Article

# Activating the D-Tagatose Production Capacity of *Escherichia coli* with Structural Insights into C4 Epimerase Specificity

Dileep Sai Kumar Palur, Jayce E. Taylor, Bryant Luu, Ian C. Anderson, Augustine Arredondo, Trevor Gannalo, Bryan A. Skorka, Pamela R. Denish, John Didzbalis, Justin B. Siegel, and Shota Atsumi*

Cite This: *J. Agric. Food Chem.* 2025, 73, 6124–6134

Read Online

| ACCESS | | Metrics & More | | Article Recommendations | | Supporting Information |
|---|---|---|---|---|---|---|

**ABSTRACT:** D-Tagatose, a rare low-calorie sweetener, is ideal for beverages due to its high solubility and low viscosity. Current enzymatic production methods from D-galactose or D-galactitol are limited by reaction reversibility, affecting the yield and purity. This study demonstrates that *Escherichia coli* harbors a thermodynamically favorable pathway for producing D-tagatose from D-glucose via phosphorylation–epimerization–dephosphorylation steps. GatZ and KbaZ, annotated as aldolase chaperones, exhibit C4 epimerization activity, converting D-fructose-6-phosphate to D-tagatose-6-phosphate. Structural analysis reveals active site differences between these enzymes and class II aldolases, indicating functional divergence. By exploiting the strains' inability to metabolize D-tagatose, carbon starvation was applied to remove sugar byproducts. The engineered strains converted 45 g L$^{-1}$ D-glucose to D-tagatose, achieving a titer of 7.3 g L$^{-1}$ and a productivity of 0.1 g L$^{-1}$ h$^{-1}$ under test tube conditions. This approach highlights *E. coli* as a promising host for efficient D-tagatose production.

**KEYWORDS:** *rare sugars, D-tagatose, metabolic engineering*

## INTRODUCTION

Growing awareness regarding the health implication of eating habits has spurred a significant demand for healthier food options.[1] In particular, there has been increased scrutiny regarding the role of dietary sugars in noncommunicable diseases.[2−4] In response to this public health problem, the food industry is adopting artificial sweeteners as a means to lower sugar caloric content without compromising consumer satisfaction, evident in the projected growth of the sugar substitute market, which is expected to reach 29.9 billion USD by 2029.[5]

The current sweetener market is predominantly comprised of nonsugar substitutes, despite recent guidelines released by the World Health Organization advising against using nonsugar sweeteners for weight loss or preventing non-communicable diseases.[6] Rare sugars, characterized by their slight chemical structure variations compared to common sugars like glucose, are increasingly sought after as alternative sweeteners.[7] D-tagatose, a rare sugar and C4 epimer of fructose, has received Generally Recognized as Safe status and is 92% as sweet as sucrose.[8] D-tagatose is defined as a low-calorie sweetener, although there remains a debate regarding its calorific value, which ranges from 1.5 kcal g$^{-1}$ to 3 kcal g$^{-1}$.[7] D-tagatose has high solubility [58% (w/w) at 21 °C] and lower viscosity than sucrose which makes it an ideal sugar substitute in beverages.[9] Moreover, its favorable flavor profile and desirable browning effects further enhance its suitability for use as a sugar substitute.[9] D-tagatose is reported to aid in weight loss and exhibits additional beneficial effects such as antiplaque, noncariogenic, antihalitosis, prebiotic, and anti-biofilm properties.[9]

D-tagatose is found in minute concentration in sterilized powdered milk, hot cocoa, cheese, yogurt, and other dairy products.[10] However, the extraction of D-tagatose from natural sources is not economically feasible due to complex purification steps and low yield.[9] Synthetic methods for D-tagatose production also suffer from similar disadvantages along with the additional use of acids, bases, and catalysts.[10] A potential solution to this problem is the "Izumoring" strategy which makes use of enzymes and serves as a framework for the interconversion of aldohexoses, ketohexoses, and hexitols.[11] In this regard, two such enzymes have been extensively investigated (Figure 1a): 1) L-arabinose isomerase (EC 5.3.1.4), which interconverts D-galactose and D-tagatose, and 2) galactitol 2-dehydrogenase (EC 1.1.1.16), which inter-converts galactitol and D-tagatose.[8,12] However, these methods encounter challenges such as the requirement for purified enzymes, cofactors like NAD$^+$, low thermal stability, or a lack of thermodynamic driving force, resulting in low yields and high cost of production.[13,14] To address these challenges, recent research encapsulated L-arabinose isomerase in *Lactobacillus plantarum*, achieving high yields of D-tagatose. However, this method relied on high-cost D-galactose as the starting material, which is needed to be separated from D-tagatose postproduction, further elevating costs.[15] Alternatively, microbial production was applied to convert lactose to
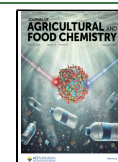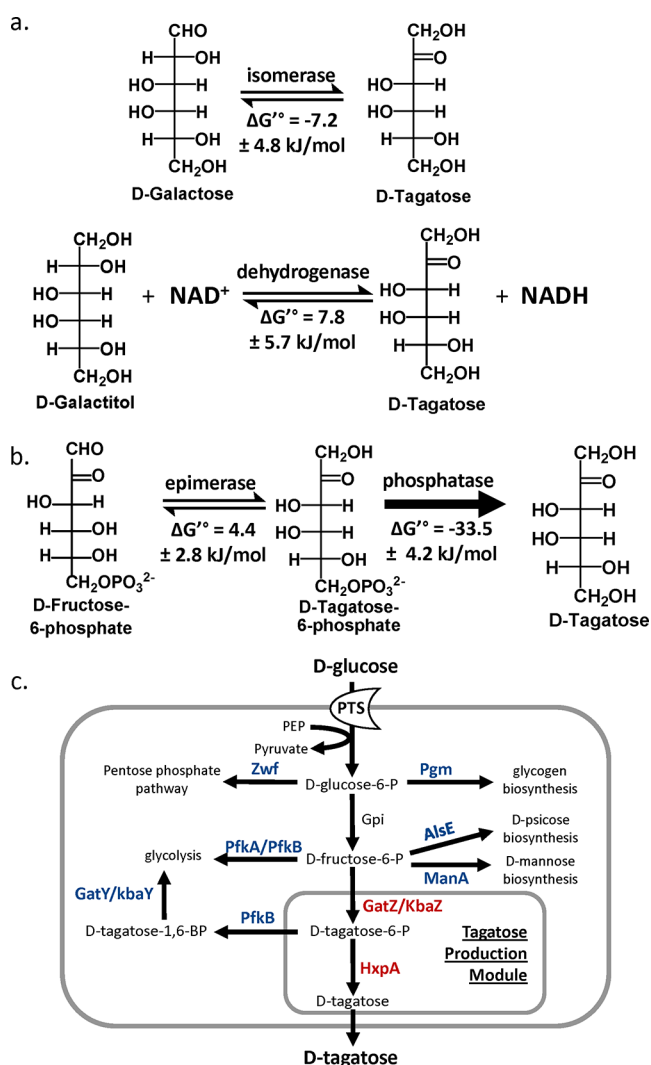
**Figure 1.** Strategies for the biosynthesis of D-tagatose. a) "Izumoring strategies" for D-tagatose production. b) Phosphorylation/dephosphorylation strategy for D-tagatose. c) The proposed biosynthetic production of D-tagatose in *E. coli*. Deleted steps are in blue. Additionally expressed steps are in red. PTS, the phosphotransferase system; GatZ, putative tagatose-1,6-bisphosphate aldolase 2 chaperone; KbaZ, putative tagatose-1,6-bisphosphate aldolase 1 chaperone; HxpA, hexitol phosphatase A; Zwf, NADP⁺-dependent glucose-6-phosphate dehydrogenase; Pgm, phosphoglucomutase; PfkA, 6-phosphofructokinase 1; PfkB, 6-phosphofructokinase 2; AlsE, D-allulose-6-phosphate 3-epimerase; ManA, mannose-6-phosphate isomerase; GatY, tagatose-1,6-bisphosphate aldolase 2; KbaY, tagatose-1,6-bisphosphate aldolase 1.

D-tagatose in *Saccharomyces cerevisiae*, but this method utilizes only D-galactose from lactose and D-glucose is used to maintain cell growth; thus the yield was low.[16] Consequently, the final media contained galactose, galactitol, and D-tagatose, resulting in higher separation and purification costs.[16]

While D-fructose holds potential as a starting material for D-tagatose production, a notable gap exists in nature: the absence of a C4 epimerase capable of directly converting D-fructose to D-tagatose.[17−19] A C4-epimerization activity of the tagaturonate-fructuronate epimerase from *Thermotoga petrophila* was identified.[19] This enzyme converted 700 g L⁻¹ of fructose into 213 g L⁻¹ of tagatose within 2 h.[19] Further optimization through directed evolution improved its thermal stability and

activity, demonstrating its potential for D-tagatose production.[20]

Recently, a phosphorylation and dephosphorylation strategy has surfaced as a promising route for rare sugar production from glucose, presenting an economically advantageous feedstock.[21,22] This pathway involves a pivotal C4 epimerization reaction, transforming D-fructose-6-phosphate (F6P) into D-tagatose-6-phosphate (T6P). Subsequent dephosphorylation of T6P provides a substantial thermodynamic driving force, facilitating efficient D-tagatose production. Utilizing this pathway, whole-cell biocatalysts have been developed, yielding 3.38 g/L of D-tagatose from 10 g/L of maltodextrin.[21] However, this approach requires high cell density and operates with a relatively low substrate loading, posing significant challenges for industrial scalability.[21] Additionally, incomplete conversion leaves residual glucose and fructose in the solution, highlighting inefficiencies in the process and the need for further optimization.[21]

In this study, we discovered that *E. coli* has all of the required enzymes to convert D-glucose to D-tagatose (Figure 1b). D-Glucose is converted to F6P via glycolysis, and the epimerization of F6P to T6P is accomplished by native enzymes: GatZ, a putative tagatose-1,6-bisphosphate aldolase 2 chaperone,[23] or KbaZ, a putative tagatose-1,6-bisphosphate aldolase 1 chaperone.[24] Although these enzymes are annotated as aldolase chaperones,[25] this study demonstrates the C4 epimerization activity of these enzymes in *E. coli*, enabling the conversion of F6P to T6P. Structural analyses of GatZ and KbaZ highlight critical differences in their active sites compared to class II fructose-1,6-bisphosphate aldolases (FBAs), including substitutions of key residues involved in metal coordination and substrate stabilization. These conserved differences across the two enzyme families provide insights into the functional divergence, enabling the unique activity of GatZ and KbaZ. Moreover, the dephosphorylation of T6P can be achieved by hexitol-phosphatase A (HxpA). Production of D-tagatose from D-glucose in *E. coli* was achieved only with the native genes of *E. coli*. D-tagatose production was further improved by additionally expressing the pathway genes and removing competing pathways, including the pentose phosphate pathway, glycogen biosynthesis, glycolysis, D-psicose production pathway, and D-mannose degradation pathway.

## ■ MATERIALS AND METHODS

**Reagents.** All enzymes involved in the molecular cloning experiments were purchased from New England Biolabs (NEB). All synthetic oligonucleotides were synthesized by Integrated DNA Technologies. Sanger Sequencing was provided by Genewiz. D-Mannitol and D-Tagatose were purchased from the Tokyo Chemical Industry. D-Psicose and D-Mannose were purchased from Sigma-Aldrich. D-Glucose and D-Fructose were purchased from Fisher Scientific.

**Strains and Plasmids.** All strains and plasmids used in this study are listed in Tables 1 and 2, respectively. All oligonucleotides are given in Table S1. Plasmids for D-tagatose production were constructed using sequence and ligation-independent cloning (SLIC).[26] The constructed plasmids were verified via sequencing. A guide to the construction of plasmids used in this study is given in Table S2.

Genome modifications such as gene deletion and gene insertion were constructed using CRISPR-Cas9-mediated homologous recombination.[27] Linear DNA repair fragments for gene deletions and insertions were constructed by

**Table 1. Strains Used in This Study**

| Strain | Genotype | Source |
|---|---|---|
| MG1655 | F-lambda-*ilvG*-*rfb*-50 *rph*-1 | 62 |
| AL1050 | MG1655, but attB:*lacI*$^q$ *tetR* spec$^R$ | 63 |
| AL3755 | AL1050, but Δ*pfkA* | This study |
| AL4240 | AL3755, but Δ*zwf* | This study |
| AL4290 | AL4240, but Δ*manA* | This study |
| AL4314 | AL4240, but Δ*alsE* | This study |
| AL4315 | AL4290, but Δ*alsE* | This study |
| AL4330 | AL4315, but Δ*pgm* | This study |
| AL4333 | AL4330, but Δ*gatY* | This study |
| AL4424 | AL4240, but Δ*gatZ* | This study |
| AL4425 | AL4240, but Δ*kbaZ* | This study |
| AL4493 | AL4330, but Δ*kbaY* | This study |
| AL4517 | AL4240, but Δ*kbaY* | This study |
| AL4532 | AL4240, but Δ*gatY* | This study |
| AL4534 | AL4493, but Δ*gatY* | This study |

**Table 2. Plasmids Used in This Study**

| Plasmid | Description | Source |
|---|---|---|
| pAL1950 | pTargetF-*pfkA*, amp$^R$, ColE1 | 22 |
| pAL1958 | pTargetF-*zwf*, amp$^R$, ColE1 | 22 |
| pAL2038 | pTargetF-*pgm*, ampR, ColE1 | 22 |
| pAL2178 | pTargetF-*manA*, ampR, ColE1 | 22 |
| pAL2233 | pTargetF-*alsE*, ampR, ColE1 | 22 |
| pAL2475 | pTargetF-*gatZ*, amp$^R$, ColE1 | This study |
| pAL2480 | $P_{LlacO1}$: *gatZ*, amp$^R$, ColE1 | This study |
| pAL2483 | pTargetF-*kbaZ*, amp$^R$, ColE1 | This study |
| pAL2490 | $P_{LlacO1}$: *gatZ*-*hxpA*, amp$^R$, ColE1 | This study |
| pAL2491 | $P_{LlacO1}$: *gatZ*-*hxpB*, amp$^R$, ColE1 | This study |
| pAL2492 | $P_{LlacO1}$: *gatZ*-*ybiV*, amp$^R$, ColE1 | This study |
| pAL2493 | $P_{LlacO1}$: *gatZ*-*yidA*, amp$^R$, ColE1 | This study |
| pAL2494 | $P_{LlacO1}$: *gatZ*-*yigL*, amp$^R$, ColE1 | This study |
| pAL2495 | $P_{LlacO1}$: *gatZ*-*yihX*, amp$^R$, ColE1 | This study |
| pAL2496 | $P_{LlacO1}$: *gatZ*-*yqaB*, amp$^R$, ColE1 | This study |
| pAL2497 | $P_{LlacO1}$: *kbaZ*-*hxpA*, amp$^R$, ColE1 | This study |
| pAL2189 | pTargetF-*kbaY*, amp$^R$, ColE1 | This study |
| pAL2190 | pTargetF-*gatY*, amp$^R$, ColE1 | This study |
| pAL2568 | pTargetF-*fbaA*, ampR, ColE1 | This study |
| pAL2574 | $P_{LlacO1}$: *gatZ*-*hxpA*\* (1G > A), amp$^R$, ColE1 | This study |
| pAL2575 | $P_{LlacO1}$: *kbaZ*-*hxpA*\* (1G > A), amp$^R$, ColE1 | This study |
| pAL2586 | $P_{LlacO1}$: *fbaA*-*hxpA*, amp$^R$, ColE1 | This study |
| pAL2587 | $P_{LlacO1}$: *gatY*-*hxpA*, amp$^R$, ColE1 | This study |
| pAL2588 | $P_{LlacO1}$: *kbaY*-*hxpA*, amp$^R$, ColE1 | This study |
| pAL2606 | $P_{gadB}$: *gatZ*-*hxpA* (1G > A), amp$^R$, ColE1 | This study |
| pAL2607 | $P_{gadB}$: *kbaZ*-*hxpA* (1G > A), amp$^R$, ColE1 | This study |
| pCas | $P_{cas}$:*cas9* $P_{araB}$:red lacI$^q$ $P_{trc}$:sgRNA pMB1 repA101(Ts) kan$^R$ | addgene #62225 |
| pTargetF | PJ23119:sgRNA-pMB1, spec$^R$, pMB1 | addgene #62226 |

amplifying genomic or plasmid DNA via PCR assembly.[27] Plasmids encoding sgRNA for CRISPR-Cas9-mediated homologous recombination were constructed using Q5 site-directed mutagenesis (New England Biolabs) using pTargetF plasmid (Addgene no. 62226) as a template. All genomic modifications were verified via PCR and sequencing. A guide to the CRISPR-Cas9-mediated gene modifications used in this study is detailed in Table S3.

**Culturing Conditions.** Overnight cultures were grown at 37 °C in 3 mL of Luria–Bertani (LB) medium with appropriate antibiotics. Antibiotic concentrations were as follows: spectinomycin (50 $\mu$g mL$^{-1}$), ampicillin (200 $\mu$g mL$^{-1}$), and kanamycin (50 $\mu$g mL$^{-1}$). M9 minimal media consist of 33.7 mM Na$_2$HPO$_4$, 22 mM KH$_2$PO$_4$, 8.6 mM NaCl, 9.4 mM NH$_4$Cl, 2 mM MgSO$_4$, 0.1 mM CaCl$_2$, A5 trace metals mix (2.86 mg L$^{-1}$ H$_3$BO$_3$, 1.81 mg L$^{-1}$ MnCl$_2$·4H$_2$O, 0.079 mg L$^{-1}$ CuSO$_4$·5H$_2$O, 49.4 $\mu$g L$^{-1}$ Co(NO$_3$)$_2$·6H$_2$O), varying concentrations of glucose, and appropriate antibiotics. M9P media for the D-tagatose production consist of M9 minimal media supplemented with 5 g L$^{-1}$ of yeast extract and appropriate antibiotics. Inducer concentrations are as follows: isopropyl-$\beta$-D-1-thiogalactopyranoside (IPTG) (1 mM). The optical density at 600 nm (OD$_{600}$) was measured with a Synergy H1 hybrid plate reader (BioTek Instruments, Inc.).

**D-Tagatose Production.** For regular cell density production experiments, overnight cultures were inoculated at 1% concentration into 3 mL of M9P media. Cells were grown at 37 °C until the OD$_{600}$ described, then induced with IPTG if necessary, and grown at 30 °C for 24 h.

For high cell density production experiments in M9P media, overnight cultures were inoculated at 1% concentration into 150 mL of M9P media. Cells were grown at 37 °C until OD$_{600}$ 0.4–0.6. Cultures were then induced with IPTG if necessary and grown for another 30 min. Cultures were centrifuged at 2,200 $g$ for 5 min and resuspended in M9P media with IPTG if necessary to a target OD$_{600}$. Cultures were grown at 30 °C.

**Structural Analysis.** To examine the conservation between the GatZ/KbaZ family and the fructose 1,6-bisphosphate aldolase (Fba) family, we performed a sequence-based analysis. First, protein sequences corresponding to each family were retrieved from UniProt.[28] The GatZ/KbaZ family was anchored to the GatZ sequence (UniProt Accession: P0C8J8), and the Fba family was anchored to the Fba sequence (UniProt Accession: A8B2U2).

To examine structural similarity, we employed the Foldseek web application[29] by searching against *E. coli* GatZ (UniProt/ AlphaFoldDB: P0C8J8). Structural similarity was assessed based on the calculated TM-score provided by Foldseek, which utilizes the standard TM-align algorithm.[30] After applying a PDB100 filter, we manually examined the resulting entries to identify structurally similar FBA crystal structures featuring ligands in catalytically relevant poses. The collected sequences were aligned using Clustal Omega, a widely used multiple sequence alignment tool, to ensure reliable and accurate alignment across all family members.[31] Conservation analyses were subsequently conducted using JalView, an integrated platform for sequence visualization and analysis.[32] JalView's conservation scoring feature was applied to calculate the degree of conservation across the alignment, with GatZ and Fba sequences serving as anchors for their respective families. These anchors ensured consistent reference points for assessing family-level conservation.

The Rosetta Molecular Suite was used to dock the ligand fructose-6-phosphate (F6P), into the epimerases GatZ, KbaZ, and the ligand tagatose-6-phosphate (T6P) into the phosphatase HxpA.[33] F6P AM1-BCC partial charges were assigned using the Antechamber suite from AMBER.[34] For the initial protein structures, we used AlphaFold version 2.3.0.[35] The AlphaFold-predicted models were then prepared for docking by using Rosetta Relax.[36] The ligand was then placed in the active site. Rosetta GALigand Dock was then used to dock F6P conformers into each enzyme.[37] Distance and angle constraints were integrated to maintain the D-fructose-6-phosphate and zinc ion in a catalytically competent geometry

for aldolase-based epimerization. A total of 2,500 simulations were run for each enzyme, and the top 10 best-scoring outputs sorted by constraint score, protein—ligand interface energy, and total system energy score were selected for analysis.

Details of the input files used, including constraints, RosettaScripts XML, and ligand parameters, can be found in the Supporting Information. Detailed instructions with sample data can be found on our GitHub at https://github.com/siegel-lab-ucd/D-tagatose-production.

**HPLC Analysis.** Analysis of D-tagatose, D-glucose, D-mannose, D-psicose, D-mannitol, and D-fructose concentrations was performed using HPLC (Shimadzu) equipped with a refractive index detector (RID) 10 A and Rezex RCU-USP sugar alcohol column (Phenomenex). The mobile phase was comprised of 100% Milli-Q water. Samples were run with an injection volume of 1 $\mu$L at a flow rate of 0.5 mL min$^{-1}$ for 7.5 min, with the column oven at 83 °C and RID cell temperature of 40 °C.

To prepare samples for HPLC analysis, 300 $\mu$L of culture was centrifuged at 17,000 $g$ for 5 min. Supernatants were applied to a 0.2 $\mu$m PVDF hydrophilic membrane 96-well filter plate and centrifuged at 2,000 rpm for 2 min into a polystyrene 96 well.

## ■ RESULTS AND DISCUSSION

**Natural D-Tagatose Production Capability of _E. coli_.** The natural ability of _E. coli_ to produce D-tagatose from D-glucose was examined. The production strain, derived from MG1655 (Table 1), did not produce D-tagatose under the production conditions (Figure 2a). To accumulate a critical intermediate, F6P, the _pfkA_ gene encoding for phosphofructokinase A was deleted, but the $\Delta pfkA$ strain did not produce D-tagatose (Figure 2a). The _zwf_ gene encoding glucose-6-phosphate dehydrogenase was deleted to further increase the F6P pool. The $\Delta pfkA$ $\Delta zwf$ strain (AL4240, Table 1) produced 0.21 g L$^{-1}$ D-tagatose in 24 h (Figure 2a) indicating that _E. coli_ naturally has all required enzymes to produce D-tagatose from D-glucose.

**Elucidating the D-Tagatose Production Pathway in _E. coli_.** While _E. coli_ has been used to produce D-tagatose, these efforts largely relied on heterologous enzymes.[21,38,39] We postulated that the natural D-tagatose pathway involves the conversion of D-glucose to F6P through glycolysis, followed by the action of an epimerase to convert F6P into T6P. Subsequently, dephosphorylation of T6P by a phosphatase yields D-tagatose (Figure 1b).

We identified five endogenous enzymes as candidates for possessing C4 epimerase activity to convert F6P into T6P. First, it has been shown that FBA encoded by _fbaA_ exhibits C4 epimerase activity _in vitro_.[18,40] Next, although GatZ is annotated as a putative tagatose-1,6-bisphosphate aldolase 2 chaperone, it was used to facilitate growth of _Agrobacterium tumefaciens_ on galactitol, suggesting GatZ can convert T6P to F6P, which is the reverse reaction used in this study.[41] Additionally, we found KbaZ annotated as a tagatose-1,6-bisphosphate aldolase 2 chaperone has high structural similarity with GatZ with a TM-Score of 0.9707.[30] Finally, the enzymes GatY (tagatose-1,6-bisphosphate aldolase 2) and KbaY (tagatose-1,6-bisphosphate aldolase 1) have high structural similarity with FbaA with TM-scores of 0.92352 and 0.92851, respectively.[25,30]

In order to elucidate the C4 epimerase involved in the production of D-tagatose, each candidate gene was deleted in
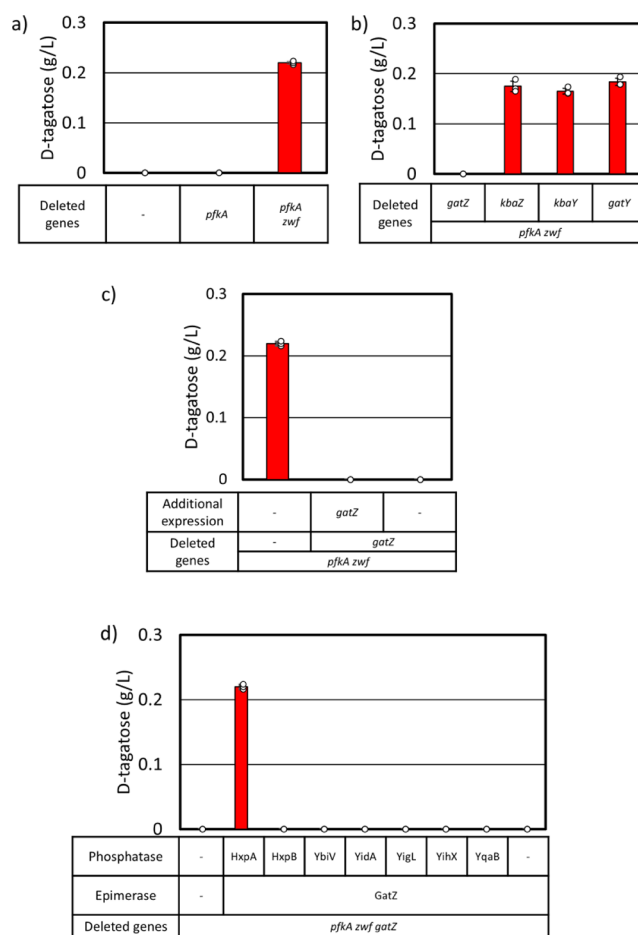


**Figure 2.** D-tagatose production capability of _E. coli_. Cells were grown in M9P media with 10 g L$^{-1}$ glucose at 37 °C to OD$_{600}$ ∼ 0.4, then grown at 30 °C for 24 h. At OD$_{600}$ ∼ 0.4, 1 mM IPTG was added (c, d). a) D-tagatose production in AL1050 (MG1655 + _lacI_$^q$ _tetR_ spec$^R$), AL3755 (AL1050 with $\Delta pfkA$), and AL4240 (AL1050 with $\Delta pfkA$ $\Delta zwf$) (Table 1). b) Effect of each epimerase gene deletion on D-tagatose production in AL4240. c) D-tagatose production in AL4424 (AL1050 with $\Delta pfkA$ $\Delta zwf$ $\Delta gatZ$) with and without additional expression of _gatZ_. (Table 1 and 2). d) Effect of additional expression of each phosphatase gene and _gatZ_ on D-tagatose production in AL4424 (AL1050 with $\Delta pfkA$ $\Delta zwf$ $\Delta gatZ$). Error bars indicate s.d. ($n$ = 3 biological replicates).

AL4240 ($\Delta pfkA$ $\Delta zwf$, Table 1). The _fbaA_ gene could not be deleted as it is essential.[42] Tagatose production was abolished only with the deletion of _gatZ_ (Figure 2b), suggesting GatZ converts F6P to T6P in _E. coli_. However, the expression of _gatZ_ under the IPTG-inducible promoter P$_{LlacO1}$,[43] from an expression plasmid in AL4424 ($\Delta pfkA$ $\Delta zwf$ $\Delta gatZ$, Table 1) did not compensate D-tagatose production (Figure 2c). It has been demonstrated that the expression of _gatZ_ enhances the activity of tagatose-1,6-bisphosphate aldolases,[25] suggesting a potential diversion of T6P away from D-tagatose production toward glycolysis. We hypothesized that additional expression of T6P phosphatase is required to redirect the carbon flux to D-tagatose production. We selected promiscuous phosphatases[44] and conducted screening by expressing hexitol phosphatase B (HxpB), sugar phosphatase YbiV, sugar phosphatase YidA, hexitol phosphatase A (HxpA), $\alpha$-D-glucose-1-phosphate phosphatase YihX, phosphosugar phosphatase YigL, or fructose-1-phosphate phosphatase YqaB along with GatZ in

AL4424 ($\Delta pfkA$ $\Delta zwf$ $\Delta gatZ$, Table 1). Coexpression of GatZ and hexitol phosphatase A (HxpA) restored D-tagatose production in AL4424 (Figure 2d).

The other C4 epimerase candidate genes may not be expressed under our culture conditions. Therefore, each epimerase candidate was expressed along with *hxpA* from an expression plasmid in AL4424 ($\Delta pfkA$ $\Delta zwf$ $\Delta gatZ$, Table 1). The additional expression of *kbaZ* restored D-tagatose production, but that of *fbaA*, *kbaY*, or *gatY* did not (Figure 3a). *gatZ* and *kbaZ* were used for further modifications.
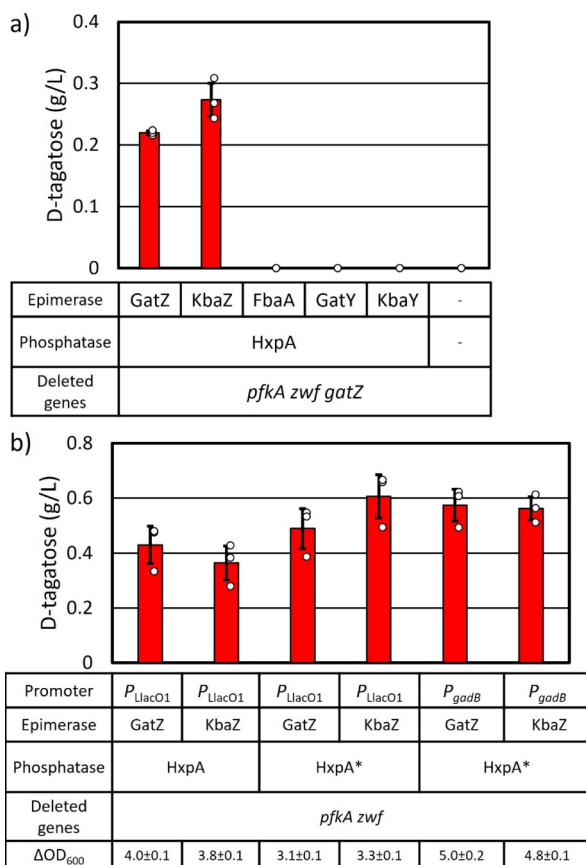
**Figure 3.** Modulating expression of genes for D-tagatose production. Cells were grown in M9P media with 10 g $L^{-1}$ glucose at 37 °C to $OD_{600} \sim 0.4$, then grown at 30 °C for 24 h. At $OD_{600} \sim 0.4$, 1 mM IPTG was added when required. a) Each candidate epimerase gene was expressed along with *hxpA* under $P_{LlacO1}$ promoter (Table 2) in AL4424 (AL1050 with $\Delta pfkA$ $\Delta zwf$ $\Delta gatZ$, Table 1). b) Either *gatZ* or *kbaZ* and *hxpA* were expressed under $P_{LlacO1}$ or $P_{gadB}$. The start codon of *hxpA* was changed from GTG to ATG (*hxpA**). Errors indicate s.d. ($n = 3$ biological replicates).

## Analysis of the Structural Similarity of the C4 Epimerase.

The search for a C4-epimerase to facilitate the conversion of D-fructose to D-tagatose has been an important effort in the sugar industry.[19] A C4-epimerization capability of the tagaturonate-fructuronate epimerase from *Thermotoga petrophila* was identified.[19] However, this reaction is thermodynamically unfavorable, leading to a mixture of D-fructose and D-tagatose. An enzyme capable of converting F6P to T6P from *Agrobacterium tumefaciens* C58 was identified.[45] This finding enabled a phosphorylation-dephosphorylation pathway for D-tagatose production.[21] Another C4 epimerase

was identified in *Sinorhizobium meliloti* which showed structural similarity to GatZ of *E. coli*.[41]

Although the GatZ/KbaZ family of proteins has been studied since 2002,[25] the reaction mechanisms have not been characterized well. Initially, these proteins were classified as fructose 1,6-bisphosphate aldolase chaperones due to their conservation to known fructose 1,6-bisphosphate aldolases, despite lacking aldolase activity under the conditions used in one study.[25] More recently, GatZ has been shown to exhibit epimerase activity, catalyzing the epimerization of T6P into F6P[41] as well as the reverse reaction from F6P to T6P.[21,46]

AlphaFold2[35] was used to predict the three-dimensional structures of GatZ and KbaZ as part of an investigation into their interactions with F6P. Rosetta[33] was used to model and simulate their interactions (Figure 4). Epimerization at the C4 position is predicted to occur through a zinc-dependent retro-aldol/aldol mechanism, similar to that of known C4 epimerases such as L-ribulose-5-phosphate 4-epimerase (AraD),[47] fructose-bisphosphate aldolase class II (FbaA),[18] and other C4 epimerases.[47] In the case of GatZ, the active site contains a $Zn^{2+}$ cofactor coordinated by two conserved histidine residues (His88 and His257) and a glutamic acid residue (Glu171) (Figure 4a,d). Additionally, an aspartate residue (Asp87) is present (Figure 4a,d). The reaction is predicted to be initiated when Asp87 deprotonates the C4 hydroxyl group of F6P, leading to C3−C4 bond cleavage to form dihydroxyacetone phosphate and glyceraldehyde intermediates. Subsequently, an aldol condensation reforms the C3−C4 bond with inverted stereochemistry at C4, producing tagatose 6-phosphate (Figure 4d).

Under the specific conditions described in one study, GatZ was unable to perform the aldolase reaction.[25] Despite this, GatZ shares an evolutionary relationship with common class II FBAs.[25,29,47] Additionally, aldol cleavage by class II fructose-1,6-bisphosphate aldolases proceeds through a similar mechanism (Figure 4c). Given the mechanistic similarities between an aldol cleavage and an aldol/retro-aldol epimerization, we sought to understand why this limitation exists. We compared the predicted catalytic pose of GatZ (Figure 4a) with the crystal structure of a known fructose 1,6-bisphosphate aldolase (FBA) (PDB: 3GAY) complexed with tagatose 1,6-bisphosphate (TBP) (Figure 4b).[48] Due to the limited availability of experimentally determined crystal structures of FBA with its native substrate fructose 1,6-bisphosphate (FBP), we opted to use 3GAY complexed with TBP. The structural similarities between TBP and FBP result in them adopting virtually identical binding poses within the enzyme's active site. Therefore, comparing the predicted catalytic pose of GatZ with the TBP-bound crystal structure is appropriate, allowing us to note several key differences.

Due to their high structural similarity, GatZ was chosen to represent the GatZ/KbaZ family within the main body for clarification. In FBA, metal coordination is carried out by three conserved histidine residues (His84, His178, and His210; Figure 4b). In contrast, in GatZ, metal coordination is carried out by two conserved histidine residues and a glutamic acid residue (His88, His257, and Glu171) (Figure 4a). Additionally in FBA, C3-OH is stabilized by hydrogen bonding carried out by Asp253 and Gln48. Hydrogen bonding stabilizing C3-OH has been lost, and the corresponding residues have been replaced with a lysine (Lys279) and a glutamic acid (Glu44) in GatZ. Finally, the FBA stabilizes the catalytic aspartic acid through an asparagine (Asn24), while in GatZ it has been
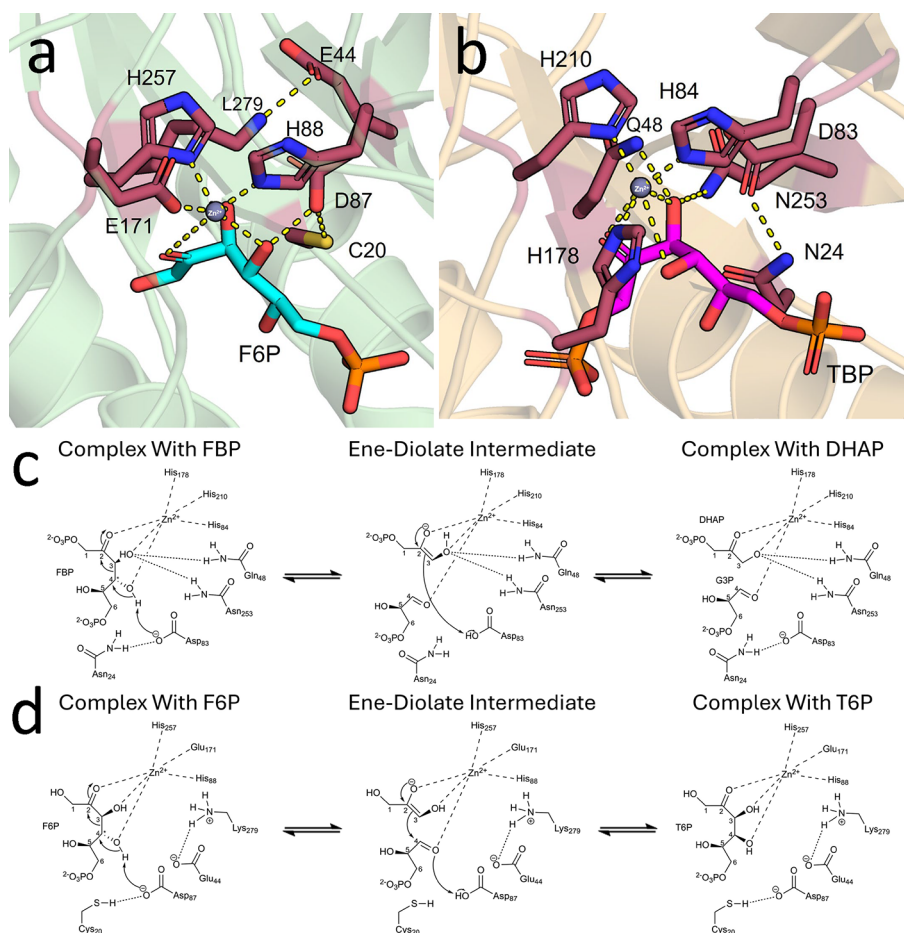
**Figure 4.** Structure analysis of GatZ. a) Predicted substrate-binding mode of the AlphaFold structure of GatZ with F6P. b) Substrate-binding model of FBA (PDB: 3GAY) with D-tagatose-1,6-diphosphate. $Zn^{2+}$ ions are shown as gray spheres in the structural models in (a) and (b). c) Reaction mechanism of FBA catalyzing the conversion of fructose-1,6-bisphosphate to dihydroxyacetone phosphate (DHAP) and glyceraldehyde-3-phosphate (G3P). d) Reaction mechanism of GatZ catalyzing the conversion of F6P to T6P.

replaced by a cysteine (Cys20). To confirm that these differences are unique to these families, sequences of fructose-1,6-bisphosphate (FBP) aldolase class II FbaA (IPR050246) and the GatZ/KbaZ-like family (IPR012062) were aligned using EMBL-EBI's Clustal Omega.[31] These differences are highly conserved across both families (Table S4) and are key differentiating factors between the two families. The observed differences between these two active sites do not clearly explain their differential preference toward one reaction or the other. Our computational analysis has generated a testable hypothesis for the GatZ/KbaZ epimerization mechanism. To evaluate the molecular determinants underlying their unique catalytic properties, further experimental validation is essential. Future studies employing techniques such as site-directed mutagenesis of key active site residues and molecular dynamics simulations will be critical for providing evidence that supports our proposed mechanism and for elucidating the evolutionary adaptations responsible for their functional divergence from canonical FBAs.[49−52]

**Modulating the Expression of the D-Tagatose Production Pathway Genes.** The native *hxpA* gene has a GTG start codon. In order to increase translation efficiency of *hxpA*, the start codon was changed to ATG (*hxpA\**).[53] This modification improved D-tagatose titers to 0.48 g $L^{-1}$ with the additional expression of *gatZ* and 0.60 g $L^{-1}$ with the

additional expression of *kabZ* (Figure 3b). The $P_{LlacO1}$ promoter was replaced with the $P_{gadB}$ promoter which is a stationary phase promoter and ∼100 times stronger than $P_{LlacO1}$.[22] The D-tagatose productivity with $P_{gadB}$ was similar to that with $P_{LlacO1}$ (Figure 3b).

**Removing Competing Pathways.** AL4240 (Δ*pfkA* Δ*zwf*, Table 1) with production plasmid produced 0.8 g $L^{-1}$ D-mannose and 1.1 g $L^{-1}$ D-psicose along with D-tagatose (Figure S1). The *manA* gene encodes mannose-6-phosphate isomerase, which interconverts F6P and mannose-6-phosphate,[54] diverting carbon flux toward D-mannose production. The *alsE* gene encodes D-allulose 6-phosphate 3-epimerase, which interconverts F6P and D-psicose-6-phosphate, is responsible for the production of D-psicose.[22] The deletion of *manA* in AL4240 (Δ*pfkA* Δ*zwf*, Table 1) resulted in the creation of AL4290 (Δ*pfkA* Δ*zwf* Δ*manA*, Table 1). This deletion improved D-tagatose production to 0.6 g $L^{-1}$ and eliminated D-mannose production (Figures 5, S1). Similarly, the deletion of *alsE* in AL4240 (Δ*pfkA* Δ*zwf*, Table 1) resulted in AL4314 (Δ*pfkA* Δ*zwf* Δ*alsE*, Table 1). This deletion improved D-tagatose production to 0.67 g $L^{-1}$ and eliminated D-psicose production (Figures 5, S1). Both of these genes were deleted in AL4240 (Δ*pfkA* Δ*zwf*, Table 1), generating AL4315 (Δ*pfkA* Δ*zwf* Δ*alsE* Δ*manA*, Table 1). AL4315 (Δ*pfkA* Δ*zwf* Δ*alsE* Δ*manA*, Table 1) with the production plasmid produced

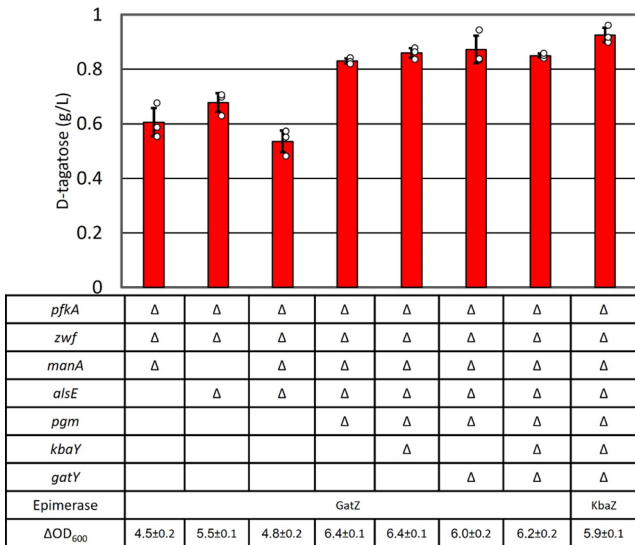| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *pfkA* | Δ | Δ | Δ | Δ | Δ | Δ | Δ | Δ |
| *zwf* | Δ | Δ | Δ | Δ | Δ | Δ | Δ | Δ |
| *manA* | Δ | | Δ | Δ | Δ | Δ | Δ | Δ |
| *alsE* | | Δ | Δ | Δ | Δ | Δ | Δ | Δ |
| *pgm* | | | | Δ | Δ | Δ | Δ | Δ |
| *kbaY* | | | | | | Δ | | Δ | Δ |
| *gatY* | | | | | | | Δ | Δ | Δ |
| Epimerase | GatZ | | | | | | | | KbaZ |
| ΔOD$_{600}$ | 4.5±0.2 | 5.5±0.1 | 4.8±0.2 | 6.4±0.1 | 6.4±0.1 | 6.0±0.2 | 6.2±0.2 | 5.9±0.1 |

**Figure 5.** The effect of gene deletions on D-tagatose production. Cells were grown in M9P media with 10 g L$^{-1}$ glucose to OD$_{600}$ ~ 0.4 at 37 °C, then grown at 30 °C for 24 h. Errors indicate s.d. (*n* = 3 biological replicates).

0.53 g L$^{-1}$ D-tagatose (Figure 5) and did not produce D-psicose and D-mannose (Figure S1).

*E. coli* can store excess glucose in the form of glycogen.[55] The *pgm* gene encoding phosphoglucomutase converts G6P to glucose-1-phosphate.[56] The *pgm* gene was deleted in AL4315 (Δ*pfkA* Δ*zwf* Δ*alsE* Δ*manA*, Table 1), generating AL4330 (Δ*pfkA* Δ*zwf* Δ*alsE* Δ*manA* Δ*pgm*, Table 1). AL4330 with the production plasmid produced 0.82 g of L$^{-1}$ D-tagatose (Figure 5).

T6P can be converted to D-tagatose 1,6-bisphosphate by tagatose-6-phoshate kinase (PfkB) that is subsequently converted to glycerone phosphate and D-glyceraldehyde 3-phoshate aldolase (GatY and KbaY).[25] The *pfkB* gene cannot be deleted as the *pfkA* gene was deleted in the strains. It has been shown that overexpression of *gatZ* enhances tagatose-1,6-bisphosphate aldolase activity in *E. coli*.[25] Activated tagatose-1,6-bisphosphate aldolase activity would divert carbon flux away from D-tagatose production and toward glycolysis. The *kbaY* gene was deleted in AL4330 (Δ*pfkA* Δ*zwf* Δ*alsE* Δ*manA* Δ*pgm*, Table 1), generating AL4493 (Δ*pfkA* Δ*zwf* Δ*alsE* Δ*manA* Δ*pgm* Δ*kbaY*, Table 1). Similarly, *gatY* was deleted in AL4330 (Δ*pfkA* Δ*zwf* Δ*alsE* Δ*manA* Δ*pgm*, Table 1), generating AL4533 (Δ*pfkA* Δ*zwf* Δ*alsE* Δ*manA* Δ*pgm* Δ*gatY*, Table 1). Both *gatY* and *kbaY* were deleted in AL4330 (Δ*pfkA* Δ*zwf* Δ*alsE* Δ*manA* Δ*pgm*, Table 1), generating AL4534 (Δ*pfkA* Δ*zwf* Δ*alsE* Δ*manA* Δ*pgm* Δ*gatY* Δ*kbaY*, Table 1). However, these deletions did not improve D-tagatose production (Figure 5).

**D-Tagatose Production Under High Culture Density Conditions.** Ensuring production performance at high cell densities is vital for obtaining a comprehensive understanding of production capabilities.[57,58] To address potential production constraints arising from glucose availability, AL4534 with pAL2606 (P$_{gadB}$: *gatZ-hxpA**) or pAL2607 (P$_{gadB}$: *kbaZ-hxpA**, Tables 1 and 2) were cultured under high cell density conditions for 24 h. Both these strains consumed 40 g L$^{-1}$ glucose in 24 h (Figure S2). Based on this result, tagatose production at high cell density condition was tested with 40 g L$^{-1}$ glucose feeding each day for 2 days. After the first 24 h,

AL4534 with pAL2606 (*gatZ*) and pAL2607 (*kbaZ*) produced 3.2 g L$^{-1}$ and 3.4 g L$^{-1}$ D-tagatose, respectively (Figure 6a).
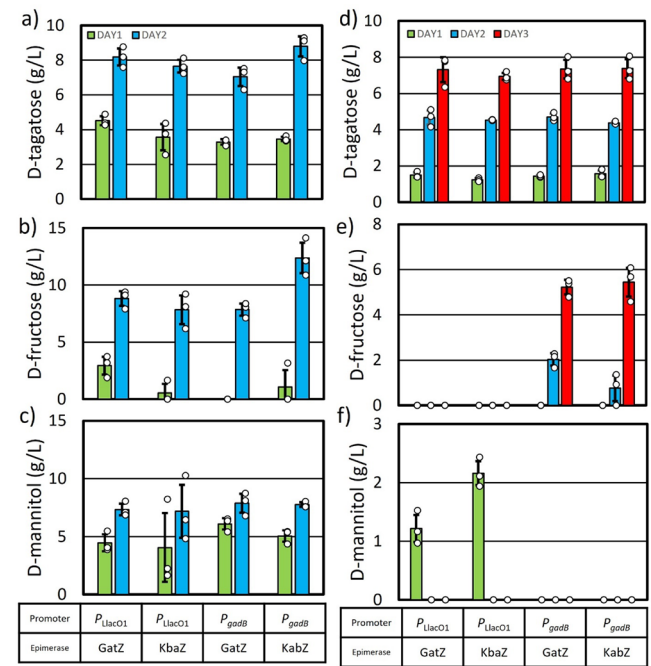


**Figure 6.** High cell density D-tagatose production. Cultures were grown in M9P media with 40 g L$^{-1}$ (a−c) and 15 g L$^{-1}$ (d−f) glucose concentration at 37 °C until an OD$_{600}$ of ~0.4−0.6. Induced with 1 mM IPTG if necessary and grown for a further 30 min. Cultures were then spun down and resuspended in M9P media with 40 g L$^{-1}$ (a−c) and 15 g L$^{-1}$ (d−f) glucose, induced with 1 mM IPTG if necessary, to an OD$_{600}$ of ~10 and grown at 30 °C. Each day M9P media with 40 g L$^{-1}$ (a−c) and 15 g L$^{-1}$ (d−f) glucose was added to the production media. (a and d) D-tagatose production, (b and e) D-fructose production, and (c and f) D-mannitol production. Error bars indicate s.d. (*n* = 3 biological replicates).

These strains also formed D-fructose and D-mannitol as side products (Figure 6b,c). A bolus of glucose was added to each culture, bringing the media glucose concentration back up to 40 g L$^{-1}$. After an additional 24 h of culturing, AL4534 with pAL2606 (*gatZ*) and pAL2607 (*kbaZ*) produced 7.0 g L$^{-1}$ and 8.7 g L$^{-1}$ D-tagatose, respectively (Figure 6a). D-fructose and D-mannitol production also increased (Figure 6a). This observation suggested the low C4 epimerase activity of GatZ and KbaZ causes the accumulated F6P to redirect the carbon flux toward D-fructose and D-mannitol. D-mannitol production is possible through the conversion of F6P to D-mannitol-1-phosphate (Mtl1P) via the enzyme mannitol-1-phosphate 5-dehydrogenase, encoded by the gene *mtlD*.

HxpA shows substrate promiscuity toward F6P and Mtl1P.[44,59] To minimize the expression of *hxpA*, we reverted to the weaker P$_{LlacO1}$ promoter plasmid system, pAL2574 (P$_{LlacO1}$: *gatZ-hxpA**) and pAL2575 (P$_{LlacO1}$: *kbaZ-hxpA**, Table 2). AL4534 with pAL2574 (*gatZ*) and pAL2575 (*kbaZ*) produced 4.5 g L$^{-1}$ and 3.5 g L$^{-1}$ D-tagatose, respectively, after 24 h (Figure 6a). After a bolus of glucose and an additional 24 h, these strains produced 8.1 g L$^{-1}$ and 7.6 g L$^{-1}$ D-tagatose, respectively. Additionally, 8.8 g L$^{-1}$ and 7.8 g L$^{-1}$ D-fructose and 7.3 g L$^{-1}$ and 7.1 g L$^{-1}$ D-mannitol were also produced (Figure 6b,c). Lowering the expression levels of *hxpA* resulted

in similar D-tagatose titers, but D-fructose and D-mannitol were still produced.

**Carbon Starvation Strategy to Eliminate Side Product Formation.** Strain AL4534 showed an inability to assimilate D-tagatose under the production conditions. Additionally, it has been shown that various *E. coli* strains cannot grow on D-tagatose,[60] although *E. coli* is capable of redirecting phosphorylated D-tagatose toward glycolysis.[61]

Carbon starvation would be an effective approach to reduce side product formation for D-tagatose production. Under D-glucose-deprived conditions, the strain cannot assimilate the produced D-tagatose, ensuring its retention in the medium. Meanwhile, it can still consume other byproducts as alternative carbon sources, reducing sugar byproduct accumulation and enhancing product purity. This approach improves process selectivity and yield, making it a valuable strategy for optimizing microbial D-tagatose production. AL4534 with pAL2574 ($P_{LlacO1}$, *gatZ*), pAL2575 ($P_{LlacO1}$, *kbaZ*), pAL2606 ($P_{gadB}$, *gatZ*), and pAL2607 ($P_{gadB}$, *kbaZ*, Table 2) were employed for D-tagatose production under high-density conditions, supplemented with daily additions of 15 g L$^{-1}$ glucose (Figure 6d−f). Within 24 h, the strains produced 1.5 g L$^{-1}$, 1.2 g L$^{-1}$, 1.4 g L$^{-1}$, and 1.5 g L$^{-1}$ of D-tagatose, respectively (Figure 6d). Notably, undetectable levels of D-fructose were formed, and AL4534 with pAL2574 ($P_{LlacO1}$, *gatZ*), pAL2575 ($P_{LlacO1}$, *kbaZ*) produced 1.2 g L$^{-1}$ and 2.1 g L$^{-1}$ of D-mannitol, respectively (Figure 6e,f). Upon additional glucose supplementation, D-tagatose titers increased to 4.6 g L$^{-1}$, 4.5 g L$^{-1}$, 4.7 g L$^{-1}$, and 4.3 g L$^{-1}$, respectively, after 48 h (Figure 6d). AL4534 with pAL2574 ($P_{LlacO1}$, *gatZ*), pAL2575 ($P_{LlacO1}$, *kbaZ*) utilized the produced D-mannitol. However, AL4534 with pAL2606 ($P_{gadB}$, *gatZ*) and pAL2607 ($P_{gadB}$, *kbaZ*) produced no D-mannitol, with 2.0 g L$^{-1}$ and 0.7 g L$^{-1}$ of D-fructose persisting in the media after 48 h (Figure 6e,f). Following another 15 g L$^{-1}$ glucose addition and 78 h of cultivation, they produced 7.3 g L$^{-1}$, 6.9 g L$^{-1}$, 7.3 g L$^{-1}$, and 7.3 g L$^{-1}$ of D-tagatose, respectively (Figure 6d). While AL4534 with pAL2574 ($P_{LlacO1}$, *gatZ*), pAL2575 ($P_{LlacO1}$, *kbaZ*) produced undetectable side products, AL4534 with pAL2606 ($P_{gadB}$, *gatZ*) and pAL2607 ($P_{gadB}$, *kbaZ*) produced 5.2 g L$^{-1}$ and 5.4 g L$^{-1}$ of D-fructose, respectively (Figure 6e). The combination of a low glucose condition (15 g L$^{-1}$)) and a weaker promoter ($P_{LlacO1}$) enabled D-tagatose production without side products by the end of the production period. Although side products were initially formed, the strain efficiently consumed them, highlighting the effectiveness of our carbon starvation strategy in optimizing product formation.

Herein, our findings reveal that *E. coli*, with only the knockout of *pfkA* and *zwf* genes, possesses the inherent capacity to convert D-glucose into D-tagatose. We further demonstrate that GatZ and KbaZ are capable of converting F6P to D-tagatose-6-phosphate. Leveraging the substrate promiscuity of native phosphatases increased the level of D-tagatose production. By eliminating competing pathways and overexpression of native *E. coli* genes, we engineered a strain capable of D-tagatose production through a thermodynamically favorable biosynthetic pathway from easily accessible feedstocks. Our strain's inability to utilize D-tagatose and their efficiency in consuming all D-glucose and side products present in the media reduce downstream purification requirements. Our results suggest that in addition to maintaining a low glucose concentration, carbon starvation during production prevents side product formation. Our final strain achieved conversion of 45 g L$^{-1}$ D-glucose to D-tagatose, reaching a titer of 7.3 g L$^{-1}$, a productivity of 0.1 g L$^{-1}$ h$^{-1}$ without formation of any major side products.

## ■ ASSOCIATED CONTENT

### Data Availability Statement

The data sets generated in this study are available from the corresponding author on reasonable request.

### ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jafc.4c12842.

Table S1, which lists the oligonucleotides used in this study; Table S2, a plasmid construction guide; and Table S3, a guide for CRISPR-Cas9-mediated gene deletions and insertions; Table S4 presents the structural differences between the aldolase and epimerase enzyme families; Figure S1 shows the effect of gene deletion on D-mannose and D-psicose production, while Figure S2 illustrates glucose consumption of AL4534 under high-cell-density conditions; additionally, Rosetta XML files are provided (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

**Shota Atsumi** − *Department of Chemistry and Biochemistry, Molecular, Cellular, and Developmental Graduate Group, University of California, Davis, Davis, California 95616, United States;* ⓞ orcid.org/0000-0002-5782-4196; Email: satsumi@ucdavis.edu

### Authors

**Dileep Sai Kumar Palur** − *Department of Chemistry, University of California, Davis, Davis, California 95616, United States*

**Jayce E. Taylor** − *Department of Chemistry, University of California, Davis, Davis, California 95616, United States*

**Bryant Luu** − *Biochemistry, Molecular, Cellular, and Developmental Graduate Group, University of California, Davis, Davis, California 95616, United States*

**Ian C. Anderson** − *Integrative Genetics and Genomics, University of California, Davis, Davis, California 95616, United States*

**Augustine Arredondo** − *Department of Chemistry, University of California, Davis, Davis, California 95616, United States; Genome Center, University of California, Davis, California 95616, United States*

**Trevor Gannalo** − *Biochemistry, Molecular, Cellular, and Developmental Graduate Group, University of California, Davis, Davis, California 95616, United States*

**Bryan A. Skorka** − *Biophysics Graduate Group, University of California, Davis, Davis, California 95616, United States*

**Pamela R. Denish** − *Department of Chemistry, University of California, Davis, Davis, California 95616, United States*

**John Didzbalis** − *Mars, Incorporated, McLean, Virginia 22101, United States*

**Justin B. Siegel** − *Department of Chemistry and Biochemistry, Molecular, Cellular, and Developmental Graduate Group, University of California, Davis, Davis, California 95616, United States; Integrative Genetics and Genomics and Biophysics Graduate Group, University of California, Davis, Davis, California 95616, United States; Genome Center, University of California, Davis, California 95616, United*

States; Department of Biochemistry and Molecular Medicine, University of California, Davis, Sacramento, California 95616, United States

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jafc.4c12842

## Author Contributions

D.S.K.P., J.E.T, B.L., A.A., P.R.D., J.D., J.B.S., and S.A. designed research; D.S.K.P., J.E.T., and B.L. performed the experiments; D.S.K.P., J.E.T., B.L., I.C.A., A.A., T.G., B.A.S., P.R.D., J.B.S., and S.A. analyzed data; D.S.K.P., B.L., I.C.A., A.A., J.D., J.B.S., and S.A. wrote the manuscript.

## Notes

The authors declare the following competing financial interest(s): The authors declare the following financial interests which may be considered as potential competing interests: Dileep Sai Kumar Palur, Jayce E. Taylor, Bryant Luu, Augustine Arredondo, Ian C. Anderson, Trevor Gannalo, Justin B. Siegel, and Shota Atsumi are inventors on the patent application related to this study. John Didzbalis is employed by Mars, Incorporated, a manufacturer of food and confectionery.

## ■ REFERENCES

(1) Alsubhi, M.; Blake, M.; Nguyen, T.; Majmudar, I.; Moodie, M.; Ananthapavan, J. Consumer Willingness to Pay for Healthier Food Products: A Systematic Review. *Obes Rev.* **2023**, *24* (1), No. e13525.

(2) Blüher, M. Obesity: Global Epidemiology and Pathogenesis. *Nat. Rev. Endocrinol.* **2019**, *15* (5), 288−298.

(3) Lustig, R. H.; Schmidt, L. A.; Brindis, C. D. The Toxic Truth about Sugar. *Nature* **2012**, *482* (7383), 27−29.

(4) Prinz, P. The Role of Dietary Sugars in Health: Molecular Composition or Just Calories? *Eur. J. Clin. Nutr.* **2019**, *73* (9), 1216−1223.

(5) *Sugar Substitutes Market Size, Share, Forecast [Latest]. MarketsandMarkets,* https://www.marketsandmarkets.com/Market-Reports/sugar-substitute-market-1134.html. (accessed 21 Novermber 2024).

(6) Rizas, K. D.; Sams, L. E.; Massberg, S. Non-Nutritional Sweeteners and Cardiovascular Risk. *Nat. Med.* **2023**, *29* (3), 539−540.

(7) Ahmed, A.; Khan, T. A.; Ramdath, D. D.; Kendall, C. W. C.; Sievenpiper, J. L. Rare Sugars and Their Health Effects in Humans: A Systematic Review and Narrative Synthesis of the Evidence from Human Trials. *Nutr. Rev.* **2022**, *80* (2), 255.

(8) Xu, Z.; Li, S.; Feng, X.; Liang, J.; Xu, H. L-Arabinose Isomerase and Its Use for Biotechnological Production of Rare Sugars. *Appl. Microbiol. Biotechnol.* **2014**, *98* (21), 8869−8878.

(9) Oh, D.-K. Tagatose: Properties, Applications, and Biotechnological Processes. *Appl. Microbiol. Biotechnol.* **2007**, *76* (1), 1−8.

(10) Roy, S.; Chikkerur, J.; Roy, S. C.; Dhali, A.; Kolte, A. P.; Sridhar, M.; Samanta, A. K. Tagatose as a Potential Nutraceutical: Production, Properties, Biological Roles, and Applications. *J. Food Sci.* **2018**, *83* (11), 2699−2709.

(11) Granström, T. B.; Takata, G.; Tokuda, M.; Izumori, K. Izumoring: A Novel and Complete Strategy for Bioproduction of Rare Sugars. *J. Biosci. Bioeng.* **2004**, *97* (2), 89−94.

(12) Jagtap, S. S.; Singh, R.; Kang, Y. C.; Zhao, H.; Lee, J.-K. Cloning and Characterization of a Galactitol 2-Dehydrogenase from *Rhizobium Legumenosarum* and Its Application in d-Tagatose Production. *Enzyme Microb. Technol.* **2014**, *58−59*, 44−51.

(13) Dai, D.; Jin, Y.-S. Rare Sugar Bioproduction: Advantages as Sweeteners, Enzymatic Innovation, and Fermentative Frontiers. *Curr. Opin. Food Sci.* **2024**, *56*, 101137.

(14) Noor, E.; Bar-Even, A.; Flamholz, A.; Lubling, Y.; Davidi, D.; Milo, R. An Integrated Open Framework for Thermodynamics of Reactions That Combines Accuracy and Coverage. *Bioinformatics* **2012**, *28* (15), 2037−2044.

(15) Bober, J. R.; Nair, N. U. Galactose to Tagatose Isomerization at Moderate Temperatures with High Conversion and Productivity. *Nat. Commun.* **2019**, *10* (1), 4548.

(16) Liu, J.-J.; Zhang, G.-C.; Kwak, S.; Oh, E. J.; Yun, E. J.; Chomvong, K.; Cate, J. H. D.; Jin, Y.-S. Overcoming the Thermodynamic Equilibrium of an Isomerization Reaction through Oxidoreductive Reactions for Biotransformation. *Nat. Commun.* **2019**, *10* (1), 1356.

(17) Jeon, E. J.; Lee, Y.-M.; Choi, E. J.; Kim, S.-B.; Jeong, K. J. Production of Tagatose by Whole-Cell Bioconversion from Fructose Using Corynebacterium Glutamicum. *Biotechnol. Bioprocess Eng.* **2023**, *28* (3), 419−427.

(18) Lee, S.-H.; Hong, S.-H.; Kim, K.-R.; Oh, D.-K. High-Yield Production of Pure Tagatose from Fructose by a Three-Step Enzymatic Cascade Reaction. *Biotechnol. Lett.* **2017**, *39* (8), 1141−1148.

(19) Shin, K.-C.; Lee, T.-E.; Seo, M.-J.; Kim, D. W.; Kang, L.-W.; Oh, D.-K. Development of Tagaturonate 3-Epimerase into Tagatose 4-Epimerase with a Biocatalytic Route from Fructose to Tagatose. *ACS Catal.* **2020**, *10* (20), 12212−12222.

(20) Liu, Z.; Guo, X.; Xu, Y.; Wu, J. Thermostability Enhancement of Tagatose 4-Epimerase through Protein Engineering and Whole-Cell Immobilization. *J. Agric. Food Chem.* **2025**, *73* (2), 1449−1457.

(21) Dai, Y.; Li, C.; Zheng, L.; Jiang, B.; Zhang, T.; Chen, J. Enhanced Biosynthesis of D-Tagatose from Maltodextrin through Modular Pathway Engineering of Recombinant *Escherichia coli.* *Biochem. Eng. J.* **2022**, *178*, 108303.

(22) Taylor, J. E.; Palur, D. S. K.; Zhang, A.; Gonzales, J. N.; Arredondo, A.; Coulther, T. A.; Lechner, A. B. J.; Rodriguez, E. P.; Fiehn, O.; Didzbalis, J.; Siegel, J. B.; Atsumi, S. Awakening the Natural Capability of Psicose Production in *Escherichia coli.* *Npj Sci Food* **2023**, *7* (1), 54.

(23) Nobelmann, B.; Lengeler, J. W. Sequence of the *Gat* Operon for Galactitol Utilization from a Wild-Type Strain EC3132 of *Escherichia coli.* *Biochim. Biophys. Acta BBA - Gene Struct. Expr.* **1995**, *1262* (1), 69−72.

(24) Brinkkötter, A.; Klöß, H.; Alpert, C.-A.; Lengeler, J. W. Pathways for the Utilization of N-Acetyl-Galactosamine and Galactosamine in *Escherichia coli.* *Mol. Microbiol.* **2000**, *37* (1), 125−135.

(25) Brinkkötter, A.; Shakeri-Garakani, A.; Lengeler, J. W. Two Class II D-Tagatose-Bisphosphate Aldolases from Enteric Bacteria. *Arch. Microbiol.* **2002**, *177* (5), 410−419.

(26) Li, M. Z.; Elledge, S. J. Harnessing Homologous Recombination in Vitro to Generate Recombinant DNA via SLIC. *Nat. Methods* **2007**, *4* (3), 251−256.

(27) Jiang, Y.; Chen, B.; Duan, C.; Sun, B.; Yang, J.; Yang, S. Multigene Editing in the *Escherichia coli* Genome via the CRISPR-Cas9 System. *Appl. Environ. Microbiol.* **2015**, *81* (7), 2506−2514.

(28) The UniProt Consortium.. UniProt: The Universal Protein Knowledgebase. *Nucleic Acids Res.* **2017**, *45* (D1), D158−D169.

(29) Van Kempen, M.; Kim, S. S.; Tumescheit, C.; Mirdita, M.; Lee, J.; Gilchrist, C. L. M.; Söding, J.; Steinegger, M. Fast and Accurate Protein Structure Search with Foldseek. *Nat. Biotechnol.* **2024**, *42* (2), 243−246.

(30) Zhang, Y.; Skolnick, J. TM-Align: A Protein Structure Alignment Algorithm Based on the TM-Score. *Nucleic Acids Res.* **2005**, *33* (7), 2302−2309.

(31) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G. Scalable Generation of High-quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7* (1), 539.

(32) Waterhouse, A. M.; Procter, J. B.; Martin, D. M. A.; Clamp, M.; Barton, G. J. Jalview Version 2—a Multiple Sequence Alignment

Editor and Analysis Workbench. *Bioinformatics* **2009**, *25* (9), 1189−1191.

(33) Leman, J. K.; Weitzner, B. D.; Lewis, S. M.; Adolf-Bryfogle, J.; Alam, N.; Alford, R. F.; Aprahamian, M.; Baker, D.; Barlow, K. A.; Barth, P.; Basanta, B.; Bender, B. J.; Blacklock, K.; Bonet, J.; Boyken, S. E.; Bradley, P.; Bystroff, C.; Conway, P.; Cooper, S.; Correia, B. E.; Coventry, B.; Das, R.; De Jong, R. M.; DiMaio, F.; Dsilva, L.; Dunbrack, R.; Ford, A. S.; Frenz, B.; Fu, D. Y.; Geniesse, C.; Goldschmidt, L.; Gowthaman, R.; Gray, J. J.; Gront, D.; Guffy, S.; Horowitz, S.; Huang, P.-S.; Huber, T.; Jacobs, T. M.; Jeliazkov, J. R.; Johnson, D. K.; Kappel, K.; Karanicolas, J.; Khakzad, H.; Khar, K. R.; Khare, S. D.; Khatib, F.; Khramushin, A.; King, I. C.; Kleffner, R.; Koepnick, B.; Kortemme, T.; Kuenze, G.; Kuhlman, B.; Kuroda, D.; Labonte, J. W.; Lai, J. K.; Lapidoth, G.; Leaver-Fay, A.; Lindert, S.; Linsky, T.; London, N.; Lubin, J. H.; Lyskov, S.; Maguire, J.; Malmström, L.; Marcos, E.; Marcu, O.; Marze, N. A.; Meiler, J.; Moretti, R.; Mulligan, V. K.; Nerli, S.; Norn, C.; ó'Conchúir, S.; Ollikainen, N.; Ovchinnikov, S.; Pacella, M. S.; Pan, X.; Park, H.; Pavlovicz, R. E.; Pethe, M.; Pierce, B. G.; Pilla, K. B.; Raveh, B.; Renfrew, P. D.; Burman, S. S. R.; Rubenstein, A.; Sauer, M. F.; Scheck, A.; Schief, W.; Schueler-Furman, O.; Sedan, Y.; Sevy, A. M.; Sgourakis, N. G.; Shi, L.; Siegel, J. B.; Silva, D.-A.; Smith, S.; Song, Y.; Stein, A.; Szegedy, M.; Teets, F. D.; Thyme, S. B.; Wang, R. Y.-R.; Watkins, A.; Zimmerman, L.; Bonneau, R. Macromolecular Modeling and Design in Rosetta: Recent Methods and Frameworks. *Nat. Methods* **2020**, *17* (7), 665−680.

(34) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25* (9), 1157−1174.

(35) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583−589.

(36) Tyka, M. D.; Keedy, D. A.; André, I.; DiMaio, F.; Song, Y.; Richardson, D. C.; Richardson, J. S.; Baker, D. Alternate States of Proteins Revealed by Detailed Energy Landscape Mapping. *J. Mol. Biol.* **2011**, *405* (2), 607−618.

(37) Park, H.; Zhou, G.; Baek, M.; Baker, D.; DiMaio, F. Force Field Optimization Guided by Small Molecule Crystal Lattice Data Enables Consistent Sub-Angstrom Protein−Ligand Docking. *J. Chem. Theory Comput.* **2021**, *17* (3), 2000−2010.

(38) Liu, W.; Zhang, Z.; Li, Y.; Zhu, L.; Jiang, L. Efficient Production of D-Tagatose via DNA Scaffold Mediated Oxidoreductases Assembly *in Vivo* from Whey Powder. *Food Res. Int.* **2023**, *166*, 112637.

(39) Zhang, G.; Zabed, H. M.; Yun, J.; Yuan, J.; Zhang, Y.; Wang, Y.; Qi, X. Two-Stage Biosynthesis of D-Tagatose from Milk Whey Powder by an Engineered *Escherichia coli* Strain Expressing L-Arabinose Isomerase from *Lactobacillus plantarum*. *Bioresour. Technol.* **2020**, *305*, 123010.

(40) Lee, S.-H.; Hong, S.-H.; An, J.-U.; Kim, K.-R.; Kim, D.-E.; Kang, L.-W.; Oh, D.-K. Structure-Based Prediction and Identification of 4-Epimerization Activity of Phosphate Sugars in Class II Aldolases. *Sci. Rep.* **2017**, *7* (1), 1934.

(41) Kohlmeier, M. G.; White, C. E.; Fowler, J. E.; Finan, T. M.; Oresnik, I. J. Galactitol Catabolism in Sinorhizobium Meliloti Is Dependent on a Chromosomally Encoded Sorbitol Dehydrogenase and a pSymB-Encoded Operon Necessary for Tagatose Catabolism. *Mol. Genet. Genomics* **2019**, *294* (3), 739−755.

(42) Goodall, E. C. A.; Robinson, A.; Johnston, I. G.; Jabbari, S.; Turner, K. A.; Cunningham, A. F.; Lund, P. A.; Cole, J. A.; Henderson, I. R. The Essential Genome of *Escherichia coli* K-12. *mBio* **2018**, *9* (1), No. e02096−17.

(43) Lutz, R.; Bujard, H. Independent and Tight Regulation of Transcriptional Units in *Escherichia coli* Via the LacR/O, the TetR/O and AraC/I1-I2 Regulatory Elements. *Nucleic Acids Res.* **1997**, *25* (6), 1203−1210.

(44) Kuznetsova, E.; Proudfoot, M.; Gonzalez, C. F.; Brown, G.; Omelchenko, M. V.; Borozan, I.; Carmel, L.; Wolf, Y. I.; Mori, H.; Savchenko, A. V.; Arrowsmith, C. H.; Koonin, E. V.; Edwards, A. M.; Yakunin, A. F. Genome-Wide Analysis of Substrate Specificities of the *Escherichia coli* Haloacid Dehalogenase-like Phosphatase Family *. *J. Biol. Chem.* **2006**, *281* (47), 36149−36161.

(45) Wichelecki, D. J.; Vetting, M. W.; Chou, L.; Al-Obaidi, N.; Bouvier, J. T.; Almo, S. C.; Gerlt, J. A. ATP-Binding Cassette (ABC) Transport System Solute-Binding Protein-Guided Identification of Novel d-Altritol and Galactitol Catabolic Pathways in Agrobacterium Tumefaciens C58 *. *J. Biol. Chem.* **2015**, *290* (48), 28963−28976.

(46) Dai, Y.; Zhang, J.; Zhang, T.; Chen, J.; Hassanin, H. A.; Jiang, B. Characteristics of a Fructose 6-Phosphate 4-Epimerase from *Caldilinea Aerophila* DSM 14535 and Its Application for Biosynthesis of Tagatose. *Enzyme Microb. Technol.* **2020**, *139*, 109594.

(47) Allard, S. T. M.; Giraud, M.-F.; Naismith, J. H. Epimerases: Structure, Function and Mechanism: *Cell. Mol. Life Sci.* **2001**, *58* (11), 1650−1665.

(48) Galkin, A.; Li, Z.; Li, L.; Kulakova, L.; Pal, L. R.; Dunaway-Mariano, D.; Herzberg, O. Structural Insights into the Substrate Binding and Stereoselectivity of *Giardia* Fructose-1,6-Bisphosphate Aldolase,. *Biochemistry* **2009**, *48* (14), 3186−3196.

(49) Li, Z.; Hu, Y.; Yu, C.; Fei, K.; Shen, L.; Liu, Y.; Nakanishi, H. Semi-Rational Engineering of D-Allulose 3-Epimerase for Simultaneously Improving the Catalytic Activity and Thermostability Based on D-Allulose Biosensor. *Biotechnol. J.* **2024**, *19* (8), 2400280.

(50) Chen, Q.; Xiao, Y.; Shakhnovich, E. I.; Zhang, W.; Mu, W. Semi-Rational Design and Molecular Dynamics Simulations Study of the Thermostability Enhancement of Cellobiose 2-Epimerases. *Int. J. Biol. Macromol.* **2020**, *154*, 1356−1365.

(51) Zhu, J.; Li, Y.; Wang, J.; Yu, Z.; Liu, Y.; Tong, Y.; Han, W. Adaptive Steered Molecular Dynamics Combined With Protein Structure Networks Revealing the Mechanism of Y68I/G109P Mutations That Enhance the Catalytic Activity of D-Psicose 3-Epimerase From Clostridium Bolteae. *Front. Chem.* **2018**, *6*, 437.

(52) Guo, Z.; Wang, L.; Su, L.; Chen, S.; Xia, W.; André, I.; Rovira, C.; Wang, B.; Wu, J. A Single Hydrogen Bond Controls the Selectivity of Transglycosylation vs Hydrolysis in Family 13 Glycoside Hydrolases. *J. Phys. Chem. Lett.* **2022**, *13* (24), 5626−5632.

(53) Sussman, J. K.; Simons, E. L.; Simons, R. W. *Escherichia coli* Translation Initiation Factor 3 Discriminates the Initiation Codon in Vivo. *Mol. Microbiol.* **1996**, *21* (2), 347−360.

(54) Gao, H.; Yu, Y.; Leary, J. A. Mechanism and Kinetics of Metalloenzyme Phosphomannose Isomerase: Measurement of Dissociation Constants and Effect of Zinc Binding Using ESI-FTICR Mass Spectrometry. *Anal. Chem.* **2005**, *77* (17), 5596−5603.

(55) Sekar, K.; Linker, S. M.; Nguyen, J.; Grünhagen, A.; Stocker, R.; Sauer, U. Bacterial Glycogen Provides Short-Term Benefits in Changing Environments. *Appl. Environ. Microbiol.* **2020**, *86* (9), No. e00049-20.

(56) Eydallin, G.; Viale, A. M.; Morán-Zorzano, M. T.; Muñoz, F. J.; Montero, M.; Baroja-Fernández, E.; Pozueta-Romero, J. Genome-Wide Screening of Genes Affecting Glycogen Metabolism in *Escherichia coli* K-12. *FEBS Lett.* **2007**, *581* (16), 2947−2953.

(57) Crater, J. S.; Lievense, J. C. Scale-up of Industrial Microbial Processes. *FEMS Microbiol. Lett.* **2018**, *365* (13), fny138.

(58) Theisen, M.; Liao, J. C. Industrial Biotechnology: *Escherichia coli* as a Host. In *Industrial Biotechnology Microorganisms*; John Wiley & Sons, Ltd, 2017; pp. 149−181. .

(59) Sévin, D. C.; Fuhrer, T.; Zamboni, N.; Sauer, U. Nontargeted in Vitro Metabolomics for High-Throughput Identification of Novel Enzymes in *Escherichia coli*. *Nat. Methods* **2017**, *14* (2), 187−194.

(60) Joo, Y.; Sung, J.-Y.; Shin, S.-M.; Park, S. J.; Kim, K. S.; Park, K. D.; Kim, S.-B.; Lee, D.-W. A Retro-Aldol Reaction Prompted the

Evolvability of a Phosphotransferase System for the Utilization of a Rare Sugar. *Microbiol. Spectrum.* **2023**, *11* (2), No. e03660−22.

(61) Ha, J.; Kim, D.; Yeom, J.; Kim, Y.; Yoo, S. M.; Yoon, S. H. Identification of a Gene Cluster for D-Tagatose Utilization in *Escherichia coli* B2 Phylogroup. *iScience* **2022**, *25* (12), 105655.

(62) Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H. Construction of *Escherichia coli* K-12 in-Frame, Single-Gene Knockout Mutants: The Keio Collection. *Mol. Syst. Biol.* **2006**, *2*, 2006−2008.

(63) Yoneda, H.; Tantillo, D. J.; Atsumi, S. Biological Production of 2-Butanone in *Escherichia coli*. *ChemSuschem* **2014**, *7* (1), 92−95.