Supporting information for

Rational design of an improved transglucosylase for production of the rare sugar nigerose

Jorick Franceus,[a] Shari Dhaene,[a] Hannes Decadt,[a] Judith Vandepitte,[a] Jurgen Caroen,[b] Johan Van der Eycken,[b] Koen Beerens[a] and Tom Desmet[*a]

[a] *Centre for Synthetic Biology (CSB), Department of Biotechnology, Ghent University. Coupure Links 653, 9000 Ghent, Belgium.*

[b] *Laboratory for Organic and Bio-Organic Synthesis, Department of Organic and Macromolecular Chemistry, Ghent University, Krijgslaan 281 S4, B-9000 Ghent, Belgium*

CONTENT

EXPERIMENTAL DETAILS

TABLES

FIGURES

MOVIES

Available as supplementary downloads

EXPERIMENTAL DETAILS

Materials and chemicals

Unless otherwise noted, all commercial chemicals were purchased from Sigma-Aldrich or Carbosynth. Primers were obtained from Integrated DNA Technologies.

Mutagenesis

All mutations were introduced into the constitutive expression plasmid pCXP34h containing the sucrose phosphorylase gene from *Bifidobacterium adolescentis*[1] by a modified two-step megaprimer based whole plasmid PCR method Sanchis protocol[2] using the primers in Table S1. A megaprimer is formed in the first cycles of the reaction, after which the whole plasmid is amplified by means of extension of the megaprimer. The PCR mix contained 0.05 U/µL PfuUltra high-fidelity DNA polymerase (Agilent), 0.2 mM dNTP mix, 2 ng/µL template and 0.1 pmol/µL of each primer (Table 1) in a total volume of 50 µL. The program started with an initial denaturation (3 min at 94 °C) followed by 5 cycles of denaturation for 30 s at 94 °C, annealing for 1 min at 55 °C and extension for 1 min/kb (size megaprimer) at 72 °C. The second stage consisted of 20 cycles of 10 s at 94 °C and extension for 2 min/kb (size entire plasmid) at 68 °C, followed by one final extension of 2 min at 72 °C. Template DNA was digested by 20 U *Dpn*I (New England Biolabs) for 1 h at 37°C. Of the remaining PCR product, 2 µL was added to 40 µL electrocompetent *E. coli* CGSC 8974 cells for transformation. Library randomisation was verified by sequencing the plasmid mixture (Macrogen) after plasmid extraction of the transformants (Qiagen Plasmid Mini Kit).

Screening by HPAEC-PAD

Individual colonies were picked from solid LB medium containing ampicillin (100 µg/mL) using an automated colony picker (QPix2, Genetix) and inoculated into sterile 96-well flat-bottomed microtiter plates (Nunc) containing 175 µL LB (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) supplemented with 100 µg/mL ampicillin. These plates were grown at 37°C and 250 rpm for 16 hours, replicated into new plates with fresh medium and grown for another 16 hours. Cells were harvested 16 hours later by centrifugation (5000 g, 20 min), the supernatans was discarded and cells were frozen at -20°C for at least 2 hours. After thawing, the cells were lysed for 30 min at 37°C by adding 100 µL of a lysis buffer containing 1 mg/mL lysozyme, 0.1 mM phenylmethylsulfonyl fluoride, 50 mM Na_2SO_4, 4 mM $MgSO_4$, 1 mM ethylenediaminetetraacetic acid and 50 mM 3-morpholinopropane-1-sulfonic acid (MOPS) buffer, pH 7. The plates were centrifuged (5000 g, 30 min), 40 µL of all crude cell extracts was transferred to a fresh plate, and 40 µL substrate solution was added to a final concentration of 100 mM sucrose and 200 mM glucose. Reactions were incubated for 24 h at 37°C, stopped by adding 160 µL 0.01 M NaOH and the plates were spun down once more. Supernatant was transferred to deep well plates and diluted to 200x with mQ water.

All reaction samples were monitored by High Performance Anion Exchange Chromatography (HPAEC; Dionex ICS-3000, Thermo Scientific) with a CarboPac PA20 pH-stable column and Pulsed Amperometric Detection (PAD). 10 µL sample was analysed at a constant flow rate of 0.5 mL/min at 30°C. For samples from screening and characterisation reactions, the eluent composition was 100 mM NaOH and 10 mM NaOAc for 9 min. For samples from the large-scale production of nigerose, the eluent composition was 30 mM NaOH during the first 10 min, then gradually changed to 100 mM NaOH and 10 mM NaOAc during one min, kept constant for 9 min, and changed again to the initial conditions within 1 min, followed by a final equilibration period of 3 min.

Characterisation of mutants

Hits from screening were expressed on larger scale and His-tag purified as previously described.[3] Expression was verified by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (10% gels) and Western blot as described in earlier work (Fig S7)[1]. For Western blot, proteins were transferred to a Hybond-ECL membrane (Amersham Biosciences) and the two antibodies used were mouse anti-polyhistidine monoclonal antibody (H1029, Sigma), followed by anti-mouse IgG alkaline phosphatase conjugate antibody (A3562, Sigma), each in a 1:2000 dilution. Bands were visualised with BCIP/NBT.
Reactions were performed at 52°C in 50 mM MOPS buffer pH 7.0. The activity and selectivity of mutants was initially compared from the initial reaction rates using purified enzyme, 100 mM sucrose and 200 mM glucose. Kinetic parameters were determined by performing reactions using ~ 0.5 mg/mL purified enzyme at 52°C using 750 mM sucrose and 75-2500 mM glucose in 50 mM MOPS buffer pH 7.0. All reactions were performed in triplicate. Samples were inactivated and analysed by HPAEC-PAD as described above.

Unit definition

One unit was defined as the amount of enzyme that produces one µmol of nigerose (or maltose) per minute under the specified conditions.

Large-scale expression of BaSP-YGQF

Recombinant expression of BaSP-YGQF (R135Y/D342G/Y344Q/Q345F) in *E. coli* CGSC 8974 was performed on 15-L scale. The strain was grown at 37°C in 500 mL LB medium (10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl) supplemented with 100 mg/L ampicillin. After overnight growth, the culture was inoculated in 15 L of double LB medium (20 g/L tryptone, 10 g/L yeast extract, 10 g/L NaCl) supplemented with 30 g/L glucose and 100 mg/L ampicillin in a 30 L BioStat C reactor (B. Braun Biotech Inc.). Temperature, pH, aeration and stirrer speed were set at 37°C, 7.0, 1.5 vvm and 800 rpm, respectively. When required, anti-foam (10% (v/v) silicone Snapsil RE 20, VWR BDH Prolabo) was added manually. Cells were harvested by centrifugation, stored at -20°C, lysed and heat purified as described earlier.[4]

Production and isolation of nigerose

Reaction at 0.5-L scale was performed at 52°C and 35 rpm in milliQ water using 1.5 M sucrose, 1.5 M glucose and 10 g/L of heat purified BaSP-YGQF in a 2-L Biostat M reactor. The reaction was stopped by heating at 95°C for 10 min. Glucoamylase and invertase (10 U/mL, Megazyme) were added and the reaction was incubated at 37°C for 30 min to degrade contaminating disaccharides, followed by inactivation at 95°C for 10 min. The resulting mixture was further purified using the Sugar Purification System from Knauer (Berlin, Germany). This preparative liquid chromatography system was equipped with an at 75°C heated Vertex Plus AX column (dimension: 250x30mm) loaded with Knauer Eurokat Na-resin (mesh 25-56 µm) and a Refractive Index Detector (RID). The setup allows separation of carbohydrates via Anion Exchange Chromatography (AEC) using only ultrapure water as solvent, in our case at a flow rate of 7.5 ml.min^{-1}. Similar AEC are industrially applied for the separation of for example glucose and fructose in high-fructose corn syrup production.[5] After 1 min column flushing with ultrapure water (7.5 ml.min-1), the glucoamylase and invertase treated sample (19.5 % brix, 2 ml) was injected on the column (2 ml.min-1), after which ultrapure water was used as solvent to wash off the carbohydrates. The nigerose fraction was collected (fraction: 11.2-13.0 min) while other frations were wasted (e.g. some Maillard reaction products eluted between approx. min 7.5 and 11.0 and glucose/fructose eluted around 13.2-16.2 min), after 24 min a new run was automatically started. Subsequently the water was evaporated using a Rotavapor R-200 (Büchi) with a 60 °C water bath and under complete vacuum. Nigerose was dried completely using the Rotavapor system, obtaining a white to off-white powder.

Analytical data

Precoated silica gel plates (Macherey-Nagel SIL G-25 UV_{254}) were used for TLC. ^{1}H-NMR and ^{13}C-NMR spectra were recorded on a Bruker Avance 400 spectrometer with chemical shifts reported in parts per million, referenced to the residual solvent signal (D_2O: 4.75 ppm). Coupling constants, *J*, are reported in hertz (Hz). Electrospray mass spectra were recorded on an Agilent 1100 series single quadrupole MS detector type VL with an APCI source and an API-ES source, provided with a Phenomenex Luna C18 (2), 5µm 250mm x 4.60mm column. Infrared spectra were recorded on a Perkin-Elmer 1000 FT-IR infrared spectrometer (horizontal attenuated total reflection (HATR)). Optical rotation was measured on a Perkin Elmer 241 polarimeter.

TLC: Rf = 0.24 in CH_3CN/H_2O 8/2; Rf = 0.49 in *n*-propanol/EtOH/H_2O 7/1/2. ^{1}H NMR (D_2O, 400MHz): anomer ratio α/β ~4/6: 5.35 (d, *J*=4.0Hz, 1H, α-anomer), 5.33 (d, *J*=3.9Hz, 1H, β-anomer), 5.21 (d, *J*=3.9Hz, 1H, α-anomer), 4.64 (d, *J*=7.9Hz, 1H, β-anomer), 3.95-4.02 (m, 1H, α+β-anomer), 3.27-3.90 (m, 11H, α+β-anomer) ppm. ^{13}C NMR (D_2O, 100MHz, APT): α+β-anomer: 99.04 (CH), 99.00 (CH), 95.94 (CH), 92.21 (CH), 82.09 (CH), 79.57 (CH), 75.62 (CH), 72.86 (CH), 72.84 (CH), 72.81 (CH), 71.72 (CH), 71.70 (CH), 71.67 (CH), 71.61 (CH), 71.16 (CH), 70.07 (CH), 70.03 (CH), 70.00 (CH), 69.39 (CH), 69.27 (CH), 60.53 (CH_2), 60.35 (CH_2), 60.34 (CH_2), 60.18 (CH_2) ppm. ESMS: 360.2 ($M+NH_4^+$, 100). IR (HATR): 3270 (s, br), 2928 (m), 1722 (w), 1646 (w), 1434 (m), 1353 (m), 1259 (m), 1199 (w), 1144 (m), 1073 (w), 1012 (vs), 917 (m), 844 (w), 799 (w), 773 (w), 750 (w), 711 (w), 639 (w) cm^{-1}. $[\alpha]_D^{20}$ +133° (c = 1.00, H_2O).

Molecular dynamics simulations

Molecular dynamics simulations were carried out on the crystal structure of *B. adolescentis* mutant Q345F (PDB ID 5MAN). The additional mutations in BaSP-YGQF (R135Y, D342G, Y344Q) were modelled in YASARA Structure.[6,7] Simulations were performed in GROMACS 4.6.5[8] in a cubic box of tip3p water providing a minimum layer of 7.5 Å of water on each side of the protein. Ion strength was brought to 0.15 mol/L. All simulations were performed with periodic boundaries to avoid edge effects and with a timestep of 1.25 fs. Under an isothermal-isochoric (NVT) ensemble at 323 K controlled by a Berendsen thermostat,[9] a steepest descent energy minimisation was performed. Then, pressure and density were stabilised under isothermal-isobaric (NPT) conditions using the Berendsen thermostat and a Parrinello-Rahman barostat.[10] The LINCS constraints[11] were used to fix the bond lengths and angles after the integration of the forces. Long-range electrostatics were calculated by the Particle Mesh Ewald algorithm.[12,13] Finally, two independent 200 ns production runs were started for each protein and snapshots of the simulations were taken at regular time intervals. Protein figures and movies were made in PyMOL 2.0.[14]

TABLES

Table S1. Primers used in this study. Mutations are underlined.

Construct	Template	Primer sequences (5'-3')
Q345F	Wild-type	CTGGATCTGTACTTCGTTAATAGTACC CAGACCGCTTCTGCGTTCTG
L341X+Q345F	Q345F	GCCGCTCTAGAAGAAGCTTG TAACGAAGTACAGATCMNNGTTCGATGCCG
D342X+Q345F	Q345F	GCCGCTCTAGAAGAAGCTTG TAACGAAGTACAGMNNCAGGTTCGATGCCG
L343X+Q345F	Q345F	GCCGCTCTAGAAGAAGCTTG TAACGAAGTAMNNATCCAGGTTCGATGCCG
Y344X+Q345F	Q345F	GCCGCTCTAGAAGAAGCTTG CTATTAACGAAMNNCAGATCCAGGTTCGATGCCGCC
D342G+Y344Q+Q345F	Q345F	GCCGCTCTAGAAGAAGCTTG CTATTAACGAACTGCAGACCCAGGTTCGATGCCGCC
L341X+D342G+Y344Q+Q345F	D342G+Y344Q+Q345F	GGCGGCATCGAACNNKGGTCTGCAGTTCGTTAATAGTAC CAAGCTTCTTCTAGAGCGGC
D342G+L343X+Y344Q+Q345F	D342G+Y344Q+Q345F	GGCGGCATCGAACCTGGGTNNKCAGTTCGTTAATAGTAC CAAGCTTCTTCTAGAGCGGC
D342G+Y344Q+Q345X	D342G+Y344Q+Q345F	GCCGCTCTAGAAGAAGCTTG GTACTATTAACMNNCTGCAGACCCAGGTTCGATGCCGCC
Y132X+D342G+Y344Q+Q345F	D342G+Y344Q+Q345F	CTGGCAGGCATTNNKCGTCCGCGTCC CAGACCGCTTCTGCGTTCTG
P134X+D342G+Y344Q+Q345F	D342G+Y344Q+Q345F	CATTTATCGTNNKCGTCCGGGTCTGCC CAGACCGCTTCTGCGTTCTG
R135X+D342G+Y344Q+Q345F	D342G+Y344Q+Q345F	CATTTATCGTCCGNNKCCGGGTCTGCC CAGACCGCTTCTGCGTTCTG
Y196X+D342G+Y344Q+Q345F	D342G+Y344Q+Q345F	GCCGCTCTAGAAGAAGCTTG CCTGCTTCTTTGGCACCMNNGCCAACCGCATCCAGG

FIGURES

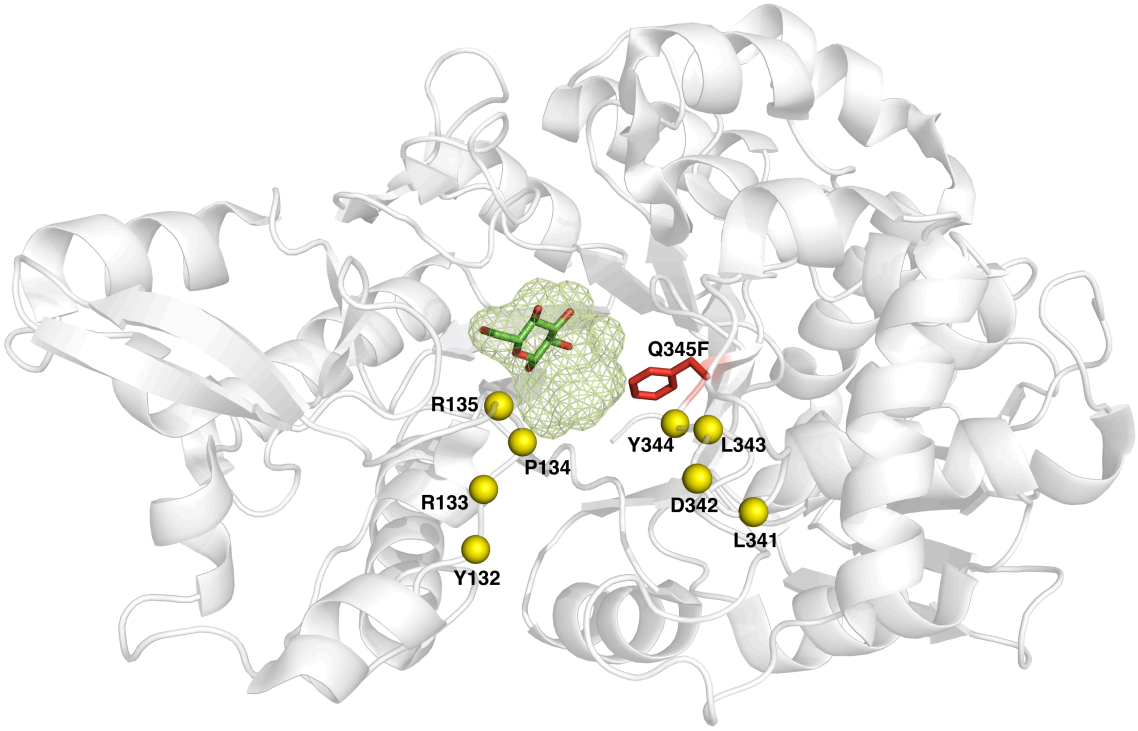


Fig. S1 Structure of *B. adolescentis* sucrose phosphorylase with the location of amino acids targeted for saturation mutagenesis (yellow spheres). Also shown is the mutation enabling synthesis of nigerose is shown (Q345F, red sticks), the glucosyl donor moiety (green sticks), and the active site pocket (green mesh).

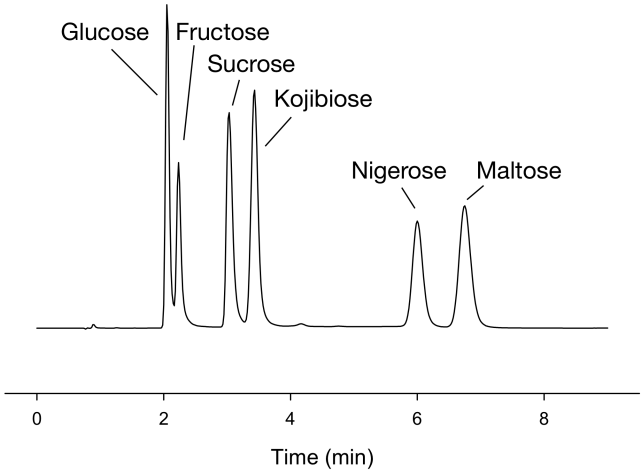


Fig. S2 High performance anion exchange chromatography profile for screening with sucrose and glucose as substrate.

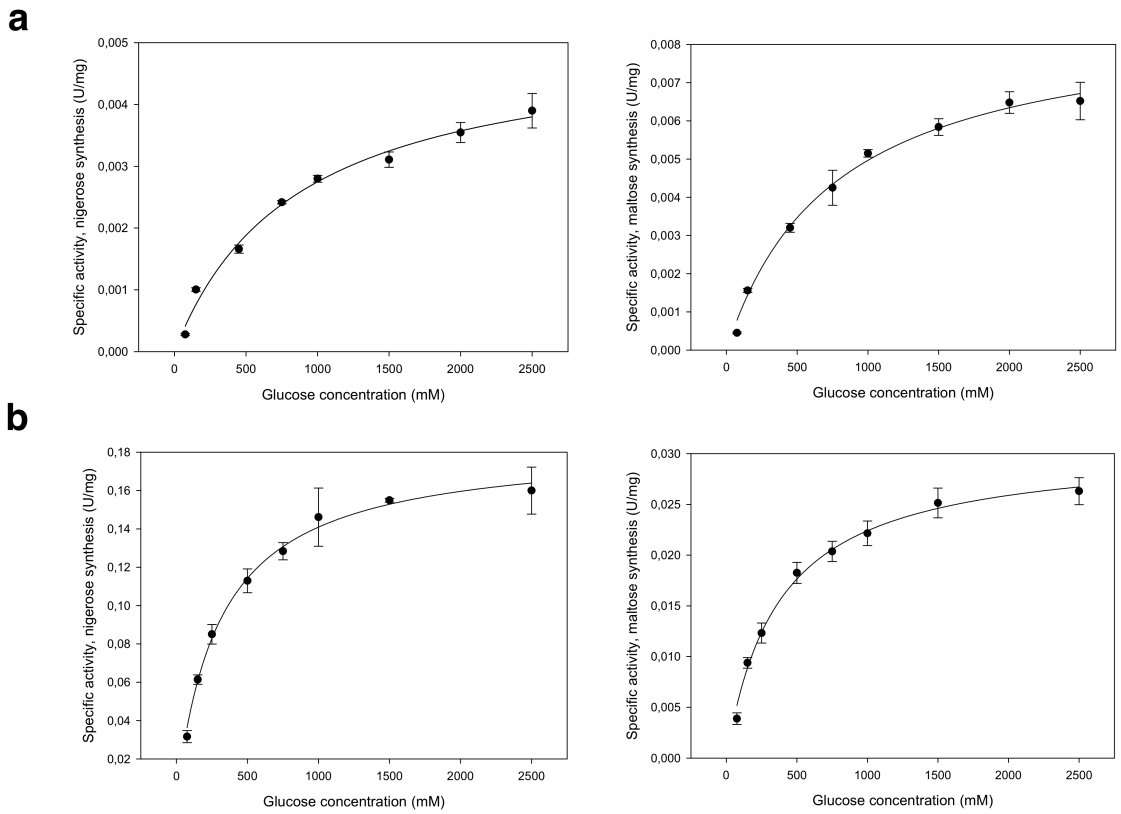


Fig. S3 Michaelis-Menten curves for glucose, for mutants (a) Q345F and (b) R135Y/D342G/Y344Q/Q345F in aqueous solution.

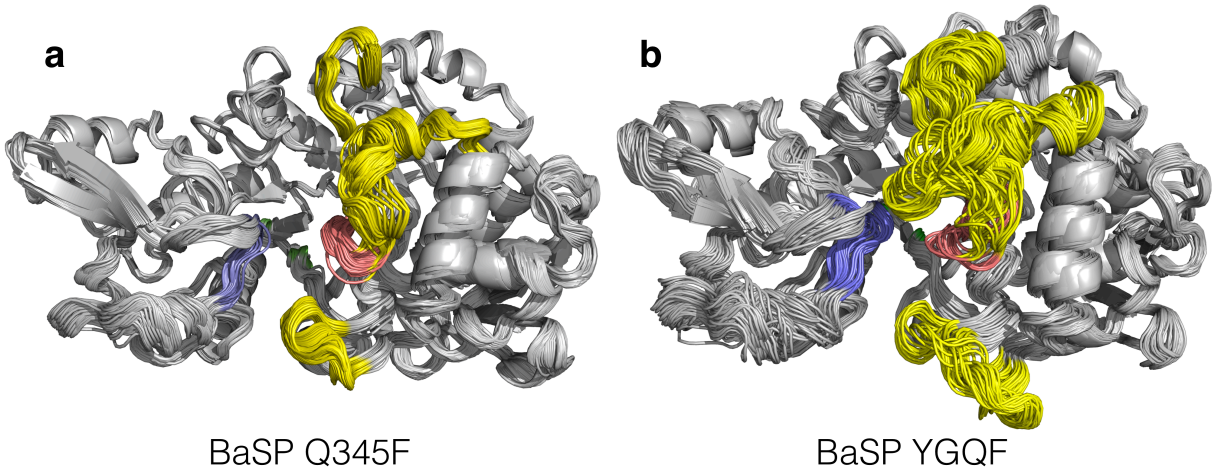


Fig. S4 Overlay of snapshots taken during molecular dynamics simulations of (a) BaSP Q345F and (b) BaSP YGQF (R135Y/D342G/Y344Q/Q345F). Colored are loop ^{132}YRPR^{135} (blue), loop ^{341}LDLYQ^{345} (red) and the surrounding loops with increased flexibility in BaSP YGQF (yellow).

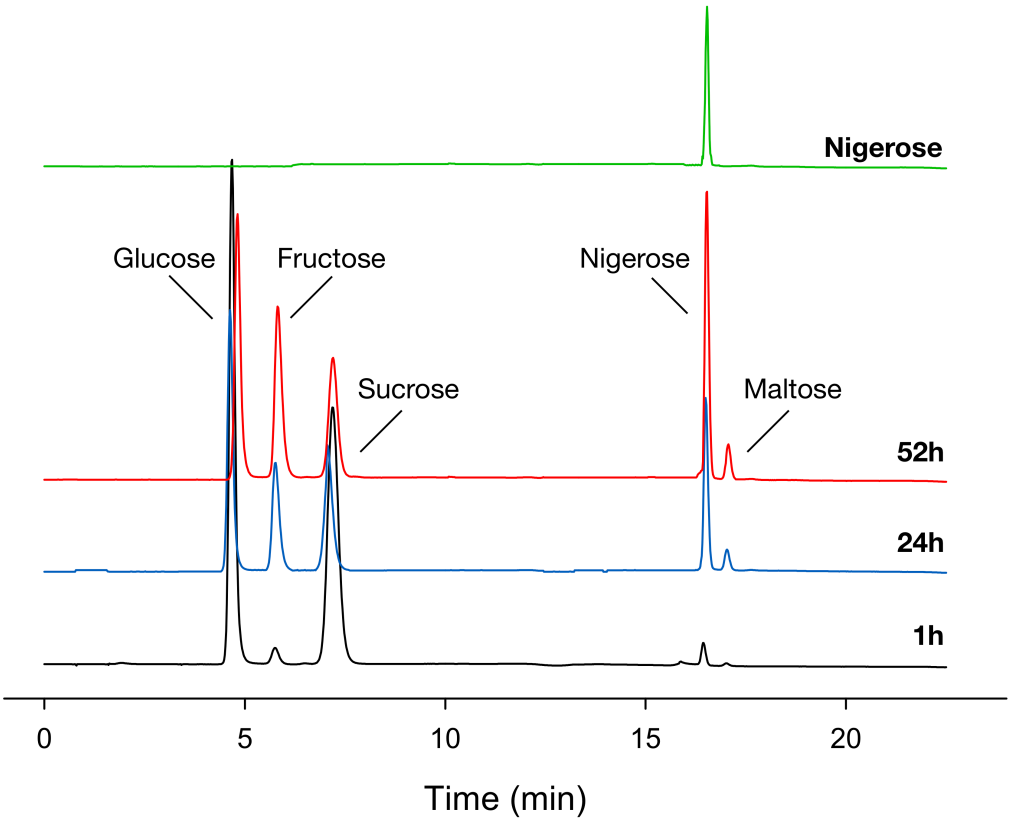


Fig. S5 HPAEC profiles of samples taken during the production of nigerose on 0.5-L scale after 1h, 24h and 52h, as well as the HPAEC profile for nigerose.

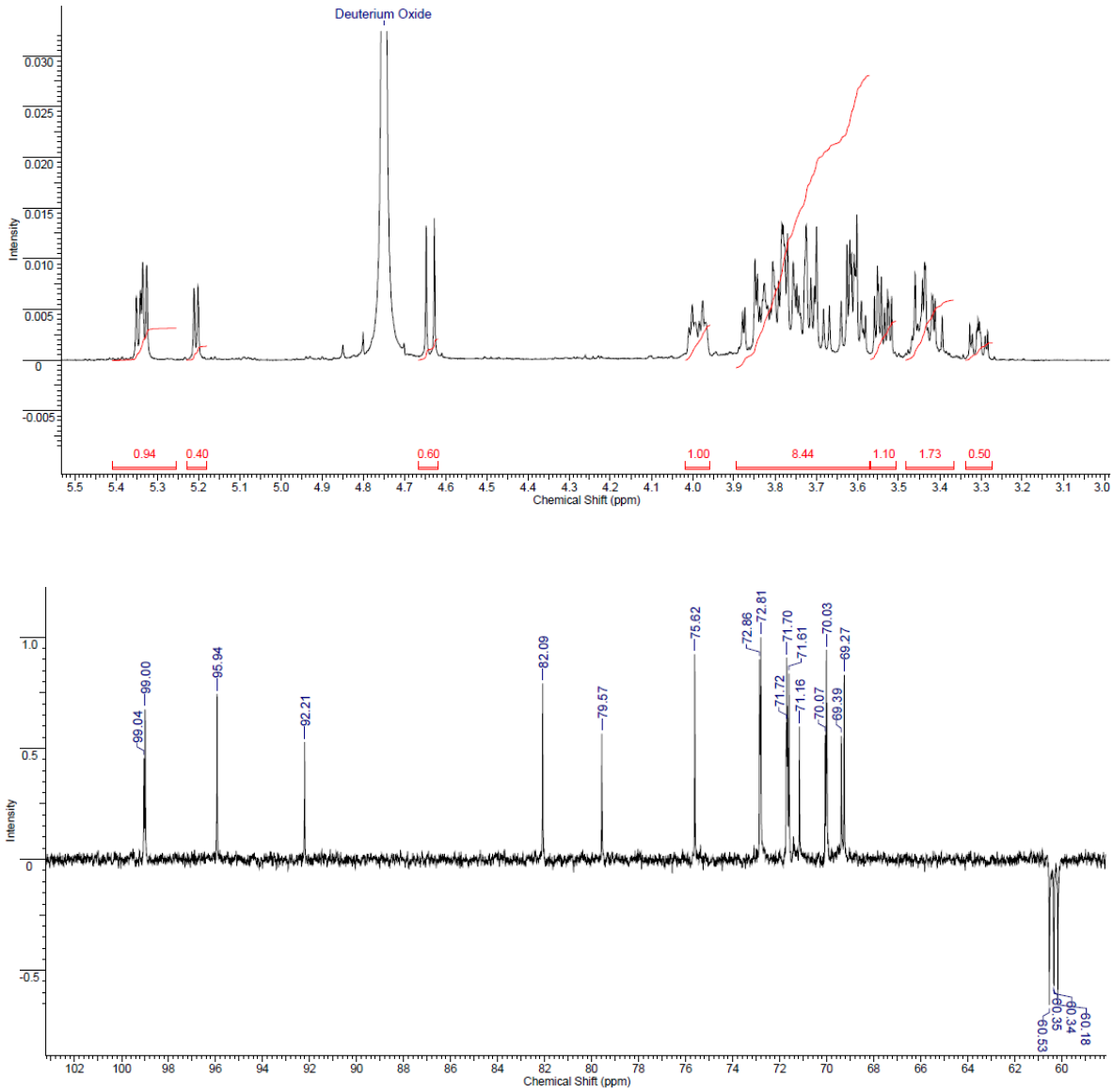


Fig. S6 ^{1}H NMR and ^{13}C NMR (APT) spectrum of nigerose.

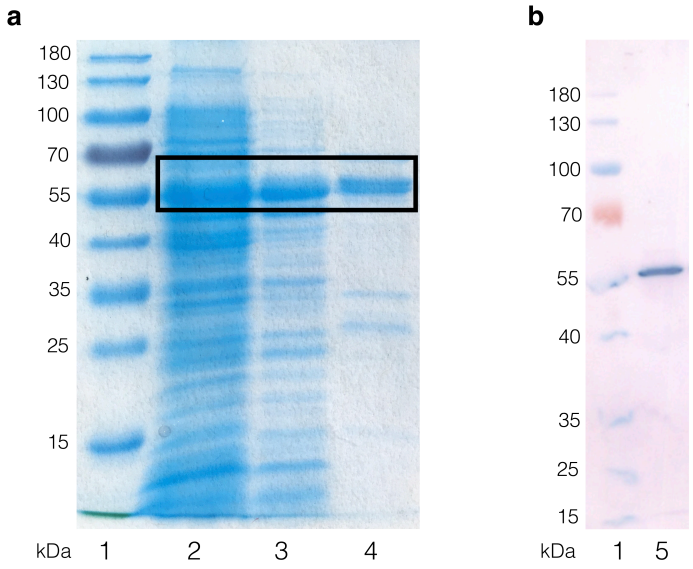
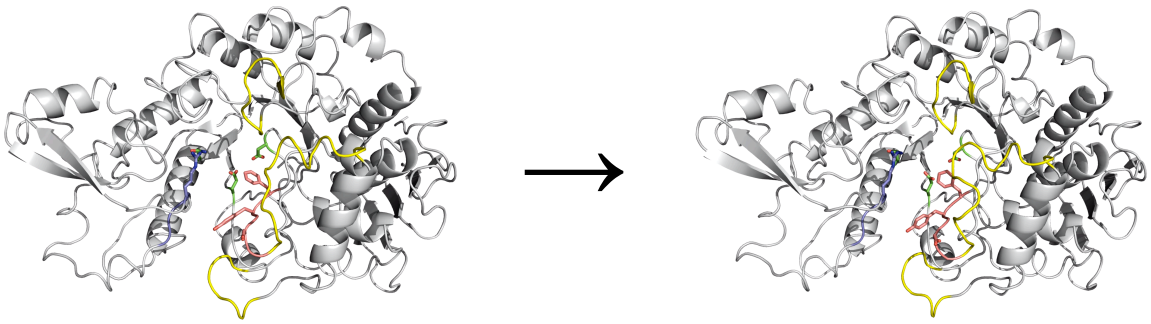


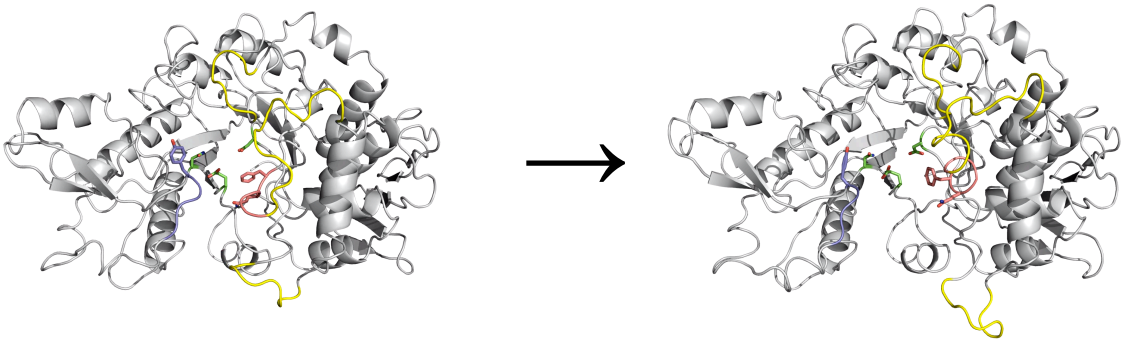
Fig. S7 (a) SDS-PAGE and (b) Western blot of protein samples. (1) PageRuler prestained protein ladder (ThermoFisher), (2) cell-free extract obtained from large-scale expression of BaSP-YGQF, (3) heat-purified extract used for production of nigerose on 0.5-L scale, (4) further purification by Ni-NTA chromatography, (5) western blot visualising the His_6-tagged protein BaSP-YGQF. The theoretical size of BaSP-YGQF is 57,6 kDa.

MOVIES

(Files available as supplementary downloads)



Supplementary movie 1 Snapshots taken during molecular dynamics simulation of BaSP Q345F. Colored are loop ^{132}YRPR^{135} (blue), loop ^{341}LDLYQ^{345} (red) and the surrounding loops with increased flexibility in BaSP-YGQF (yellow). The positions that are different in BaSP-YGQF (Arg135, Asp342, Tyr344) are shown in sticks. Additionally, the catalytic nucleophile, acid-base catalyst and transition state stabilizer are shown in green sticks.



Supplementary movie 2 Snapshots taken during molecular dynamics simulation of BaSP-YGQF. Colored are loop ^{132}YRPY^{135} (blue), loop ^{341}LGLQF^{345} (red) and the surrounding loops with increased flexibility in BaSP-YGQF (yellow). The positions that are different in BaSP Q345F (Tyr135, Gly342, Gln344) are shown in sticks. Additionally, the catalytic nucleophile, acid-base catalyst and transition state stabilizer are shown in green sticks.

REFERENCES

1 D. Aerts, T. Verhaeghe, M. De Mey, T. Desmet and W. Soetaert, *Eng. Life Sci.*, 2011, **11**, 10–19.

2 J. Sanchis, L. Fernández, J. D. Carballeira, J. Drone, Y. Gumulya, H. Höbenreich, D. Kahakeaw, S. Kille, R. Lohmer, J. J.-P. Peyralans, J. Podtetenieff, S. Prasad, P. Soni, A. Taglieber, S. Wu, F. E. Zilly and M. T. Reetz, *Appl. Microbiol. Biotechnol.*, 2008, **81**, 387–397.

3 T. Verhaeghe, K. De Winter, M. Berland, R. De Vreese, M. D'hooghe, B. Offmann and T. Desmet, *Chem. Commun.*, 2016, **52**, 3687–3689.

4 K. Beerens, K. De Winter, D. Van De Walle, C. Grootaert, S. Kamiloglu, L. Miclotte, T. Van De Wiele, J. Van Camp, K. Dewettinck and T. Desmet, *J. Agric. Food Chem.*, 2017, **65**, 6030–6041.

5 C. G. Huber and G. K. Bonn, *J. Chromatogr. Libr.*, 1995, **58**, 147–177.

6 E. Krieger, G. Vriend and C. Spronk, .

7 G. Vriend, *J. Mol. Graph.*, 1990, **8**, 52–56.

8 E. Lindahl, B. Hess and D. Van Der Spoel, *J. Mol. Model.*, 2001, 306–317.

9 H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.

10 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182–7190.

11 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.

12 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.

13 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577–8592.

14 The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.