nature chemistry

ARTICLES

PUBLISHED ONLINE: 10 AUGUST 2015 | DOI: 10.1038/NCHEM.2321

Asymmetric assembly of aldose carbohydrates from formaldehyde and glycolaldehyde by tandem biocatalytic aldol reactions

Anna Szekrenyi[1†], Xavier Garrabou[1†], Teodor Parella[2], Jesús Joglar[1], Jordi Bujons[1] and Pere Clapés[1*]

The preparation of multifunctional chiral molecules can be greatly simplified by adopting a route via the sequential catalytic assembly of achiral building blocks. The catalytic aldol assembly of prebiotic compounds into stereodefined pentoses and hexoses is an as yet unmet challenge. Such a process would be of remarkable synthetic utility and highly significant with regard to the origin of life. Pursuing an expedient enzymatic approach, here we use engineered D-fructose-6-phosphate aldolase from *Escherichia coli* to prepare a series of three- to six-carbon aldoses by sequential one-pot additions of glycolaldehyde. Notably, the pertinent selection of the aldolase variant provides control of the sugar size. The stereochemical outcome of the addition was also altered to allow the synthesis of L-glucose and related derivatives. Such engineered biocatalysts may offer new routes for the straightforward synthesis of natural molecules and their analogues that circumvent the intricate enzymatic pathways forged by evolution.

Carbohydrates and glycoconjugates play crucial roles in various intercellular and intracellular functions such as cell–cell recognition, cell adhesion and cell signalling processes involved in bacterial and viral infection, cancer metastasis and inflammatory reactions[1–6]. In addition, carbohydrates contribute to the creation of structural diversity when conjugated with other biomolecules such as lipids, proteins and metabolites[7,8]. Hence, there is growing interest in the preparation of functionally and/or stereochemically diverse carbohydrates and analogues that could be used as biological probes and in the study of their biological roles. This has led to the emergence of a vast array of assays that aim to establish their therapeutic potential.

Carbohydrate synthesis is generally approached via the modification of suitable chiral starting materials. However, this requires cumbersome successive protection and deprotection steps[9]. Alternatively, straightforward and efficient bottom-up strategies for the *de novo* construction of polyoxygenated scaffolds may be devised through the consecutive connection of small building blocks, by C–C bond-forming reactions with the concomitant installation of hydroxylated chiral positions[3,10]. The catalytic asymmetric aldol addition reaction has proved to be an extremely useful carboligation procedure for the generation of polyoxygenated molecules[10–12]. When aiming to recreate the prebiotic synthesis of sugars, simple amino acids and inorganic catalysts were found to assemble glycolaldehyde (**1a**) and formaldehyde (**1b**) into diverse carbohydrates, thereby offering partial control over the sugar size, configuration and enantiomeric ratio[13,14]. Pursuing a synthetically useful bottom-up approach, Northrup and MacMillan synthesized partially protected hexoses by the sequential aldol addition of three synthetic glycolaldehyde equivalents, using direct organocatalytic aldolization and a subsequent metal-catalysed directed Mukaiyama aldol reaction (Fig. 1a)[15,16]. Despite these pioneering contributions, the expedient synthesis of carbohydrates from simple unprotected substrates remains an as yet unmet challenge.

Suitably engineered aldolases are promising candidates in the search for a route towards the direct catalytic assembly of formaldehyde and glycolaldehyde. Aldolases are gaining recognition among chemists as highly efficient asymmetric C–C bond-forming catalysts that act on minimally protected substrates in environmentally benign aqueous media[17,18]. Adequately engineered aldolases may catalyse successive $C_1+C_2+C_2$ or $C_2+C_2+C_2$ connections, thereby affording control over the size and configuration of the sugar. In an early realization of this approach, Wong and co-workers described a sequential double addition of acetaldehyde to 2-chloroacetaldehyde using 2-deoxy-D-ribose 5-phosphate aldolase (DERA), an extremely useful and beneficial reaction in the preparation of an atorvastatin precursor[19,20]. However, the installation of consecutive hydroxyl functions needs a glycolaldehyde donor, which is not how aldolases function in nature. As an alternative, biomimetic approaches based on dihydroxyacetone phosphate (DHAP)-dependent aldolases have been devised to provide C_n+C_3 connections. These require the assistance of a number of enzymes to generate DHAP from C_3 achiral starting materials and further transform the phosphorylated ketose obtained into an aldose by appropriate isomerases (Fig. 1b)[21,22].

An excellent candidate for our goal of synthesizing carbohydrates from simple unprotected substrates is the D-fructose-6-phosphate aldolase from *Escherichia coli* (FSA), a versatile class I aldolase of unknown physiological function. FSA exhibits an unmatched breadth of nucleophile selectivity, accepting glycolaldehyde (**1a**) as well as a number of hydroxylated ketones[23]. Wild-type FSA catalyses stereocontrolled homo- and cross-aldol additions of glycolaldehyde, affording small sugars (for example, L-glyceraldehyde and D-threose) and functionally diverse analogues[24]. Tight control of a single glycolaldehyde addition was invariably observed, hampering the preparation of higher-carbon-content aldoses. The substrate specificity of FSA has been tailored in our laboratory by a structure-guided programme of active-site engineering to expand

[1]Biotransformation and Bioactive Molecules Group, Instituto de Química Avanzada de Cataluña, IQAC-CSIC, Jordi Girona 18-26, 08034 Barcelona, Spain. [2]Depatment Química. Servei de Ressonància Magnètica Nuclear, Universitat Autònoma de Barcelona, Bellaterra, Spain. †Present addresses: Institut für Organische Chemie und Biochemie, Technische Universität Darmstadt, Alarich-Weiss-Strasse 4, 64287 Darmstadt, Germany (A.S.); Laboratory of Organic Chemistry, HCI F322, ETH Hönggerberg, 8093 Zurich, Switzerland (X.G.). *e-mail: pere.clapes@iqac.csic.es

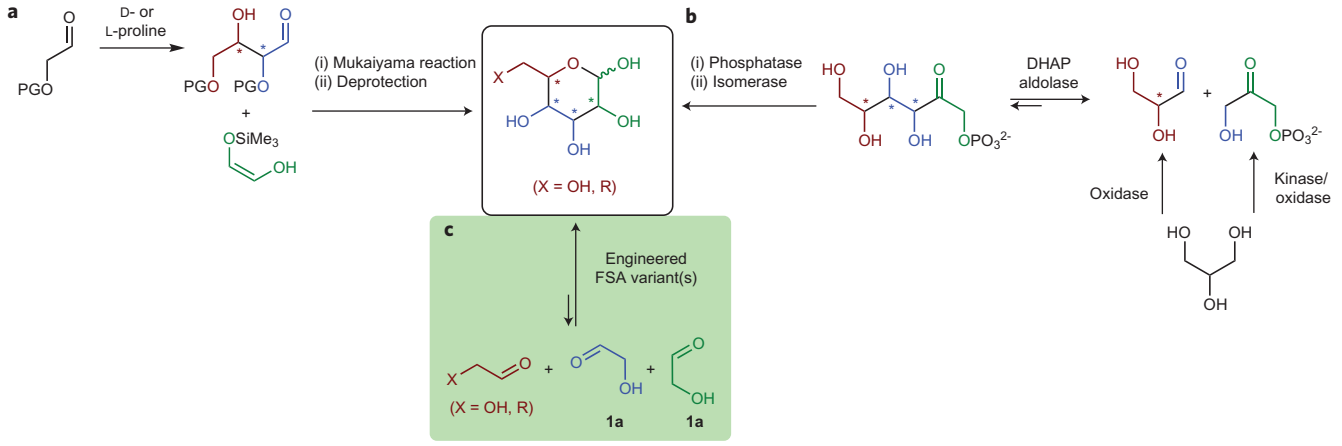


Figure 1 | Bottom-up synthesis of hexoses from achiral starting materials via chemical or biocatalytic routes. a, Combination of organocatalysis and a metal-assisted Mukaiyama aldol addition. **b**, Biomimetic procedure using a combination of kinase, phosphatase, oxidoreductase, aldolase and isomerase activities. **c**, Novel approach based on engineered variants of D-fructose-6-phosphate aldolase from *E. coli* (FSA) as catalysts for the stereoselective trimerization of glycolaldehyde.

its scope of application[25,26]. Mutations at positions L107 and particularly A129 were found to be crucial for its donor activity and selectivity, whereas the A165G mutation enhances the reactivity of poorly reactive aldehydes[25–27]. These studies have provided a toolbox of FSA variants with different synthetic capabilities as a function of the acceptor and donor substrates. In the present work, we capitalize on knowledge gained regarding the activity and selectivity of FSA along the route towards the synthesis of aldose sugars and analogues, via stereoselective glycolaldehyde trimerization as well as its double addition to other acceptor aldehydes (Fig. 1c). Additional engineering efforts have been devoted to expanding the substrate scope and altering the stereochemical course of the additions. In this way, the simple carboligations that possibly enabled sugar synthesis in the prebiotic world could be governed with high selectivity by a set of tailored enzymes.

Results

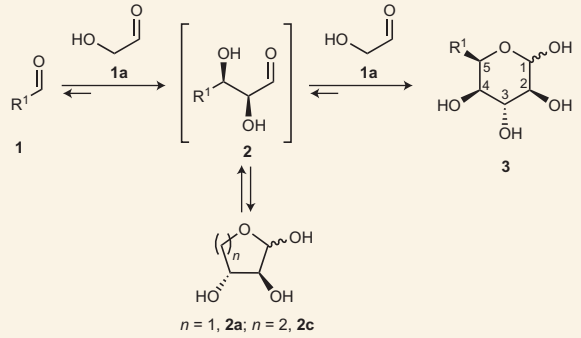
A set of FSA variants with enhanced activity and selectivity towards glycolaldehyde dimerization was initially assayed in the successive C_1+C_2+C_2 connection of 2 equiv. of glycolaldehyde (**1a**) to 1 equiv. of formaldehyde (**1b**). The formation of a double addition product, L-xylose (**3b**), was observed with different variants (Table 1, entries 7–10), and the most advantageous, $FSA^{A129T/A165G}$, was selected for further studies. Under optimized experimental conditions, $FSA^{A129T/A165G}$ catalysed the formation of L-xylose (**3b**) with full stereoselectivity in 98% conversion (62% isolated yield) (Table 1, entry 7). Slow addition of **1a** facilitated the sequential formation of L-glyceraldehyde (**2b**) and L-xylose (**3b**) and prevented the dimerization of **1a** (Fig. 2a,c). Next, the panel of enzymes was assayed in the C_2+C_2+C_2 connection of **1a**. A number of FSA mutants (Table 1, entries 4–6) catalysed the stereoselective trimerization of **1a** to furnish the monosaccharide D-idose (**3a**), achieving 85% conversion (48% isolated yield) with $FSA^{A129T/A165G}$ (Table 1, entry 4). Molecular models of intermediate reaction steps catalysed by this double mutant show that it is well suited to bind D-threose at the acceptor site and to promote its reaction with the third unit of glycolaldehyde (Fig. 3). Thus, this is the first bottom-up, one-pot synthesis of a hexose with a single enzyme and full control of the four stereogenic centres. In this synthesis, glycolaldehyde (**1a**) is rapidly dimerized into D-threose (**2a**) (Fig. 2b,d). Then, a subsequent second addition of **1a** to the acyclic form of **2a**, in equilibrium with the stable hemiketal furanoid form[28], produces **3a**. The **1a** needed for the second addition is obtained solely from the *in situ* retro-aldolization of **2a**. This fact, in addition to the low concentration of acyclic **2a** present in solution, made the trimerization of **1a** slower than its dimerization. In contrast, FSA variants that dramatically accelerated glycolaldehyde dimerization[26], such as FSA^{A129G}, did not significantly catalyse the sugar extension to D-idose (Supplementary Table 4). In this manner, the pertinent choice of FSA variants allows for the preparation of a complete, stereochemically consistent series of three- to six-carbon aldose sugars, namely L-glyceraldehyde (**2b**), D-threose (**2a**), L-xylose (**3b**) and D-idose (**3a**), from achiral precursors (Table 1, entries 1, 2, 4 and 7).

Acetaldehyde (**1d**) and glycolaldehyde analogues (that is, **1e–j**, Table 1) were assayed as acceptors for the double addition of glycolaldehyde (that is, with the C_2+C_2+C_2 connection) to generate 6-deoxy-D-idose and 6-substituted D-idose derivatives (Table 1, entries 11–22). The selectivity of FSA variants towards glycolaldehyde as a donor ensures that the other aldehyde component will invariably act as an acceptor substrate. Poor results were obtained with the available FSA variants, prompting us to further engineer their acceptor binding site. The proximity of residue S166 to the acceptor site (Fig. 3) led us to consider the substitution of S166 by Gly to alleviate the steric hindrance at that position. Variants $FSA^{A129T/S166G}$ and $FSA^{A129T/A165G/S166G}$ were found to tolerate these aldehydes satisfactorily, producing the synthesis of 6-deoxy- (**3d**), 6-*O*-methyl- (**3e**), 6-*O*-benzyl- (**3f**), 6-(benzylthio)- (**3g**), 6-*O*-phenyl (**3h**), 6-deoxy-6-azido (**3i**) and 6-deoxy-6-chloro- (**3j**) D-idose (Table 1, entries 11–20).

Other simple C3 aldehydes (**1c** and **1k–l**) were tested for the double addition of glycolaldehyde (C_3+C_2+C_2 connection) to generate high-carbon-content aldoses. Using $FSA^{A129T/S166G}$, 6-deoxy-6-methyl-D-idose (**3k**) and 6-deoxy-6-(phenylmethyl)-D-idose (**3l**) were synthesized (Table 1, entries 21 and 22). 3-Hydroxypropanal (**1c**) only produced a single glycolaldehyde addition, most probably because of the high stability of the 4-deoxy-β-D-*threo*-pentopyranose hemiacetal (**2c**) formed after the first **1a** addition (Table 1, entry 3).

Minimal active-site engineering was also envisioned to gain access to alternative hexose stereoisomers. Notably, inversion of the stereochemistry at C3 (that defined during the approach of the acceptor aldehyde to the FSA–donor intermediate) in the first addition of **1a** would afford L-glucose, a non-caloric sweetener of complex preparation by alternative procedures[32]. Such C3 inversion appeared feasible, because the stereochemical approach of the aldehyde to the aldolase–donor complex is often compromised when

Table 1 | Single and double additions of glycolaldehyde (1a) catalysed by FSA variants.



a: $R_1 = CH_2OH$ **e**: $R_1 = CH_2OCH_3$ **i**: $R_1 = CH_2N_3$
b: $R_1 = H$ **f**: $R_1 = CH_2OBn$ **j**: $R_1 = CH_2Cl$
c: $R_1 = CH_2CH_2OH$ **g**: $R_1 = CH_2SBn$ **k**: $R_1 = CH_2CH_3$
d: $R_1 = CH_3$ **h**: $R_1 = CH_2OPh$ **l**: $R_1 = CH_2CH_2Ph$

Entry	Product	FSA variant*	Conversion (%) (Reaction time, h)	Isolated yield[¶] (%)
1	2a	A129G[†]	98 (24)[§]	67
2	2b	Wild-type[†]	99 (24)[§]	-[#]
3	2c	L107Y/A129G[†,‡]	90 (5)[‖]	43
4	3a	A129T/A165G[‡]	69 (24), 85 (96)[‖]	48
5	3a	A129T/S166G	53 (24)	-[#]
6	3a	A129T/A165G/S166G	50 (24)	-[#]
7	3b	A129T/A165G[‡]	98 (24)[‖]	62
8	3b	A129T[‡]	80 (24)	-[#]
9	3b	A129V[‡]	95 (24)	-[#]
10	3b	L107Y/A129G[‡]	90 (24)	-[#]
11	3d	A129T/A165G/S166G	70 (72)[‖]	7
12	3e	A129T/A165G[‡]	69 (24)	-[#]
13	3e	A129T/A165G/S166G	79 (24)[‖]	39
14	3f	A129T/A165G/S166G	90 (24), 94 (48)[‖]	46
15	3f	A129T/A165G[†]	80 (24)	-[#]
16	3g	A129T/S166G	77 (24), 80 (52)[‖]	39
17	3g	A129T[†]	81 (24)	-[#]
18	3h	A129T/S166G	65 (52)[‖]	43
19	3i	A129T/S166G	80 (27)[‖]	62
20	3j	A129T/S166G	67 (52)[‖]	11
21	3k	A129T/S166G	51 (48)[‖]	31
22	3l	A129T/S166G	58 (48)[‖]	47

*Variants used that provided the highest product formation; [†]one single **1a** addition; [‡]FSA variant constructed as in ref. 26; [§]data from refs 24 and 26; [‖]under scaled-up conditions (Supplementary Tables 4–26); [¶]scaled-up reactions were performed with the respective most productive FSA variant. In entries 4 to 22 intermediate products **2** were not isolated. Purification procedures were not optimized. The isolated carbohydrates were unequivocally characterized by NMR. Unbiased kinetic stereocontrol for *si* face addition of the FSA variant-bound nucleophile to the corresponding *si* face of the acceptor, yielding 2*S*,3*R*,4*R*,5*R* (for 3a, 3d–e, 3f, 3h–i and 3k–l), 2*S*,3*R*,4*R*,5*S* (for 3g and 3j) and 2*S*,3*R*,4*S* (for 3b) configured carbohydrates. No other diastereoisomers were found within the limits of NMR detection (>96:4). Negligible NMR signals, which might be attributed to any other potential stereoisomer, could not be assigned unequivocally. To ensure that no potential stereoisomers were eliminated during purification, all spots visible in the thin-layer chromatography analysis were separated by chromatography and carefully analysed by NMR. No other stereoisomers were detected. For additional information concerning FSA stereochemistry see refs 24–26 and 28–30. For NMR data and structural characterization see Supplementary pages 16–91; [#]not isolated.

dealing with non-natural substrates[26,33] and it has been successfully inverted by engineering several aldolases[34]. In a preliminary exploration, we observed that dimerization of **1a** with FSA^{A165G} provided an ~4:1 mixture of D-threose (**2a**) and L-erythrose (**4a**), as inferred from chromatographic and NMR analysis (Supplementary Table 28 and Supplementary Fig. 28b,c). FSA^{A165G} was then subjected to saturation mutagenesis at position S166, producing the variant $FSA^{A165G/S166P}$ with a highly enhanced preference for L-erythrose formation (Supplementary Table 28 and Supplementary Fig. 28a–c). Attempts to engineer the most active FSA variants[26] to form L-erythrose (for example, library $FSA^{L107Y/A129G/A165G/\mathbf{S166X}}$) were unsuccessful. In a second step, a set of FSA variants was assayed for selective aldol addition of **1a** to L-erythrose (**4a**) (Supplementary Table 28). FSA^{A129G} was the most successful catalyst for the formation of L-glucose (**5a**) as the major product (81% conversion, 38% isolated yield) whereas D-idose (**3a**) was not detected (Table 2, entry 1). The accumulation of L-glucose (**5a**) was favoured by two circumstances: (1) the higher thermodynamic stability of L-glucose (that is, all-*trans* substituted) relative to D-idose (that is, *cis*/*trans* configured) and (2) the inability of FSA^{A129G} to cleave L-erythrose (**4a**) and to use D-threose (**2a**) as an acceptor. Thus, the synthesis of L-glucose (**5a**) was accomplished in the two-step one-pot trimerization of **1a** using a tandem of engineered FSA variants.

Pertinent selection of FSA variants allowed the synthesis of L-glucose derivatives. Tandem addition of **1a** to **1e** catalysed by $FSA^{A129T/A165G}$ and FSA^{A129T} produced 6-*O*-methyl-L-glucose (**5e**), and 6-deoxy-6-chloro-L-glucose (**5j**) was obtained from **1a** and **1j** using $FSA^{A129T/A165G}$ in both additions (Table 2, Supplementary Tables 30 and 32 and Supplementary Figs 30 and 32). In both cases, the first glycolaldehyde addition yielded *syn*/*anti* mixtures of products, and the second addition used the *anti* adducts selectively (that is, **4e** and **4j**) to produce L-glucose derivatives as the main products.

Discussion

Life takes advantage of the modular assembly of simple building blocks as an essential strategy to create molecular complexity. In the case of monosaccharides, sequential addition of their most simple precursors (**1a** and **1b**) through enzymatic catalysis ('systems biocatalysis') may be envisioned. This was not the path followed by natural evolution, which instead forged highly intricate pathways from central cellular metabolites. In this Article, we show that the direct enzymatic assembly of small aldehydes and glycolaldehyde into aldoses and analogues, with control of the sugar size and configuration, is feasible with engineered variants of D-fructose-6-phosphate aldolase (FSA). Using *syn*-selective FSA variants, the alternation of formaldehyde or glycolaldehyde as the initial

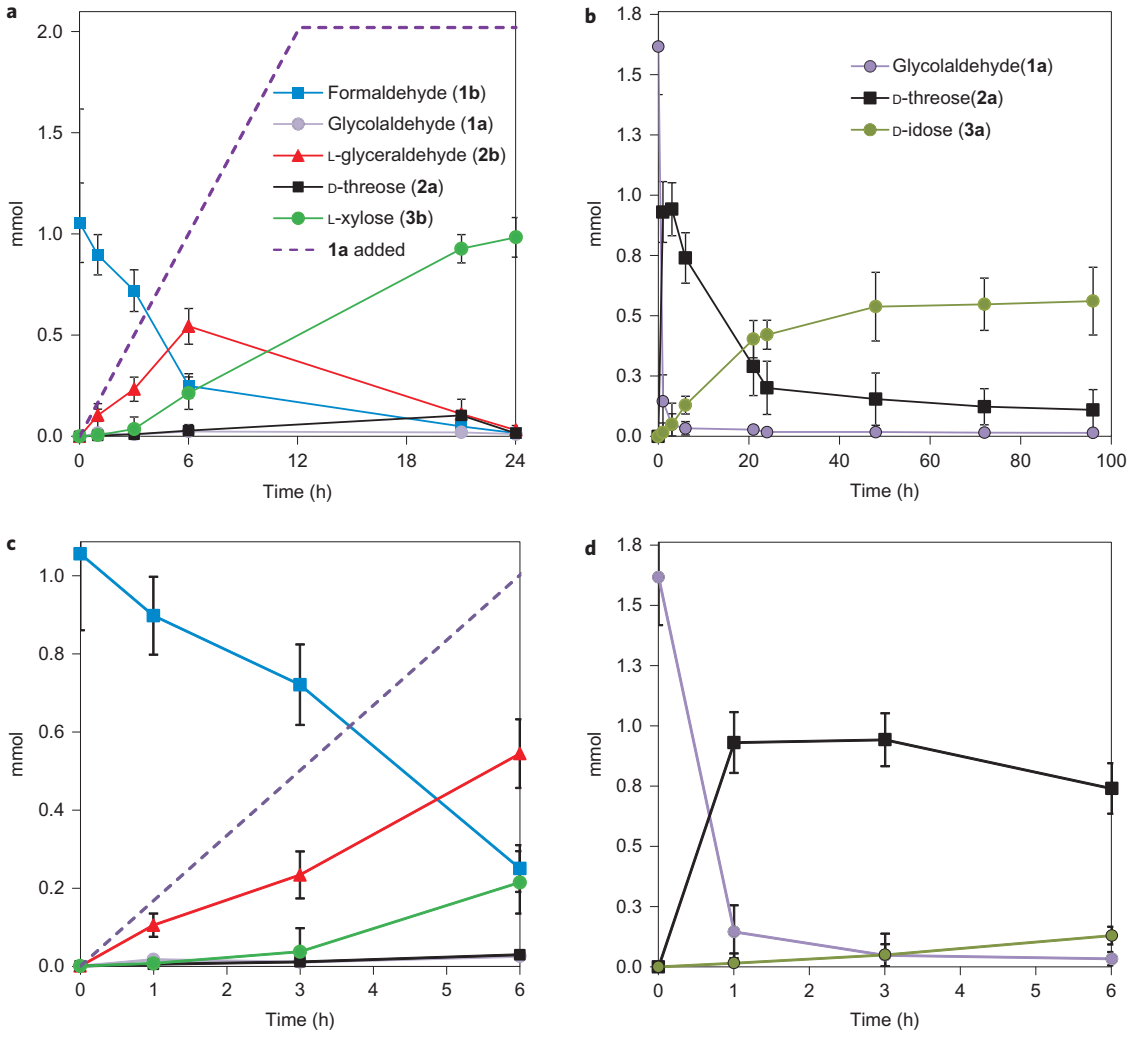


Figure 2 | Time progression curves for the enzymatic synthesis of L-xylose (3b) and D-idose (3a). a–d, Time progression curves for L-xylose (3b) (**a**,**c**) and D-idose (3a) (**b**,**d**). In **a** and **c**, the addition of 2 equiv. of **1a** to 1 equiv. **1b** is presented. The first addition produced **2b** followed by an ensuing addition of **1a** to produce 3b. Slow addition of **1a** prevented the formation of **2a**. Conditions: to a solution of **1b** (1 mmol) in triethanolamine buffer (10 ml, 50 mM, pH 8), $FSA^{A129T/A165G}$ (20 mg, 5.2 U) was added. The reaction was started by the fed-batch addition with a syringe pump of a 400 mM **1a** solution (5.0 ml, 0.42 ml h^{-1}, 2 mmol) for 12 h. In **b** and **d**, the enzymatic trimerization of **1a** is presented. Glycolaldehyde (**1a**) rapidly dimerizes into **2a** and then a second addition of **1a** to the acyclic form of **2a** produces 3a. Conditions: **1a** (1.5 mmol) was dissolved in triethanolamine buffer (50 mM, pH 8, 10 ml). The reaction was started by adding $FSA^{A129T/A165G}$ (30 mg, 7.8 U). Error bars are the values of the standard error of the mean of three independent experiments under the same reaction conditions.

acceptors produces odd and even carbon products with D and L configurations, respectively. This is in contrast with the prevalence of D-carbohydrates in nature, which is a consequence of the use of D-glyceraldehyde-3-phosphate as a building block. FSA is thus an unprecedented biocatalyst for carbohydrate synthesis because of the simplicity of the substrates required and the high stereoselectivity displayed. This compares favourably to other aldolases involved in sugar metabolism, which mainly act on ketoses, catalysing C_n+C_3 connections using dihydroxyacetone phosphate.

Formose sugar synthesis is commonly regarded as the source of carbohydrates necessary for the establishment of a prebiotic world based on RNA molecules as self-replicating, evolvable information units[35]. More feasible alternatives have been suggested, such as direct access to activated pyrimidine ribonucleosides, which would bypass the separate synthesis of the free sugar and the nucleobase moieties[36]. However, these approaches lead to a plethora of different products with detrimental effects on RNA replication. It has been suggested that self-replicating prebiotic polymers based on chiral building blocks could have preceded RNA synthesis and provided a chemical scenario suitable for the creation and combination of nucleotides[37]. The enzymatic reactions presented in this work could show that a sufficiently 'complex polymer' can assemble enantiopure carbohydrates, including thermodynamically unfavourable configurations (for example, D-idose) in aqueous conditions, from formaldehyde and glycolaldehyde. It might therefore be hypothesized that chiral polymers exhibiting defined tertiary structures could have enabled selective carbohydrate syntheses in a prebiotic system. Remarkably, significant acceleration of the retroaldol reaction has been achieved with reported simple β-amino acid foldamers[38].

The engineering of FSA variants with altered substrate preferences and stereochemical outcomes was achieved through minimal modifications of the catalytic site. Indeed, this is consistent with the fact that during protein evolution all the functional sequences of a protein are connected by trajectories involving changes in only one single amino-acid residue[39]. Different substrates require different FSA variants, which always result from a combination of a few (that is, one or two) residue variations near the catalytic site. Selected mutations at positions A165 and S166 of the active site (Fig. 3) fostered the

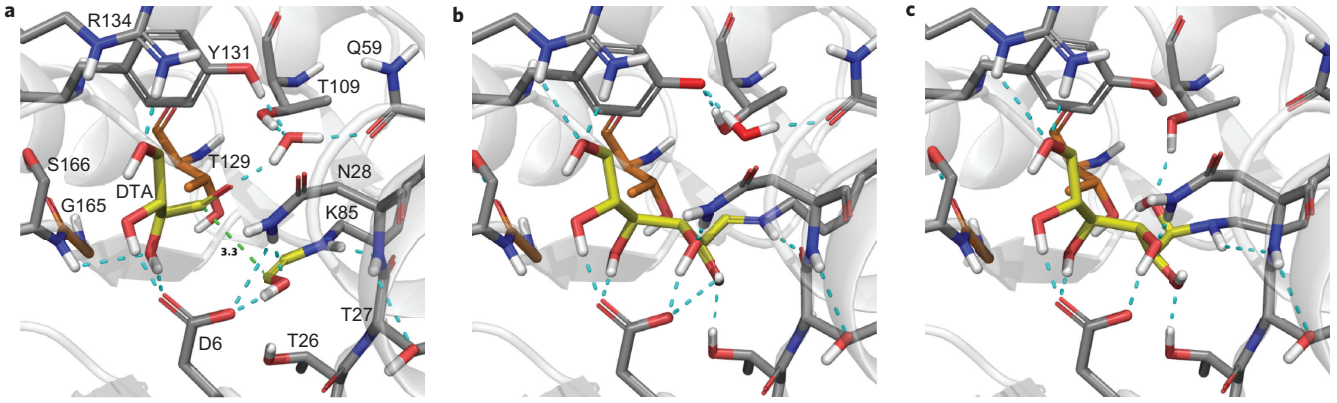
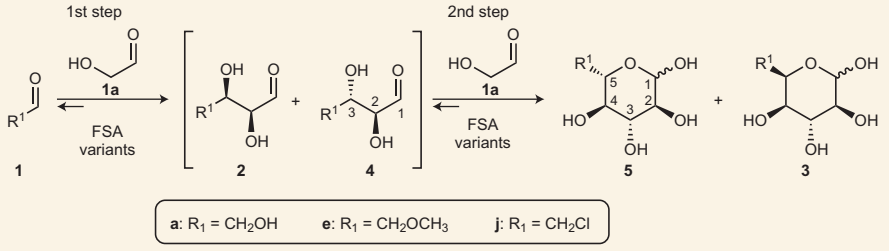


Figure 3 | Molecular models of FSA-bound intermediates during formation of D-idose (3a) catalysed by FSA A129T/A165G. a–c, Active site of $FSA^{A129T/A165G}$ complexed with D-threose (acyclic form, DTA) and the enamine of glycolaldehyde (**a**), the corresponding imine adduct (**b**) and the subsequent hemiaminal intermediate (**c**). Mutated protein residues and non-protein moieties are shown in orange and yellow, respectively, and the rest of the protein is shown in grey. In **a**, the two reactive carbon atoms are shown in a pre-reactive disposition at a distance of 3.3 Å and the D-threose carbonyl oxygen is forming a hydrogen bond with an essential water molecule, which is also hydrogen-bonded to residues Q59, T109 and Y131. It is postulated that this water molecule mediates proton transfer from the phenol group of residue Y131 to the acceptor aldehyde group[26], once the C–C bond is formed, to generate the corresponding iminium cation (**b**). Subsequent nucleophilic attack of the water on the iminium-activated carbon atom, promoted by proton abstraction by the Y131 phenolate, would yield the hemiaminal precursor of D-idose (**c**). All the structures shown were modelled and energy minimized using a previously described protocol[26]. The Schrödinger Suite 2014-3 was used to carry out all calculations[31].

Table 2 | Synthesis of L-glucose and C6 derivatives using one or two FSA variants.



Entry	Product	FSA variant	Conversion 5/3 (%) (Reaction time, h)	Isolated yield[‖] of 5 (%)
1	5a	1st step: A165G/S166P*; 2nd step: A129G[†]	81/ND[§] (240)	38[¶]
2	5e	1st step: A129G/A165G[†]; 2nd step: A129T[†]	41 / 2 (29)	4[¶]
3	5j	A129T/A165G[†,‡]	60 / 2 (52)	14

*FSA variant in this work; [†]FSA variant constructed as in ref. 26; [‡]one-step one-pot; [§]not detected; [‖]purification procedures were not optimized. The isolated carbohydrates were unequivocally characterized by NMR yielding 2*S*,3*R*,4*R*,5*S* (for 5a and 5e) and 2*S*,3*R*,4*R*,5*R* (for 5j) configured carbohydrate compounds. No other diastereoisomers were observed within the limits of NMR detection (>96:4). For NMR data and structural characterization see Supplementary pages 92–118; [¶]isolated yield including both steps. Intermediate products **2** and **4** were not isolated.

double addition of **1a** and provided inversion at the C3 carbon in the first **1a** addition. Remarkably, a single variant can be engineered to produce opposite selectivity at each aldol step, as in the synthesis of 6-chloro-L-glucose (5j). Such active site malleability demonstrates the great evolutionary potential for the functional diversification of FSA, which nature has not exploited, as inferred from the reduced sequence variability observed within the FSA family[23,40]. Application of higher-throughput screening techniques can be expected to boost the fast evolution of FSA and its variants towards higher activity and stereoselectivity, and enable unprecedented non-biological reactivity. It is also noteworthy that diverse FSA variants from the toolbox were selective for a single glycolaldehyde addition (Supplementary Tables 14, 16, 18 and 20). This is highly important because controlling glycolaldehyde addition by biocatalyst selection allows the synthesis of the mono addition adducts as final products or as intermediate acceptors for further additions with other aldolases with alternative donor specificity or stereochemical outcomes.

In summary, we have developed a procedure for the preparation of a number of aldose carbohydrates and derivatives from simple formaldehyde, glycolaldehyde and other aldehydes by biocatalytic tandem reactions using engineered FSA variants. Remarkably, the preparation of these intricate molecules is achieved with optimal energetic and atomic economy, avoiding both the functionalization and activation required in alternative chemical or biocatalytic approaches. Overall, our results illustrate the utility of engineered enzymes in the construction of complex molecules, harnessing simple substrates and biological transformations that circumvent the complexity of the natural enzymatic pathways forged by evolution.

Received 7 March 2015; accepted 8 July 2015;
published online 10 August 2015

References

1. Weymouth-Wilson, A. C. The role of carbohydrates in biologically active natural products. *Nat. Prod. Rep.* **14,** 99–110 (1997).
2. Bertozzi, C. R. & Kiessling, L. L. Chemical glycobiology. *Science* **291,** 2357–2364 (2001).
3. Stallforth, P., Lepenies, B., Adibekian, A. & Seeberger, P. H. Carbohydrates: a frontier in medicinal chemistry. *J. Med. Chem.* **52,** 5561–5577 (2009).
4. Wong, C.-H. *Carbohydrate-based Drug Discovery* Vols 1 & 2 (Wiley, 2003).
5. Campo, V. L., Aragao-Leoneti, V., Teixeira, M. B. M. & Carvalho, I. Carbohydrates and glycoproteins: cellular recognition and drug design. *New Dev. Med. Chem.* **1,** 133–151 (2010).

6. Furukawa, K. *et al.* Fine tuning of cell signals by glycosylation. *J. Biochem.* **151,** 573–578 (2012).
7. Galan, M. C., Benito-Alifonso, D. & Watt, G. M. Carbohydrate chemistry in drug discovery. *Org. Biomol. Chem.* **9,** 3598–3610 (2011).
8. Grunwald, P. in *Carbohydrate-Modifying Biocatalysts* (ed. Grunwald, P.) 29–118 (Pan Stanford, 2012).
9. Hudlicky, T., Entwistle, D. A., Pitzer, K. K. & Thorpe, A. J. Modern methods of monosaccharide synthesis from non-carbohydrate sources. *Chem. Rev.* **96,** 1195–1220 (1996).
10. Mlynarski, J. & Gut, B. Organocatalytic synthesis of carbohydrates. *Chem. Soc. Rev.* **41,** 587–596 (2012).
11. Markert, M. & Mahrwald, R. Total syntheses of carbohydrates: organocatalyzed aldol additions of dihydroxyacetone. *Chem. Eur. J.* **14,** 40–48 (2008).
12. Mahrwald, R. *Modern Methods in Stereoselective Aldol Reactions* (Wiley, 2013).
13. Hein, J. E. & Blackmond, D. G. On the origin of single chirality of amino acids and sugars in biogenesis. *Acc. Chem. Res.* **45,** 2045–2054 (2012).
14. Ruiz-Mirazo, K., Briones, C. & De La Escosura, A. Prebiotic systems chemistry: new perspectives for the origins of life. *Chem. Rev.* **114,** 285–366 (2014).
15. Northrup, A. B. & MacMillan, D. W. C. Two-step synthesis of carbohydrates by selective aldol reactions. *Science* **305,** 1752–1755 (2004).
16. Northrup, A. B., Mangion, I. K., Hettche, F. & MacMillan, D. W. C. Enantioselective organocatalytic direct aldol reactions of α-oxyaldehydes: step one in a two-step synthesis of carbohydrates. *Angew. Chem. Int. Ed.* **43,** 2152–2154 (2004).
17. Clapés, P. & Garrabou, X. Current trends in asymmetric synthesis with aldolases. *Adv. Synth. Catal.* **353,** 2263–2283 (2011).
18. Clapés, P. & Joglar, J. in *Modern Methods in Stereoselective Aldol Reactions* (ed. Mahrwald, R.) 475–528 (Wiley, 2013).
19. Wong, C.-H. *et al.* Recombinant 2-deoxyribose-5-phosphate aldolase in organic synthesis: use of sequential two-substrate and three-substrate aldol reactions. *J. Am. Chem. Soc.* **117,** 3333–3339 (1995).
20. Jennewein, S. *et al.* Directed evolution of an industrial biocatalyst: 2-deoxy-D-ribose 5-phosphate aldolase. *Biotechnol. J.* **1,** 537–548 (2006).
21. Durrwachter, J. R., Drueckhammer, D. G., Nozaki, K., Sweers, H. M. & Wong, C.-H. Enzymic aldol condensation/isomerization as a route to unusual sugar derivatives. *J. Am. Chem. Soc.* **108,** 7812–7818 (1986).
22. Alajarin, R., Garcia-Junceda, E. & Wong, C.-H. A short enzymic synthesis of L-glucose from dihydroxyacetone phosphate and L-glyceraldehyde. *J. Org. Chem.* **60,** 4294–4295 (1995).
23. Samland, A. K., Rale, M., Sprenger, G. A. & Fessner, W.-D. The transaldolase family: new synthetic opportunities from an ancient enzyme scaffold. *ChemBioChem* **12,** 1454–1474 (2011).
24. Garrabou, X. *et al.* Asymmetric self- and cross-aldol reaction of glycolaldehyde catalyzed by D-fructose-6-phosphate aldolase. *Angew. Chem. Int. Ed.* **48,** 5521–5525 (2009).
25. Gutierrez, M., Parella, T., Joglar, J., Bujons, J. & Clapés, P. Structure-guided redesign of D-fructose-6-phosphate aldolase from *E. coli*: remarkable activity and selectivity towards acceptor substrates by two-point mutation. *Chem. Commun.* **47,** 5762–5764 (2011).
26. Szekrenyi, A. *et al.* Engineering the donor selectivity of D-fructose-6-phosphate aldolase for biocatalytic asymmetric cross-aldol additions of glycolaldehyde. *Chem. Eur. J.* **20,** 12572–12583 (2014).
27. Castillo, J. A. *et al.* A mutant D-fructose-6-phosphate aldolase (Ala129Ser) with improved affinity towards dihydroxyacetone for the synthesis of polyhydroxylated compounds. *Adv. Synth. Catal.* **352,** 1039–1046 (2010).
28. Concia, A. L. *et al.* D-Fructose-6-phosphate aldolase in organic synthesis: cascade chemical-enzymatic preparation of sugar-related polyhydroxylated compounds. *Chem. Eur. J.* **15,** 3808–3816 (2009).
29. Soler, A. *et al.* Sequential biocatalytic aldol reactions in multistep asymmetric synthesis: pipecolic acid, piperidine and pyrrolidine (homo)iminocyclitol derivatives from achiral building blocks. *Adv. Synth. Catal.* **356,** 3007–3024 (2014).
30. Rale, M., Schneider, S., Sprenger, G. A., Samland, A. K. & Fessner, W.-D. Broadening deoxysugar glycodiversity: natural and engineered transaldolases unlock a complementary substrate space. *Chem. Eur. J.* **17,** 2623–2632 (2011).
31. Schrödinger Suite 2014-3 (2014).
32. Guaragna, A., D'Alonzo, D., Paolella, C., Napolitano, C. & Palumbo, G. Highly stereoselective *de novo* synthesis of L-hexoses. *J. Org. Chem.* **75,** 3558–3568 (2010).
33. Fessner, W. -D. *et al.* Enzymes in organic synthesis. Part 1. Diastereoselective, enzymatic aldol addition with L-rhamnulose- and L-fuculose-1-phosphate aldolases from *E. coli*. *Angew. Chem. Int. Ed.* **30,** 555–558 (1991).
34. Bolt, A., Berry, A. & Nelson, A. Directed evolution of aldolases for exploitation in synthetic organic chemistry. *Arch. Biochem. Biophys.* **474,** 318–330 (2008).
35. Orgel, L. E. Prebiotic chemistry and the origin of the RNA world. *Crit. Rev. Biochem. Mol. Biol.* **39,** 99–123 (2004).
36. Powner, M. W., Gerland, B. & Sutherland, J. D. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **459,** 239–242 (2009).
37. Brewer, A. & Davis, A. P. Chiral encoding may provide a simple solution to the origin of life. *Nature Chem.* **6,** 569–574 (2014).
38. Müller, M. M., Windsor, M. A., Pomerantz, W. C., Gellman, S. H. & Hilvert, D. A rationally designed aldolase foldamer. *Angew. Chem. Int. Ed.* **48,** 922–925 (2009).
39. Currin, A., Swainston, N., Day, P. J. & Kell, D. B. Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently. *Chem. Soc. Rev.* **44,** 1172–1239 (2015).
40. Samland, A. K. & Sprenger, G. A. Transaldolase: from biochemistry to human disease. *Int. J. Biochem. Cell Biol.* **41,** 1482–1494 (2009).

Acknowledgements

This work was supported by Spanish MINECO grants CTQ2012-31605 and CTQ2012-32436, the Generalitat de Catalunya (2009 SGR 00281), ERA-IB MICINN, PIM2010EEI-00607 (EIB.10.012. MicroTechEnz-EIB, www.fkit.unizg.hr/miten) and COST Action CM1303 Systems Biocatalysis. A.S. acknowledges the CSIC for a JAE predoctoral contract programme.

Author contributions

P.C. and X.G. designed the study. A.S. and X.G. performed mutagenesis, library screening, activity measurements and synthesis of the compounds. J.B. performed the molecular docking experiments and designed the mutations. T.P. performed and supervised the NMR experiments and structural assignation of compounds. J.J., J.B. and P.C. supervised the scientific work. All authors contributed to writing the paper.

Additional information

Supplementary information and chemical compound information are available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.C.

Competing financial interests

The authors declare no competing financial interests.