselectivity corresponds to a relatively small difference in activation free energy (~0.8 kcal/mol, vide infra), and so these differences in diastereocontrol may equally be caused by other secondary effects.

As previously reported, the mechanism of the wild-type SsKDG-aldolase-catalyzed reaction involves direct interactions between the Y130 phenol residue and the pyruvate and glyceraldehyde substrates (see Figure 8A).[25,33] This Y130 residue also forms a direct hydrogen bond with the T43 residue, which interacts with the Y103 residue of another aldolase monomer, thereby establishing a stabilizing hydrogen bonding network of tyrosine/threonine residues that has been observed in other aldolases. These amino acids are well defined in the wild-type aldolase structure, with each residue exhibiting relatively low mobility to produce an effective hydrogen bonding network that facilitates catalysis (Figure 8A). Two of these tyrosine residues (103 and 130) have been mutated to phenylalanine residues in Variant 4, with the residual F130 residue observed to adopt two conformations: a conformation similar to the one seen in the wild-type enzyme and a new conformation stabilized through a π-stacking interaction with F39 (Figure 8B). Movement of the F130 side chain out of the active site, together with replacement of Y103 by a hydrophobic phenylalanine residue, results in relaxation of the active site, which allows the Y99 residue to adopt a new conformation that places its phenolic group in a similar position to that previously occupied by the Y130 residue of the wild-type enzyme. Therefore, this conformational rearrangement in the Variant 4 active site appears to be sufficient to enable Y99 to participate in the aldol reaction of pyruvate and L-Glyceraldehyde to selectively afford L-KDGal, albeit at a significantly reduced rate when compared to the corresponding aldol reaction of the wild-type enzyme. Although Y99 can interact directly with T43 and the C1 carbonyl of L-KDGal, it is likely to be too remote from the 4-OH group to provide a stabilizing interaction. Consequently, the 4-OH group of L-KDGal rotates to make interactions with the hydroxyl group of T157 and the backbone amide group of G179 (Figure 8C), with this nonoptimal orientation also likely to contribute to the considerably lower reaction rate of Variant 4.

Factors Affecting the Diastereoselectivities of the Aldol Reactions of Wild-Type KDG-Aldolase and Its Variants. The preceding sections describe how subtle changes in active site interactions of Schiff-base-linked Lys155-bound KDGlc and KDGal aldol products of the wild-type aldolase and its mutant variants have been used to rationalize the

https://doi.org/10.1021/acscatal.2c03285
ACS Catal. 2022, 12, 11444−11455

diastereoselectivities of the respective aldol reactions responsible for their formation. The drs of the aldol products produced by the variants could potentially be caused by differences in aldehyde substrate-binding geometries prior to carbon−carbon bond formation and/or differences in the free energy barriers of the transition states leading to formation of the KDGlc and KDGal diastereomers.[26] Consequently, it was decided to explore the binding preferences of the wild-type and variant aldolases using molecular dynamic simulations of each variant containing a Lys155-pyruvate Schiff base and a bound glyceraldehyde substrate. Simulations were performed using D- and L-glyceraldehydes as substrates for the wild-type enzyme, with D-glyceraldehyde used for Variants 1 and 2, and L-glyceraldehyde used for Variants 3 and 4 (see the Supporting Information for details). None of these simulations produced an excess of pre-*S* or pre-*R* glyceraldehyde-bound conformations that could explain the experimentally determined dr levels of either the wild-type enzyme or its variants (Figures S6 and S7). Therefore, it follows that the observed differences in the diastereoselectivities of each variant's aldol reaction must arise from differences in the transition-state energies of the *anti*- and *syn*-aldol reactions that lead to their respective KDGlc and KDGal products [as rationalized by the X-ray crystallographic analyses (vide supra)]. This conclusion is not unexpected, as the structures of the docked variant-aldol products are likely to resemble the transition states of the aldol reactions that lead to their formation quite closely, with the C3−C4 bond distance in their transition states expected to be ∼2.1 Å.[27] The differences in the energy barriers of the transition states that produce the observed dr levels of each variant are relatively small, corresponding to a difference of 0.8 to 1.8 kcal/mol depending on the dr value of the variant (determined from $\Delta G = -RT\ln(Z)$, where Z is the diastereomeric ratio at 70 °C).

Examination of the structures of the wild-type aldolase bound to D-KDGlc and D-KDGal reveals that their 4-OH groups both make hydrogen bonds to the OH groups of Y130 and T157. Assuming that similar interactions occur in the transition states leading to their formation, similar two-point hydrogen bond interactions with the electron lone pairs of the carbonyl group of D-glyceraldehyde will activate it toward nucleophilic attack by the K155-bound pyruvate nucleophile. The nonselective aldol reaction of the wild-type enzyme requires that the incoming K155-bound pyruvate nucleophile has approximately equal access to both the *Re*- and *Si*-faces of the bound aldehyde group (as indicated by the MD simulations) and approximately equal energy barriers for reaction. This is achieved by binding D-glyceraldehyde in different pre-*R* and pre-*S* conformations that are stabilized by different hydrogen bonding interactions between the 2-OH and 3-OH residues of D-glyceraldehyde (5-OH and 6-OH in KDGlc and KDGal) and different amino acid residues and/or bound water molecules within the active site. Therefore, the promiscuity of the wild-type enzyme is due to its ability to access different hydrogen bond networks that stabilize transition states with the pyruvate nucleophile attacking either the *Si*- or *Re*-face of the aldehyde group that generate D-KDGlc and D-KDGal, respectively. For the wild-type enzyme, the energy differences between these pre-*R* and pre-*S* derived transition states leading to D-KDGlc and D-KDGal are negligible, which leads to formation of mixtures of aldol products in low 55:45 dr favoring D-KDGlc.

Examination of the structure of D-KDGlc-bound Variant 1 indicates that the aldehyde group of D-glyceraldehyde is predicted to be activated through a new hydrogen bond to C157, with its 5-OH and 6-OH residues then making different interactions with active site hydrogen bond acceptors that favor D-KDGlc formation. No ligand-bound structures could be obtained for Variant 2; however, the structure of Variant 3 bound to L-KDGlc indicates that the aldehyde group of L-glyceraldehyde in the transition state leading to L-KDGlc (major aldol product) should be activated toward nucleophilic attack through hydrogen bonding to the amide bonds of G179 and V196. Alternatively, the structure of Variant 4 bound to L-KDGal reveals that the aldehyde group of L-glyceraldehyde in the transition state leading to L-KDGal (major aldol product) should be activated by hydrogen bonding to the hydroxyl group of T157 and the amide bond of G179. Therefore, Variants 3 and 4 bind the carbonyl group of L-glyceraldehyde in different regions of their active sites when compared to the bound carbonyl group of D-glyceraldehyde in the active sites of the wild-type enzyme and Variant 1.

The ligand binding structures shown in Figure 2 reveal that the wild-type KDGlc aldolase can contribute multiple hydrogen bond acceptor residues and bridging water molecules to enable its D- and L-glyceraldehyde substrates to be bound effectively. Therefore, different networks of hydrogen bond acceptors can be used to bind either glyceraldehyde in different pre-*R* and pre-*S* conformations that can each react with the K155-bound pyruvate nucleophile through transition states that lead to KDGlc and KDGal products. In the wild-type enzyme, the difference in the energies of the pre-*R* and pre-*S* derived transition states must be small, which means that the KDGlc and KDGal diastereomers are produced in similar amounts. Selective mutation of active-site residues to generate Variants 1−4 results in mutants whose active sites make different H-bonding interactions with the carbonyl and hydroxyl groups of their bound glyceraldehyde substrates. The different hydrogen bonding networks of each variant then serve to perturb the relative energy differences of the transition states leading to their L/D-KDGlc and L/D-KDGal products, thus allowing selective production of a different major aldol diastereomer in each case.

We believe that our strategy of selectively introducing or deleting active site residues that make selected hydrogen bond interactions with the polar substituents of an aldehyde substrate may prove to be a generally useful approach for improving the specificity and stereoselectivity profiles of other types of aldolase. Although we have shown that this approach leads to significant decreases in activity, the generally high activities of many aldolases indicates that variants with lower activities but good levels of stereocontrol can be employed for the synthesis of useful quantities of enantiopure aldol products. While this study has been directed toward evolving the stereoselectivity of the KDG aldolase toward polar hydroxy-aldehyde substrates, we anticipate that a similar active site editing strategy may be useful for conferring aldol reactivity and stereoselectivity toward more hydrophobic aldehyde substrates.

CONCLUSIONS

Toscano et al.[15] have remarked that "enzyme active sites provide highly optimized microenvironments for the catalysis of biologically useful chemical transformations. Consequently, changes at these centers can have large effects on enzyme activity." In their review, they describe how minimal modifications to enzyme active sites can expand their catalytic abilities to give both enhanced and new activities, noting that such an approach needs high-resolution structures of an enzyme's active site to be successful, preferably with substrates

https://doi.org/10.1021/acscatal.2c03285
ACS Catal. 2022, 12, 11444−11455

bound. In contrast to previous reports that have aimed to widen the catalytic activity and stereospecificity of highly stereoselective enzymes, our study has employed a stereochemically promiscuous aldolase enzyme as a template to develop stereoselective variants,[30] using structurally informed site-directed mutagenesis to generate a suite of mutant aldolases that can be used to prepare all four possible aldol stereoisomers. Structural determination of the variant enzymes has allowed us to rationalize many of the observed changes in catalytic activity and stereoselectivity, with simulations indicating that stereoselectivity likely arises during carbon−carbon bond formation. Notably, some of the mechanisms that produce the observed structural changes were not as originally predicted, once again demonstrating the challenge of modifying enzyme stereoselectivity through a structure-guided approach.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscatal.2c03285.

Parameters for the pyruvate-lysine Schiff base (KPI) used in the molecular dynamics simulations (ZIP)

HPLC traces for aldolase-catalyzed reactions of Variants 3 and 4, images of the connected density and bound product in Variant 1, of the bound product in Variant 2, and around the bound substrate in Variant 3, table of crystallographic data, details of the enzyme−substrate molecular dynamics simulations, synthetic procedures, characterization data for intermediates and products, and relevant NMR spectra (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Susan J. Crennell − *Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K.;* orcid.org/0000-0003-1697-8508; Phone: +44-1225-384302; Email: bsssjc@bath.ac.uk

Authors

Sylvain F. Royer − *Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K.;* Present Address: University of Oxford, Chemistry Research Laboratory, Mansfield Road, Oxford, OX1 3TA, U.K. (S.F.R.)

Xuan Gao − *School of Biochemistry, University of Bristol, Bristol BS8 1TD, U.K.;* orcid.org/0000-0003-2643-5726

Robin R. Groleau − *Department of Chemistry, University of Bath, Bath BA2 7AY, U.K.*

Marc W. van der Kamp − *School of Biochemistry, University of Bristol, Bristol BS8 1TD, U.K.;* orcid.org/0000-0002-8060-3359

Steven D. Bull − *Department of Chemistry, University of Bath, Bath BA2 7AY, U.K.;* orcid.org/0000-0001-8244-5123

Michael J. Danson − *Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, U.K.*

Complete contact information is available at: https://pubs.acs.org/10.1021/acscatal.2c03285

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the BBSRC (S.F.R.; M.W.v.d.K, BB/M026280/1) and EPSRC CDT in Catalysis (R.R.G., EP/L016443/1) for funding. We acknowledge the valuable contributions made to this project by many undergraduate and masters project students: Luke Haslett, Emily Gostling, Marina Angelopoulou, Laurence Devesse, Leigh Knight, Marina Vabistsevits, Jennifer Frewin, Pooja Jhugroo, and Jennifer Gurney. The authors would like to thank Diamond Light Source for beamtime (proposal mx1226) and the staff of beamlines I03 and I04 for assistance with data collection. Characterization facilities were provided by the Material and Chemical Characterization Facility (MC^2) at the University of Bath (https://doi.org/10.15125/mx6j-3r54), and we wish to thank Dr. Kathryn Proctor for her help and expertise. MD simulations were conducted using the computational facilities of the Advanced Computing Research Centre, University of Bristol.

■ ABBREVIATIONS

SsKDG-aldolase, *Sulfolobus solfataricus* 2-keto, 3-deoxygluconate aldolase; D-KDGlc, D-2-keto-3-deoxy gluconate; D-KDGal, D-2-keto-3-deoxy galactonate; L-KDGlc, L-2-keto-3-deoxy gluconate; L-KDGal, L-2-keto-3-deoxy galactonate.

■ REFERENCES

(1) Brovetto, M.; Gamenara, D.; Saenz Méndez, P.; Seoane, G. A. C-C Bond-Forming Lyases in Organic Synthesis. *Chem. Rev.* **2011**, *111*, 4346−4403.

(2) Clapés, P.; Fessner, W.-D.; Sprenger, G. A.; Samland, A. K. Recent progress in stereoselective synthesis with aldolases. *Curr. Opin. Chem. Biol.* **2010**, *14*, 154−167.

(3) Windle, C. L.; Berry, A.; Nelson, A. Aldolase-catalyzed steroselective synthesis of fluorinated small molecules. *Curr. Opin. Chem. Biol.* **2017**, *37*, 33−38.

(4) Clapés, P.; Garrabou, X. Current Trends in Asymmetric Synthesis with Aldolases. *Adv. Synth. Catal.* **2011**, *353*, 2263−2283.

(5) Alcantara, A. R.; Pace, V.; Hoyos, P.; Sandoval, M.; Holzer, W.; Hernaiz, M. J. Chemoenzymatic synthesis of carbohydrates as antidiabetic and anticancer drugs. *Curr. Top. Med. Chem.* **2014**, *14*, 2694−2711.

(6) Schmidt, N. G.; Eger, E.; Kroutil, W. Building Bridges: Biocatalytic C-C bond formation toward multifunctional products. *ACS Catal.* **2016**, *6*, 4286−4311.

(7) For an example of a wild-type enzyme that shows promiscuity towards non-natural enolate donors see Roldán, R.; Sanchez-Moreno, I.; Scheidt, T.; Hélaine, V.; Lemaire, M.; Parella, T.; Clapés, P.; Fessner, W.-D.; Guérard-Hélaine, C. Breaking the dogma of aldolase specificity: Simple aliphatic ketones and aldehydes are nucleophiles for fructose-6-phosphate aldolase. *Chem. − Eur. J.* **2017**, *23*, 5005−5009.

(8) Saravanan, T.; Reif, M.-L.; Yi, D.; Lorilliere, M.; Charmantray, F.; Hecquet, L.; Fessner, W.-D. Engineering a thermostable transketolase for arylated substrates. *Green Chem.* **2017**, *19*, 481−489.

(9) Adams, J. P.; Brown, M. J. B.; Diaz-Rodriguez, A.; Lloyd, R. C.; DoruRoiban, G. Biocatalysis: A pharma perspective. *Adv Synth Catal* **2019**, *361*, 2421−2432.

(10) For recent reports where the selectivity profiles of aldolases have been modified by protein engineering to accept non-natural enolate donors, see: (a) Hernández, K.; Szekrenyi, A.; Clapés, P. Nucleophile Promiscuity of Natural and Engineered Aldolases. *ChemBioChem* **2018**, *19*, 1353−1358. (b) Saravanan, T.; Junker, S.; Kickstein, M.; Hein, S.; Link, M.-K.; Ranglack, J.; Witt, S.; Lorillière, M.; Hecquet, L.; Fessner, W.-D. Donor Promiscuity of a Thermostable Transketolase by Directed Evolution: Efficient Complementation of 1-Deoxy-d-xylulose-5-

phosphate Synthase Activity. *Angew. Chem., Int. Ed.* **2017**, *56*, 5358−5362. (c) Hernández, K.; Joglar, J.; Bujons, J.; Parella, T.; Clapés, P. Nucleophile Promiscuity of Engineered Class II Pyruvate Aldolase YfaU from E. Coli. *Angew. Chem., Int. Ed.* **2018**, *57*, 3583−3587. (d) Junker, S.; Roldan, R.; Joosten, H.-J.; Clapés, P.; Fessner, W.-D. Complete Switch of Reaction Specificity of an Aldolase by Directed Evolution In Vitro: Synthesis of Generic Aliphatic Aldol Products. *Angew. Chem., Int. Ed.* **2018**, *57*, 10153−10157.

(11) Hélaine, V.; Gastaldi, C.; Lemaire, M.; Clapés, P.; Guérard-Hélaine, C. Recent Advances in the Substrate Selectivity of Aldolases. *ACS Catal.* **2022**, *12*, 733−761.

(12) Fong, S.; Machajewski, T. D.; Mak, C. C.; Wong, C.-H. Directed evolution of D-2-keto-3-deoxy-6-phosphogluconate aldolase to new variants for the efficient synthesis of D- and L-sugars. *Chem. Biol.* **2000**, *7*, 873−883.

(13) Jaeger, K.-E.; Eggert, T. Enantioselective biocatalysis optimized by directed evolution. *Curr. Opin. Biotechnol.* **2004**, *15*, 305−313.

(14) Hsu, C.-C.; Hong, Z.; Wada, M.; Franke, D.; Wong, C.-H. Directed evolution of d-sialic acid aldolase to l-3-deoxy-*manno*-2-octulosonic acid (l-KDO) aldolase. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 9122−9126.

(15) Toscano, M. D.; Woycechowsky, K. J.; Hilvert, D. Minimalist active-site redesign: teaching old enzymes new tricks. *Angew. Chem., Int. Ed.* **2007**, *46*, 3212−3236.

(16) For recent reports on the use of wild-type and engineered 2-oxo acid dependent aldolases with broad specificity profiles for their enolate and/or aldehyde donors that have been used for the synthesis of a wide range of aldol products, see: (a) Laurent, V.; Gourbeyre, L.; Uzel, A.; Hélaine, V.; Nauton, L.; Traïkia, M.; de Berardinis, V.; Salanoubat, M.; Gefflaut, T.; Lemaire, M.; Guérard-Hélaine, C. Pyruvate Aldolases Catalyze Cross-Aldol Reactions between Ketones: Highly Selective Access to Multi-Functionalized Tertiary Alcohols. *ACS Catal.* **2020**, *10*, 2538−2543. (b) de Berardinis, V.; Guerard-Helaine, C.; Darii, E.; Bastard, K.; Helaine, V.; Mariage, A.; Petit, J.-L.; Poupard, N.; Sanchez-Moreno, I.; Stam, M.; Gefflaut, T.; Salanoubat, M.; Lemaire, M. Expanding the reaction space of aldolases using hydroxypyruvate as a nucleophilic substrate. *Green Chem.* **2017**, *19*, 519−526. (c) Moreno, C. J.; Hernández, K.; Charnok, S. J.; Gittings, S.; Bolte, M.; Joglar, J.; Bujons, J.; Parella, T.; Clapés, P. Synthesis of γ-Hydroxy-α-amino Acid Derivatives by Enzymatic Tandem Aldol Addition−Transamination Reactions. *ACS Catal.* **2021**, *11*, 4660−4669. (d) Laurent, V.; Uzel, A.; Hélaine, V.; Nauton, L.; Traïkia, M.; Gefflaut, T.; Salanoubat, M.; de Berardinis, V.; Lemaire, M.; Guérard-Hélaine, C. Exploration of Aldol Reactions Catalyzed by Stereoselective Pyruvate Aldolases with 2-Oxobutyric Acid as Nucleophile. *Adv. Synth. Catal* **2019**, *361*, 2713−2717. (e) Marín-Valls, R.; Hernández, K.; Bolte, M.; Parella, T.; Joglar, J.; Bujons, J.; Clapés, P. Biocatalytic Construction of Quaternary Centers by Aldol Addition of 3,3-Disubstituted 2-Oxoacid Derivatives to Aldehydes. *J. Am. Chem. Soc.* **2020**, *142*, 19754−19762. (f) Baker, P.; Seah, S. Y. K. Rational Design of Stereoselectivity in the Class II Pyruvate Aldolase Bphl. *J. Am. Chem. Soc.* **2012**, *134*, 507−513. (g) Fang, J.; Hait, D.; Head-Gordon, M.; Chang, M. C. Y. Chemoenzymatic Platform for Synthesis of Chiral Organofluorines Based on Type II Aldolases. *Angew. Chem., Int. Ed.* **2019**, *58*, 11841−11845.

(17) Marín-Valls, R.; Hernández, K.; Bolte, M.; Joglar, J.; Bujons, J.; Clapés, P. Chemoenzymatic Hydroxymethylation of Carboxylic Acids by Tandem Stereodivergent Biocatalytic Aldol Reaction and Chemical Decarboxylation. *ACS Catal.* **2019**, *9*, 7568−7577.

(18) Windle, C. L.; Muller, M.; Nelson, A.; Berry, A. Engineering aldolases as biocatalysts. *Curr. Opin. Chem. Biol.* **2014**, *19*, 25−33.

(19) Danson, M. J., Lamble, H. J.; Hough, D.W. Central metabolism. In *Archaea: Molecular and Cellular Biology*. Cavicchioli, R., Ed., Chapter 12; ASM Press: Washington DC, USA, 2007; pp 260−287.

(20) Nunn, C. E. M.; Johnsen, U.; Schönheit, P.; Fuhrer, T.; Sauer, U.; Hough, D. W.; Danson, M. J. Metabolism of pentose sugars in the hyperthermophilic archaea *Sulfolobus solfataricus* and *Sulfolobus acidocaldarius*. *J. Biol. Chem.* **2010**, *285*, 33701−33709.

(21) Buchanan, C. L.; Connaris, H.; Danson, M. J.; Reeve, C. D.; Hough, D. W. An extremely thermostable aldolase from *Sulfolobus solfataricus* with specificity for non-phosphorylated substrates. *Biochem. J.* **1999**, *343*, 563−570.

(22) Archer, R. M.; Royer, S. F.; Mahy, W.; Winn, C. L.; Danson, M. J.; Bull, S. D. Syntheses of 2-keto-3-deoxy-D-xylonate and 2-keto-3-deoxy-L-arabinonate as stereochemical probes for demonstrating the metabolic promiscuity of *Sulfolobus solfataricus* towards D-xylose and L-arabinose. *Chem. − Eur. J.* **2013**, *19*, 2895−2902.

(23) Lamble, H. J.; Heyer, N. I.; Bull, S. D.; Hough, D. W.; Danson, M. J. Metabolic pathway promiscuity in the Archaeon *Sulfolobus solfataricus* revealed by studies on glucose dehydrogenase and 2-keto-3-deoxygluconate aldolase. *J. Biol. Chem.* **2003**, *278*, 34066−34072.

(24) Lamble, H. J.; Danson, M. J.; Hough, D. W.; Bull, S. D. Engineering stereocontrol into an aldolase-catalysed reaction. *Chem. Commun.* **2005**, 124−126.

(25) Theodossis, A.; Walden, H.; Westwick, E. J.; Connaris, H.; Lamble, H. J.; Hough, D. W.; Danson, M. J.; Taylor, G. L. The structural basis for substrate promiscuity in 2-keto-3-deoxygluconate aldolase from the Entner-Doudoroff pathway in *Sulfolobus solfataricus*. *J. Biol. Chem.* **2004**, *279*, 43886−43892.

(26) Ahmed, H.; Ettema, T. J. G.; Tjaden, B.; Geerling, A. C. M.; van der Oost, J.; Siebers, B. The semi-phosphorylative Entner-Doudoroff pathway in hyperthermophilic Archaea: a re-evaluation. *Biochem. J.* **2005**, *390*, 529−540.

(27) Lamble, H. J.; Theodossis, A.; Milburn, C. C.; Taylor, G. L.; Hough, D. W.; Bull, S. D.; Danson, M. J. Promiscuity in the part-phosphorylative Entner-Doudoroff pathway of the archaeon *Sulfolobus solfataricus*. *FEBS Lett.* **2005**, *579*, 6865−6869.

(28) Azido-functionalised aldehydes have also been used as acceptors for the wild-type SsKDG-aldolase to generate aldol products with useful levels of stereocontrol, see: Schurink, M.; Wolterink-van Loo, S.; van der Oost, J.; Sonke, T.; Franssen, M.C.R. Substrate specificity and stereoselectivity of two *Sulfolobus* 2-keto-3-deoxygluconate aldolases towards azido-substituted aldehydes. *ChemCatChem* **2014**, *6*, 1073−1081.

(29) Royer, S. F.; Haslett, L.; Crennell, S. J.; Hough, D. W.; Danson, M. J.; Bull, S. D. Structurally informed site-directed mutagenesis of a stereochemically promiscuous aldolase to afford stereochemically complementary biocatalysts. *J. Am. Chem. Soc.* **2010**, *132*, 11753−11758.

(30) Long, F.; Vagin, A.; Young, P.; Murshudov, G. N. BALBES: a molecular replacement pipeline. *Acta Crystallogr., D: Biol. Crystallogr.* **2008**, *D64*, 125−132.

(31) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., D: Biol. Crystallogr.* **1997**, *D53*, 240−255.

(32) Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkóczi, G.; Chen, V. B.; Croll, T. I.; Hintze, B.; Hung, L. W.; Jain, S.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R. D.; Poon, B. K.; Prisant, M. G.; Read, R. J.; Richardson, J. S.; Richardson, D. C.; Sammito, M. D.; Sobolev, O. V.; Stockwell, D. H.; Terwilliger, T. C.; Urzhumtsev, A. G.; Videau, L. L.; Williams, C. J.; Adams, P. D. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr., D: Struct. Biol.* **2019**, *75*, 861−877.

(33) Daniels, A. D.; Campetto, I.; van der Kamp, M. W.; Bolt, A. H.; Trinh, C. H.; Phillips, S. E. V.; Pearson, A. R.; Nelson, A.; Mulholland, A. J.; Berry, A. Reaction mechanism of *N*-acetylneuraminic acid lyase revealed by a combination of crystallography, QM/MM simulation and mutagenesis. *ACS Chem. Biol.* **2014**, *9*, 1025−1032.

https://doi.org/10.1021/acscatal.2c03285
ACS Catal. 2022, 12, 11444−11455