Supporting Information

Hill et al. 10.1073/pnas.1311990110

SI Materials and Methods

Protein Expression and Purification. Wild-type murine β-galactocerebrosidase (mGALC) was expressed and purified as described previously with minor modifications detailed below (1). The E258Q mutation was made by mutagenesis PCR in the previously described pSecTag2B–mGALC plasmid using forward primer 5′-GGAAGAAGCTGTGGTCATCTCAAGATTTTAGCACTATCAAC-3′ and reverse primer 5′-GTTGATAGTGCTAAAATCTTGAGATGACCACAGCTTCTTCC-3′. The resultant construct was confirmed by sequencing. A HEK 293T cell line stably expressing E258Q mGALC was then established and cultured as described previously (1).

HEK 293T cells stably expressing and secreting His_6-tagged GALC were cultured in 2,125-cm^2 roller bottles (Greiner) using Freestyle 293 expression media (Gibco). After 5–7 d, conditioned medium was harvested and 0.2 μm-filtered. Per 100 mL of conditioned medium, 0.5 mL of Ni-NTA resin was preequilibrated in wash buffer (30 mM imidazole, PBS, pH 7.4) before protein binding (120 min; 4 °C). Each 0.5 mL volume of resin was pelleted by centrifugation (600 × *g*; 7 min; 4 °C) and washed three times in 50 mL of wash buffer. Washed resin was applied to a gravity column, and protein was eluted with 300 mM imidazole and PBS (pH 7.4). Eluted protein was buffer-exchanged by dialysis (10-kDa molecular mass cutoff; Thermo Scientific) into either PBS for storage at 4 °C or 150 mM NaCl and 10 mM Hepes (pH 7.4) for crystallization. When stored at 4 °C, purified GALC was stable and retained full enzymatic activity for at least 2 wk.

Crystallization. GALC was concentrated to 2.5 mg/mL in 150 mM NaCl and 10 mM Hepes (pH 7.4) using Amicon Ultra 10-kDa molecular mass cutoff centrifugal concentrators (Millipore). Crystallization experiments were set up in 96-well sitting drop vapor diffusion plates using an Innovadyne Screenmaker microfluidic handling platform. Microseeding (Hampton Research SeedBead) was necessary for diffraction-quality, untwinned crystals (2). Crystals were grown in 600-nL drops (200 nL of protein, 200 nL of precipitant, and 200 nL of seed) at 20 °C by equilibration against an 80-μL reservoir of 0.2 M sodium acetate, 0.1 M sodium cacodylate (pH 6.8), and 34% wt/vol polyethylene glycol 8000.

Crystal Soaks with Small Molecules. Crystals of GALC used for soaking experiments were small (50 × 50 × 43 μm; 5.4×10^{-5} mm^3) and belonged to space group *R*32 with high solvent content (62%; $V_m = 3.2$ $Å^3 \cdot Da^{-1}$). The 4NβDG (Sigma; N1252) and 1,2-dideoxy-D-lyxo-hex-1-enopyranose (D-galactal) (Acros Organics 370442500) solutions were prepared at 100 mM in water and diluted to 20 mM in crystallization reservoir solution. Soaks were initiated by adding 0.5 μL of ligand solution directly to the crystallization drop, giving an effective concentration of 10–15 mM depending on the final equilibrated volume of the crystallization drop (~ 250–500 nL). Immediately after adding ligand solution, 0.5 μL of perfluoropolyether oil (Hampton Research) was layered over the drop. In this way, crystals were soaked at room temperature for 10 min with 4NβDG to produce enzyme–product complex and for 2 h with D-galactal to produce covalent intermediate complex. Crystals were cryoprotected by removal from the drop through perfluoropolyether oil before flash-cooling in liquid nitrogen. The GALC–4NβDG complex was isolated by a diffusion trapping method incorporating rapid crystal harvesting and freezing (3). Following addition of substrate solution and perfluoropolyether oil, individual crystals were immediately removed from the drop and flash-cooled in liquid nitrogen, resulting in substrate soak times of between 20 and 40 s. Two different datasets were collected on crystals soaked for 20 and 40 s and structure analysis (see below) revealed both to possess substrate at full occupancy. The dataset from the 40-s soak was chosen for refinement as it was of marginally higher resolution.

Data Collection, Structure Determination, and Refinement. Diffraction data were recorded at Diamond Light Source beamline I04-1 on a Pilatus 2M detector (Dectris). Datasets were collected at λ = 0.98 Å at a temperature of 100 K. Diffraction data were indexed and integrated using iMOSFLM (4) and then scaled and merged using AIMLESS (5, 6). Resolution cutoff was decided by $CC_{1/2}$ value of >0.5 and $I/\sigma I$ of >2.0 in the outer resolution shell (7). The reflections excluded from refinement ("FreeR") were identical to those used previously for the unliganded GALC structure (1). Unbiased electron-density maps of GALC–ligand complexes were generated following rigid-body refinement in phenix.refine (8) using the native unliganded GALC as the input model (3ZR5). Protein Data Bank files and CIF restraints for ligands were generated from alphanumeric SMILES strings using eLBOW (9). All further refinement was performed iteratively using COOT (10) and phenix.refine. Refinement of the occupancy of the substrate, either as a single group or as two groups consisting of the galactose moiety and the leaving group, with fixed or variable *B* factors, in all cases showed the substrate to be fully occupied. MolProbity (11) was consulted throughout the refinement process. Structural figures and movies were rendered using PyMOL (Schrödinger LLC). For clarity, electron-density maps are displayed within 1.5–2.0 Å of highlighted residues. Hydrogen-bonding graphical representations were created using LIGPLOT+ (12, 13).

Enzymatic Activity Assays. The low solubility limit of the fluorogenic substrate 4-methylumbelliferyl-β-D-galactopyranoside used in previous studies meant that the available concentration range was insufficient for Michaelis–Menten experiments. Instead, the chromogenic substrate 4-nitrophenyl-β-D-galactopyranoside (4NβDG) was used to determine a pH activity profile and enzyme kinetic parameters for GALC. Product 4-nitrophenol has an absorbance peak at 410 nm when fully deprotonated at pH > 9.5, hence endpoint assays were designed in which reactions were terminated by addition of alkaline stopping buffer (360 mM NaOH, 280 mM glycine, pH 10.6) before colorimetric quantification.

For pH-profile experiments, 10 mM 4NβDG was prepared in a range of citrate/phosphate buffers (pH 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, and 7.0) supplemented with 50 mM NaCl. Before mixing, enzyme and substrate solutions were pre-equilibrated to 37 °C. GALC was added to a final concentration of 13.4 nM per 1.5-mL reaction and incubated with shaking at 37 °C. At eight time points, 200 μL from each reaction was removed, transferred to a transparent 96-well plate and terminated with 50 μL of stopping buffer. Absorbance at 410 nm (A_{410}) was measured with a SpectraMax M5 plate reader using SoftMax Pro software (Molecular Devices). Absorbance data were converted to 4-nitrophenol concentration by using a standard curve, performed in triplicate (Fig. S1*A*). Mean data from all time points are presented as $\mu M\ min^{-1}$.

For steady-state kinetic experiments, 4NβDG was dissolved to concentrations of 50, 25, 12.5, 6.25, 3.13, 1.56, and 0.78 mM in 20 mM sodium acetate, 150 mM NaCl, 0.1% vol/vol Nonidet P-40, pH 4.6. Purified GALC was dialyzed into an identical buffer. Before mixing, enzyme and substrate solutions were pre-equilibrated to

37 °C. GALC was added to final concentrations of 13.4, 10.1, 7.54, 5.66, 4.24 and 3.18 nM per 700-μL reaction and incubated with shaking at 37 °C. At three time points, 200 μL from each reaction was removed, transferred to a transparent 96-well plate and terminated with 50-μL stopping buffer. Absorbance at 410 nm was measured as above.

For a valid kinetic analysis, it was important to ensure that Michaelis-Menten steady-state assumptions were satisfied (14). In all experiments, 4-nitrophenol production was plotted against time to confirm linearity (Fig. S1*B*), indicative of initial velocity conditions. At each substrate concentration, observed initial velocity was plotted against enzyme concentration to check linearity (Fig. S1*C*). K_m and V_{max} parameters were obtained from plots of initial velocity against substrate concentration by nonlinear curve fitting to the Michaelis-Menten equation using Prism 5 (GraphPad). k_{cat} was determined as the gradient of the linear plot of V_{max} against enzyme concentration. Experiments were performed in triplicate.

To study D-galactal inhibition, steady state kinetic experiments were performed as described above, using 3.18 nM GALC and D-galactal at concentrations of 5000, 1000, 500, 100, 50 and 10 μM. K_i was obtained from plots of initial velocity against substrate concentration by curve-fitting to a competitive-inhibition model using Prism 5 (GraphPad). K_i is represented graphically as the x intercept of the linear plot of $K_m obs$ against D-galactal concentration. (Fig. S5). Experiments were performed in triplicate.

Fluorescence-Based Thermal-Shift Assay. To characterize thermal stability of the E258Q GALC protein, melting experiments were performed using a Bio-Rad MiniOpticon RT-PCR thermal cycler (20 °C to 95 °C, 1 °C per step, 20-s equilibration time per step). Reactions of 50 μL comprised 5.0 μg of purified GALC and "5×" SyPRO Orange dye (Invitrogen molecular probes S6651) in 20 mM sodium acetate (pH 4.6). The protein T_m was the inflection point of the sigmoidal melting curve, obtained by curve fitting using Prism 5 (GraphPad). All denaturation experiments were performed in triplicate.

Circular Dichroism Spectroscopy. To ensure that E258Q GALC protein was correctly folded, circular dichroism (CD) spectra were recorded at 20 °C on a Jasco J-810 spectropolarimeter equipped with a Jasco PTC-348WI temperature controller. Purified wild-type and mutant GALC were dialyzed into 20 mM sodium acetate, pH 4.6. CD spectra were collected over the wavelength range 195–260 nm and with a resolution of 0.1 nm, a bandwidth of 1 nm, and a response time of 1 s. Final spectra were the sum of 20 scans accumulated at a speed of 50 nm/min.

1. Deane JE, et al. (2011) Insights into Krabbe disease from structures of galactocerebrosidase. *Proc Natl Acad Sci USA* 108(37):15169–15173.
2. Walter TS, et al. (2008) Semi-automated microseeding of nanolitre crystallization experiments. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 64(Pt 1):14–18.
3. Hajdu J, et al. (2000) Analyzing protein functions in four dimensions. *Nat Struct Biol* 7(11):1006–1012.
4. Leslie AGW (1992) Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* (Biotechnology and Biological Sciences Research Council, Wiltshire, UK), No. 26.
5. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* 62(Pt 1):72–82.
6. Evans PR (2011) An introduction to data reduction: Space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):282–292.
7. Karplus PA, Diederichs K (2012) Linking crystallographic model and data quality. *Science* 336(6084):1030–1033.
8. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
9. Moriarty NW, Grosse-Kunstleve RW, Adams PD (2009) Electronic Ligand Builder and Optimization Workbench (eLBOW): A tool for ligand coordinate and restraint generation. *Acta Crystallogr D Biol Crystallogr* 65(Pt 10):1074–1080.
10. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
11. Chen VB, et al. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12–21.
12. Wallace AC, Laskowski RA, Thornton JM (1995) LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* 8(2):127–134.
13. Laskowski RA, Swindells MB (2011) LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* 51(10):2778–2786.
14. Michaelis L, Menten ML, Johnson KA, Goody RS (2011) The original Michaelis constant: Translation of the 1913 Michaelis-Menten paper. *Biochemistry* 50(39):8264–8269.

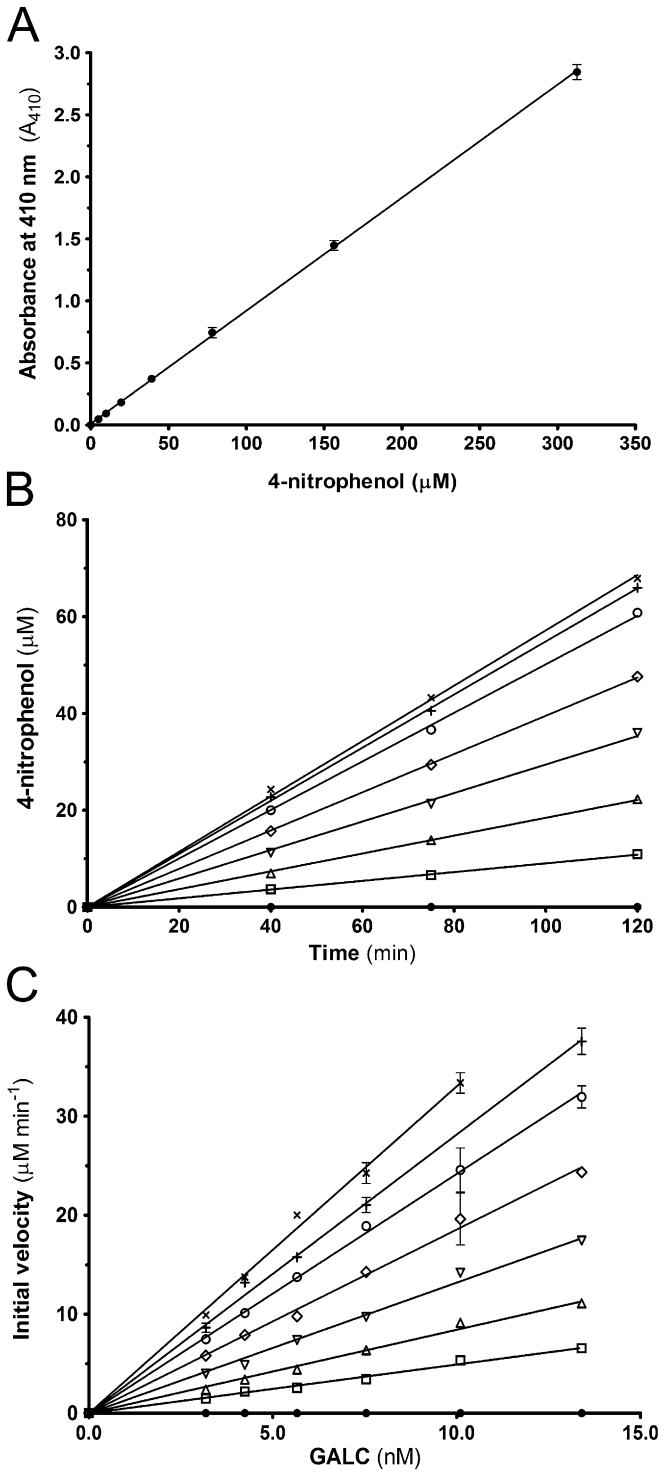


Fig. S1. Validation of Michaelis–Menten steady-state conditions. (*A*) Standard curve used for the conversion of A_{410} to 4-nitrophenol concentration. SEM error bars are shown. (*B*) Plot confirming initial velocity conditions at 4NβDG concentrations of 50 mM (×), 25 mM (+), 12.5 mM (○), 6.25 mM (◇), 3.13 mM (▽), 1.56 mM (△), 0.78 mM (□), and no substrate (●). (*C*) Plot showing a linear relationship between observed initial velocity and enzyme concentration at 4NβDG concentrations of 50 mM (×), 25 mM (+), 12.5 mM (○), 6.25 mM (◇), 3.13 mM (▽), 1.56 mM (△), 0.78 mM (□), and no substrate (●). SEM error bars are shown.

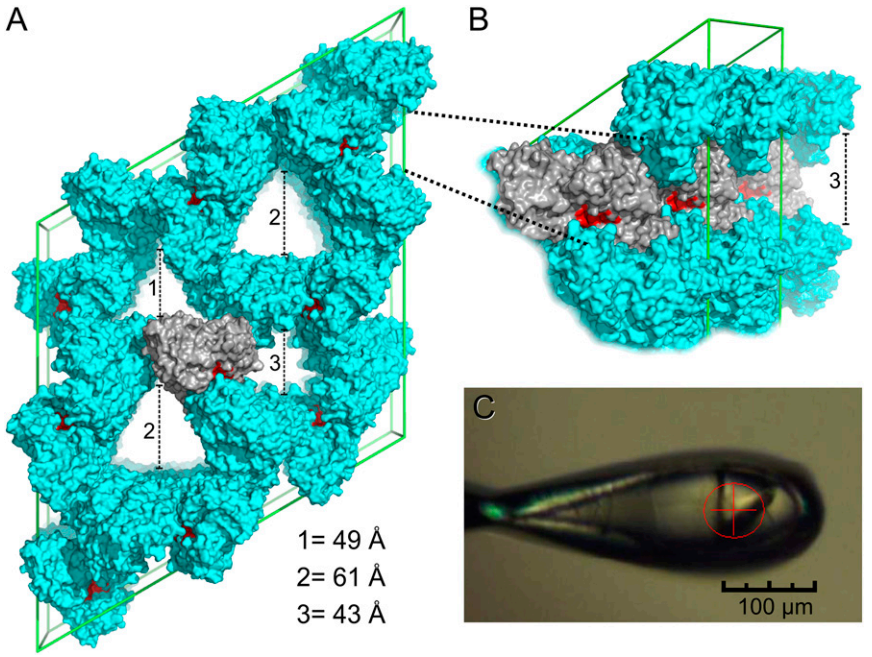


Fig. S2. Details of solvent channels and crystal dimensions. (*A*) Packing of *R*32 unit cell viewed along *c* axis. One GALC molecule is present per asymmetric unit (gray). Active sites are highlighted (red). Crystals are 61.9% solvent with a Matthews coefficient of 3.229 $A^3 \cdot Da^{-1}$. Each unit cell contains three copies of three major solvent channels denoted 1, 2, and 3. Minor solvent channels are also present along *a* and *b* axes. (*B*) View of solvent channel 3, rotated around the vertical axis by 45° relative to *A*. GALC active sites are accessible to solvent via channel 3. (*C*) Image of the crystal used to collect the GALC–4NβDG dataset. The beam diameter (50 μm) is shown in red. Crystals are triangular prisms measuring 50 × 50 × 43 μm with a volume of 5.375×10^{-5} mm^3.

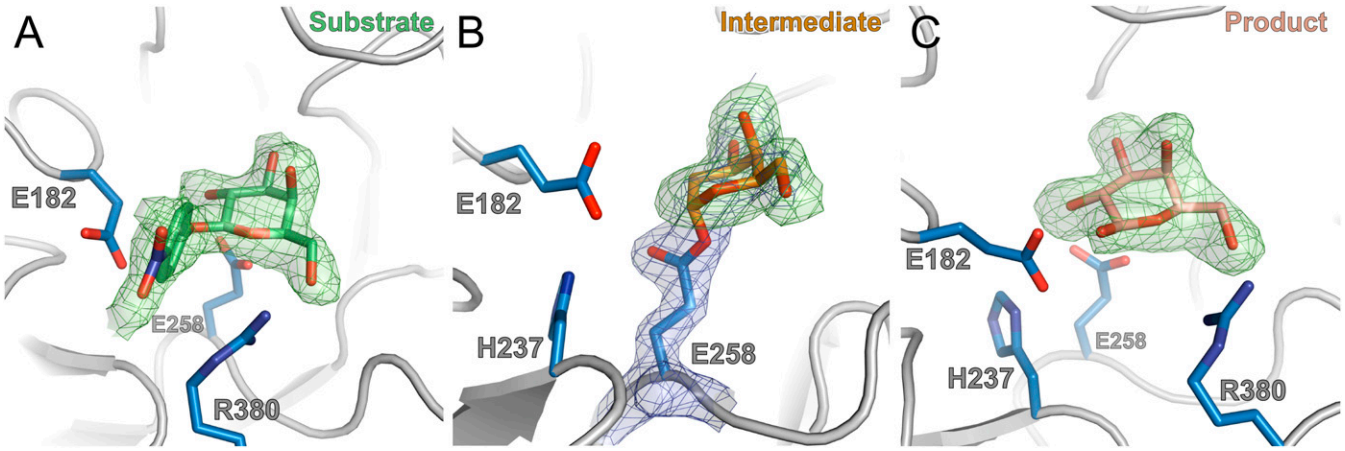


Fig. S3. Electron-density maps before ligand modeling. (*A*) Unbiased difference electron-density maps (F_O-F_C, 3.0σ, green) for substrate 4NβDG at the active site of GALC. (*B*) Electron-density maps (F_O-F_C, 3.0σ, green; 2F_O-F_C, 0.25 $e^-/Å^3$, blue) illustrating the covalent bond formed between D-galactal and the side chain of residue E258. (*C*) Unbiased difference electron-density maps (F_O-F_C, 3.0σ, green) for product β-D-galactose at the active site of GALC.

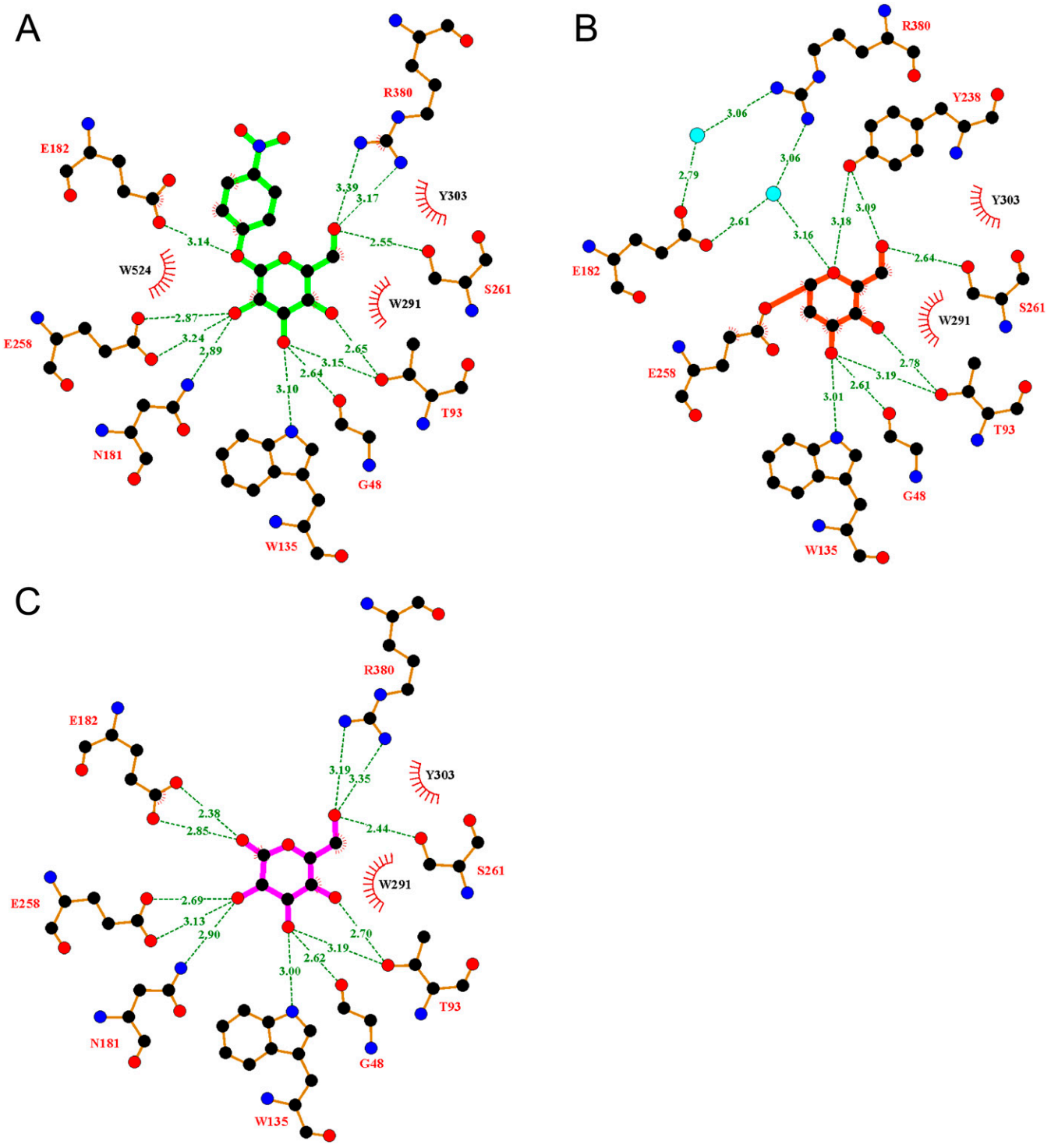


Fig. S4. Schematic representation of GALC amino acid side chain interactions with the 4NβDG substrate (*A*), the D-galactal covalent intermediate (*B*), and the galactose product (*C*). Hydrogen bonds (green dashed lines) and hydrophobic interactions (red fans) are illustrated.

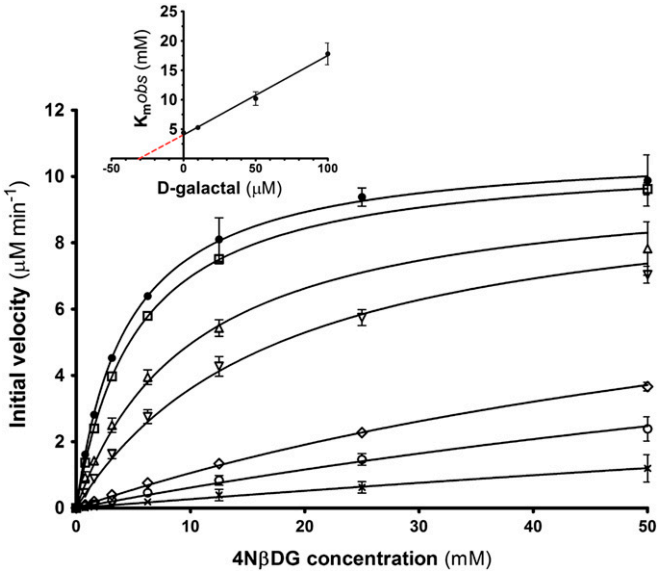


Fig. S5. Competitive-inhibition kinetics of D-galactal. Shown is a plot of initial velocity vs. substrate concentration at D-galactal concentrations of 5,000 μM (×), 1,000 μM (○), 500 μM, (◇), 100 μM, (▽), 50 μM, (△), 10 μM, (□) and no inhibitor (●). (*Inset*) Plot of $K_m obs$ vs. D-galactal concentration showing K_i (32 ± 1.8 μM) as the *x* intercept. SEM error bars are shown.

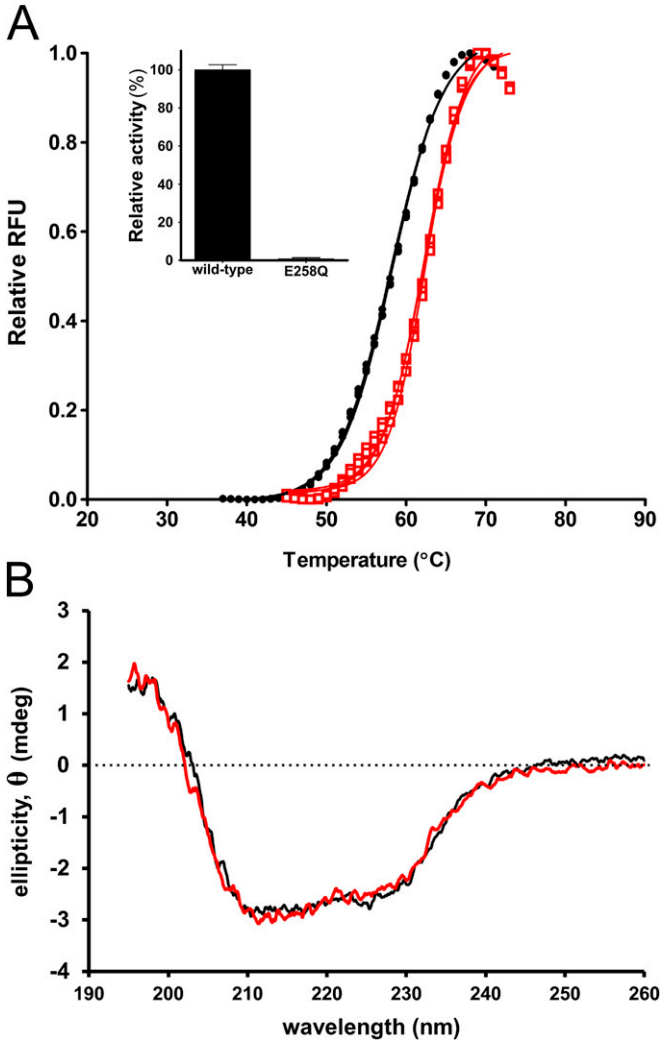


Fig. S6. E258Q protein activity, stability, and folding. (*A*) Fluorescence based thermal shift assay for wild-type (●) and E258Q (□) GALC in sodium acetate buffer (pH 4.6). Triplicate data are shown. Wild-type, $T_m = 58.2 \pm 0.1$ °C; E258Q, $T_m = 62.2 \pm 0.1$ °C. (*Inset*) Relative activity of E258Q mutant compared with wild-type GALC. SEM error bars are shown. (*B*) Far-UV CD spectra of GALC (black) and GALC E258Q (red) in sodium acetate buffer (pH 4.6). The averages of 20 scans are shown.

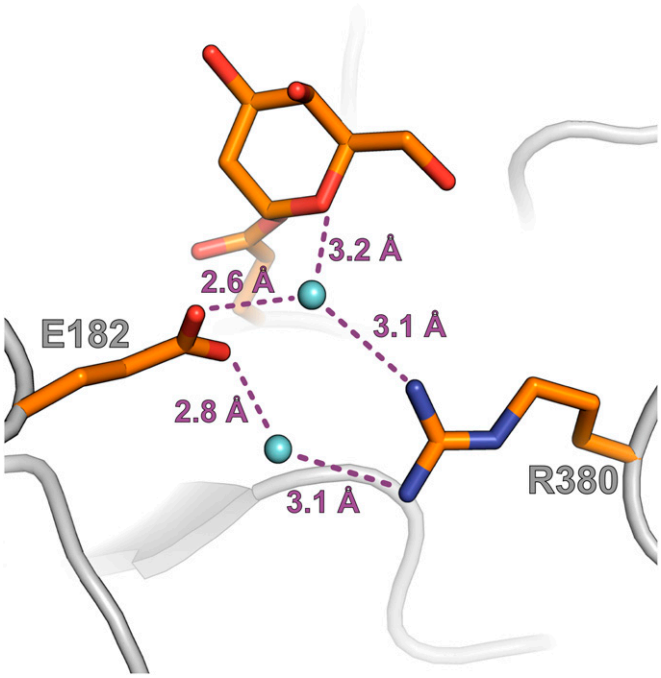
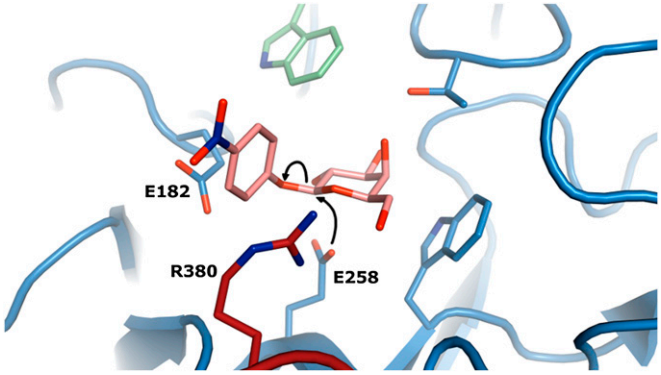


Fig. S7. Hydrogen bonding (pink dashed lines) of two water molecules (light blue spheres) in the active site upon formation of the covalent glycosyl-enzyme intermediate.



Movie S1. Annotated animation of a complete double-displacement reaction cycle. Conformational changes are illustrated by morphs between individual structures, interpolated using rigimol in PyMOL. All movements of atoms depicted are hypothetical; refined electron-density maps ($2F_O$-F_C, blue) appear during the movie to indicate the structures used as "key frames" for illustrating conformational changes.

Movie S1