Chiral resolution through stereoselective transglycosylation by sucrose phosphorylase: application to the synthesis of a new biomimetic compatible solute, (*R*)-2-*O*-α-D-glucopyranosyl glyceric acid amide

Patricia Wildberger, Lothar Brecker and Bernd Nidetzky

ELECTRONIC SUPPLEMENTARY INFORMATION

Materials

Glyceric acid amide was prepared from 2,3-dihydroxypropanoic acid (40% solution in water; TCI Europe nv, Antwerp, Belgium), which was dissolved in methanol and methylated with an excess of etheral solution of diazomethane. Methyl 2,3-dihydroxypropanoate was isolated from the concentrated reaction mixture and purified by column chromatography on silica gel with chloroform-methanol mixture as eluent. Then, methyl 2,3-dihydroxypropanoate was dissolved in methanol and an equal volume of 25 - 28% aq. NH_3 was added to this solution. The reaction mixture was kept overnight at room temperature, concentrated in vacuo, and the residue was crystallised from water (from concentrated solution). No impurities were determined in the ^1H NMR spectrum, indicating a glyceric acid amide purity of > 97%.

Glyceric acid amide: ^1H NMR (D_2O, 600.13 MHz): δ 4.16 (dd, J = 4.8, 3.6 Hz, 1H, H-2), 3.75 (dd, J = 12.0, 3.6 Hz, 1H, H-3a), 3.73 (dd, J = 12.0, 4.8 Hz, 1H, H-3b). ^{13}C NMR (D_2O, 150.90 MHz): 177.5 (C-1), 72.0 (C-2), 63.1 (C-3).

All other chemicals used were of highest purity available from Sigma-Aldrich (Vienna, Austria) or Roth (Karlsruhe, Germany).

Enzyme production

Recombinant sucrose phosphorylase from *Leuconostoc mesenteroides*, encoded in pASK-IBA7+ (IBA GmbH, Goettingen, Germany),[1] was produced in *Escherichia coli* BL21-Gold(DE3) and purified to apparent homogeneity (by SDS PAGE) according to reported procedures.[2] Activity of the purified preparation was determined at 30°C using a continuous coupled enzymatic assay with phosphoglucomutase and glucose 6-phosphate dehydrogenase.[3] Protein concentration of the purified preparation was measured with the BioRad dye-binding method referenced against BSA.

Product identification by NMR analysis

Equimolar concentrations of sucrose and racemic glyceric acid amide (500 mM each) were incubated in the presence of sucrose phosphorylase (30 U/mL). The reaction was carried out as described in

Table S1, reaction #1. The content of the mixture was lyophilised and analysed as described under NMR measurements.

The resulting spectra indicate a complete transformation of sucrose. Analysis indicates the presence of D-fructose (1.00 eq., mixture of both isomers), D-glucose (0.50 eq., mixture of both isomers), four not further identified disaccharides (in sum < 0.05 eq.), residual glyceric acid amide (0.55 eq.), and (*R*)-2-*O*-α-D-glucopyranosyl glyceric acid amide (0.45 eq.).

(*R*)-2-*O*-α-D-glucopyranosyl glyceric acid amide (determined from crude reaction mixture): ^{1}H NMR (D$_2$O, 400.13 MHz): δ 5.08 (d, J = 3.8 Hz, 1H, H-1), 4.21 (dd, J = 4.2, 3.7 Hz, 1H, H-2'), 3.97 (dd, J = 12.2, 2.5 Hz, 1H, H-6a), 3.94 (m, 1H, H-5), 3.88 (dd, J = 12.4, 3.7 Hz, 1H, H-3a'), 3.82 (dd, J = 12.4, 4.2 Hz, 1H, H-3b'), 3.75 (dd, J = 9.8, 9.9 Hz, 1H, H-3), 3.65 (dd, J = 12.2, 4.5 Hz, 1H, H-6b), 3.56 (dd, J = 9.8, 3.8 Hz, 1H, H-2), 3.39 (dd, J = 9.9, 9.5 Hz, 1H, H-4). ^{13}C NMR (D$_2$O, 100.61 MHz): 174.9 (C-1'), 98.2 (C-1), 77.9 (C-2'), 72.8 (C-3), 71.6 (C-2), 69.6 (C-5), 69.5 (C-4), 63.7 (C-6), 60.8 (C-3').

Isolation of residual glyceric acid amide from the reaction mixture

Equimolar concentrations of sucrose and racemic glyceric acid amide (400 mM each) were incubated in the presence of sucrose phosphorylase (7 U/mL). The reaction was carried out as described in Table S1, reaction #3, with the exception that incubation was prolonged to 65 h. The residual glyceric acid amide (≥ 35% conversion) was isolated from the reaction mixture on a Shimadzu CTO-20AC system equipped with an Aminex HPX-87C column (Bio-Rad, Vienna, Austria) that was operated at 75°C. Deionised water was the mobile phase and was applied at a constant flow rate of 0.3 mL/min. Refractive index detection (Shimadzu RID-10A) was used for all compounds. Under the conditions applied, the following retention times were obtained: 17.2 min for sucrose, 20.7 min for D-glucose, 26.9 min for D-fructose and 36.8 min for the residual glyceric acid amide. The reaction product co-eluted with sucrose. Samples containing the purified residual glyceric acid amide were pooled. The presence of the purified residual glyceric acid amide and the removal of all other components were shown by HPAE-PAD and further proved by NMR analysis. A ^{1}H NMR spectrum was recorded to confirm that no organic impurities have been present. Hence no further optical active impurities influenced the subsequent chiroptic measurements (optical rotatory dispersion and circular dichroism spectroscopy).

Reaction optimisation for chiral synthesis and chiral resolution

To find conditions in which glucosyl transfer to (*R*)-glyceric acid amide was favoured over hydrolysis, three reaction time courses were recorded in which sucrose and racemic glyceric acid amide were present in the following concentrations: 1) 400 mM sucrose and 1000 mM racemic glyceric acid amide, 2) equimolar concentrations of sucrose and racemic glyceric acid amide (400 mM each) and 3) 1000 mM sucrose and 400 mM racemic glyceric acid amide. The reactions were started by addition of

sucrose phosphorylase (7 U/mL) and performed as described in Table S1, reaction #2 - #4. Quantification of substrate consumption and product formation was done using HPAE-PAD; the presence of the reaction product was verified by NMR analysis.

Unlike chiral synthesis of (R)-2-O-α-D-glucopyranosyl glyceric acid amide that can be performed using glyceric acid amide in molar excess over sucrose, chiral resolution of racemic glyceric acid amide requires different conditions. We therefore applied equimolar concentrations of sucrose and racemic glyceric acid amide (500 mM each) to an enzymatic conversion performed in the presence of a high concentration of sucrose phosphorylase (40 U/mL). The reaction was carried out as described in Table S1, reaction #5. Quantification of substrate consumption and product formation was done using HPAE-PAD. About 50% of the initial glyceric acid amide was utilised in the reaction. Considering that only (R)-glyceric acid amide was glucosylated by sucrose phosphorylase whereas (S)-glyceric acid amide was inactive, the enzymatic reaction proceeded up to the theoretically possible yield. Therefore, this implies that enzymatic glucosylation of (R)-glyceric acid amide from sucrose is useful not only for chiral synthesis of (R)-2-O-α-D-glucopyranosyl glyceric acid amide, but it can also effectively applied to chiral resolution of racemic glyceric acid amide, enabling recovery of (S)-glyceric acid amide.

NMR measurements

The reaction mixture was dissolved in D_2O (99.95% D, 0.7 mL) and transferred into 5 mm NMR sample tubes (Promochem, Wesel, Germany). Spectra were measured with Topspin 3.1 on Bruker DRX-400 AVANCE or AV-III 600 (Bruker, Rheinstetten, Germany). Measurement frequencies were 400.13 MHz and 100.61 MHz for ^1H and ^{13}C, respectively, or 600.13 MHz and 150.90 MHz for ^1H and ^{13}C, respectively. Measurement temperature was 298.1 ± 0.1 K. For the 1D proton spectrum 32k data points were recorded and Fourier transformed to spectra with a range of 6000 Hz (^1H). 2D DQF-COSY, TOCSY, NOESY, HMQC, and HMBC spectra were measured by 256 experiments with 1024 data points each. Appropriate linear forward prediction, sinusoidal multiplication and Fourier transformation led to 2D-spectra with a range of 4800 Hz and 20000 Hz for ^1H and ^{13}C, respectively. Acetone was used as external standard for ^1H (δH 2.225) and ^{13}C (δC 30.89) spectra.

High performance anion exchange chromatography with pulsed amperometric detection (HPAE-PAD)

HPAE-PAD was performed using a Dionex BioLC system (Dionex Corporation, Sunnyvale, USA) equipped with a CarboPac PA10 column (4 x 250 mm) and an Amino Trap guard column (4 x 50 mm). D-Glucose, D-fructose, sucrose, (R)- and (S)-glyceric acid amide and (R)-2-O-α-D-glucopyranosyl glyceric acid amide were detected with an ED50A electrochemical detector using a gold working electrode and a silver/silver chloride reference electrode by applying the predefined waveform for carbohydrates. Elution was carried out at a flow rate of 0.3 mL/min at an isocratic flow

of 52 mM NaOH for 50 min. Under the conditions applied, the following retention times were obtained: 5.4 min for (*R*)- and (*S*)-glyceric acid amide, 18.0 min for D-glucose, 20.2 min for D-fructose and 33.9 min for sucrose. Authentic standards for D-glucose, D-fructose and sucrose were used for peak identification, and quantification was based on peak area that was suitably calibrated with standards of known concentration dissolved in 50 mM MES buffer, pH 7.0. The two enantiomers of glyceric acid amide co-eluted in the analysis and could neither separated nor quantified by using the described procedure. The formation of (*R*)-2-*O*-α-D-glucopyranosyl glyceric acid amide was indicated by the appearance of a shoulder (on the right) of the acceptor peak, but could not be quantified directly from the HPAE-PAD chromatogram. The concentration of (*R*)-2-*O*-α-D-glucopyranosyl glyceric acid amide was therefore determined indirectly from the difference in the concentration of D-fructose and D-glucose. Release of D-fructose paralleled the consumption of sucrose; the formation of D-glucose indicated that glucosylated sucrose phosphorylase had also reacted with water. The portion of sucrose utilised for transglucosylation was assumed to have resulted in glucosylation of (*R*)-glyceric acid amide. Results of NMR analysis support the assumption. However, minor formation of disaccharide by-products ($< 5\%$) was not considered in the analysis of the HPAE-PAD data.

Optical rotatory dispersion and circular dichroism (CD) spectroscopy

The purified residual glyceric acid amide was subjected to an optical rotation measurement. For that purpose, the purified residual glyceric acid amide was lyophilised, dissolved in methanol and the optical rotation was determined by using a Perkin Elmer Automatic Polarimeter 241 (Waltham, USA). Data were recorded at the sodium D line using a 100 mm path length cell. Results are reported as $[\alpha]_D^T$.

Far-UV CD spectra were recorded at 25°C on an Applied Photophysics Chirascan Plus system (Surrey, UK). The purified residual glyceric acid amide was diluted in acetonitrile to a final concentration of 50 mM. CD spectra were collected at a scan speed of 20 nm/min at 1 nm bandwidth and response time of 4 s. All spectra were recorded in a 0.1 cm cuvette between 180 and 300 nm. The data shown are the average of 10 recorded spectra.

Molecular Docking

AutoDock 4.2[4] as implemented in Yasara V 11.11.21 was used for the enzyme-ligand docking. The AMBER03 force field[5] and the default parameters provided by the standard docking macro were used, except that the number of runs was increased to 40. The X-ray crystal structure of sucrose phosphorylase from *Bifidobacterium adolescentis* having a β-glucosyl enzyme intermediate linked to Asp-192 (PDB-entry 2gdv, molecule A)[6] was used as macromolecule in a molecular docking experiment that employed (*R*)- and (*S*)-glyceric acid amide, the protonated form of (*R*)- and (*S*)-glyceric acid and (*R*)- and (*S*)-3-methoxy-1,2-propanediol, respectively, as ligand. The 3D coordinates for the ligands were generated from SMILES strings using Chimera

(http://www.cgl.ucsf.edu/chimera). The ligands were flexibly placed into the active site of sucrose phosphorylase. A search space of $10 \times 10 \times 10$ Å around the C1 atom of the glucosyl moiety of the β-glucosyl enzyme intermediate was used. The docking algorithm resulted in 3 binding modes for (R)-glyceric acid amide, 2 for (R)- and (S)-glyceric acid as well as for (R)- and (S)-3-methoxy-1,2-propanediol, and in 1 binding mode for (S)-glyceric acid amide. In the case that the docking algorithm resulted in more than 1 binding mode, the binding modes had comparable energies. Best-fit binding modes were selected on the basis of meaningful reactive-group arrangement. Distances between the C2-OH of the ligand to the anomeric carbon of the β-glucosyl residue and distances between the C2-OH of the ligand to the oxygen of the ionised side chain of Glu-232 were determined in all cases. Based on the distances determined, the binding modes in which the accommodation of the ligand appeared to be highly plausible for catalysis were selected. The criterions for the selection comprised that the distance of the reactive 2-OH of the acceptor to the catalytic base (Glu-232) was ≤ 3.2 Å and that the distance of the reactive 2-OH of the acceptor to the anomeric carbon of the β-glucosyl residue was ≤ 3.1 Å. PyMOL (http://pymol.sourceforge.net) was used for visualisation.

Homology modeling

The homology model of sucrose phosphorylase from *Leuconostoc mesenteroides* was obtained by the Automated Modeling tool of Swiss Model web service http://swissmodel.expasy.org/[7, 8] using the crystal structure of sucrose phosphorylase from *Bifidobacterium adolescentis* (PDB-entry 2gdv)[6] as template.
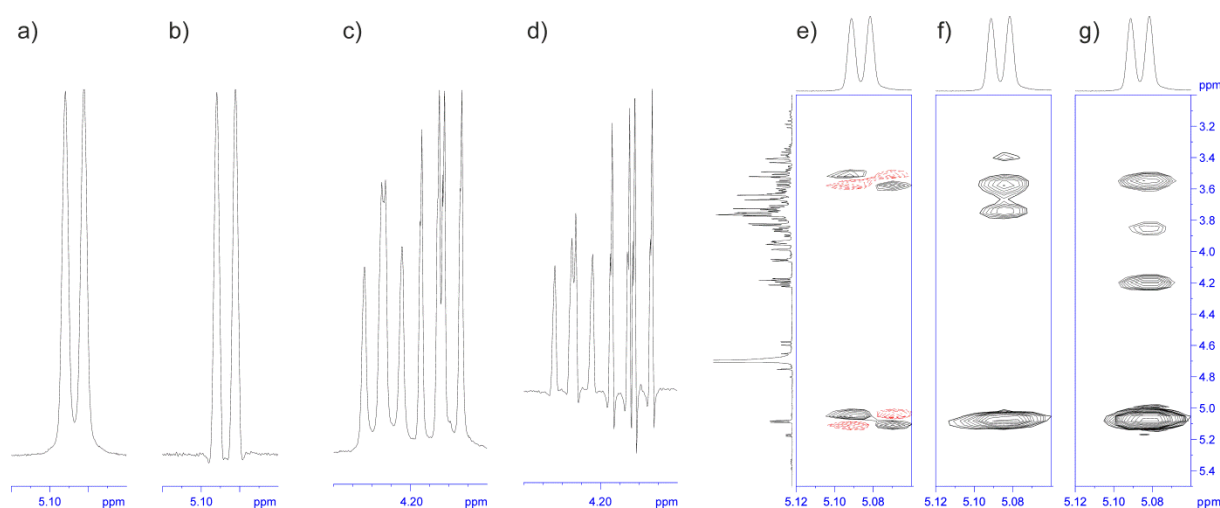
Figure S1. 1D and 2D homonuclear proton spectra of the reaction mixture obtained upon conversion of equimolar concentrations of sucrose and racemic glyceric acid amide (500 mM each). The conversion of glyceric acid amide was about 45%. The data are consistent with exclusive formation of one diastereomer resulting from glucosylation of glyceric acid amide.

All signals of the product recorded from crude reaction mixture do not show any shoulder. Additionally no further signals are present. Hence there is no indication for the presence of a second diastereomer. A: Proton signal H-1 of the glycoside after direct Fourier transformation showing Lorentzian line shape signal group without presence of any additional signals and shoulders, which would indicate the presence of a second diastereomer. B: Proton signal H-1 of the glycoside after Gaussian multiplication (LB = -0.3; GB = 0.2) and Fourier transformation leading to narrower line width at half height. There is also no indication for a second signal. C: Proton signal H-2' of the glycoside as well as proton signal of the free glyceric acid amide after direct Fourier transformation showing Lorentzian line shape signal group without any additional signals and shoulders, which would indicate a second diastereomer. D: Proton signal H-2' of the glycoside as well as proton signal of the free glyceric acid amide after Gaussian multiplication (line broadening factor (LB) = -0.3; Gaussian broadening factor (GB) = 0.2) and Fourier transformation leading to narrower line width at half height. There is also no indication for a second signal. E: DQF-COSY trace of proton signal H-1 of the glycoside indicating the diagonal peak and the cross peak, which shows the coupling to H-2. F: TOCSY (100 ms mixing time) trace of proton signal H-1 of the glycoside, which indicates all protons in the glycoside spin system. G: NOESY (800 ms mixing time) trace of proton signal H-1 of the glycoside. It indicates the spatial closeness to other protons, in particular to H-2' of the aglycon moiety.
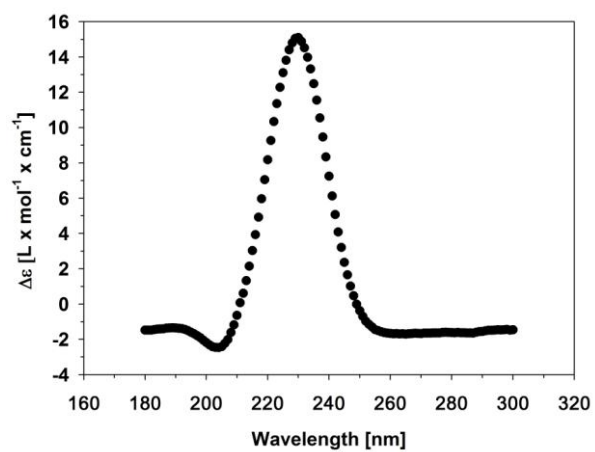
Figure S2. The CD-spectrum of the purified residual glyceric acid amide is shown.
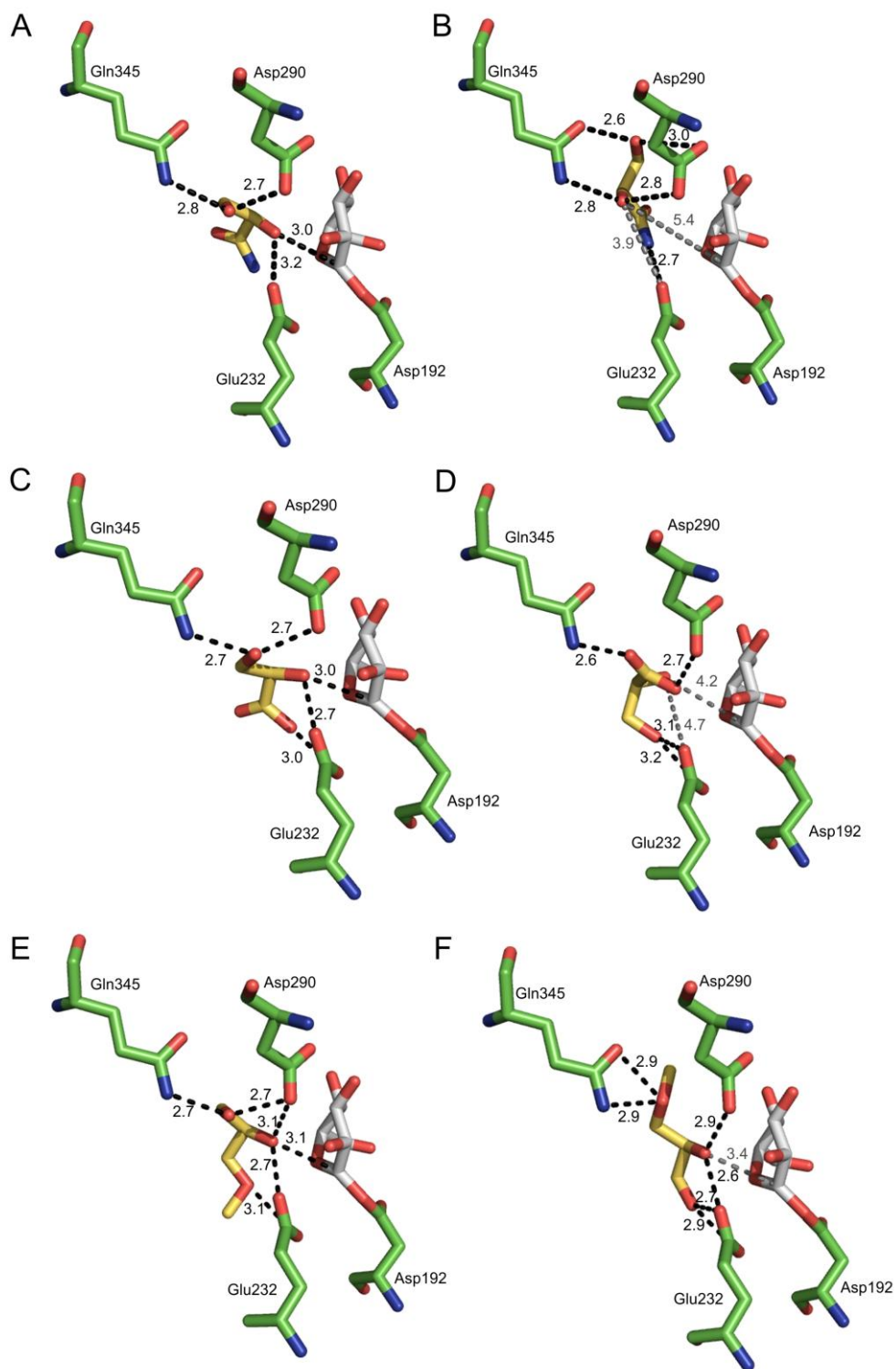
Figure S3. Close-up view of predicted best-fit molecular docking poses of (*R,S*)-glyceric acid amide (A, B), (*R,S*)-glyceric acid (C, D) and (*R,S*)-3-methoxy-1,2-propanediol (E, F) with the β-glucosyl enzyme intermediate of sucrose phosphorylase (PDB-entry 2gdv, molecule A). Hydrogen bonds (≤ 3.2 Å) are shown as black-dashed lines. Interactions potentially relevant for catalysis are shown as gray-dashed lines. Distances are given in Å.

Note: When glyceric acid was employed as ligand, the docking experiment was performed with the protonated and deprotonated form of the *R* and *S* enantiomer of glyceric acid. The protonated form of both enantiomers resulted in more favorable binding energies (up to 1.9 kcal/mol) than the corresponding unprotonated (negatively charged) forms. The *R* enantiomer of unprotonated glyceric acid did not result in a plausibly reactive binding pose. The *S* enantiomer gave multiple binding poses. Overall, the results of docking studies suggest that the protonated form of glyceric acid is the preferred form of the ligand bound by the enzyme. It is of course unlikely that glyceric acid is protonated at neutral pH where the enzymatic reactions are performed. However, it is possible that the enzyme provides an additional proton during binding and the overall result is that the bound form of glyceric acid is protonated. Variable protonation state of the enzyme cannot be represented well in the docking experiment.
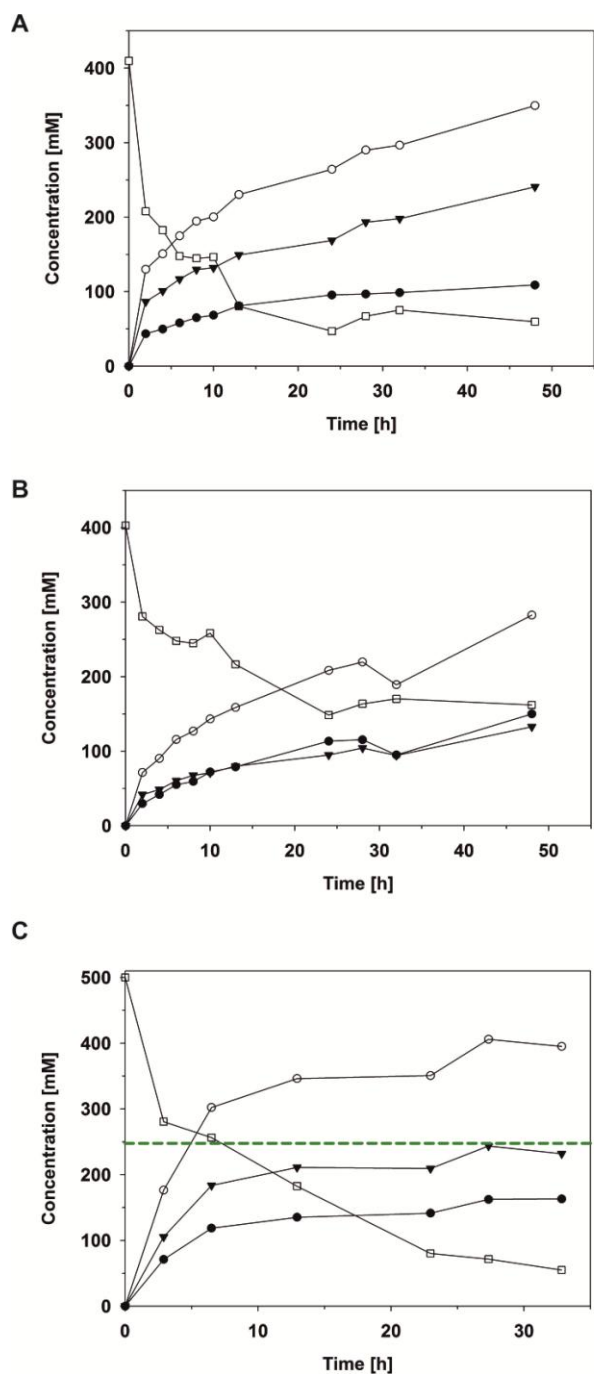
Figure S4. Time courses for the enzymatic synthesis of (*R*)-2-*O*-α-D-glucopyranosyl glyceric acid amide from 400 mM sucrose and 1000 mM racemic glyceric acid amide (A) and equimolar concentrations of sucrose and racemic glyceric acid amide (400 mM, B; 500 mM, C). Reactions in panels A and B were performed using an enzyme concentration of 7 U/mL. In panel C, an enzyme concentration of 40 U/mL was used. The green line in panel C shows the maximum theoretical yield of (*R*)-2-*O*-α-D-glucopyranosyl glyceric acid amide assuming that only (*R*)-glyceric acid amide is glucosylated by the enzyme. The symbols indicate sucrose (□), D-fructose (○), D-glucose (●) and (*R*)-2-*O*-α-D-glucopyranosyl glyceric acid amide (▼).
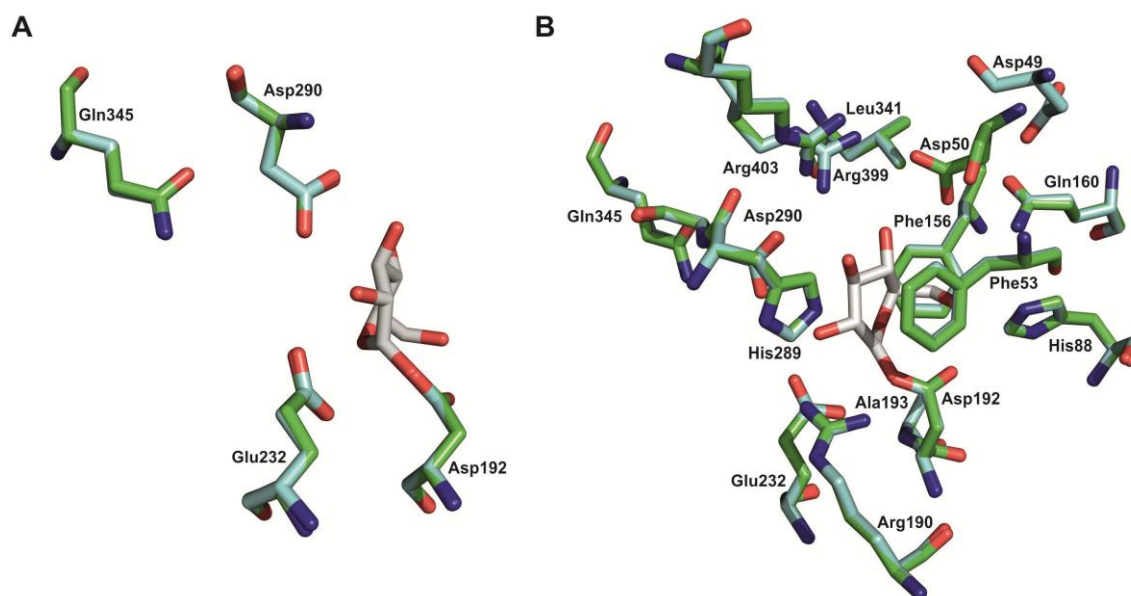
Figure S5. The homology model of sucrose phosphorylase from *Leuconostoc mesenteroides* (blue) superimposes almost perfectly on the experimental X-ray crystal structure of the glucosyl enzyme intermediate of sucrose phosphorylase from *Bifidobacterium adolescentis* (PDB-entry 2gdv; green). Amino acid sequences of the two enzymes are 34.14% identical overall. All residues known to participate in substrate binding and catalysis are completely conserved in the two sequences. A: The orientation of amino acid residues directly important for the proposed binding of glyceric acid amide (or 1,2-diols) is completely conserved in the homology model as compared to the experimental protein structure. B: Conservation of residue orientation extends to a second shell of amino acids at a 5-Å distance to the bound glucosyl moiety, except for slight differences in orientation of Asp-50 in the homology model as compared to Asp-49 in the experimental structure. However, the respective aspartic acid residue is not supposed to have a role for acceptor binding.

Table S1. Reaction conditions, substrate concentrations used, and product yields achieved in reactions performed for product identification by NMR analysis (reaction #1), isolation of residual glyceric acid amide (reaction #3), reaction optimisation for chiral synthesis of (*R*)-2-*O*-α-D-glucopyranosyl glyceric acid amide (reaction #2 - #4), and chiral resolution of racemic glyceric acid amide (reaction #5).

Reaction	Sucrose [mM]	Glyceric acid amide [mM]	Sucrose phosphorylase [Units of activity/mL]	Reaction time [h]	Theoretical product[1] [mM]	Product[2] [mM]	Product[3] [%]	Glucose[4] [mM]	Product:glucose[5]
#1	500	500	30	24	250	225	90%	250	0.9
#2	400	1000	7	48	400	241	60%	109	2.2
#3	400	400	7	48	200	133	67%	150	0.9
#4	1000	400	7	48	200	100	50%	85	1.2
#5	500	500	40	33	250	238	95%	163	1.5

In all cases, the reaction was performed in 50 mM MES buffer, pH 7.0, at 30°C and 550 rpm using purified sucrose phosphorylase. Samples were taken after appropriate time intervals, heat-treated (at 99°C for 5 min) to inactivate the enzyme and centrifuged at 13000 rpm for 10 min to remove precipitated protein prior to further analysis.

[1] Maximum theoretical concentration of (*R*)-2-*O*-α-D-glucopyranosyl glyceric acid amide, considering that only (*R*)-glyceric acid amide was glucosylated by sucrose phosphorylase whereas (*S*)-glyceric acid amide was inactive.

[2] Experimentally determined concentration of (*R*)-2-*O*-α-D-glucopyranosyl glyceric acid amide at the end of the reaction time.

[3] Yield of experimentally determined (*R*)-2-*O*-α-D-glucopyranosyl glyceric acid amide based on the maximum theoretical concentration of (*R*)-2-*O*-α-D-glucopyranosyl glyceric acid amide.

[4] Experimentally determined concentration of glucose at the end of the reaction time.

[5] The column shows the ratio of the concentration of (*R*)-2-*O*-α-D-glucopyranosyl glyceric acid amide and the concentration of glucose at the end of the reaction time. A value > 1 reflects that transglucosylation is preferred over hydrolysis under the conditions used.

1. P. Wildberger, C. Luley-Goedl and B. Nidetzky, *FEBS Lett*, 2011, **585**, 499-504.
2. P. Wildberger, A. Todea and B. Nidetzky, *Biocatal Biotransform*, 2012, **30**, 326-337.
3. C. Goedl, A. Schwarz, A. Minani and B. Nidetzky, *J Biotechnol*, 2007, **129**, 77-86.
4. G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew and A. J. Olson, *J Comput Chem*, 1998, **19**, 1639-1662.
5. Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang and P. Kollman, *J Comput Chem*, 2003, **24**, 1999-2012.
6. O. Mirza, L. K. Skov, D. Sprogøe, L. A. van den Broek, G. Beldman, J. S. Kastrup and M. Gajhede, *J Biol Chem*, 2006, **281**, 35576-35584.
7. K. Arnold, L. Bordoli, J. Kopp and T. Schwede, *Bioinformatics*, 2006, **22**, 195-201.
8. F. Kiefer, K. Arnold, M. Kunzli, L. Bordoli and T. Schwede, *Nucleic Acids Res*, 2009, **37**, D387-392.