Supplementary Information

Structural and Functional Characterization of YdjI, an Aldolase of Unknown Specificity in *Escherichia coli* K12

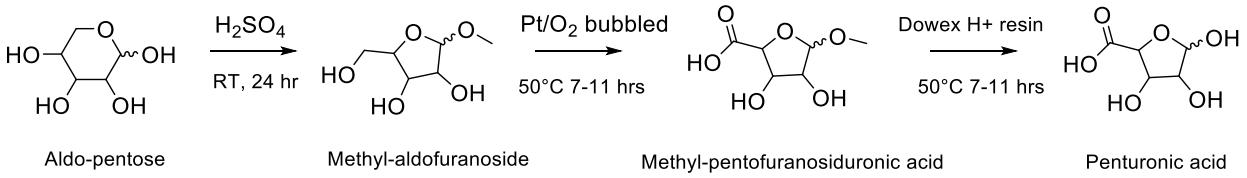
Jamison P. Huddleston[‡], James B. Thoden[δ], Brandon J. Dopkins[δ], Tamari Narindoshvili[‡], Blair J. Fose[‡], Hazel M. Holden[δ], and Frank, M. Raushel[‡*]

[‡]Department of Chemistry, Texas A&M University, College Station, Texas, 77843, United States.

[δ]Department of Biochemistry, University of Wisconsin-Madison, Madison, Wisconsin, 53706, United States

Synthesis of Penturonic Acids. Penturonic acids (D-arabinuronate, L-arabinuronate, D-xyluronate, L-lyxuronate, and D-riburonate) were synthesized using modifications to a previously published method (*1*, *2*). Generally, 1.5 grams of the pentose sugar (D-arabinose, L-arabinose, D-xylose, L-lyxose, or D-ribose) were stirred in 60 mL of methanol and 0.25 mL of concentrated H_2SO_4 overnight at room temperature. The resulting solution was passed through a column of Dowex OH^- resin to neutralize the solution and then concentrated under reduced pressure. The resulting methylfuranoside was resuspended in 50 mL of water with 300-400 mg of platinum on activated carbon (10%). The solution was stirred at 50 °C and oxygen gas was bubbled through the solution. Periodically, 100 mg of sodium bicarbonate was added to keep the reaction solution near neutral pH. In total, 300-500 mg of sodium bicarbonate was added during the reaction. Typically, the reaction was allowed to proceed for 7-16 h. The reaction solution was centrifuged, filtered, concentrated to approximately ~20 mL, and then characterized by ^{13}C NMR spectroscopy. The solution was loaded onto a DEAE Sephadex A25 column (3.8 x 13 cm) that had been pre-equilibrated with 500 mM NH_4HCO_3, pH 8.5, and washed with water until the pH was neutral. The penturonic methylglycosides were eluted from the column using a 600 mL gradient of 10-100 mM NH_4HCO_3. Typically, the penturonic acids eluted at ~50 mM HCO_3^-. Fractions were monitored using the phenol-sulfuric acid test for carbohydrates (*3, 4*) with the most concentrated fractions pooled and concentrated under reduced pressure to a final volume of 10 mL. ^{13}C NMR spectroscopy was used to verify the structures of the reaction products. The methyl protecting group was removed by incubating the compound with Dowex H^+ resin, pH 2-3, overnight at 40-50 °C. The Dowex resin was removed by filtration and the resulting solution was passed through a Dowex Na^+

column until neutral. The resulting solution was dried under reduced pressure yielding the sodium salt of the penturonic acids, which appeared as slightly yellow, hydroscopic, solids. ^{13}C NMR and ^{1}H NMR spectroscopy were used to verify the final products of the reactions compared to values previously published (*1*). The net overall yield was approximately 60% for each reaction based on dry weight of final product.

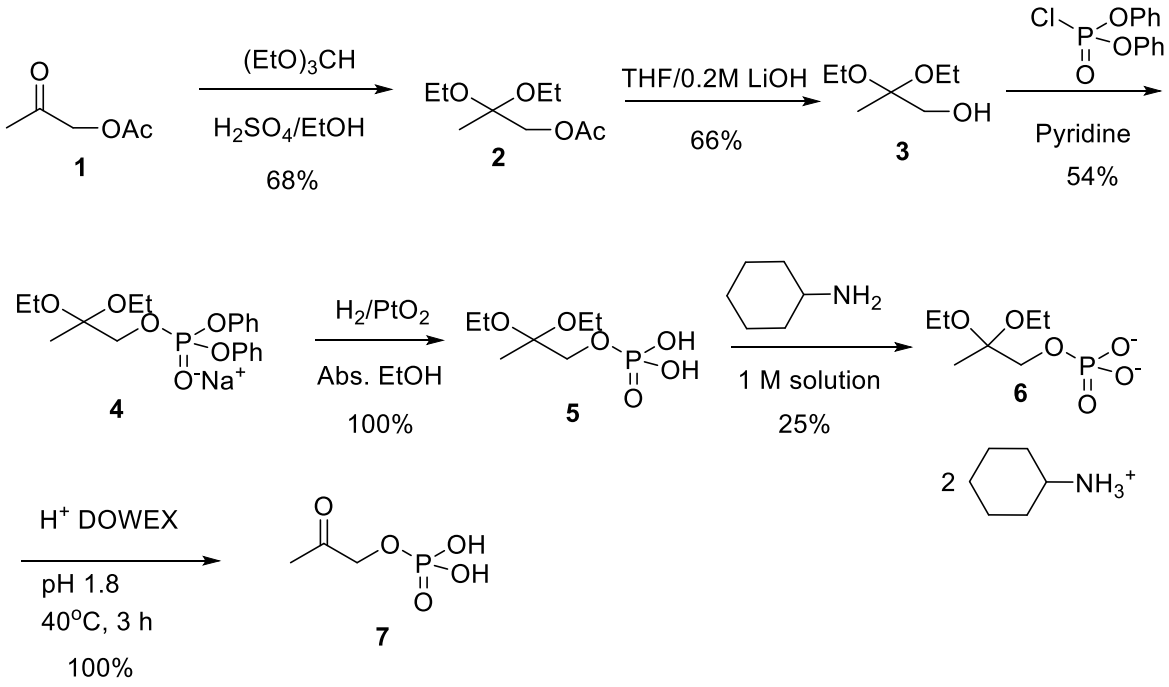


Scheme S1: General synthesis of penturonic acids

Synthesis of Dihydroxyacetone Phosphate. DHAP was synthesized enzymatically using dihydroxyacetone, glycerol kinase, and ATP (*5*). Typically, 0.23 mmol of dihydroxyacetone (~21 mg) was mixed with a 10% excess of ATP, 0.25 mmol (~143 mg), 0.1 mmols $MgCl_2$, and 4 U of glycerol kinase from *Cellulomonas* sp. (Sigma Aldrich) in 20 mM imidazole, pH 7.4, to a total reaction volume of 2.0 mL. Before the addition of the enzyme, KOH (5 M) was added to adjust the pH to 7.4. The reaction was allowed to proceed for 5 h. Activated charcoal was used to remove the ADP. Typically, ~50 mg of charcoal was added to the filtered reaction solution, mixed by inversion, and then removed by centrifugation and filtration. This is repeated, usually about five times, until the absorbance at 260 nm of a 250-fold diluted sample was less than 0.1. Glycerol kinase was removed by passage through a PALL Nanosep 10K omega centrifugation filter. Aliquots were frozen and stored at -20 °C. The final product, dihydroxyacetone phosphate, was

confirmed by ^{31}P and ^{1}H NMR spectroscopy and compared to commercial standard. The isolated yield after the addition of charcoal is about 50%.

Synthesis of Hydroxyacetone Phosphate. Hydroxyacetone phosphate was synthesized following a previous published procedure with modifications (**Scheme S2**) (*6*).



Scheme S2: Synthesis of hydroxyacetone phosphate. Yields for each step are shown under the arrows.

The diethyl ketal of acetol acetate (**2**) was prepared as described previously by mixing 10 g (86.2 mmol) of acetol acetate (**1**, Sigma Aldrich), triethyl orthophormate (16.7 mL, 87.4 mmol), 1.2 mL of absolute ethanol and 1 drop of concentrated sulfuric acid at room temperature for two days. The resulting solution was neutralized to pH 7.2 by adding of 2N NaOH and fractionated by vacuum distillation; 6.8 g of **2** was collected (b.p. 58-59 ^{o}/4mm). To break acetate ester linkage, **2** (3.0 g) was dissolved in 20 mL of THF and

110 mL 0.2 M LiOH was added, reaction was stirred for 3 h at room temperature. The product was extracted with ether (3 X 40 mL), dried on Na_2SO_4 and ether was removed using a water aspirator yeidling 2.0 g of pure diethyl ketal of hydroxyacetone **3.** **3** is a volatile oil under high pressure (b.p. 62-64 ^oC/8.5 mm) and must be handled carefully while removing solvent (*6*).

Compound **3** (0.72 g, 4.8 mmol, 1 equiv) was dissolved in 5 mL of dry pyridine and chilled on ice to 4 °C. Diphenyl phosphorochloridate (1.3 mL, 6.24 mmol, 1.3 equval) was added dropwise and reaction stored at 4 °C for 12 h. Water was added to destroy excess phosphorochloridate and the solution was concentrated under *vacuo*. The resulting syrup was dissolved in dichloromethane and washed with water, then 0.1 N HCl , and 1 M $NaHCO_3$, and dried on Na_2SO_4. After concentration, the product **4** was purified by silica gel column chromatography (Hexanes/EtOAc, 5:1) yielding 1.0 g (54 %) of diphenyl phosphoryl derivative of acetol diethyl ketal as colorless oil. Fractions were tested by thin-layer chromatography using cerium molybdate stain (Hanessian's Stain) (*7*).

4: ^1H NMR (400 MHz, $CDCl_3$) δ 7.38 - 7.33 (m, 4H), 7.27 - 7.25 (m, 4H), 7.22-7.17 (m, 2H), 4.15 (d, 2H, J = 5.9 Hz), 3.56-3.45 (m, 4H) 1.38 (s, 3H), 1.16 (t, 6H, J = 7.1 Hz). ^{31}P NMR (160 MHz, $CDCl_3$) δ – 11.49 (s). MS: 403.13 Da (M^- plus Na^+). $MS_{expected:}$ 403.13 Da (M^- plus Na^+)

Compound **4** (1.0 g) was resuspended in 45 mL of absolute ethanol and PtO_2 (360 mg) and kept under 1 atm H_2 pressure for 48 h. The reaction was monitored by `^1H NMR. The catalyst was removed by filtration and 5 mL of water was added to the filtrate and the solution was then brought to pH 9.5 with 1M aqueous cyclohexylamine. The solution was stored overnight at -20 °C. Ethanol was removed and crude product was crystalized from acetone/ether solution (25 :1). 250 mg of white crystals of cyclohexylamine salt of acetol phosphate diethyl ketal (**6**) were collected (m.p. 147-149 ^oC (*6*)). Finally, 50 mg of **6** was

dissolved in 5 mL of distilled H_2O and 0.5 g of Dowex H^+ was added to a pH of 1.9. This solution was heated to 40 °C for 3 h. DOWEX beads were removed by filtration and solution was lyophilized. The final product, hydroxyacetone phosphate (**7**), was confirmed by ^{31}P and ^{1}H NMR spectroscopy. Net yield is approximately 6%.

7: ^{1}H NMR (400 MHz, D_2O) δ 4.59 (d, 2H, J = 8.2 Hz), 2.18 (s, 3H).
^{31}P NMR (160 MHz, D_2O) δ + 0.46 (s).

Synthesis of Teturonic acids. D-erythroteturonic acid and D- and L-threoteturonic acid were synthesized following the procedure outlined previously (*8*). This procedure uses lead acetate to catalyze an oxidative decarboxylation of the hexuronic acids (D-glucuronic acid, D-galacturonic acid, and L-galacturonic acid) to yield the respective teturonic acids. D-glucuronic acid and D-galacturonic acids were purchased from Sigma Aldrich. Yields for the decarboxylation reaction are 80%. ^{13}C NMR and ^{1}H NMR spectroscopy were used to verify the final products of the reactions. L-galacturonic acid was unavailable for purchase. To generate L-galacturonic acid, a procedure for the synthesis of D-galacturonic acid was followed with slight modifications (*9, 10*). L-galactose (1 g) was converted to 1,2:3,4-di-*O*-isopropylidene-L-galactopyranose by mixing with 2.4 g of $CuSO_4$ and 0.11 mL concentrated H_2SO_4 in 22 mL of anhydrous acetone. This reaction was allowed to proceed at room temperature for 24 h. The resulting solution was filtered. To the filtrate, calcium hydroxide was added until neutral (about 1 g) and stirred for 5 h. 1,2:3,4-Di-*O*-isopropylidene-L-galactopyranose was purified by flash chromatography using ethyl acetate and hexanes (1:1) as the mobile phase. 1.14 grams of 1,2:3,4-di-*O*-isopropylidene-L-galactopyranose was obtained. 1.14 g of 1,2:3,4-di-*O*-isopropylidene-L-galactopyranose was dissolved into 50 mL of H_2O and 1.18 mL of KOH (7.5 N) was added.

To this solution, 4.31 g of $KMnO_4$ was added over a period of 2 h with the final solution allowed to stir for 24 h. Excess $KMnO_4$ was destroyed by the addition of H_2O_2 and MnO_2 was removed by filtration. After purification, 800 mg of the potassium salt of 1,2:3,4-di-*O*-isopropylidene-L-galacturonate was obtained. 400 mg of 1,2:3,4-di-*O*-isopropylidene-L-galacturonate was hydrolyzed by adding 7 mL of H_2O and Dowex H^+ resin to a pH of 2.0 and heated to 90 °C for 2 h in a pressure vial. Dowex resin was removed by filtration and the solution was neutralized by adding $NaHCO_3$. 400 mg of the sodium L-galacturonate was obtained. ^{13}C NMR and ^1H NMR spectroscopy were used to verify the final products of the reactions. Net yield of the reaction is approximately 37%.

Synthesis of Tartronate Semialdehyde and Glycoaldehyde. Tartronate semialdehyde and glycoaldehyde were synthesized by decarboxylation of dihydroxyfumaric acid following a previous published procedure with slight modifications (*11*). Dihydroxyfumaric acid (250 mg) was added to 10 mL of H_2O. Magnesium chloride (1 mg) was added with stirring and heated to 37 °C for 25 min under nitrogen gas. Unreacted starting material was removed by filtration and the flow through was applied to a 10 mL Dowex (H^+) column and eluted with water. KOH dissolved in anhydrous ethanol was added, dropwise, to the solution until neutral. The precipitate was collected by vacuum filtration and dried by lyophilization. Tartronate semialdehyde was confirmed by ^1H NMR spectroscopy . . The resulting mixture was treated the same as tartronate semialdehyde yielding a yellow syrup. was confirmed by ^1H NMR spectroscopy.

Synthesis of L-threose-4-P, L-arabinose-5-P and D-xylose-5-P. L-threose-4-P was synthesized as previously described for D-threose-4-P by substituting diethyl-L-tartaric acid (Sigma Aldrich) as the starting material for the L- version. (*12*) Synthesis of L-

arabinose-5-P and D-xylose-5-P was followed as described previously with slight modifications (*13*). For the synthesis of D-xylose-5-P, the starting material, 1,2-*O*-isopropylidene-α-D-xylofuranose, was purchased from Sigma-Aldrich. 0.95 g of 1,2-*O*-isopropylidene-α-D-xylofuranose dissolved into 12 mL of anhydrous pyridine. To this solution, 1.1 equivalents of dibenzyl phosphorochloridate (1.9 grams) was slowly added and left at 4 °C for 24 h. 2.0 mL of H_2O was added to destroy the excess dibenzyl phosphorochoridate, followed by 15 mL of toluene and dried under reduced pressure. 15 mL of toluene was added and dried a second time. The product was purified by flash chromatography using ethyl acetate and hexanes (2.5:1) as the mobile phase. 1.1 g of pure product, the dibenzyl protected-5-phospho-1,2-*O*-isopropylidene-α-D-xylofuranose was obtained. To remove the benzyl protecting groups, all of the product (1.1 g) was dissolved into 40 mL of methanol and 150 mg of Pd on carbon was added. The reaction was kept at 10 atm for 6 h. This solution was filtered and was dried under vacuum. For the final deprotection step, 1,2-*O*-isopropylidene-α-D-xylofuranose-5-P (370 mg) was dissolved in 20 mL of H_2O and heated to 45 °C for 24 h, followed by neutralization with $NaHCO_3^-$. The final solution was dried under reduced pressure yielding 350 mg of D-xylose-5-P.

L-arabinose-5-P was synthesized in the same manner except the starting material, 1,2-O-isopropylidene-α-L-arabinofuranose, was unavailable for purchase and needed to be synthesized. The description for synthesis of the D-arabino product was previously published (*14*) and followed exactly has described using L-arabinose, resulting in 1.4 grams of 1,2-*O*-isopropylidene-α-L-arabinofuranose from 5 g of L-arabinose starting material.

Synthesis of D-arabinose-5-P. Synthesis of D-arabinose-5-P closely followed the protocol outlined previously (*15*). In general, 20.6 mg arabinose (0.137 mmol) and 18.7 mg

phosphoenolpyruvate (0.09 mmol) was added to solution containing 2.0 mM ATP, 2.0 mM $MgCl_2$, and 20 U hexokinase and 10 U pyruvate kinase in 50 mM HEPES buffer, pH 7.0. The reaction was allowed to proceed for 18 h and determined to be complete by the disappearance of the resonance for PEP(-0.5 ppm) in the ^{31}P NMR spectrum and appearance of a new resonance for D-arabinose-5-P (4.2 ppm). The resulting D-arabinose-5-P was purified using a DEAE column using the same method as described above. Mass spectrometry showed a peak at 229.01 Da (expected at 229.02 Da).

Table S1. Comparison of activity with FBP aldolase[a]

Substrate	turnover ($\mu M\ s^{-1}$)	% Activity compared to F16BP
Fructose-1,6-bisphoshate[b,c]	23	-
D-xylulonate-1-P	0.2	0.9%
L-*glycero*-L-*galacto*-octulonate-1-P	1.02	4.3%
D-glycero-L-galacto-octulono-1-P	0.32	1.4%
L-glycero-D-altro-octulono-1-Phos	1.1	4.8%
D-*erythro*-L-*galacto*-nonlulonate-1-P	0.22	1.0%
L-threo-L-galacto-nonulono-1-Phos	0.42	1.8%

[a]Reactions were carried out with 0.5 U of FBPA in 50 mM Hepes, pH 7.4 with 1 mM substrate monitored by coupled assay. [b]Reaction was done in triplicate with 1 mM F16BP and 0.05 U of FBPA. [c]Reported activity by Sigma Aldrich for 1 U of aldolase is 67 $\mu M\ s^{-1}$.

Table S2. Substrates screened for Activity with YdjI and DHAP

# of Carbons	Substrate	# of Carbons	Substrate	# of Carbons	Substrate
2	acetaldehyde	**5**	D-ribose	**6**	D-mannose
	glycoaldehyde		D-arabinose		D-rhamnose
	glyoxylate		D-xylose		D-glucose
	glycoaldehyde-P		L-lyxose		D-galactose
3	D/L-glyceraldehyde		L-arabinose		L-idose
	D-glyceraldehyde-3-P		L-xylose		D-gulose
	L-glyceraldehyde-3-P		D-ribose-5-P		D-altrose-6-P
	propanal		D-arabinose-5-P		D-glucose-6-P
	(D/L) tartronate semialdehyde		L-arabinose-5-P		D-galactose-6-P
			D-xylose-5-P		D-mannose-6-P
	D-lactaldehyde		D-riburonic		D-glucuronic
	L-lactaldehyde		D-araburonic		D-galacturonic
	methyl glyoxal		D-xyluronic		D-mannuronic
4	L-threose		L-lyxuronic		L-iduronic
	D-erythrose		L-araburonic		L-guluronic
	L-erythrose				
	L-threoteturonate				
	D-threoteturonate				
	D-erythroteturonate				
	L-threose-4-P				

Activities with YdjI are color-coded. Blue shows substrates with kinetic constants determined. Black shows substrates that have activity but do not have kinetic values. Red shows substrates that did not show measurable activity with YdjI

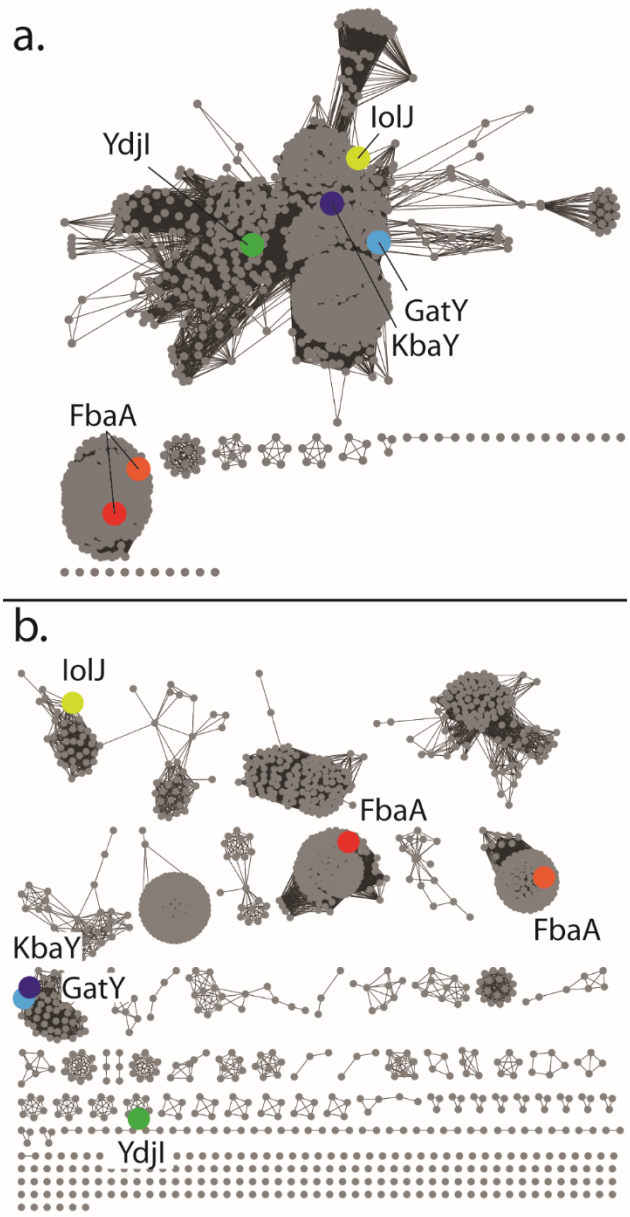


Figure S1: (a) Sequence similarity network of cog0191 at an alignment score of 55. (b) Sequence similarity network of cog0191 at an alignment score of 100. This is considered isofunctional as all known functions have been separated into single clusters. All clusters contain enzymes that share ~60% sequence identity. FbaA (red) is the sequence from *E. coli*. FbaA (orange) is the sequence from *Mycobacterium tuberculosis*. At an alignment score of 100 there are found separate clusters.

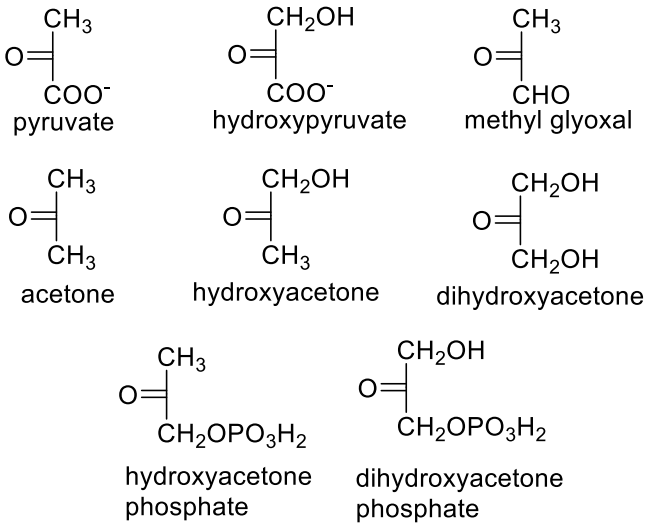


Figure S2: Keto-sugars tested with YdjI for deuterium exchange.

Figure S3: The structures of the 49 aldoses carbohydrates screened with YdjI and DHAP for activity. These are color coded to indicate activity matched in **Table S2**. Blue indicates activity that has been quantified. Red indicates no observed activity. Black indicates activity, but has not been quantified due to slow rate of turnover and/or instability of substrates or products.

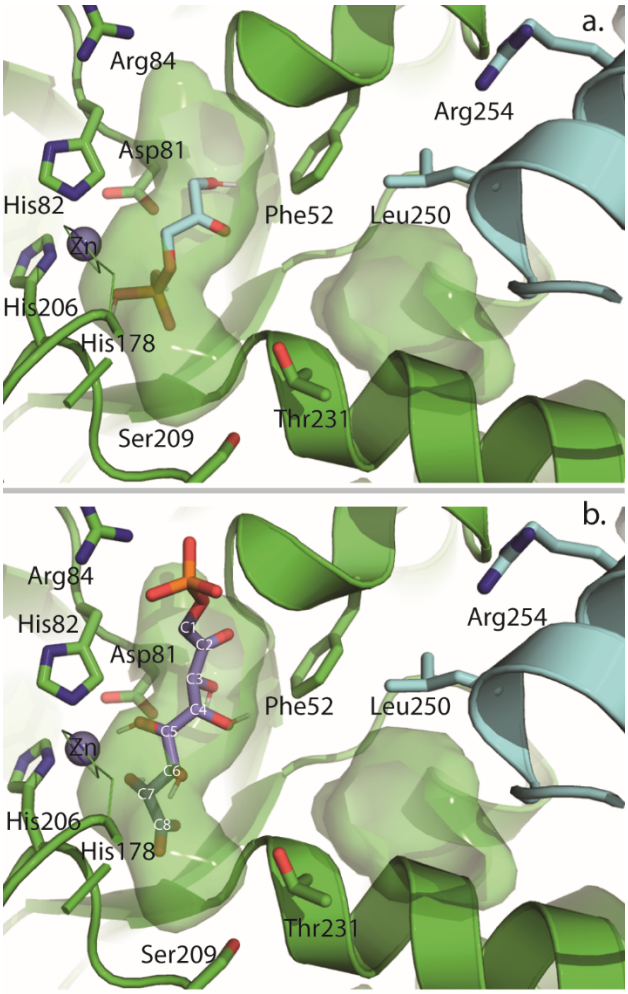


Figure S4: (a) The top pose of DHAP docked using standard AutoDock Vina into the structure of YdjI. The carbonyl and the hydroxyl groups of C2 and C3 are positioned 5.7 and 6.4 Å away from the zinc ion, respectively. (b) The top pose of L-*glycero*-L-*galacto*-octulonate-1-P using standard AutoDock Vina into the structure of YdjI. The carbonyl and the hydroxyl groups of C2 and C3 are positioned 7.5 and 6.3 Å away from the zinc ion, respectively.

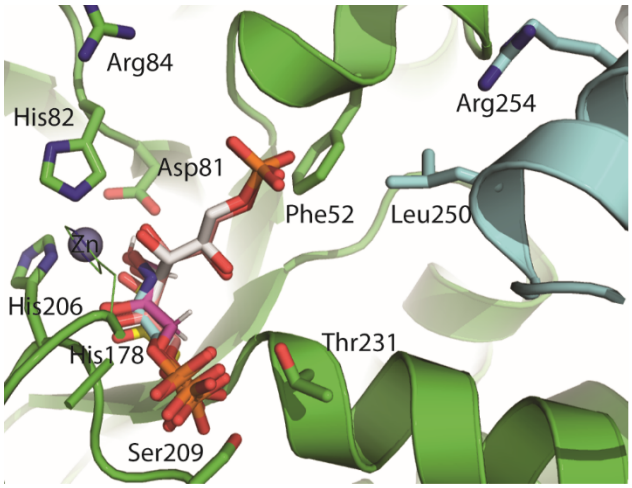


Figure S5. An overlay of bound substrates from five FBPA structures into the YdjI structure. Shown in light blue sticks is a DHAP analog, phosphoglycolohydroxamate bound to FBPA from *E. coli* (Pdb code: 1b57). Shown in purple sticks is DHAP bound to FBPA from *E. coli* (Pdb code: 5vjd). Shown in yellow sticks is DHAP bound to FBPA from *B. anthracis* (Pdb code: 3q94). Shown in pink sticks is FBP bound to FBPA from *G. lamblia* (Pdb code: 3gb6). Shown in white sticks is FBP bound to FBPA from *M. tuberculosis* (Pdb code: 3elf). On average, the distance between the C2 carbonyl group and the zinc ion is 2.7 Å and between the C3 hydroxyl group and the zinc ion is 2.2 Å.

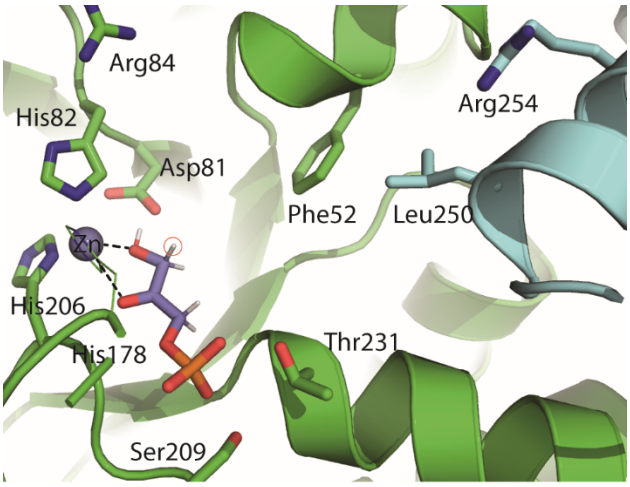


Figure S6. The manual positioning of DHAP into the structure of YdjI such that the carbonyl of C2 and the hydroxyl of C3 are 2.7 Å and 2.2 Å from the zinc, respectively. The phosphate is located 2.5 Å and 2.6 Å from the Thr231 and Ser209 residues. Shown in the red circle of DHAP is the *proS*-H that will be removed during the catalytic mechanism.

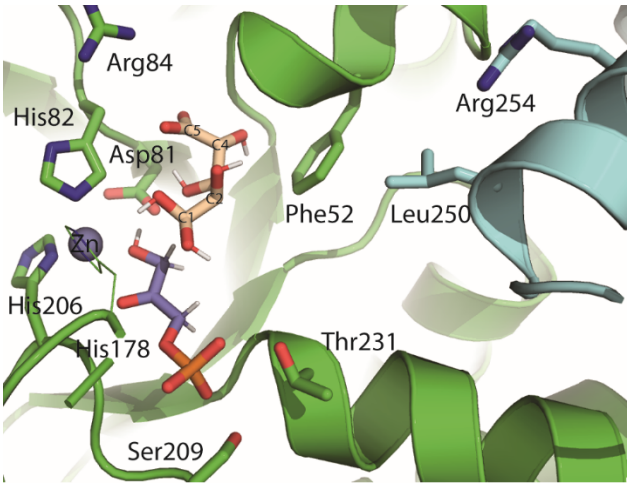


Figure S7. The best pose of the aldo-hydrate of L-araburonate into the structure of YdjI.

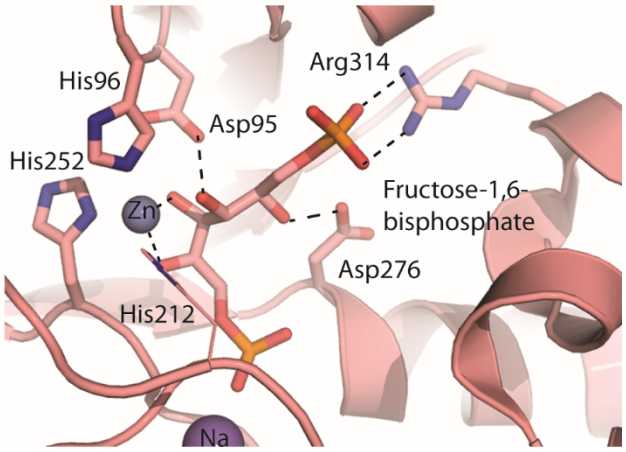


Figure S8. Crystal structure of FBP bound to *M. tuberculosis* FBP aldoalse (PDB: 3elf). The P6 phosphate interacts with Arg314. O4 interacts with Asp95 and O5 hydroxyl interacts with Asp276.

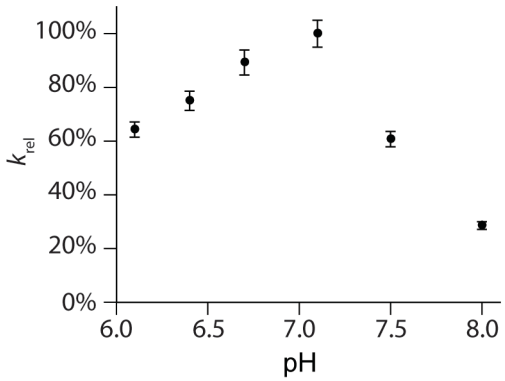


Figure S9. Relative rates of deuterium-exchange from solvent of DHAP (5.0 mM) by YdjI (2 μM) at various pH values (25 mM imidazole/HCl, pH 6.1-8.0).

REFERENCES

(*1*) Wu, J., and Serianni, A. S. (1991) D-penturonic acids: solution studies of stable-isotopically enriched compounds by ^{1}H- and ^{13}C-NMR spectroscopy, *Carbohydr. Res., 210*, 51-70.

(*2*) Brusa, C., Muzard, M., Rémond, C., and Plantier-Royon, R. (2015) β-Xylopyranosides: synthesis and applications, *RSC Advances, 5*, 91026-91055.

(*3*) Masuko, T., Minami, A., Iwasaki, N., Majima, T., Nishimura, S.-I., and Lee, Y. C. (2005) Carbohydrate analysis by a phenol–sulfuric acid method in microplate format, *Anal. Biochem., 339*, 69-72.

(*4*) Nielsen, S. S. (2010) Phenol-sulfuric acid method for total carbohydrates, In *Food Analysis Laboratory Manual* (Nielsen, S. S., Ed.), pp 47-53, Springer US, Boston, MA.

(*5*) Wong, C. H., and Whitesides, G. M. (1983) Synthesis of sugars by aldolase-catalyzed condensation reactions, *J. Org. Chem., 48*, 3199-3205.

(*6*) Sellinger, O. Z., and Neal Miller, O. (1958) The metabolism of acetol phosphate: I. Synthesis, properties and enzymic reduction, *Biochim. Biophy.s Acta, 29*, 74-81.

(*7*) C. Pirrung, M. (2006) Appendix 3: recipes for TLC stains, In *The Synthetic Organic Chemist's Companion*, pp 171-172.

(*8*) Gorin, P. A. J., and Perlin, A. S. (1956) The preparation of D-arabopenturonic acid and D-threo- and D-erythro-teturonic acids, *Can. J. Chem, 34*, 693-700.

(*9*) Niemann, C., and Link, K. P. (1934) Synthesis of the hexuronic acids: IV. The synthesis of D-galacturonic acid from D-galactose, *J. Biol. Chem., 104*, 195-204.

(*10*) Sell, H. M., and Link, K. P. (1938) The preparation of D-galacturonic acid from D-galactose, *J. Am. Chem. Soc., 60*, 1813-1814.

(*11*) Fukunaga, K. (1960) Studies on the metabolism of dihydroxyfumarate, hydrxypyruvate, and their related compounds I. Enzymatic formation of xylulose in rat liver, *J. Biochem., 47*, 741-754.

(*12*) Ahn, M., Pietersma, A. L., Schofield, L. R., and Parker, E. J. (2005) Mechanistic divergence of two closely related aldol-like enzyme-catalysed reactions, *Org. Biomol. Chem., 3*, 4046-4049.

(*13*) Ahn, M., Cochrane, F. C., Patchett, M. L., and Parker, E. J. (2008) Arabinose 5-phosphate analogues as mechanistic probes for *Neisseria meningitidis* 3-deoxy-d-manno-octulosonate 8-phosphate synthase, *Bioorg. Med. Chem., 16*, 9830-9836.

(*14*) Wang, L., Feng, S., An, L., Gu, G., and Guo, Z. (2015) Synthetic and immunological studies of *Mycobacterial* lipoarabinomannan oligosaccharides and their protein conjugates, *J Org Chem, 80*, 10060-10075.

(*15*) Bednarski, M. D., Crans, D. C., DiCosimo, R., Simon, E. S., Stein, P. D., Whitesides, G. M., and Schneider, M. J. (1988) Synthesis of 3-deoxy-D-manno-2-octulosonate-8-phosphate (KDO-8-P) from D-arabinose: generation of D-arabinose-5-phosphate using hexokinase, *Tetrahedron Lett, 29*, 427-430.