SUPPLEMENTARY INFORMATION

Deciphering the enzymatic mechanism of sugar ring contraction in UDP-apiose biosynthesis

Simone Savino[a,b]§, Annika J. E. Borg[c]§, Alexander Dennig[b,c]§, Martin Pfeiffer[c], Francesca de Giorgi[a,c], Hansjörg Weber[d], Kshatresh Dutta Dubey[e], Carme Rovira[e,f], Andrea Mattevi[a,b]*, and Bernd Nidetzky[b,c]*

[a] Department of Biology and Biotechnology, University of Pavia, Via Ferrata 1, 27100, Pavia, Italy

[b] Austrian Centre of Industrial Biotechnology, Petersgasse 14, 8010 Graz, Austria

[c] Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, NAWI Graz, Petersgasse 12, 8010 Graz, Austria

[d] Institute of Organic Chemistry, Graz University of Technology, NAWI Graz, Stremayrgasse 9, 8010 Graz, Austria

[e] Department of Inorganic and Organic Chemistry (Organic Chemistry Section) & Institute of Computational and Theoretical Chemistry (IQTCUB), University of Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain

[f] Catalan Institution for Advanced Studies (ICREA), Passeig Lluís Companys 23, 08010 Barcelona.

* Corresponding authors (B.N., bernd.nidetzky@tugraz.at; A.M., andrea.mattevi@unipv.it)

§ Equally contributing first authors

Table of Contents

Supplementary Methods

Chemicals, enzymes and strains

Uridine monophosphate (UMP, 98% purity) was from Carbosynth (Compton, UK). Uridine diphosphate (UDP) and UDP-D-glucuronic acid (UDP-GlcA (**2**); >98% purity) was from Sigma-Aldrich (Vienna, Austria). NAD^+ (>98% purity) was from Roth (Karlsruhe, Germany). Deuterium

oxide (99.96% ^{2}H) was from Euriso-Top (Saint-Aubin Cedex, France). Polyethylene glycols in various lengths (200, 300, 400, 1000, 4000, 6000 and 8000) and ammonium sulphate were from Sigma-Aldrich. Protease inhibitors (leupeptin and pepstatin-A), DNase, PMSF (protease inhibitor), and lysozyme were also from Sigma-Aldrich. All other reagents and chemicals were of highest available purity. For plasmid DNA isolation the GeneJET Plasmid Miniprep Kit (Thermo Scientific; Waltham, MA, USA) was used. DpnI and Q5® High-Fidelity DNA polymerase were from New England Biolabs (Frankfurt am Main, Germany). Oligonucleotide primers were from Sigma-Aldrich. Sparse matrix and systematic screening kits for crystallization of UAXS were from Molecular Dimensions (Newmarket, Suffolk, UK), Qiagen (Hilden, Germany) and Jena Bioscience (Jena, Germany).

Expression and purification of UAXS for biochemical characterization

UAXS (cloned into pET26 vector) was expressed as N-terminally histidine tagged protein in *E. coli* BL21 (DE3) cells.[1] For this, cells were grown in lysogeny broth media (37 °C, 120 rpm; 50 µg/mL kanamycin) until reaching an OD_{600} of 0.6. Then IPTG (250 µM) was added followed by incubation of cells for 24 h (120 rpm, 18 °C). Disruption of cells and purification of UAXS via His-tag affinity chromatography (GE Healthcare 5 mL His trap column; loaded with Cu^{2+} instead of Ni^{2+}) was done as described elsewhere in detail.[1] Fractions containing pure protein (confirmed by SDS-PAGE; see Supplementary Figure 3) were pooled and concentrated to a volume of 200 to 300 µL using a Vivaspin Turbo 30 kDa cut-off centrifugation tube (Sartorius, Göttingen, Germany; 4000 × g, 4 °C). For removal of imidazole, 5 mL storage buffer (50 mM potassium phosphate buffer, pH 7.0; 5% (v/v) glycerol; 1 mM dithiothreitol) were added followed by centrifugation until reaching a volume of 200 to 300 µL. The washing step was repeated three times. For in-situ ^{1}H-NMR analysis of enzymatic reactions, the storage buffer was prepared in D_2O (pD 7.0). For long-term storage, purified enzymes were frozen in liquid N_2 and stored at -20 °C until further use. The protein concentration of purified UAXS was determined by UV spectroscopy (λ = 280 nm) on a DeNovix DS-11+ micro volume spectrophotometer (DeNovix, Wilmington, DE, USA). The molar extinction coefficient of 48360 $M^{-1}cm^{-1}$ and a molecular mass of 44703 Da were used for calculating the UAXS concentration.[1]

Expression and purification of UAXS for crystallization experiments

Expression of UAXS was done as described above using the identical *E. coli* strain and plasmid. Transformed cells were centrifuged (4000 x g for 15 minutes, 4 °C) and stored as pellet at -20°C. For purification of UAXS, cells were re-suspended in buffer A (20 mM Tris-HCl; pH 7.5; 50 mM NaCl) containing protease inhibitors (leupeptin and pepstatin-A; 10 µg/mL each), DNase (1

mg/mL), PMSF (2 mM), and lysozyme (10 µg/mL). Cells were disrupted by sonication and solid particles removed by centrifugation (70000 x g for 45 minutes, 4 °C). Clear supernatants were collected, pooled, and filtered through a 0.45 µM syringe filter before loading on a 5 mL His trap excel column (GE Healthcare). Prior to loading the protein, the column was connected to a single wavelength ÄKTA system (GE Healthcare) and equilibrated with buffer A. The filtered lysate was loaded onto the His trap column while collecting the flow through. Unbound *E. coli* proteins were removed by washing the column with 5 column volumes buffer A. Elution of UAXS was achieved at 50% buffer B (buffer A containing 400 mM imidazole). Dispensed 0.5 mL fractions were collected and samples were analysed by SDS-PAGE to determine purity. A single, abundant band at the expected molecular mass of 45 kDa was found in correspondence of the elution peak. Fractions containing pure UAXS were pooled and concentrated by centrifugation in 30 kDa cut-off Amicon centrifugal filter units (Merck Millipore, Darmstadt, Germany). The concentrated UAXS protein was loaded on a Superdex 200 Increase 10/300 size exclusion column (GE Healthcare). A single sharp peak obtained by size exclusion chromatography indicated a unique oligomeric state of UAXS as dimer (Supplementary Figure 5), making the sample suitable for crystallization experiments.

Crystallization and X-ray diffraction analysis of UAXS

Crystallization plates (Swissci, Molecular Dimensions for robot dispensing; Cryschem (Hampton Research, Aliso Viejo, CA, USA) for handmade conditions) were prepared containing 10 to 100 mg/mL purified UAXS protein. The experiments were reproduced multiple times without ligands or by adding 1 to 2 mM NAD^+ and/or 1 to 2 mM UDP. Systematic handmade screens of varying concentrations and varying types of polyethylene glycols and $(NH_4)_2SO_4$ were dispensed in sitting and hanging drop configurations. Further, crystallization screening kits (sparse matrix and systematic) were tested, dispensing them manually in hanging drop or with a Douglas Instruments Oryx 8 robot in 96-wells plates. The only positive hit yielding diffracting protein crystals was found in well D10 of Qiagen Classic Suite, composed of 0.8 M sodium phosphate, 0.8 M potassium phosphate, and 0.1 M sodium-HEPES (all pH 7.5). Initial crystals grew when supplementing 1 mM NAD^+ and 1 mM UDP to a 25 mg/mL protein solution. The obtained transparent crystals appeared as regularly shaped hexagonal based prisms (Supplementary Figure 6). Crystals were found to be reproducible in an identical handmade reservoir condition. Under those conditions, crystals could be grown also in absence of NAD^+ and UDP. The crystal structure of C100A variant was obtained under identical conditions in co-crystallization experiments with UDP-GlcA (**2**) (2 mM) and NADH (2 mM). The crystals were harvested and cryo-cooled in their respective crystallization solution added with 20% v/v glycerol, followed by X-ray diffraction analysis at the SLS facility (Villigen,

Switzerland) or ESRF (Grenoble, France). Diffracting crystals were displaying low anisotropy and, after XDS indexing[2], space group was found to be $P3_221$ (number 154 of International Tables of Crystallography) with defined cell parameters (a = 146.03 Å; b = 146.03 Å; c = 132.85 Å; α = 90°; β = 90°; γ = 120°). Data were scaled using the Aimless program from CCP4 suite[3] and molecular replacement was performed with Phaser-MR[4] using a 29% sequence-identity model (PDB entry: 3SLG) trimmed with Chainsaw.[5] A Phenix Autobuild run[6] was performed to trace the main chain followed by a Fem run[7] to improve map density. At this point manual tracing of sidechains and poorly traced surface loops was performed using Coot[8] alternated to *REFMAC*5 cycles[9]. Non-crystallographic symmetry (NCS) was applied, taking advantage of the dimeric arrangement in the unit cell; translation-libration-screw-rotation (TLS) was also used. Molprobity[10] was used for model validation.

Molecular dynamics (MD) simulations and Quantum Mechanical/Molecular Mechanical (QM/MM) calculations

System Setup: Molecular Dynamics (MD) simulations were performed with the Amber 18 GPU code[11], using the Amber ff14SB force field for the protein, TIP3P model[12] for water and ions and GLYCAM06[13] for the UDP-GlcA substrate. Established force field parameters from the Amber database were used for the NAD^+ cofactor (http://research.bmh.manchester.ac.uk/bryce/amber/). Of the two UAXS-chains present in the experimental structures, chain A had no gaps or unresolved loops and was therefore used to generate a dimer by aligning two copies of chain A using UCSF Chimera.[14] The resulting dimer was protonated using Leap module of Amber18 and used as input for PROPKA (version 3.1) to ensure that all ionizable amino acids were correctly protonated and all His, Asn and Gln residues had their most probable orientations, respectively. The conserved Tyr185, which in other enzymes of this superfamily functions as catalytic acid-base[15], was simulated in the deprotonated or protonated state. MODELLER[16] was used to generate atom coordinates for missing side-chains of the protein. Since the nicotinamide of NAD^+ in the original structure of the C100A variant was only partially resolved, the missing coordinates were modelled by aligning each chain with the experimental structure of human UXS (PDB: 2B69)[15], which contains the complete structure of NAD^+. The resulting protein dimer including two copies of NAD^+ and UDP-GlcA (**2**) (Michaelis complex; Figure 2a,c) was placed into a truncated octahedral solvation box of water molecules extending to 10 Å from the protein surface, and then the complex was neutralized with Na^+ ions.

MD simulations: After proper parameterizations and setup, the resulting structure was minimized (5000 steps of steepest descent and 10000 steps of conjugate gradient) to remove poor contacts and

allow the system to relax. Afterwards, the system was gently annealed from 10 to 300 K under the isothermal (NVT) ensemble for 50 ps with a weak restraint of 5 kcal mol^{-1} $Å^{-2}$ applied on protein and ligands, while keeping water and ions relaxed. Subsequently, one ns of density equilibration in the isothermal-isobaric (NPT) ensemble at a target temperature of 300 K and a target pressure of 1.0 atm was performed. The Langevin thermostat[17] and the Barendsen barostat[18] with collision frequency of 2 ps, a pressure relaxation time of 1 ps with a weak restraint of 1 kcal mol^{-1} $Å^{-2}$ were used with a time step of 2 fs. It should be noted that the 1 ns of density equilibration is not identical with conformational equilibration, but rather a weakly restrained MD simulation, in which the system is slowly relaxed to achieve a uniform density after heating dynamics under periodic boundary conditions. Thereafter, we removed all of the restraints applied during heating and density dynamics and further equilibrated the system. Two replicas were considered. In each one, we equilibrated the system for ~3 ns to get a well-settled pressure and temperature for conformational and chemical analysis. The hydrogen bonds were constrained using SHAKE[19] while Particle Mesh Ewald[20] was used to treat the long rang electrostatic interactions in all simulations, with a cut off distance of 12 Å for the van der Waals interactions. Long-time scale MD production simulations were performed for 0.6 μs (300 ns for each replica). Analysis of the sugar ring conformations and the RMSD evolution along the simulation is provided in Supplementary Figure 15 (see below). All analyses of the trajectories were done with Cpptraj Module of the Amber MD package.[11]

QM/MM calculations: Two representative snapshots of active-site conformation were chosen from the MD trajectories and QM/MM (quantum mechanics/molecular mechanics) optimization was performed on those.[21] For QM/MM calculations, a snapshot from the most persistent conformation from each replica was taken. To find the most frequent conformation in each replica, we first performed a population analysis and subsequently, by visual inspection of the most populated trajectories, a conformation was chosen with suitable substrate-enzyme interactions for catalysis. NAD^+, GlcA, the phosphate group of UDP and protein residues (Cys100, Tyr105, Thr139, Glu141, Tyr185) were kept in the QM region. All other protein residues and water molecules were kept in the MM region. All atoms within 8 Å of GlcA were treated as active residues and their effect due to polarization on the QM atoms were considered in the QM/MM calculations. All QM/MM calculations were performed using ChemShell,[22,23] combining Turbomole[24] for the QM part and DL_POLY[25] for the MM part. The electronic embedding scheme was used to account for the polarizing effect of the enzyme environment on the QM region. The link atom model for the frontier covalent bonds on the QM and MM interfaces was applied to treat the QM/MM boundary. In QM/MM geometry optimizations, the QM region was treated by the hybrid B3LYP functional with

def2-SVP as basis set. The same force-fields as in the MD simulations previously described were employed throughout this study to describe the atoms of the MM region.

Detection and quantification of NADH in UAXS preparations

NADH fluorescence measurements were performed in a Hellma fluorescence cuvette (80 µL sample volume; Suprasil quartz, 3×3 mm light path) placed in a Hitachi F-4500 (Tokyo, Japan) fluorescence spectrophotometer (excitation slit: 5 nm, emission slit: 5 nm). NADH was quantified via its specific fluorescence using excitation wavelength (λ_{Ex}) of 340 nm and emission wavelength (λ_{Em}) of 460 nm, respectively. Defined concentrations of NADH (2.5 to 25 µM) were prepared in double distilled water and directly used for measurements. NADH bound in purified wild-type UAXS was determined by denaturation of the enzyme (10 µL, 19.6 and 23.5 µM) in methanol (90 µL) at 30 °C for 3 h. Solids were removed by centrifugation (16000 × g, 60 min, 4 °C). The supernatant was collected, and the fluorescence emission of the free cofactor measured as described above. The occupancy of NADH in UAXS was calculated in % based on the concentration of released NADH (in µM) divided by the total protein concentration (in µM) applied to the experiment (Supplementary Figure 12). To determine the possible exchange of NADH by NAD^+ in the active site of wild-type UAXS, 1.3 µM of purified enzyme was incubated in presence of 100 µM NAD^+ in potassium phosphate buffer (50 mM, pH 7). The decrease of the fluorescence emission signal (i.e. cofactor exchange from NADH to NAD^+; protein bound NADH has a higher quantum yield than free/unbound NADH[26]) was followed as described above (Supplementary Figure 12). In order to reduce the enzyme-bound NAD^+, UAXS (20 mg/mL, 412 µM) was incubated with $NaBH_4$ (300 mM final concentration). $NaBH_4$ was added in five steps (60 mM each) every 1 h and the final mixture was incubated at 37 °C for 16 h. The enzyme solution was transferred into a centrifugation filter tube (Vivaspin 500, 10 kDa cut-off) and washed five times (centrifugation at 16000 × g, 4 °C) thoroughly with buffer (25 mM Tris-HCl pH 7.0; 1 mM dithiothreitol) to remove residual $NaBH_4$. A sample of 12.5 µl was taken and the enzyme denatured by incubating it in methanol (75% (v/v) final concentration), followed by centrifugation (16000 x g, 4 °C, 1 h) to remove the solid residues. The supernatant was analyzed on a HPLC-UV system (see details for HPLC analysis below). The amount of released NADH was calculated based on a calibration curve for free NADH (Supplementary Figure 42). The efficiency of the reduction was confirmed by comparing the NADH content after $NaBH_4$ reduction to the NADH content of the same amount of denatured UAXS before reduction. Representative HPLC chromatograms before and after reduction are shown in Supplementary Figure 43.

Site-directed mutagenesis

Site-directed mutagenesis was performed using a modified version of the QuikChange protocol and Q5 DNA polymerase for DNA amplification. PCRs were performed in 50 µL reaction volume according to supplier's recommendation using 40 ng of plasmid DNA as template and 0.2 µM forward primer or reverse primer, respectively. The DNA oligonucleotide primers used are shown below. The underlined nucleotides highlight the mutations introduced by PCR.

Name	Mutation	DNA primer sequence (5'-3')
C100A_Fw	C100A	CAATCTGGCGGCCATTGCTACGCCTGCTGATTATAAC
C100A_Rv	C100A	GTTATAATCAGCAGGCGTAGCAATGGCCGCCAGATTG
C100S_Fw	C100S	CAATCTGGCTGCAATTAGCACGCCTGCTGATTATAAC
C100S_Rv	C100S	GTTATAATCAGCAGGCGTGCTAATTGCAGCCAGATTG
Y105F_Fw	Y105F	CATTTGTACGCCTGCTGATTTTAACACCCGTCCGCTGGAC
Y105F_Rv	Y105F	GTCCAGCGGACGGGTGTTAAAATCAGCAGGCGTACAAATG
Y105A_Fw	Y105A	CATTTGTACGCCTGCTGATGCTAACACCCGTCCGCTGGAC
Y105A_Rv	Y105A	GTCCAGCGGACGGGTGTTAGCATCAGCAGGCGTACAAATG
T139V_Fw	T139V	GATCCATTTTTCCGTGTGTGAAGTGTACGGC
T139V_Rv	T139V	GCCGTACACTTCACACACGGAAAAATGGATC
C140S_Fw	C140S	CATTTTTCCACCAGCGAAGTGTACGGC
C140S_Rv	C140S	GCCGTACACTTCGCTGGTGGAAAAATG
C140A_Fw	C140A	CATTTTTCCACCGCTGAAGTGTACGGC
C140A_Rv	C140A	GCCGTACACTTCAGCGGTGGAAAAATG
Y185F_Fw	Y185F	CAGCGTTGGTCTTTTGCATGTGCTAAAC
Y185F_Rv	Y185F	GTTTAGCACATGCAAAAGACCAACGCTG
E141A_Fw	E141A	CATTTTTCCACCTGTGCAGTGTACGGCAAAACGATTGGTTC
E141A_Rv	E141A	CACTGCACAGGTGGAAAAATGGATCAGGCGTTTGTTGTTTTCACTGC

First, three cycles of linear PCR amplification were done with separate forward and reverse primers (initial denaturation: 1 min/94°C; 3 cycles amplification: 30 sec/94°C for denaturation, 30 sec/55°C for primer annealing and 15 min/72°C extension; final extension: 72°C/10 min). Both reaction solutions were mixed (combining forward/reverse primers) and the PCR program was restarted for 15 cycles of exponential DNA amplification (initial denaturation: 1 min/94°C; 3 cycles amplification: 30 sec/94°C for denaturation, 30 sec/55°C for primer annealing and 6 min/72°C extension; final extension: 72°C/10 min). PCR products were separated by agarose gel electrophoresis and visualized by DNA staining. Residual template DNA was removed by addition of 2 U DpnI and following the recommended procedure by the supplier. Without further purification, PCR products were transformed via heat-shock into chemical competent *E. coli* Top 10 cells (New England Biolabs, Frankfurt am Main, Germany). Plasmid DNA was extracted and

sequenced with T7prom/T7term primers provided by LGC Genomics (Berlin, Germany) to verify mutations. Correct constructs were transformed into *E. coli* BL21 cells followed by expression of UAXS as described above.

Product analysis by HPLC

Reaction samples were analyzed by reversed phase HPLC-UV (λ = 262 nm) on a Shimadzu Prominence HPLC system (Shimadzu, Korneuburg, Austria). Separation was achieved on a Kinetex C18 analytical HPLC column (5 µm; 4.6 × 50 mm; Phenomenex, Aschaffenburg, Germany) using isocratic elution (35 °C; flow rate of 2 mL/min; 5 µL injection volume; mobile phase: 87.5% 40 mM tetra-n-butylammonium bromide in 20 mM potassium phosphate buffer (pH 5.9) and 12.5% acetonitrile). Separation of UDP-xylose (**3**) and UDP-apiose (**1a**) is not possible under the described conditions.[1,27] Therefore, the formation of UDP-apiose (**1a**) can be referred to the release of UMP (formed during decomposition of UDP-apiose).[1,27] Authentic standards of UMP and UDP-GlcA (**2**) were used as analytical references (Supplementary Figures 17, 18). UMP is a known decomposition product derived from self-cyclization of instable UDP-apiose.[28] An optimized protocol for HPLC-UV detection of UDP-apiose (**1a)** is described below in the section describing the conversion of UDP-4-keto-xylose (**5**) with UAXS and human UXS.

Kinetic characterization of enzymatic conversion of UDP-GlcA

Enzymatic rates (k_{cat}) were measured for wild-type UAXS and all generated variants. We used purified human UXS to generate UDP-xylose (**3**) as analytical reference (Supplementary Figure 19).[1,29] The reaction mixtures contained (if not stated otherwise) 2 mM UDP-GlcA (**2**), 100 µM NAD^+ and 14 mg/mL enzyme (except: wild-type UAXS and C100S/C140A variant, 2 mg/mL; C100S variant, 7 mg/mL; Y105A variant, 3.5 mg/mL) in a final reaction volume of 1 mL potassium phosphate buffer (pH 7.0, 50 mM). Reactions with the T139V variant contained 0.5 mM UDP-GlcA (**2**), 1.0 mM NAD^+ and 3.3 mg/mL enzyme. In all cases, UDP-GlcA (**2**) was added to initiate the reaction followed by incubation at 30 °C (no agitation). In order to stop the reaction at defined time points, 50 µL reaction sample were taken and mixed with equal volume of a 50% acetonitrile. Samples were incubated for at least one hour on ice, to allow denaturation of the protein. Insoluble compounds were removed by centrifugation (16000 × g, 15 min, 4 °C). Depletion of UDP-GlcA (**2**) over time was used to calculate enzyme activity (k_{cat}) (see Supplementary Figures 23-32). Initial rates were obtained from the linear part of the time course, as shown in Supplementary Figures 23-32. The k_{cat} was calculated from the initial rate, recorded in nmol/(min mg protein), with the appropriate molecular mass of the functional enzyme monomer (wild-type UAXS: 44703 g/mol).

The activity of wild-type UAXS for conversion of UDP-GlcA (**2**) was further recorded in D_2O (Supplementary Figure 36). Therefore, purified wild-type UAXS was re-buffered against phosphate buffer prepared in D_2O (50 mM potassium phosphate; pD 7.0) containing 1 mM DTT, 1% glycerol (w/v). The enzyme (2 mg/mL, 44.7 µM), NAD^+ (100 µM) and UDP-GlcA (**2**) (1 mM) were mixed together in the corresponding buffers (prepared in H_2O or D_2O; pH/pD 7.0) followed by incubation at 30 °C for 2 h. Samples were taken at defined time points and the reaction was quenched by addition of methanol (50%). After centrifugation the supernatant was collected and analyzed by HPLC.

^{1}H-NMR analysis of enzymatic reactions

Enzyme selectivity was determined by ^{1}H-NMR analysis (Supplementary Figures 33-35). For this, all reagents including UAXS were prepared using exclusively D_2O as solvent. The reaction mixtures contained 2 mM UDP-GlcA (**2**), 100 µM NAD^+, 2 mg/mL purified UAXS in a final volume of 700 µL of potassium phosphate buffer (50 mM; pD 7.0). UDP-GlcA (**2**) conversions were carried out for 24 h at 30°C. Samples were centrifuged (15 min, 4 °C, 16000 g) and 700 µL filled into NMR tubes followed by measurement in a Varian INOVA 500-MHz NMR spectrometer (Agilent Technologies, Santa Clara, California, USA). The VNMRJ 2.2D software was used for all measurements. ^{1}H-NMR spectra (499.98 MHz) were recorded on a 5 mm indirect detection PFG-probe with pre-saturation of the water signal by a shaped pulse. The following standard pre-saturation sequence was performed: 2 s relaxation delay; 90° proton pulse; 2.048 s acquisition time; 8 kHz spectral width; number of points 32 k. Signals from anomeric C1 of the sugar were used for integration and quantification of products (percentage of each proton signal relative to the sum of all C1 proton signals in a sample) in order to determine selectivity of wild-type UAXS and mutants. ACD/NMR Processor Academic (Edition 12.0, Advanced Chemistry Development Inc., Toronto, CA) was used for evaluation of spectra.

Synthesis and isolation of UDP-xylose (3) from deuterium wash-in experiments

For preparative isolation of UDP-xylose (**3**), 2 mM UDP-GlcA (**2**) were converted in total volume of 9 ml buffer prepared in D_2O (50 mM potassium phosphate, pD 7.0). The reactions were started by adding 100 µM NAD^+ and enzyme (2 mg/ml for wild-type UAXS or 5 mg/ml in case of variants C100A and C100S). All reactions were carried out overnight at 30°C (no agitation) until no further substrate depletion was detectable by HPLC. Residual enzyme was removed by centrifugal filtration (4000 x g, 4 °C) using a 10 kDa cut-off centrifugation tube. For purification, the flow-through was loaded at 2 mL/min flowrate via a 10 ml sample loop onto an ÄKTA FPLC system (GE Healthcare) connected to a 125 ml Toyopearl SuperQ-650M anion exchange column (GE

Healthcare) equilibrated with 20 mM sodium acetate solution (pH 4.3). Flow-rates between 3 and 5 mL/min were used during the purification process depending on back pressure of the column. After washing the column with 250 mL NaOAc (20 mM), compounds were eluted with a step-wise gradient of 1 M sodium acetate buffer (pH 4.3). The following elution steps were used during compound purification: 150 mL of 200 mM NaOAc, 250 mL of 350 mM NaOAc, 150 mL of 500 mM NaOAc, 330 mL of 1 M NaOAc. After eluting all compounds, the column was re-equilibrated in 20 mM NaOAc. Fractions containing the desired products were identified by UV absorption (λ = 254 nm). A representative chromatogram from the ion-exchange purification is shown in Supplementary Figure 37. The fractions containing UDP-xylose (**3**) (confirmed by HPLC) were combined and concentrated under reduced pressure (40 °C) to a final volume of 5 to 10 mL using a Laborota 4000 rotary evaporator (Heidolph, Schwabach, Germany). At next NaOAc was removed by size-exclusion chromatography. Therefore, the concentrated sample was loaded via a 5 mL sample loop onto an ÄKTA FPLC system connected to a Superdex G-10 column (GE Healthcare). Filtered and deionized water was used as mobile phase and elution of UDP-xylose (**3**) was detected by UV absorption (λ = 254 nm). A representative ÄKTA chromatogram from the desalting step is shown in Supplementary Figure 38. The fractions containing UDP-xylose (**3**) were combined and concentrated under reduced pressure (40 °C) to a final volume of approximately 20 mL. Desalted UDP-xylose (**3**) was frozen in liquid N_2 and residual H_2O was removed by lyophilization (Christ Alpha 1-4 lyophilizer, B. Braun Biotech International, Melsungen, Germany), yielding the nucleotide sugar (>95% pure) as white powder. The product was re-dissolved in D_2O (700 µL) prior to NMR analysis.

Enzymatic synthesis and isolation of UDP-4-keto-xylose

UDP-4-keto-xylose (**5**) was synthesized by converting UDP-GlcA (**2**) with the C-terminal subunit of ArnA from *E. coli*.[30] The enzyme was obtained as synthetic gene and expressed in *E. coli* BL21 (DE3). After affinity chromatography utilizing a C-terminal Strep-tag, the protein was obtained in >90% purity (Supplementary Figure 39). The obtained protein was re-buffered (25 mM Tris-HCl pH 8.0 and 1 mM DTT) and stored at -20 °C. The reaction for synthesis of UDP-4-keto-xylose (**5**) contained 2 mM UDP-GlcA (**2**), 2 mM NAD^+ and 5 mg/mL purified ArnA. The sample was incubated for 5 min at 30 °C (no agitation; 70% conversion, >99% selectivity for **5**). Longer reaction times resulted in significant formation of the unwanted side product UDP-xylose (**3**).[31] The conversion of UDP-GlcA (**2**) was constantly followed by HPLC-UV (see Supplementary Figure 40). When reaching the highest amount of UDP-4-keto-xylose (**5**) ArnA was removed from the mixture by centrifugal filtration (Vivaspin Turbo centrifugal filter tube, 30 kDa cut-off). The desired product was separated from the other components by anion exchange and size-exclusion

chromatography using ÄKTA FPLC system (see above purification of UDP-Xyl).[15] The purity and structure of UDP-4-keto-xylose (**5**) was confirmed by ^{1}H-NMR (Supplementary Figure 41).

Conversion of UDP-4-keto-xylose (5) with UAXS and UXS

The UDP-4-keto-xylose (**5**) (1 mM) and NADH (10 mM) were dissolved in KPi buffer (50 mM, pH 7.0). The reaction was started by addition of the respective SDR catalyst (20 mg/mL wild-type UAXS (reduced or non-reduced), 15 mg/mL C100S variant or 5 mg/mL human UXS) followed by incubation at 30 °C (no agitation). The reaction was quenched with methanol (75% (v/v) final concentration) at desired time points and the precipitated enzyme was removed by centrifugation (16000 x g, 4 °C, 30 min) prior to HPLC analysis. The products were separated on a Kinetex C18 column (5 µm, 100 Å, 150 x 4.6 mm) using 3% methanol and 97% tetra-n-butylammonium bromide (TBAB) buffer (40 mM, 20 mM KPi, pH 5.9) as mobile phase. UDP-sugars and NAD(H) were detected via UV absorbance at 262 nm. Examples of representative HPLC chromatograms for wild-type UAXS, C100S variant and human UXS are shown in Supplementary Figure 45. Wild-type UAXS showed similar activity and behavior for conversion of UDP-4-keto-xylose (**5**) using either reduced (0.06 mU/mg) or non-reduced (0.055 mU/mg) enzyme. For deuterium wash-in experiments, UDP-4-keto-xylose (**5**) (2 mM) and NADH (10 mM) were dissolved in KPi buffer (50 mM, pD 7.0). All solutions and UAXS were prepared using D_2O as bulk solvent as stated above. The reaction was started by addition of UAXS (11 mg/mL in case of wild-type UAXS; 3.8 mg/mL in case of C100A variant). The reactions were run for 6 h followed by removal of enzymes centrifugal filtration (Vivaspin 500, 10 kDa cut-off). Clear supernatants were collected, frozen in liquid N_2 and freeze-dried overnight. Solid residues were dissolved in D_2O (600 µL) and used for NMR (HSQC) analysis.

Supporting Tables and Figures

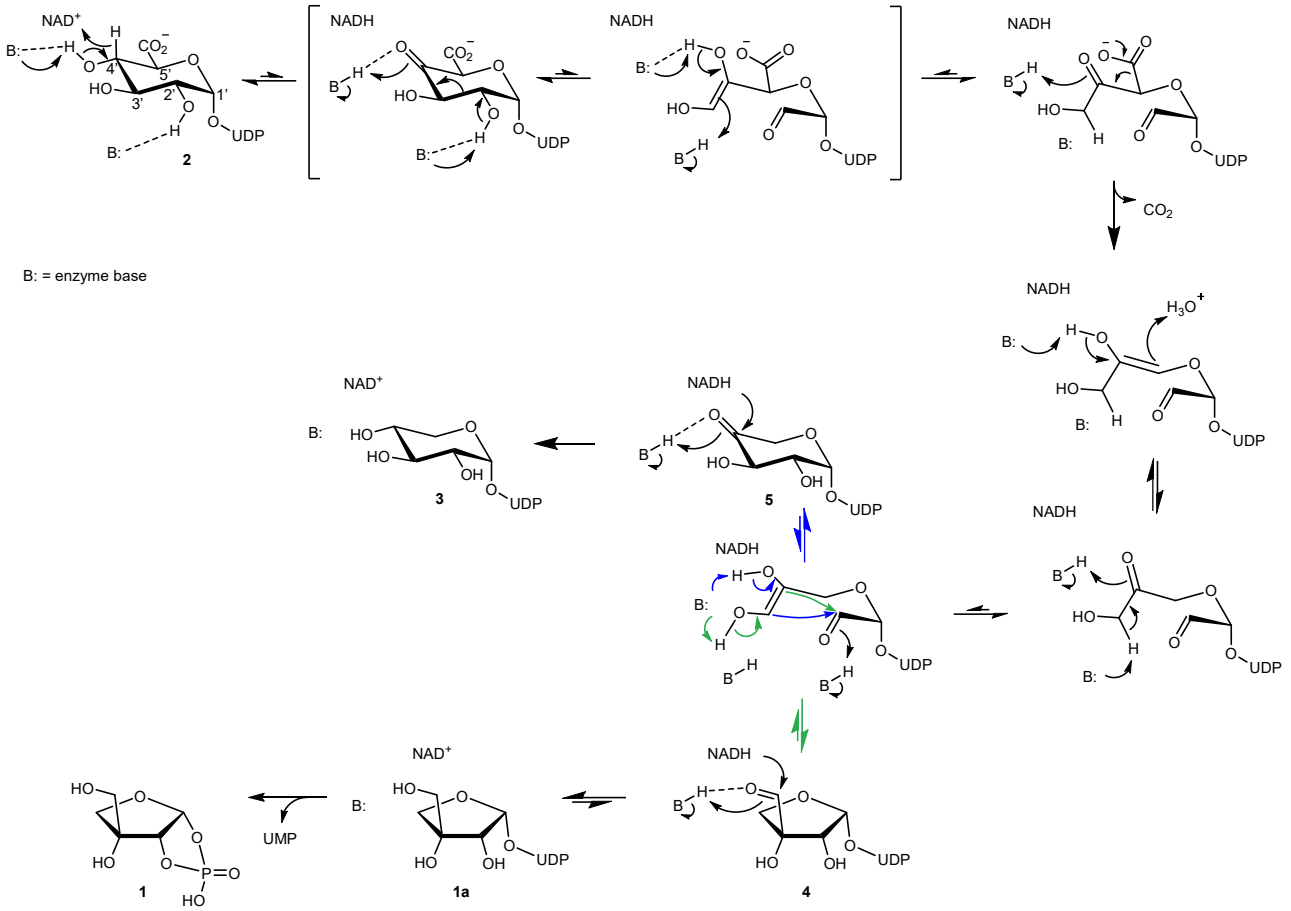
Supplementary Table 1: Crystallographic table for structures of wild-type UAXS in complex with UDP and NAD^+ and the C100A variant in complex with UDP-GlcA (**2**) and NADH, respectively.[a]

	Wild-type, PDB: 6H0N	C100A, PDB: 6H0P
Resolution range	49.1 - 3.0 (3.2 - 3.0)	48.4 - 3.5 (3.8 - 3.5)
Space group	$P3_221$	$P3_221$
Unit cell (Å), (°)	146.0 146.0 132.8 90 90 120	144.6 144.6 130.5 90 90 120
Total reflections	143316 (20101)	173716 (38422)
Unique reflections	32264 (4488)	20833 (4867)
Multiplicity	4.4 (4.5)	8.3 (7.9)
Completeness (%)	98.9 (95.4)	99.8 (99.2)
Mean I/sigma (I)	9.2 (1.0)	13.1 (1.9)
R-merge (%)	0.146 (1.291)	0.151 (1.138)
CC1/2[b]	0.992 (0.378)	0.998 (0.844)
R-work (%)	0.207 (0.369)	0.200 (0.360)
R-free	0.253 (0.377)	0.243 (0.317)
Number of non-hydrogen atoms	6009	5870
Protein	5848	5723
Ligands	138	147
Waters	23	-
RMS (bonds) (Å)	0.013	0.013
RMS (angles) (°)	1.82	1.91

Ramachandran favoured (%)	94.44	91.85
Ramachandran allowed (%)	4.75	7.34
Ramachandran outliers (%)	0.81	0.82
Average B-factor	77.85	125.51
Macromolecules	80.35	131.0
NAD^+/NADH	74.30	125.8
UDP/ UDP-GlcA	70.00	124.1

[a] Statistics for the highest-resolution shell are shown in parentheses

[b] A cut-off criterion for resolution limits was applied on the basis of the mean intensity correlation coefficient of half-subsets of each dataset

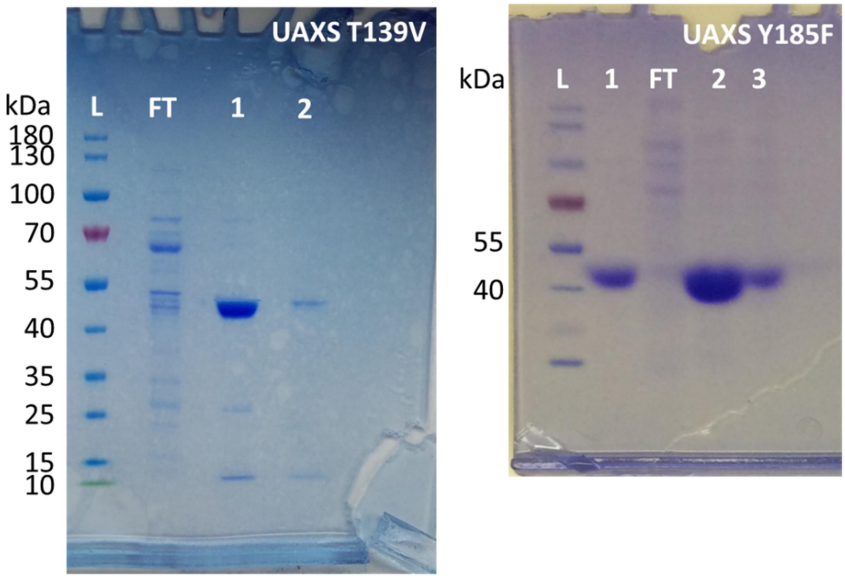


Supplementary Figure 1. Proposed mechanism of UAXS based on evidence from earlier studies. The steps of substrate oxidation, aldol cleavage for ring opening, and decarboxylation are shown in full detail here (see Figure 1 of main text). Kinetic and mechanistic studies of UAXS[1] suggest that substrate oxidation and aldol cleavage are strongly coupled steps of the reaction but it is not known how UAXS promotes the multi-step catalytic conversion of UDP-GlcA (**2**). The reaction of UXS (UDP-xylose synthase) does not involve ring opening. Like the UAXS reaction, it proceeds via an initial oxidation of UDP-GlcA (**2**) at the C4'. Decarboxylation of the incipient UDP-4-keto-GlcA leads to UDP-4-keto-xylose (**5**) which is reduced to UDP-xylose (**3**). **1** = α-D-apiofuranosyl-1,2-cyclic phosphate; **1a** = UDP-apiose; **2** = UDP-GlcA; **3** = UDP-xylose; **5** = UDP-4-keto-xylose.

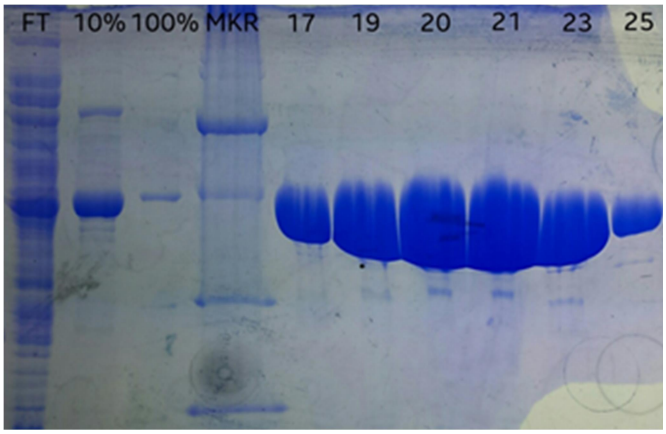
[Uniprot accession number of UAXS: Q9ZUY6]

```
>sp|Q9ZUY6|AXS1_ARATH  UDP-D-apiose/UDP-D-xylose  synthase  1  OS=Arabidopsis
thaliana GN=AXS1 PE=1 SV=1
MANGANRVDLDGKPIQPLTICMIGAGGFIGSHLCEKLLTETPHKVLALDVYNDKIKHLLEPDTVEWSGRIQFHRINIKHD
SRLEGLVKMADLIINLAAICTPADYNTRPLDTIYSNFIDALPVVKYCSENNKRLIHFSTCEVYGKTIGSFLPKDHPLRDD
PAFYVLKEDISPCIFGSIEKQRWSYACAKQLIERLVYAEGAENGLEFTIVRPFNWIGPRMDFIPGIDGPSEGVPRVLACF
SNNLLRREPLKLVDGGESQRTFVYINDAIEAVLLMIENPERANGHIFNVGNPNNEVTVRQLAEMMTEVYAKVSGEGAIES
PTVDVSSKEFYGEGYDDSDKRIPDMTIINRQLGWNPKTSLWDLLESTLTYQHRTYAEAVKKATSKPVAS

sp|Q9M0B6|GAE1_ARATH      QVR-----------TSAQIHRSGGISVLVTGATGFVGSHVSLALRKRGDGV-VGLDNFNN  122
sp|Q8NBZ7|UXS1_HUMAN      KIRDLEKSFTQKYPPVKFLSEKDRKRILITGGAGFVGSHLTDKLMMDGHEVTVVDNFFTG  124
sp|Q9ZUY6|AXS1_ARATH      ----MANG--ANR-VDLDGKPIQPLTICMIGAGGFIGSHLCEKLLTETPHKVLALDVYN-  52
sp|P77398|ARNA_ECOLI      TLG-LVQG--SRLNSQPACTARRRTRVLILGVNGFIGNHLTERLLREDHYEVYGLDIGS-  350
                                              : : *  **:*.*:   *          :  .

sp|Q9M0B6|GAE1_ARATH      YYDPSLKRARRSLLSSRGIFVVEGDLNDAKLLAKLFD-----VVAFTHVMHLAAQAGVRY  177
sp|Q8NBZ7|UXS1_HUMAN      RK-----RNVEHWIGHENFELI-----------NHDVVEPLYIEVDQIYHLASPASPPN  167
sp|Q9ZUY6|AXS1_ARATH      -------DKIKHLLEPDTVEW-SGRIQFHRINIKHDSRLEGLVKMADLIINLAAICTPAD  104
sp|P77398|ARNA_ECOLI      -------DAISRFLNHPHFHFVEGDISI------HSEWIEYHVKKCDVVLPLVAIATPIE  397
                                     :    .               .          :  *.: .

sp|Q9M0B6|GAE1_ARATH      ALENPQSYVHSNIAGLVNLLEICKAANPQPAIVWASSSSVYGLNEKVPFSESD-------  230
sp|Q8NBZ7|UXS1_HUMAN      YMYNPIKTLKTNTIGTLNMLGLAKRVG--ARLLLASTSEVYGDPEVH------PQSED--  217
sp|Q9ZUY6|AXS1_ARATH      YNTRPLDTIYSNFIDALPVVKYCSENN--KRLIHFSTCEVYGKTIGSFLPKDHPLRDDPA  162
sp|P77398|ARNA_ECOLI      YTRNPLRVFELDFEENLRIIRYCVKYR--KRIIFPSTSEVYGMCSDKYFDE---------  446
                             .*   .  :    : ::  .        ::  *:..***

sp|Q9M0B6|GAE1_ARATH      ----------------RTDQPASLYAATKKAGEEITHTYNHIYGLAITGLRFFTVYGPWG  274
sp|Q8NBZ7|UXS1_HUMAN      -----------YWGHVNPIGPRACYDEGKRVAETMCYAYMKQEGVEVRVARIFNTFGPRM  266
sp|Q9ZUY6|AXS1_ARATH      FYVLKEDISPCIFGSI--EKQRWSYACAKQLIERLVYAEGAENGLEFTIVRPFNWIGPRM  220
sp|P77398|ARNA_ECOLI      ------DHSNLIVGPV--NKPRWIYSVSKQLLDRVIWAYGEKEGLQFTLFRPFNWMGPRL  498
                                          *   *:  : :  :     *: .   * *.  **
```

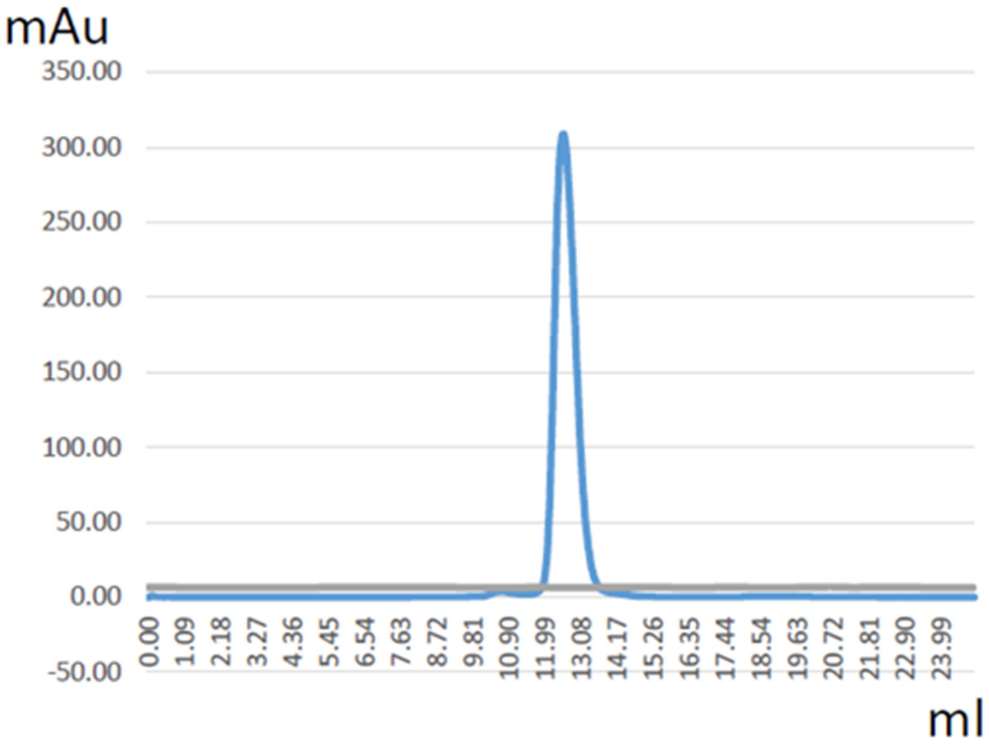
Supplementary Figure 2. Amino acid sequence alignment of UDP-GlcA (**2**) converting enzymes. (Top) The complete sequences of UAXS1 from *Arabidopsis thaliana* (Uniprot: Q9M0B6) used for this study. (Bottom) Sequence alignment shows partial sequences of epimerase GAE1 from *Arabidopsis thaliana* (Uniprot: Q9M0B6), UXS1 from *Homo sapiens* (Uniprot: Q8NBZ7), UAXS1 from *Arabidopsis thaliana* (Uniprot: Q9M0B6; this study) and ArnA from *E. coli* (Uniprot: P77398). Yellow highlighted amino acids correspond to active site residues targeted in UAXS for mutagenesis in this study as well as consensus amino acids in GAE1, UXS1, ArnA and UAXS1 (C100 and C140). Numbers indicate amino acid positions in the respective protein sequence.



Supplementary Figure 3. Enzyme purification documented by SDS-PAGE, shown for variants T139V and Y185F. The wild-type UAXS and other variants yielded soluble protein with similar purity and quantity. L, molecular mass markers; FT, flow through during sample loading on His trap column; other lanes: protein after purification loaded in different quantities.

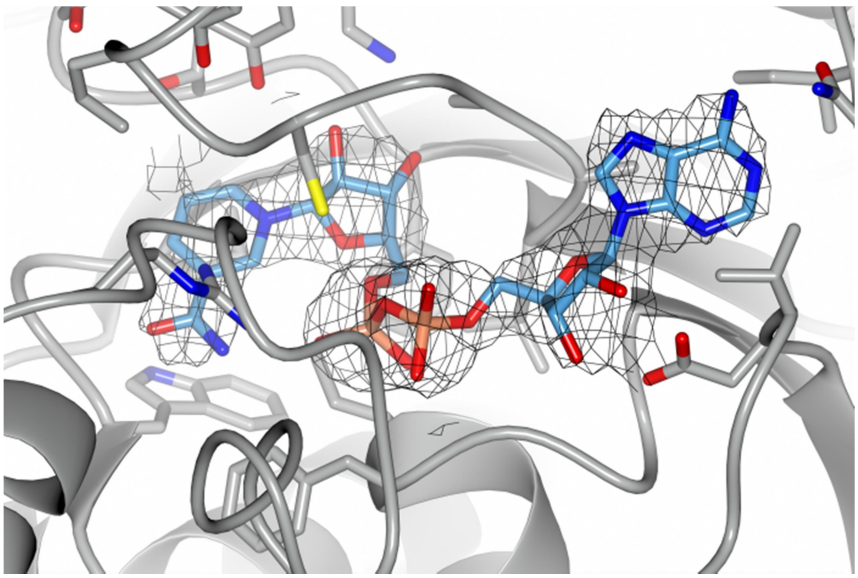


Supplementary Figure 4. SDS polyacrylamide gel showing purified wild-type UAXS used for crystallization experiments. From left to right: FT, flow through during sample loading on His trap column; 10% and 100%, wash steps with 10% and 100% buffer B, containing 400 mM imidazole; MKR, molecular mass marker; 17-25, fractions containing purified UAXS (elution with 50% buffer B).

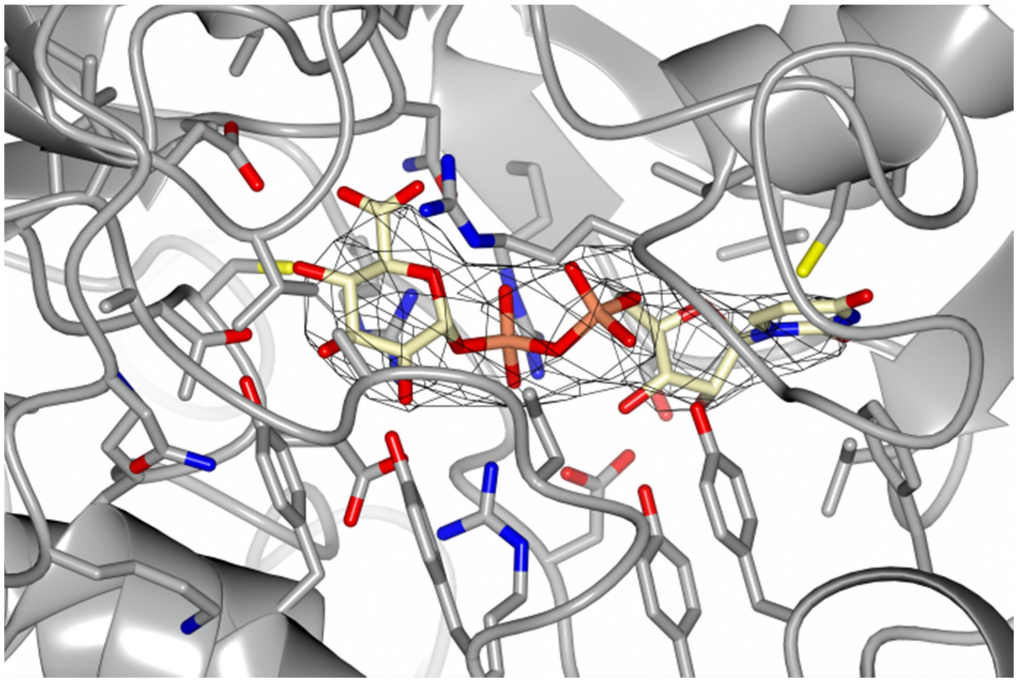


Supplementary Figure 5. Absorbance trace from size exclusion chromatography performed with purified wild-type UAXS. A single monodispersed peak consistent with a 90 kDa object (dimeric UAXS; 12 to 14 mL) is eluted from the column.

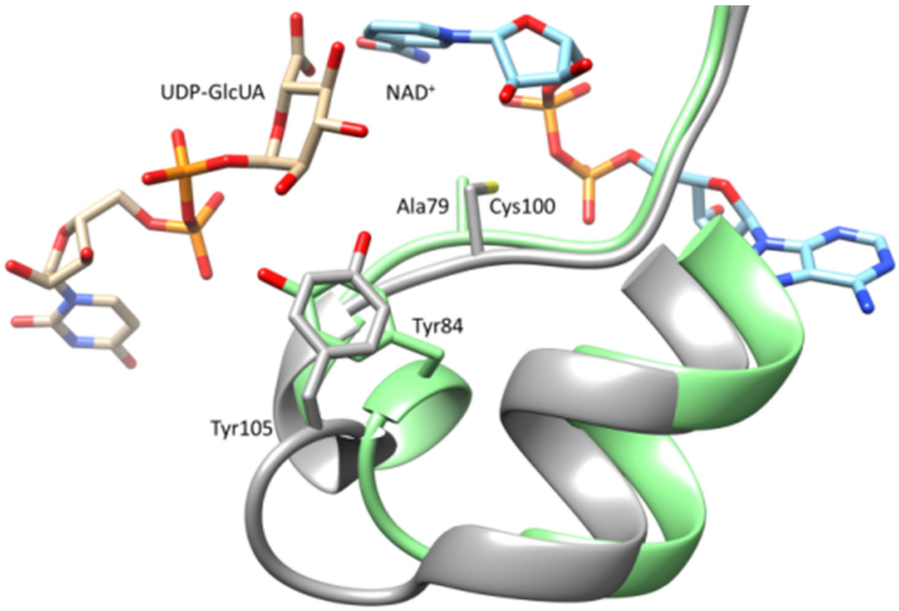
Supplementary Figure 6. Crystals of wild-type UAXS in a 4 μL sitting drop.



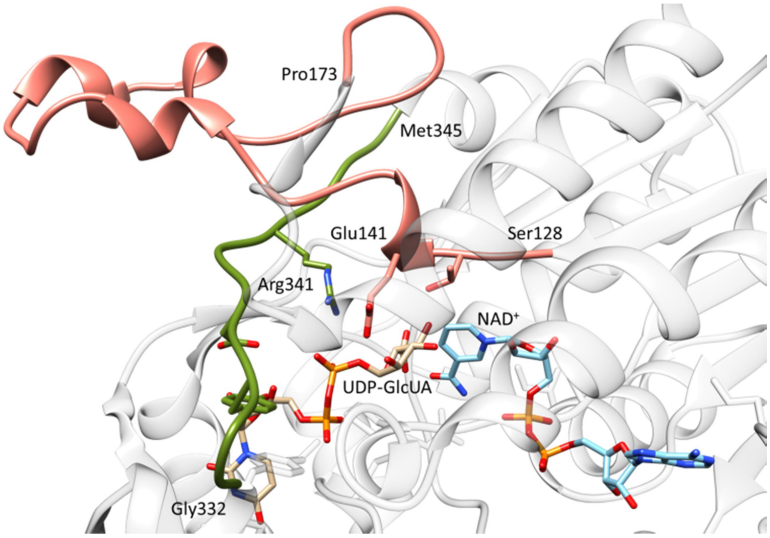
Supplementary Figure 7. The final weighted 2Fo-Fc electron density map with fully traced NAD^+ in wild-type UAXS. The contour level is 1.2 σ.



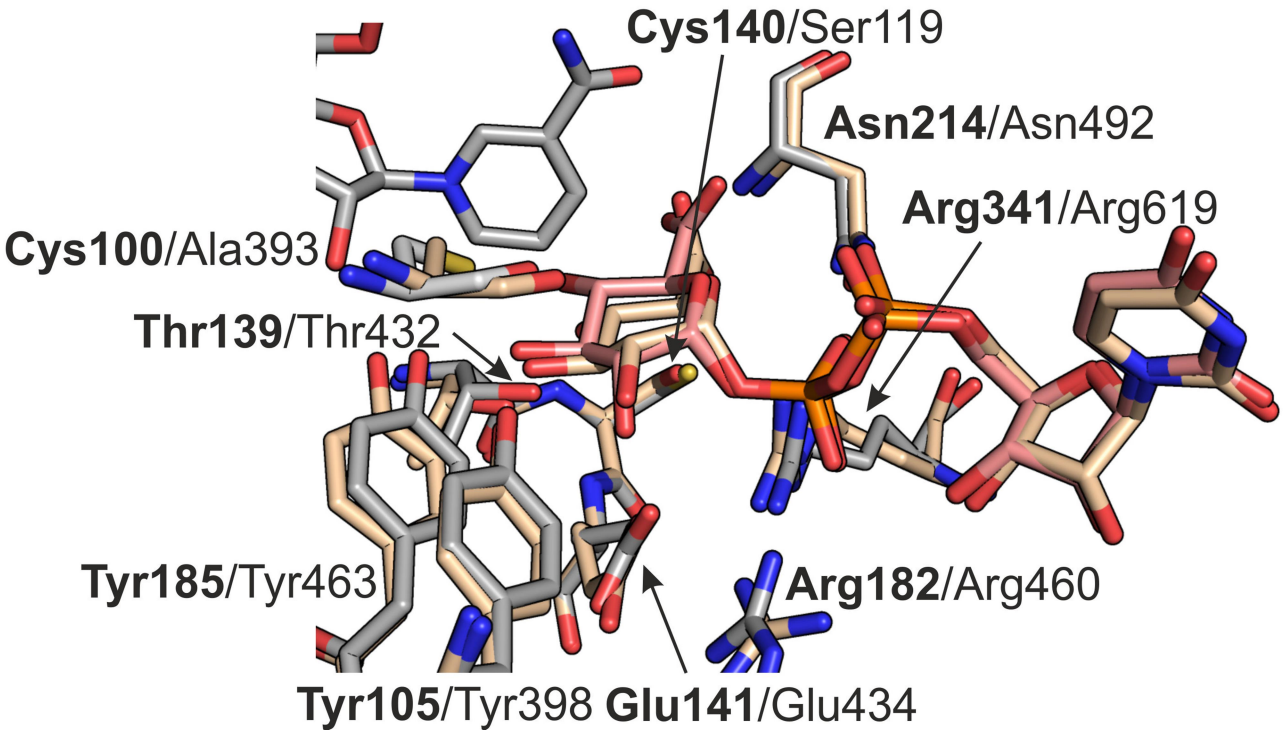
Supplementary Figure 8. The final weighted 2Fo-Fc electron density map for UDP-GlcA (**2**) within the structure of C100A variant co-crystallized with NADH and UDP-GlcA (**2**). The contour level is 1.2 σ.



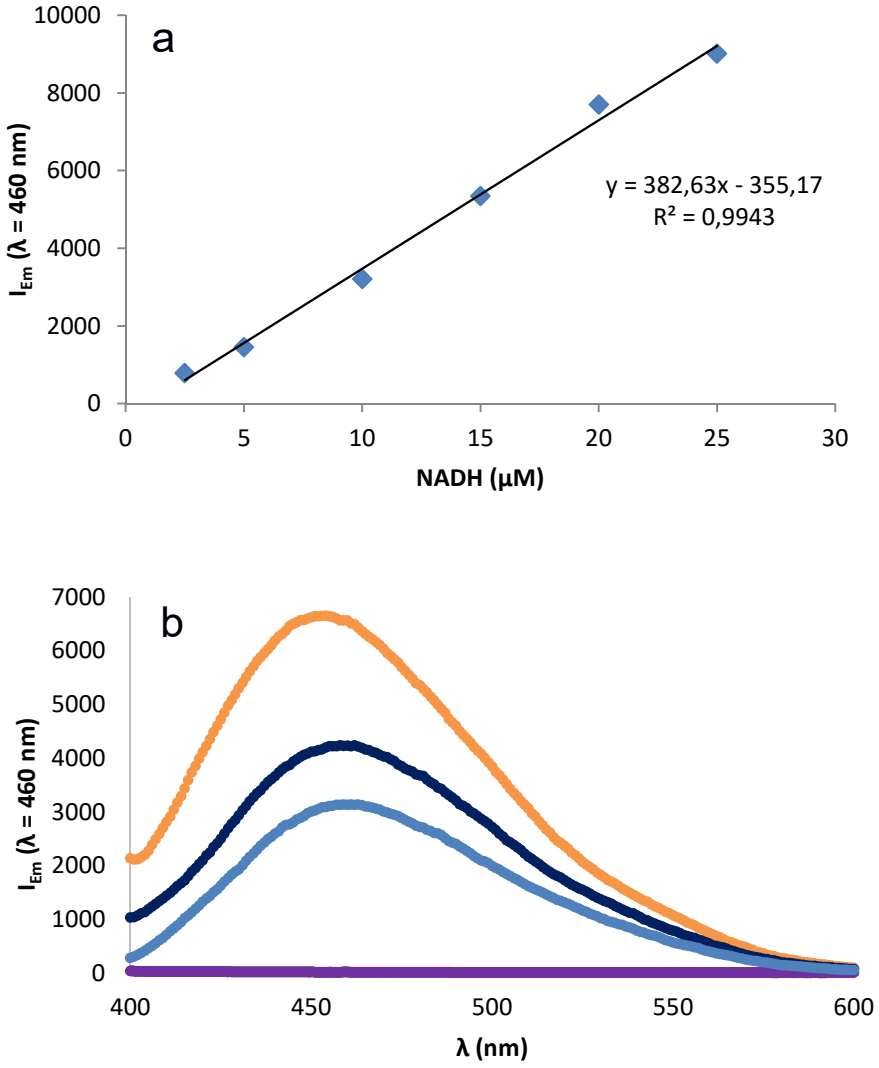
Supplementary Figure 9. Comparison of the loop proximal to the active site of UAXS (amino acid residues 97-115) and human UXS (amino acid residues 76-94; PDB: 2B69). In human UXS (light green ribbon)[22], the shorter loop results in Tyr84 pointing away from the active site (distance between the Tyr84-OH and the C2'-OH and C3'-OH is 3.1 and 5.5 Å, respectively), while in UAXS (grey ribbon) Tyr105-OH point towards the C2'-OH and C3'-OH of UDP-GlcA (**2**) with distances of 3.0 and 3.4 Å, respectively. The UAXS model was generated by combining two experimental structures (wild-type UAXS and C100A variant).



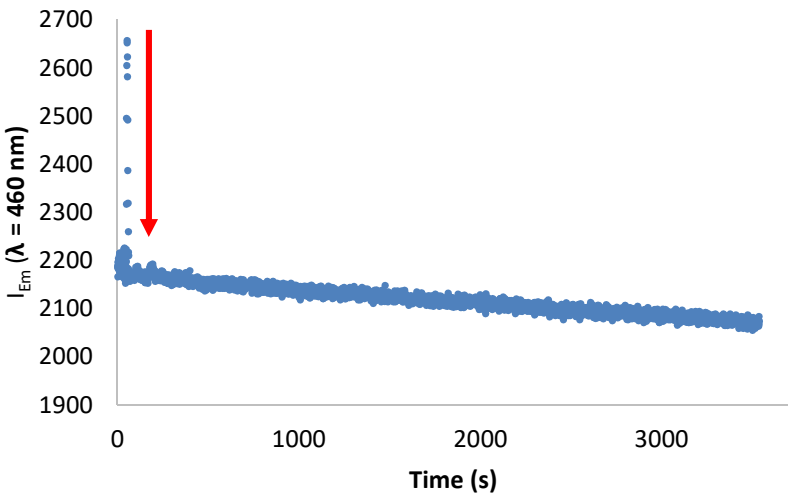
Supplementary Figure 10. Crystal structure of UAXS highlighting loops harbouring Glu141 (salmon ribbon, amino acid residues 128-173) and Arg341 (olive ribbon, amino acid residues 332-345).



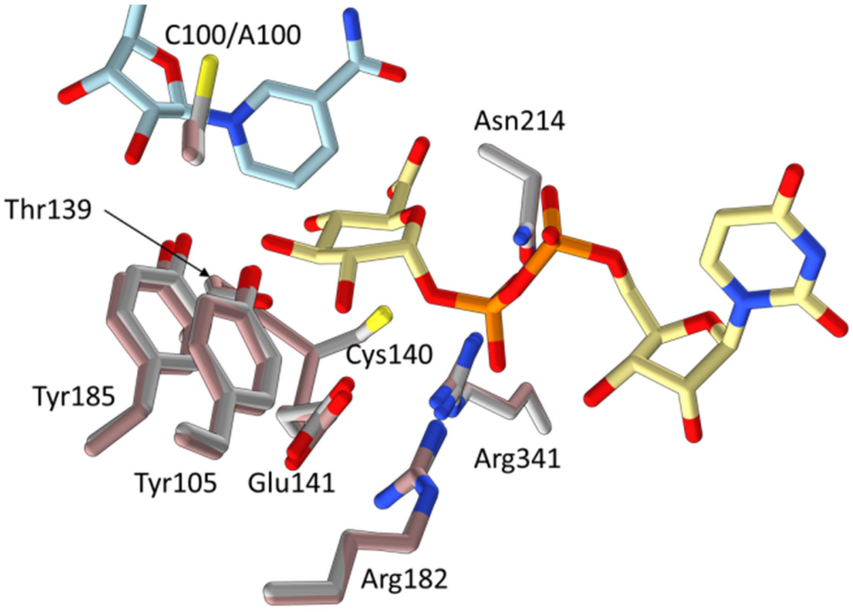
Supplementary Figure 11. Structural comparison of the active sites of wild-type UAXS (grey carbon atoms) in complex with UDP-GlcA (**2**) and NAD^+ and ArnA in complex with UDP-GlcA (**2**) (PDB entry 1Z7E; beige carbon atoms).[32] For comparison, the UAXS structure also contains the glucuronyl moiety of UDP-GlcA (**2**) in the conformation observed in the structure of the C100A variant (Supplementary Figure 8). The major differences in the active site are residues Cys100 and Cys140 of UAXS as compared to Ala393 and Ser119 of ArnA.



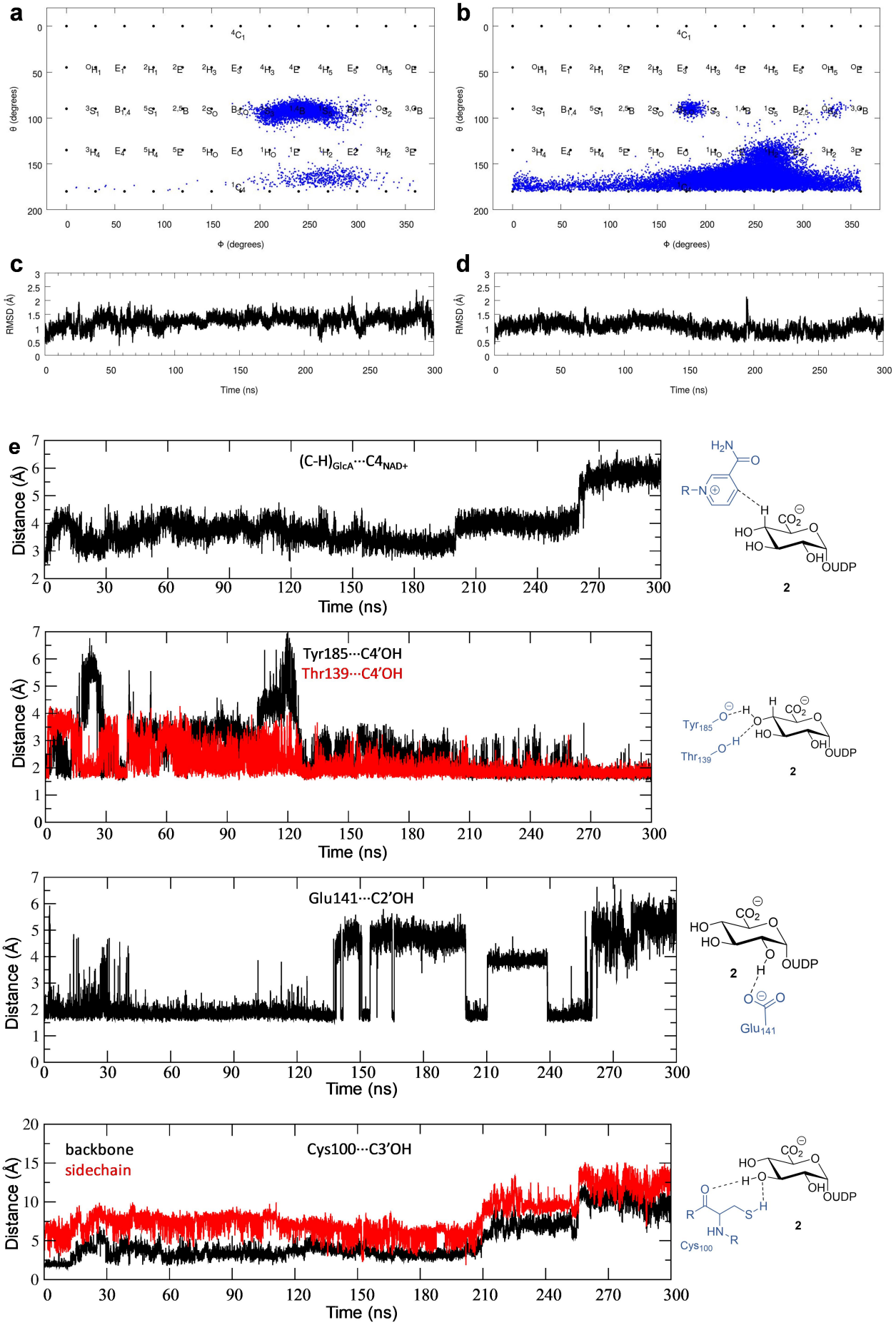
Supplementary Figure 12. NAD(H) binding to UAXS. The top panel (**a**) shows the calibration for NADH fluorescence. Standard solutions (80 µL each) of NADH were prepared in double distilled water and the fluorescence emission intensity was measured with λ_{Ex} at 340 nm and λ_{Em} at 460 nm. The bottom panel (**b**) shows the fluorescence emission spectra of released NADH from two independently purified UAXS preparations. Batch 1 (orange line): 18.3 µM NADH was released from 23.5 µM of UAXS (78% occupancy by NADH). Batch 2 (dark blue line): 12.0 µM NADH was released from 19.6 µM of UAXS (61% occupancy by NADH). Standard solution of 10 µM NADH in methanol (light blue line) was used as a reference. Purple line: storage buffer used in experiments with UAXS.

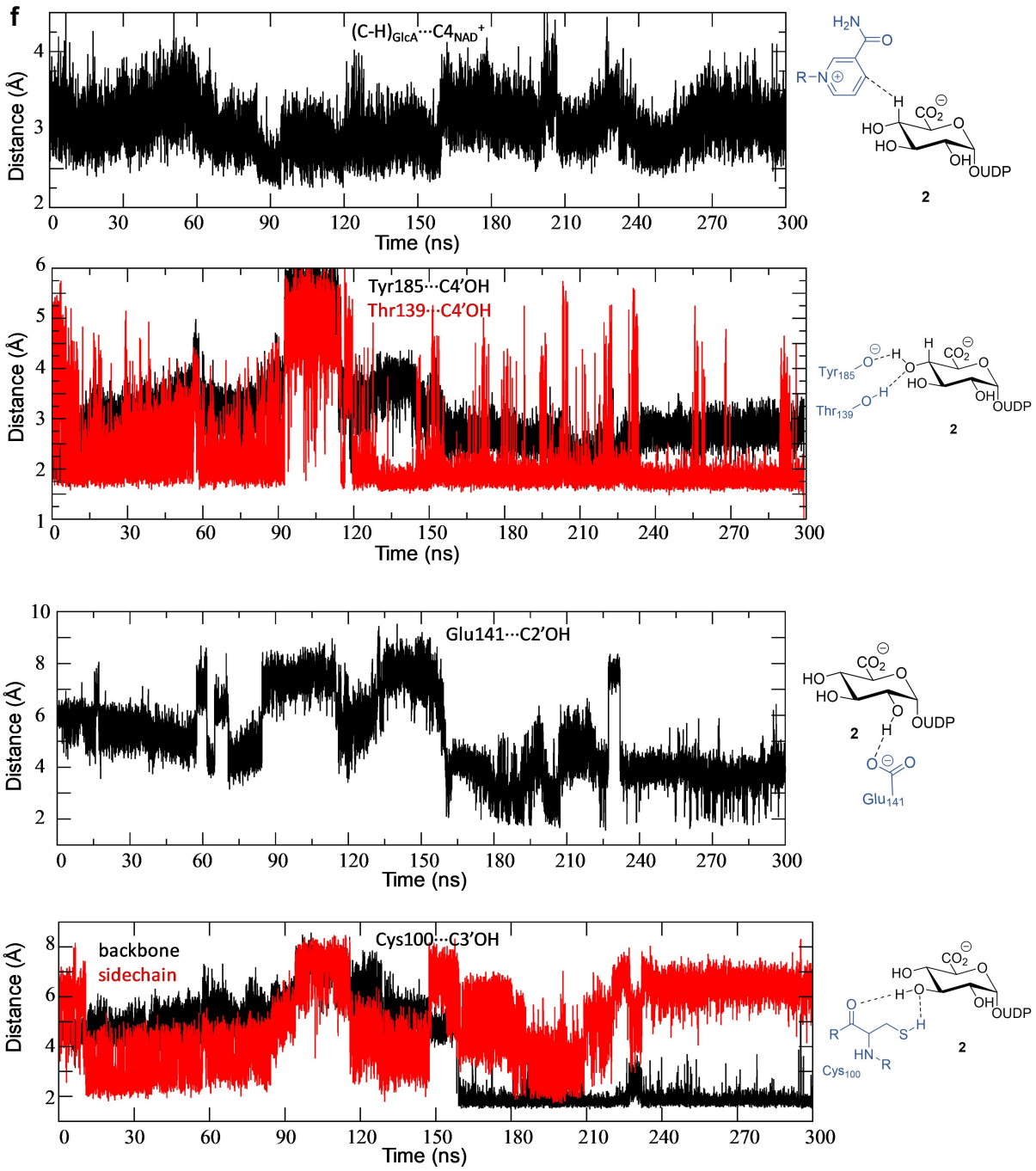


Supplementary Figure 13. Exchange of NADH by NAD^+ in the active site of wild-type UAXS. Note that there is a small spectral difference between the free and bound form of NADH.[26] NADH exists in (at least) two forms: a free state (unbound) with low quantum yield and a protein bound state, where the quantum yield can be substantially higher. The red arrow indicates the rapid drop in emission directly after addition of NAD^+ to the sample.

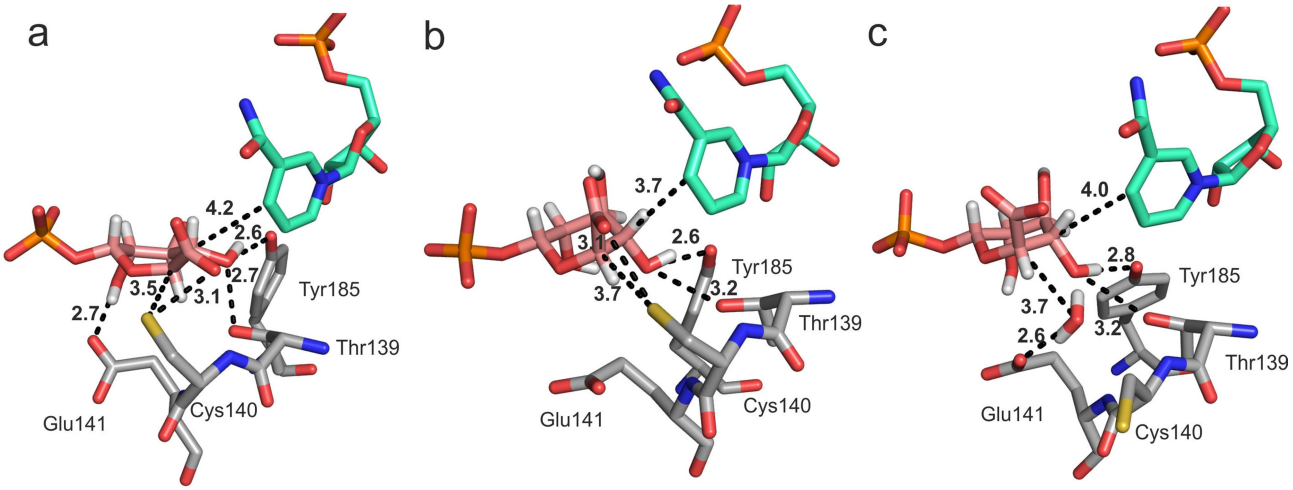


Supplementary Figure 14. Superposition of active-site residues of wild-type UAXS (light grey carbon atoms) and C100A variant (light brown carbon atoms).

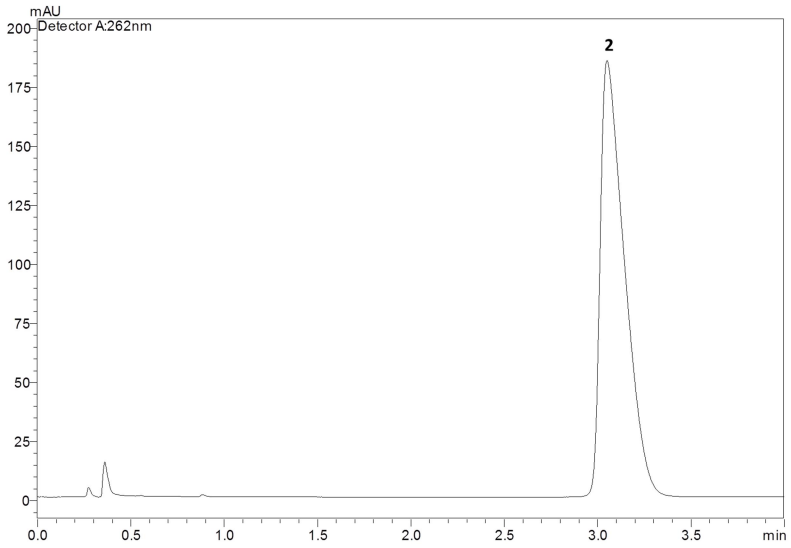

a
b
θ (degrees)
Φ (degrees)
c
d
RMSD (Å)
Time (ns)
e
(C-H)GlcA···C4NAD+
Distance (Å)
Time (ns)
Tyr185···C4'OH
Thr139···C4'OH
Glu141···C2'OH
backbone
sidechain
Cys100···C3'OH
OUDP
2



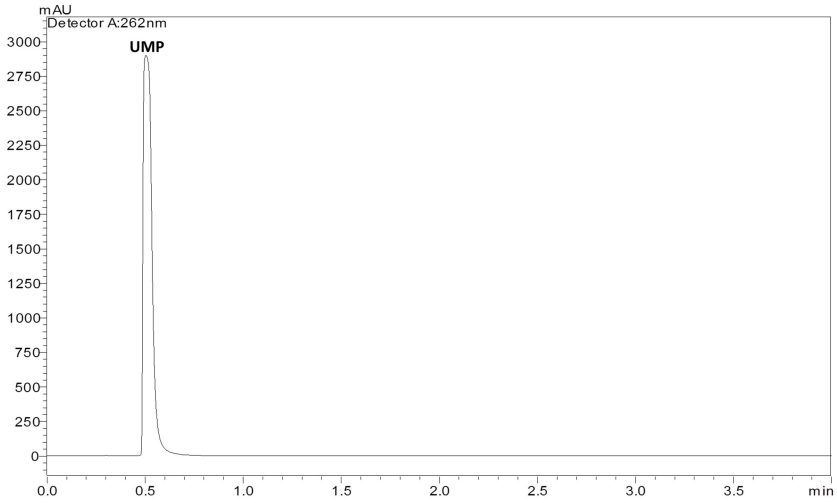
Supplementary Figure 15. Maps of the ring puckering geometries and representation of the root mean square displacement (RMSD) of the protein backbone with respect to the initial structure and selected distances along the MD simulations. Panels (**a**), (**c**) and (**e**) show data from Replica 1. Panels (**b**), (**d**) and (**f**) show data from Replica 2. The distance plots for Cys100 are shown to highlight the conformational flexibility of this residue and its proximity to the C3/C3OH of substrate. However, note: Cys100 is not involved in the initial steps of oxidation and ring opening.



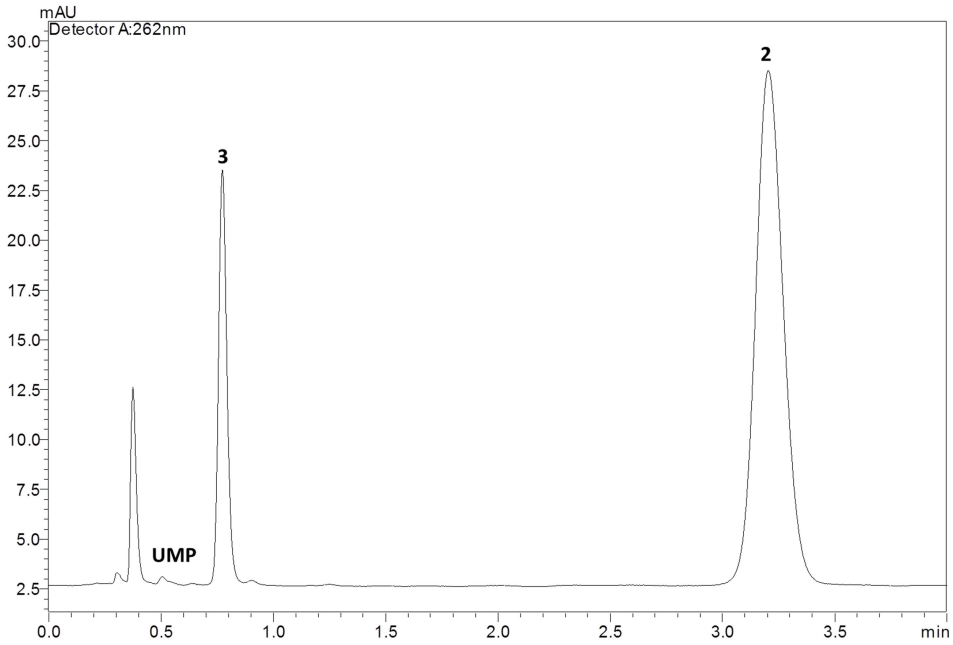
Supplementary Figure 16. Results of MD simulation suggest a possible role of Cys140 in the obligatory proton uptake to substrate C5' in the decarboxylation step of the UAXS reaction. The panels **a** - **c** show representative snapshots from the trajectories of the MD simulations (panel **a**, Replica 1; panels **b** and **c**, Replica 2). The side chain of Cys140 is generally quite flexible in the MD simulations. (**a**) An enzyme-substrate complex from Replica 1 that has Glu141 interacting with the 2'-OH and Cys140 placed close to the C5' of UDP-GlcA (**2**). The Cys140 would thus be positioned for stereospecific (*Si* side) protonation of the C5' from below the sugar ring. Note: due to side chain flexibility, the position of Cys140 is slightly different from that in Figure 3a of the main text. The results show that Cys140 could easily adopt a conformation suitable for proton transfer during the decarboxylation. (**b**) An enzyme-substrate complex from Replica 2 that has Cys140 positioned side-on of the sugar ring. Although the side chain of Cys140 points toward the C5', it is not in a position suitable for proton transfer. Glu141 is not interacting with the 2'-OH. (**c**) An enzyme-substrate complex from Replica 2 that would establish an UXS-like[15] proton conduit to the C5'. This proton conduit involves a Glu141-coordinated water that is placed below the sugar ring for stereospecific protonation. The analogous arrangement of groups (involving the homologous glutamate residue) is observed in human UXS.[15]



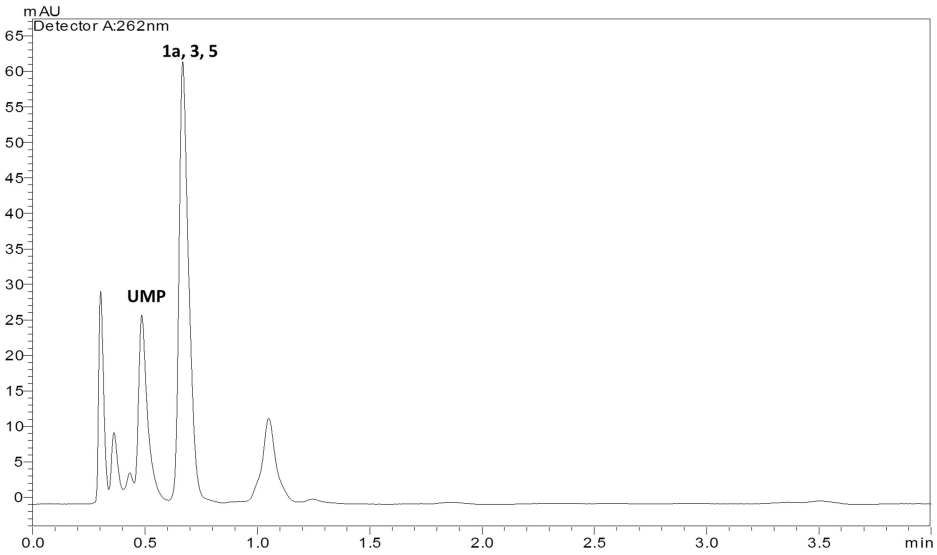
Supplementary Figure 17. HPLC trace for commercial standard of UDP-GlcA (**2**). The sample contained no enzyme (negative control).



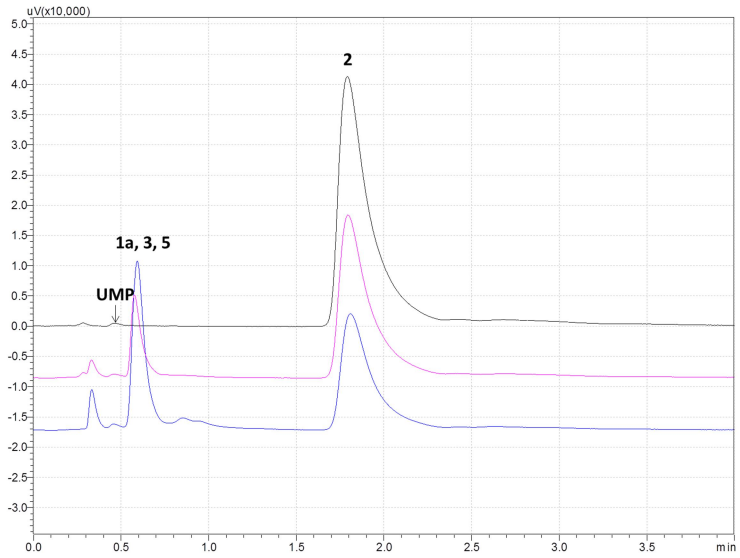
Supplementary Figure 18. HPLC trace for commercial standard of UMP. The sample contained no enzyme.



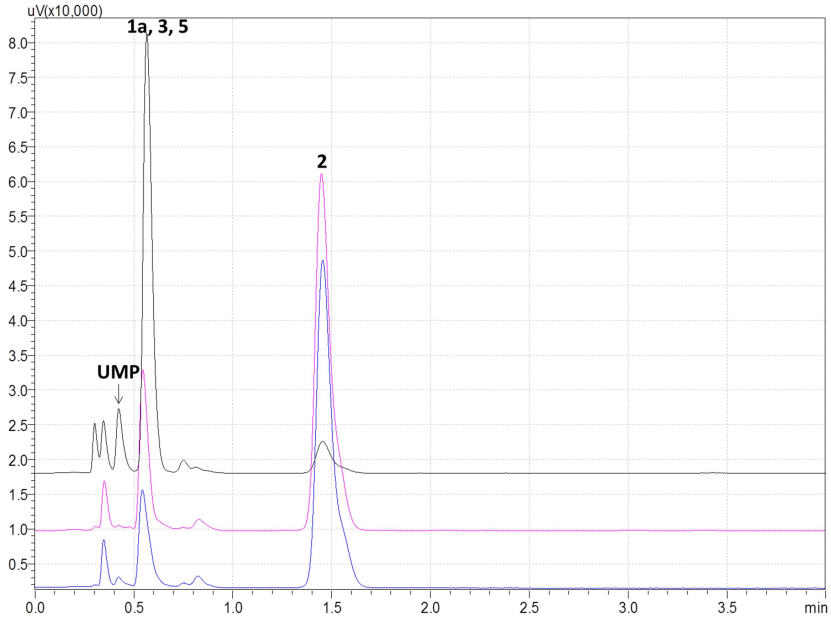
Supplementary Figure 19. HPLC trace for conversion of UDP-GlcA (**2**) by human UXS. **3** = UDP-xylose.



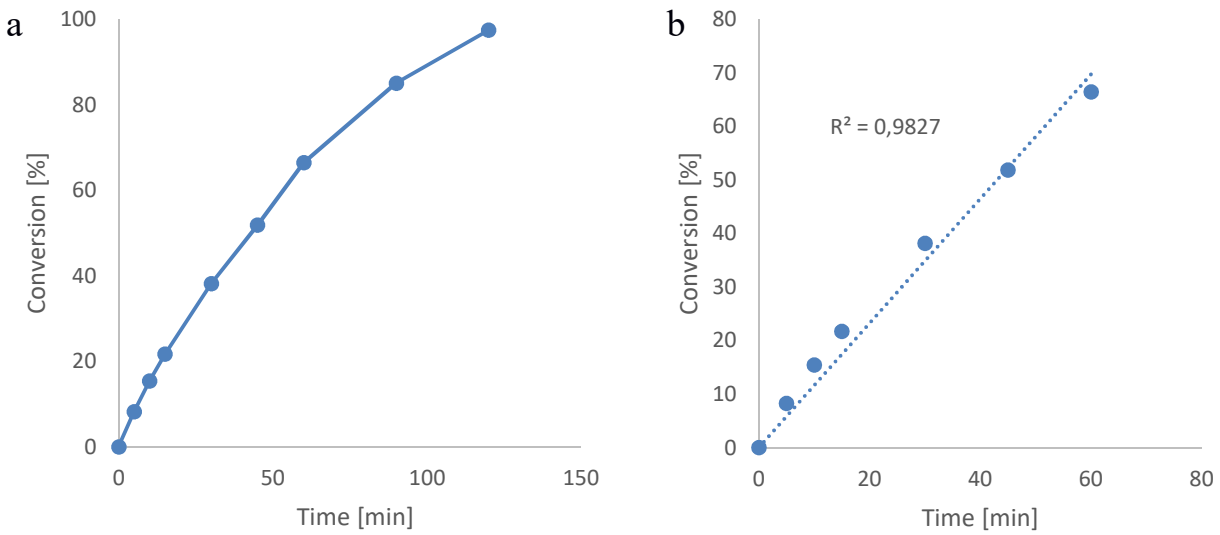
Supplementary Figure 20. HPLC trace showing conversion of UDP-GlcA (**2**) by wild-type UAXS. **1a** = UDP-apiose; **3** = UDP-xylose; **5** = UDP-4-keto-xylose. Products cannot be separated efficiently using HPLC under the described conditions. The peak at ~1 min corresponds to UDP.



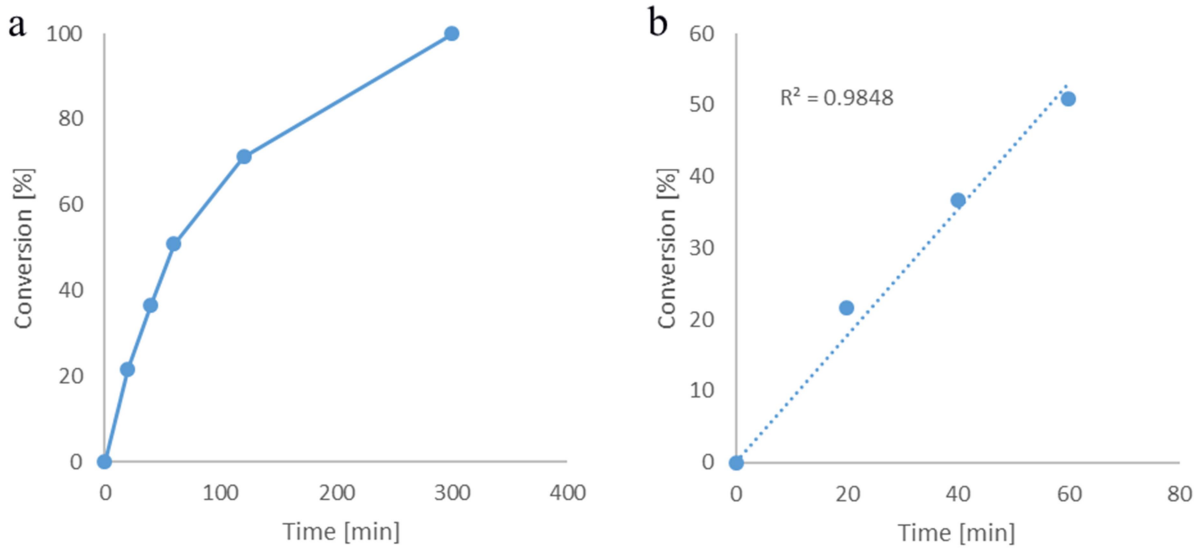
Supplementary Figure 21. Representative conversion of UDP-GlcA (**2**) by Y105F variant. Black lane = no enzyme. Pink lane = human UXS. Blue lane = Y105F variant. **1a** = UDP-apiose; **2** = UDP-GlcA; **3** = UDP-xylose; **5** = UDP-4-keto-xylose. Products cannot be separated efficiently using HPLC under the described conditions. Retention times were shifted slightly due to minor changes in mobile phase composition, flow-rate, and HPLC column performance. Reference compounds (UMP, UDP-GlcA, **2**) and control reactions (UDP-GlcA (**2**) conversion with wild-type UAXS and human UXS) were injected to ensure reliable assignment of products.



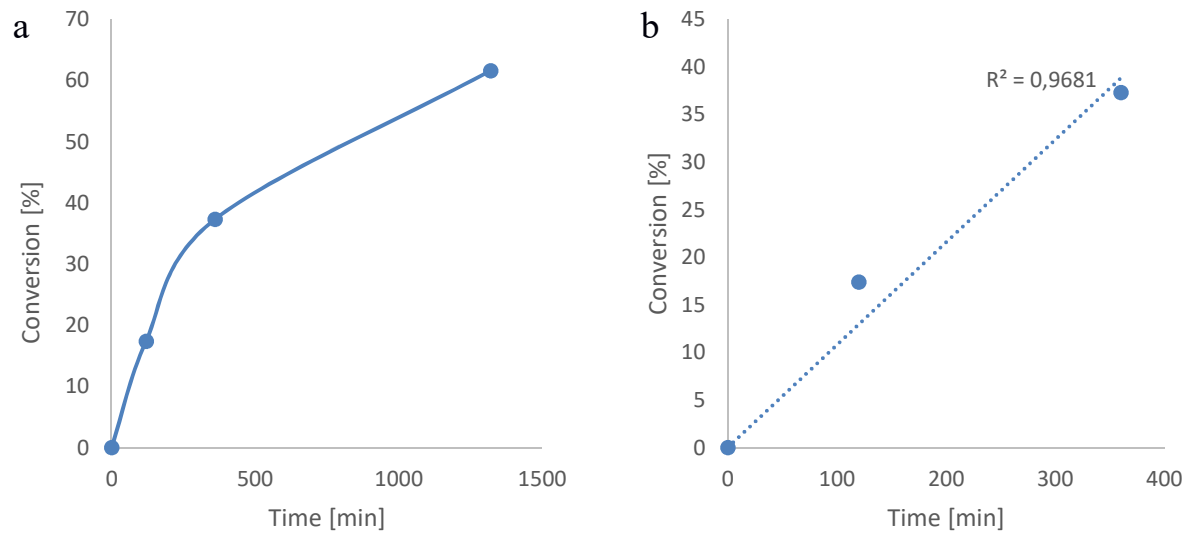
Supplementary Figure 22. HPLC traces for conversion of UDP-GlcA (**2**) by T139V variant. Black lane: UAXS pH 7.0; Pink lane: T139V variant (reaction at pH 7.0); Blue lane: T139V variant (reaction at pH 8.5). **1a** = UDP-apiose; **2** = UDP-GlcA; **3** = UDP-xylose; **5** = UDP-4-keto-xylose. Products cannot be separated efficiently using HPLC under the described conditions. Retention times were shifted slightly due to minor changes in mobile phase composition, flow-rate, and HPLC column performance. Reference compounds (UMP, UDP-GlcA (**2**)) and control reactions (UDP-GlcA (**2**) conversion with wild-type UAXS and human UXS) were injected to ensure reliable assignment of products.



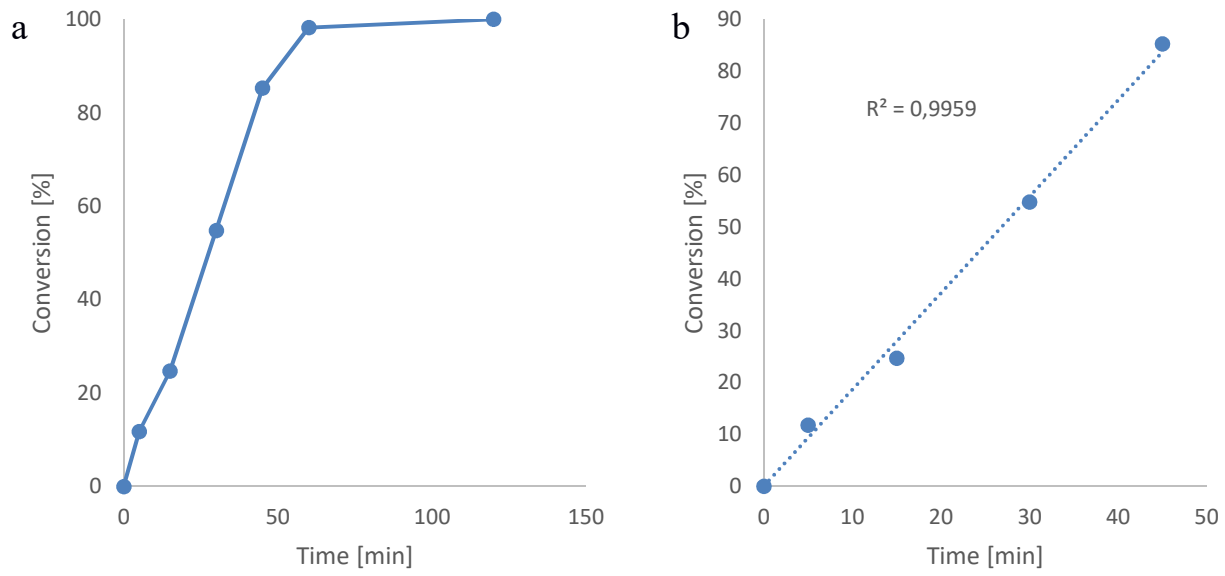
Supplementary Figure 23. Time course for conversion of 2 mM UDP-GlcA (**2**) by wild-type UAXS (2 mg/mL). (**a**) Full time course recorded in the experiment. (**b**) Part of the time course used for initial rate determination. The data shown are averages from two independent replicates. The initial rate determined from linear fit of the data in panel b was 11 nmol/(min mg protein). This value is used to calculate a k_{cat} of 0.49 (± 0.02) min^{-1}.



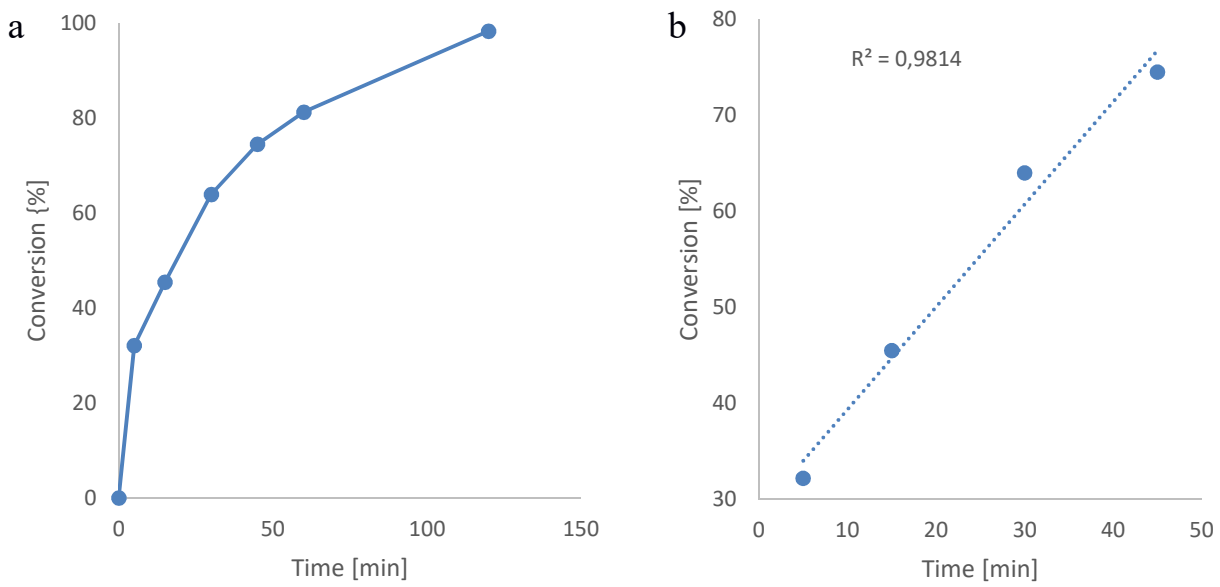
Supplementary Figure 24. Time course for conversion of 2 mM UDP-GlcA (**2**) by Y105F variant (14.3 mg/mL). (**a**) Full time course recorded in the experiment. (**b**) Part of the time course used for initial rate determination. The data shown are averages from two independent replicates. The initial rate determined from linear fit of the data in panel b was 1.2 nmol/(min mg protein). This value is used to calculate a k_{cat} of 0.053 (± 0.017) min^{-1}.



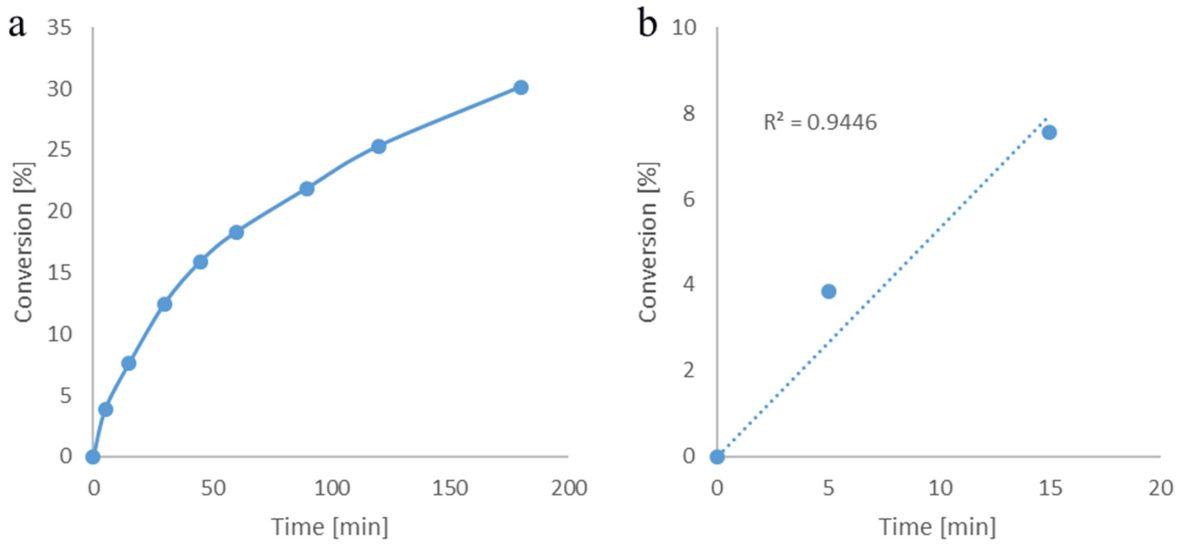
Supplementary Figure 25. Time course for conversion of 0.5 mM UDP-GlcA (**2**) by T139V variant (3.3 mg/mL). Due to the very low activity and the suppressed reduction activity of this variant, 1 mM NAD^+ were supplemented to the reaction. (**a**) Full time course recorded in the experiment. (**b**) Part of the time course used for initial rate determination. The data shown are averages from two independent replicates. The initial rate determined from linear fit of the data in panel b was 6.6×10^{-5} nmol/(min mg protein). This value is used to calculate a k_{cat} of 3×10^{-3} ($\pm 9.6 \times 10^{-5}$) min^{-1}.



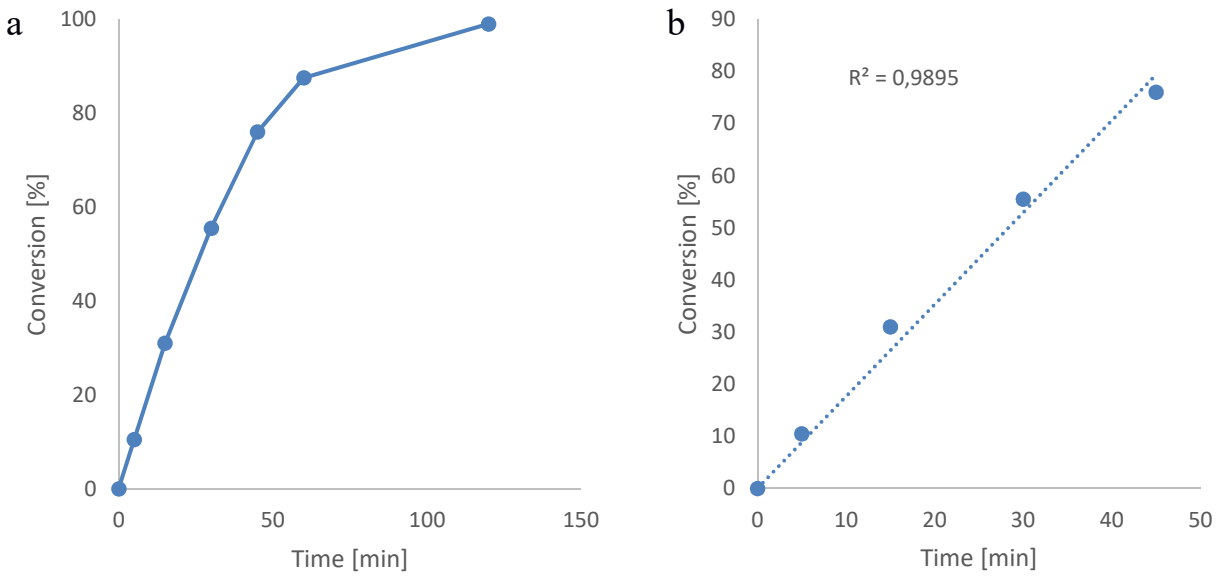
Supplementary Figure 26. Time course for conversion of 2 mM UDP-GlcA (**2**) by C100A variant (14.3 mg/mL). (**a**) Full time course recorded in the experiment. (**b**) Part of the time course used for initial rate determination. The data shown are averages from two independent replicates. The initial rate determined from linear fit of the data in panel b was 2.6 nmol/(min mg protein). This value is used to calculate a k_{cat} of 0.12 ($\pm$ 4.9 $\times$ 10^{-4}) min^{-1}.



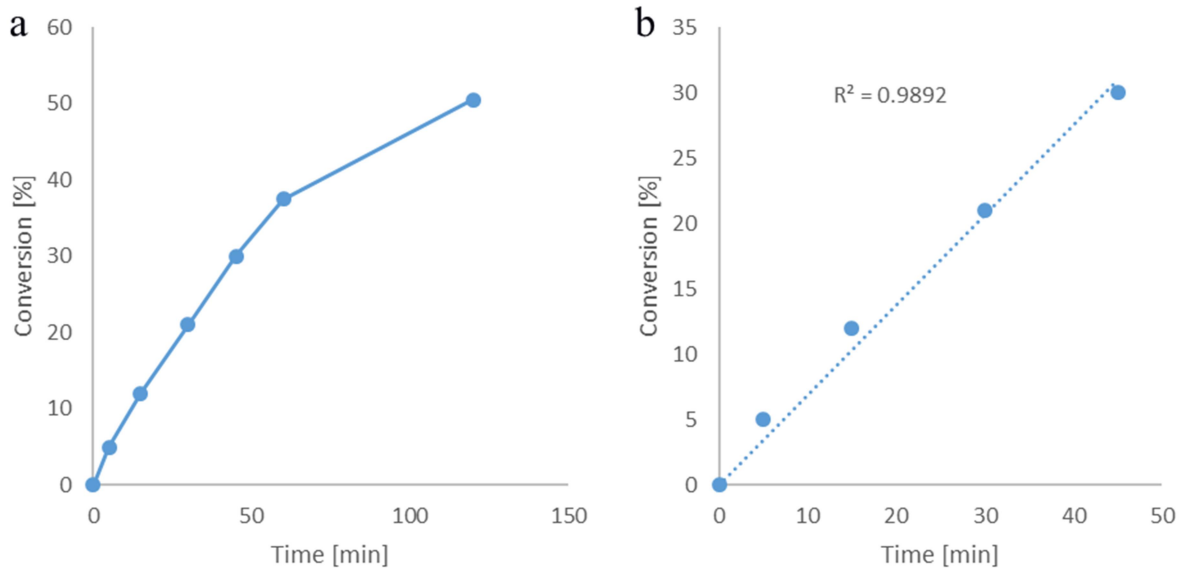
Supplementary Figure 27. Time course for conversion of 2 mM UDP-GlcA (**2**) by C140S variant (14.3 mg/mL). (**a**) Full time course recorded in the experiment. (**b**) Part of the time course used for initial rate determination. The data shown are averages from two independent replicates. The initial rate determined from linear fit of the data in panel b was 1.5 nmol/(min mg protein). This value is used to calculate a k_{cat} of 0.07 ($\pm$ 1.2 $\times$ 10^{-3}) min^{-1}.



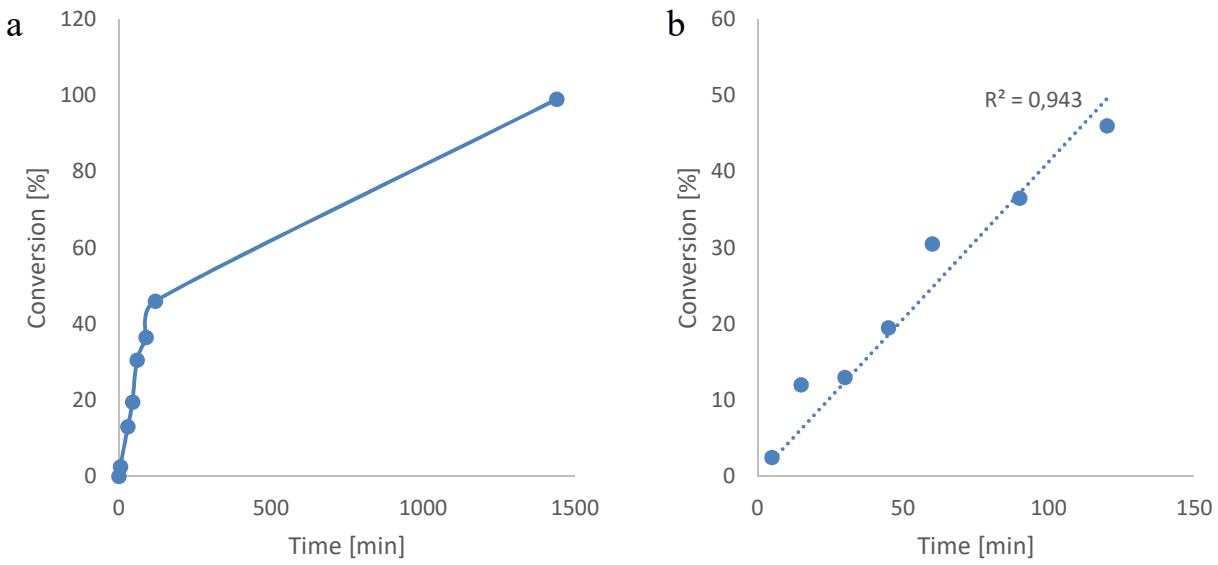
Supplementary Figure 28. Time course for conversion of 2 mM UDP-GlcA (**2**) by C100A/C140S variant (2 mg/mL). (**a**) Full time course recorded in the experiment. (**b**) Part of the time course used for initial rate determination. The data shown are averages from two independent replicates. The initial rate determined from linear fit of the data in panel b was 5 nmol/(min mg protein). This value is used to calculate a k_{cat} of 0.23 (± 0.01) min^{-1}.



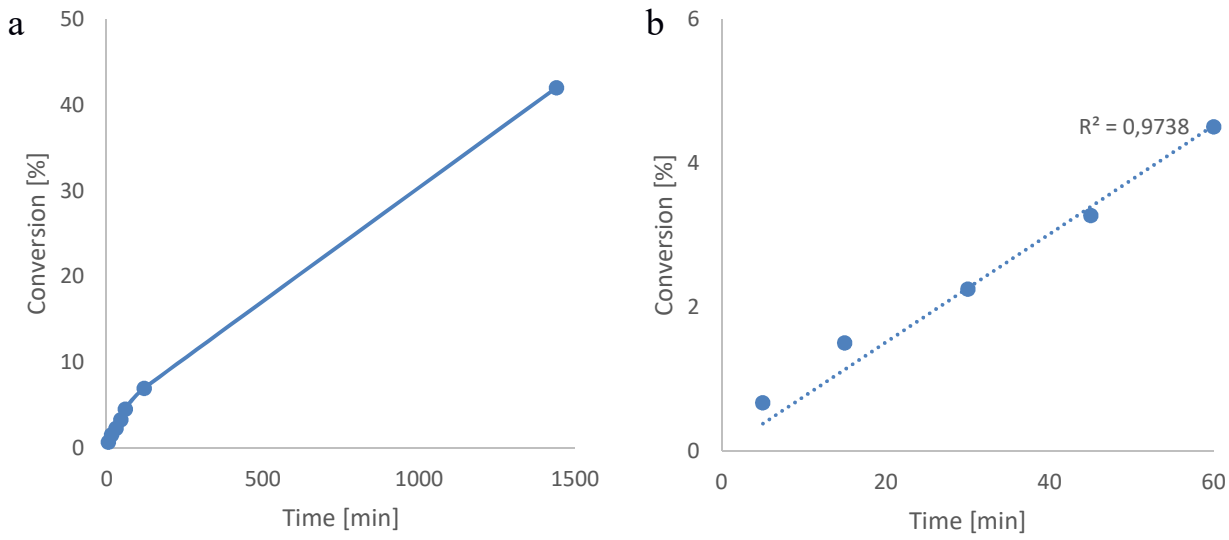
Supplementary Figure 29. Time course for conversion of 2 mM UDP-GlcA (**2**) by C140A variant (14.3 mg/mL). (**a**) Full time course recorded in the experiment. (**b**) Part of the time course used for initial rate determination. The data shown are averages from two independent replicates. The initial rate determined from linear fit of the data in panel b was 2.4 nmol/(min mg protein). This value is used to calculate a k_{cat} of 0.11 ($\pm$ 1.1 $\times$ 10^{-3}) min^{-1}.



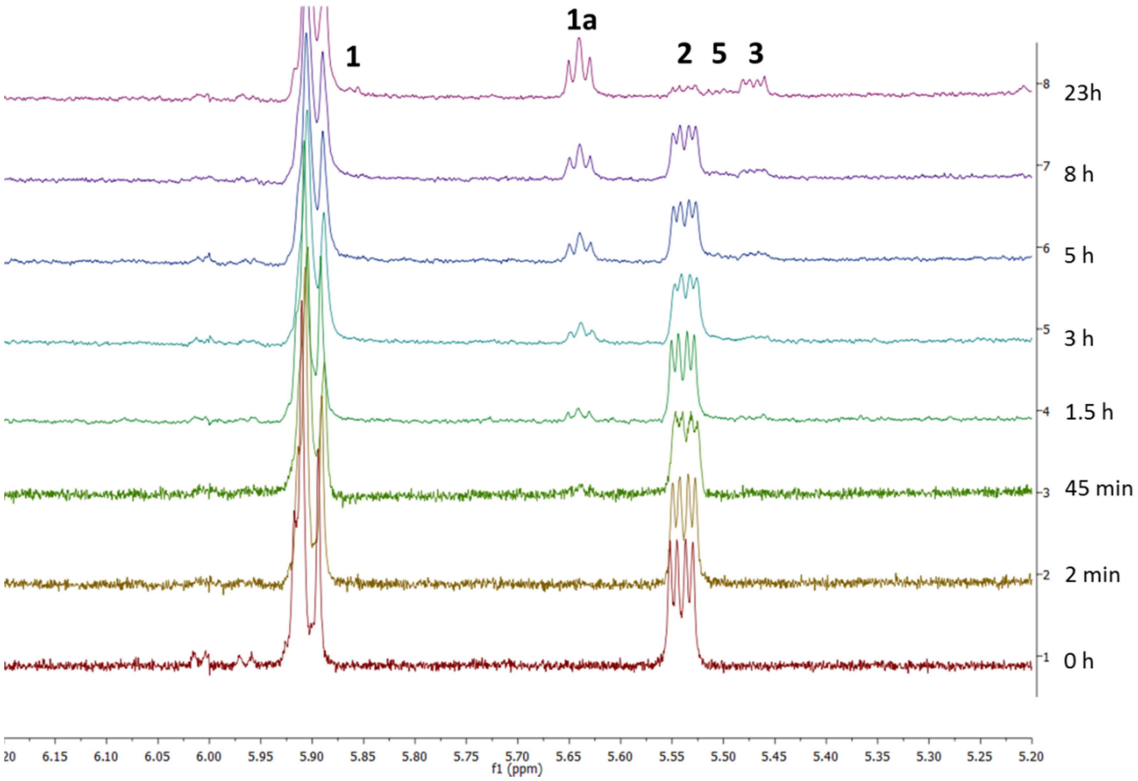
Supplementary Figure 30. Time course for conversion of 2 mM UDP-GlcA (**2**) by C100S variant (7 mg/mL). (**a**) Full time course recorded in the experiment. (**b**) Part of the time course used for initial rate determination. The data shown are averages from two independent replicates. The initial rate determined from linear fit of the data in panel b was 1.9 nmol/(min mg protein). This value is used to calculate a k_{cat} of 0.09 ($\pm$ 9.2 $\times$ 10^{-4}) min^{-1}.



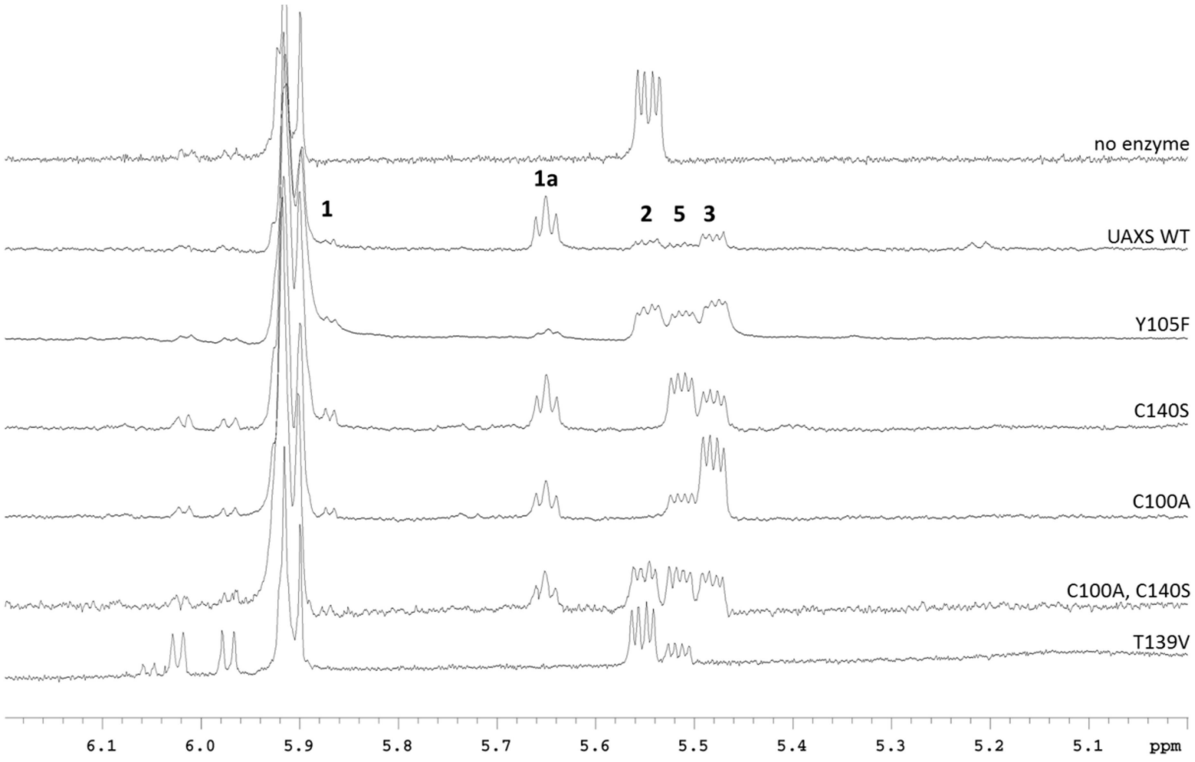
Supplementary Figure 31. Time course for conversion of 2 mM UDP-GlcA (**2**) by Y105A variant (3.5 mg/mL). (**a**) Full time course recorded in the experiment. (**b**) Part of the time course used for initial rate determination. The data shown are averages from two independent replicates. The initial rate determined from linear fit of the data in panel b was 2.9 nmol/(min mg protein). This value is used to calculate a k_{cat} of 0.13 ($\pm$ 7.4 $\times$ 10^{-3}) min^{-1}.



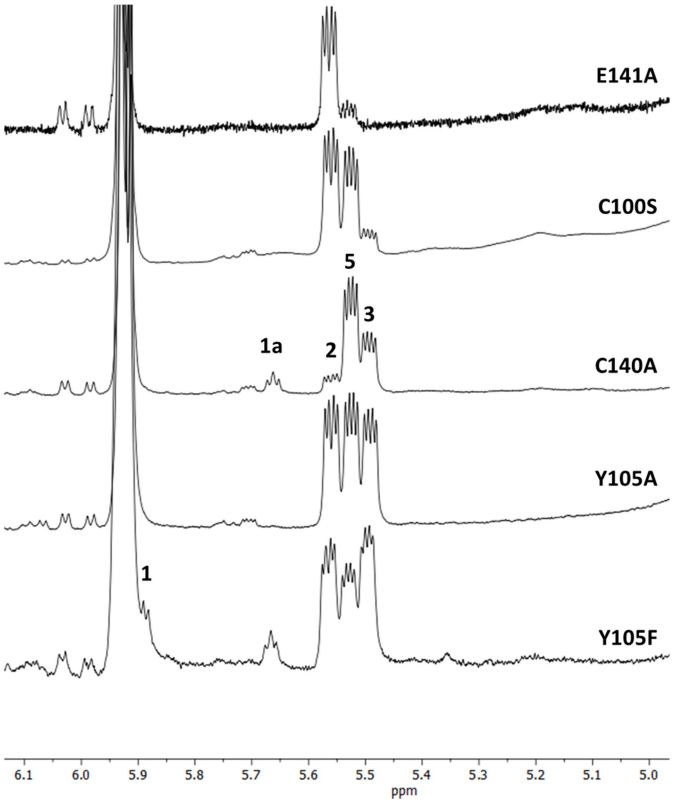
Supplementary Figure 32. Time course for conversion of 2 mM UDP-GlcA (**2**) by E141A variant (14.8 mg/mL). (**a**) Full time course recorded in the experiment. (**b**) Part of the time course used for initial rate determination. The data shown are averages from two independent replicates. The initial rate determined from linear fit of the data in panel b was 6.6×10^{-5} nmol/(min mg protein). This value is used to calculate a k_{cat} of 3.0×10^{-3} ($\pm$ 7.8×10^{-5}) min^{-1}.



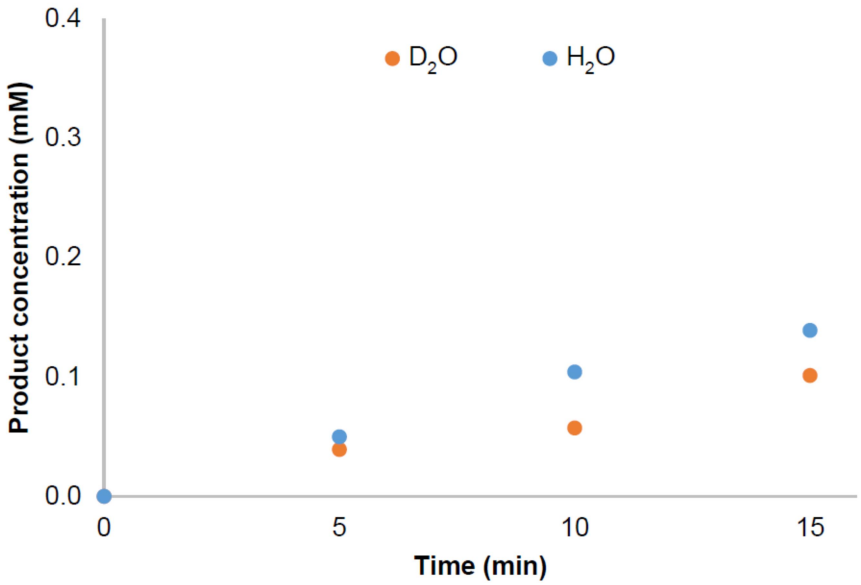
Supplementary Figure 33. ^{1}H-NMR time course for conversion of 2 mM UDP-GlcA (**2**) by wild-type UAXS (2 mg/mL). **1** = α-D-apiofuranosyl-1,2-cyclic phosphate; **1a** = UDP-apiose; **2** = UDP-GlcA; **3** = UDP-xylose; **5** = UDP-4-keto-xylose.



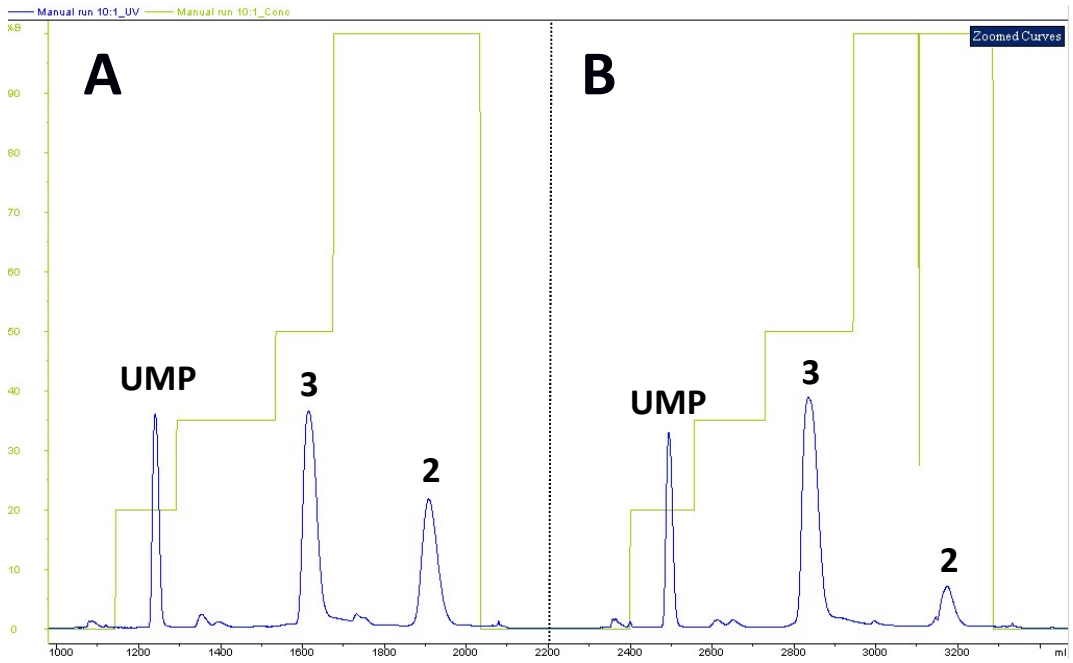
Supplementary Figure 34. Alignment of ^{1}H-NMR signals from conversion of 2 mM UDP-GlcA (**2**) with wild-type UAXS and enzyme variants generated by site-directed mutagenesis. As control, a reaction without enzyme is shown (lane 1). **1** = α-D-apiofuranosyl-1,2-cyclic phosphate; **1a** = UDP-apiose; **2** = UDP-GlcA; **3** = UDP-xylose; **5** = UDP-4-keto-xylose. The doublets around 6 ppm are signals from the uracil moiety of UMP and UDP.



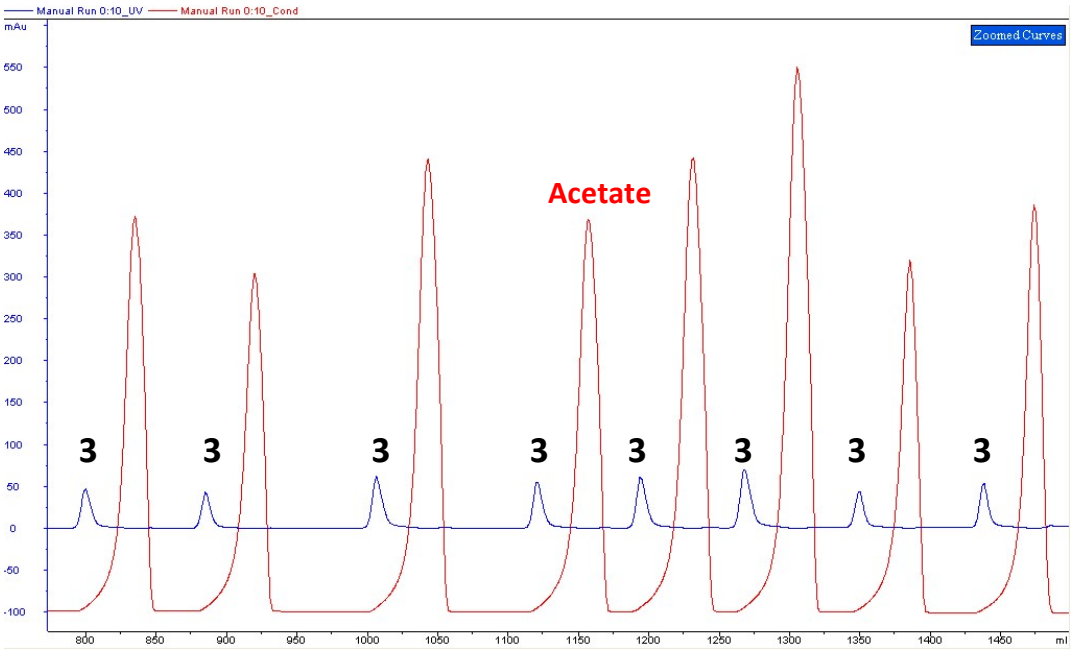
Supplementary Figure 35. Alignment of ^{1}H-NMR signals from conversion of 2 mM UDP-GlcA (**2**) with enzyme variants generated by site-directed mutagenesis. **1** = α-D-apiofuranosyl-1,2-cyclic phosphate; **1a** = UDP-apiose; **2** = UDP-GlcA; **3** = UDP-xylose; **5** = UDP-4-keto-xylose.



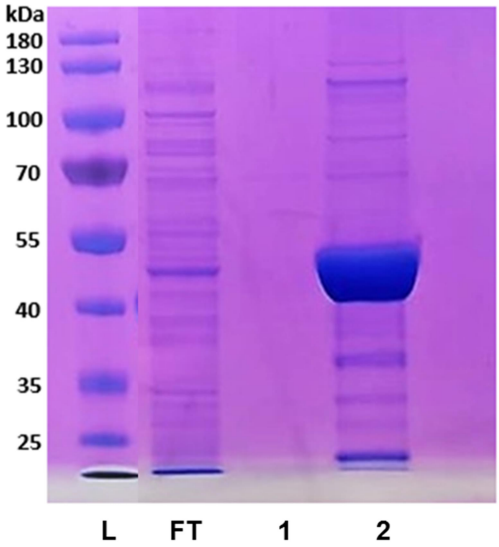
Supplementary Figure 36. Time course for conversion of 1 mM UDP-GlcA (**2**) by wild-type UAXS (2 mg/mL) using D_2O or H_2O as bulk solvent. Reactions were performed in duplicates. Linear fits of the data give a specific activity of 4.7 nmol/(min mg protein) in H_2O and 3.2 nmol/(min mg protein) in D_2O, respectively.



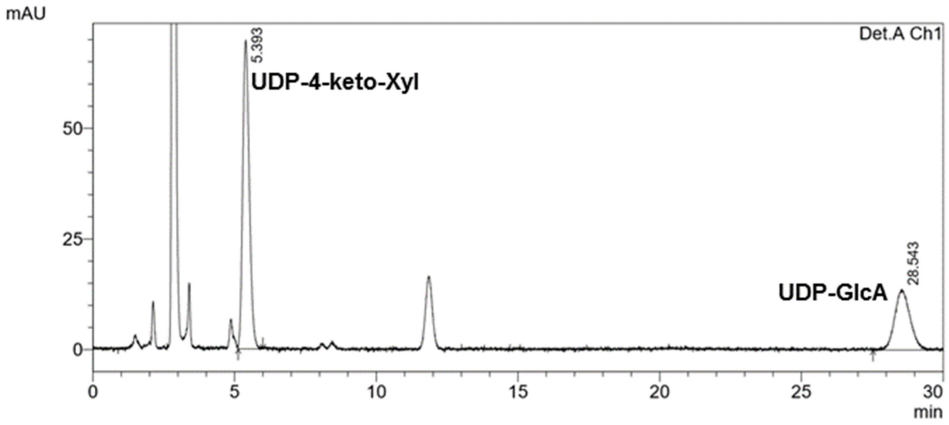
Supplementary Figure 37. Chromatogram recorded during the isolation of UDP-xylose (**3**) from conversions of UDP-GlcA (**2**) by C100S (A) and C100A (B) performed in D_2O (deuterium wash-in experiment). Blue line = UV signal. Green line = salt gradient used for elution. UMP = uridine monophosphate.



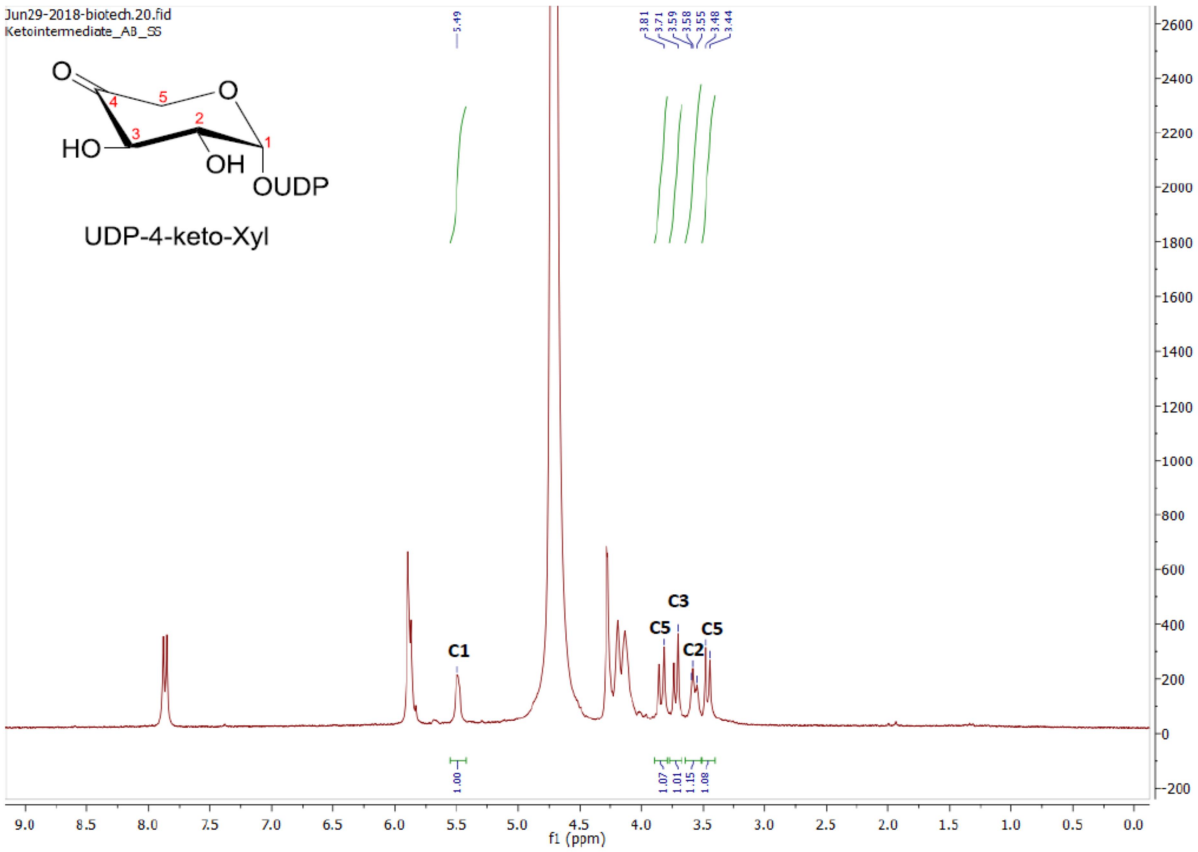
Supplementary Figure 38. Chromatogram recorded during the desalting of isolated UDP-xylose (**3**). Blue line = UV signal. Red line = conductivity signal of acetate.



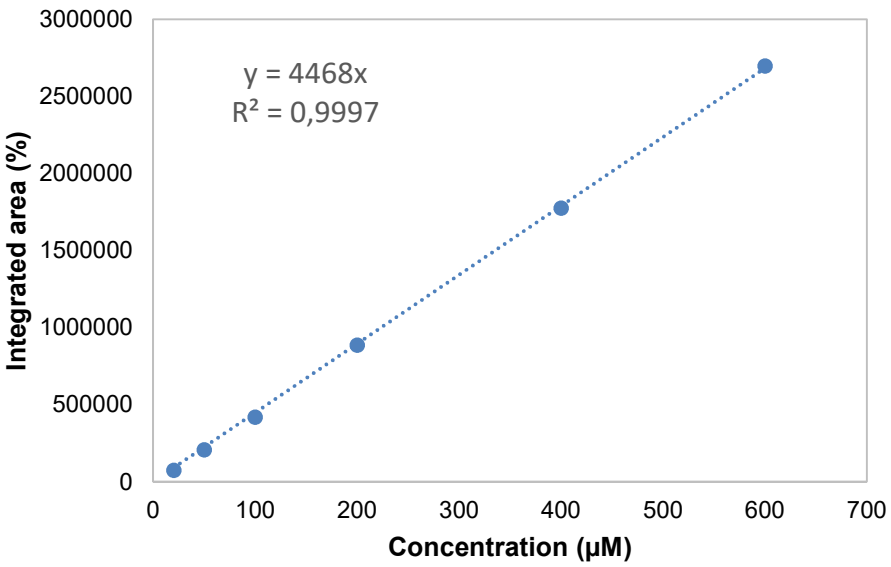
Supplementary Figure 39. SDS polyacrylamide gel showing the Strep-tag purified C-terminal domain of ArnA. From left to right: L, molecular mass marker; FT, flow through during sample loading; 1, washing fraction; 2, elution fraction containing purified ArnA (molecular mass = 43 kDa).[30,31]



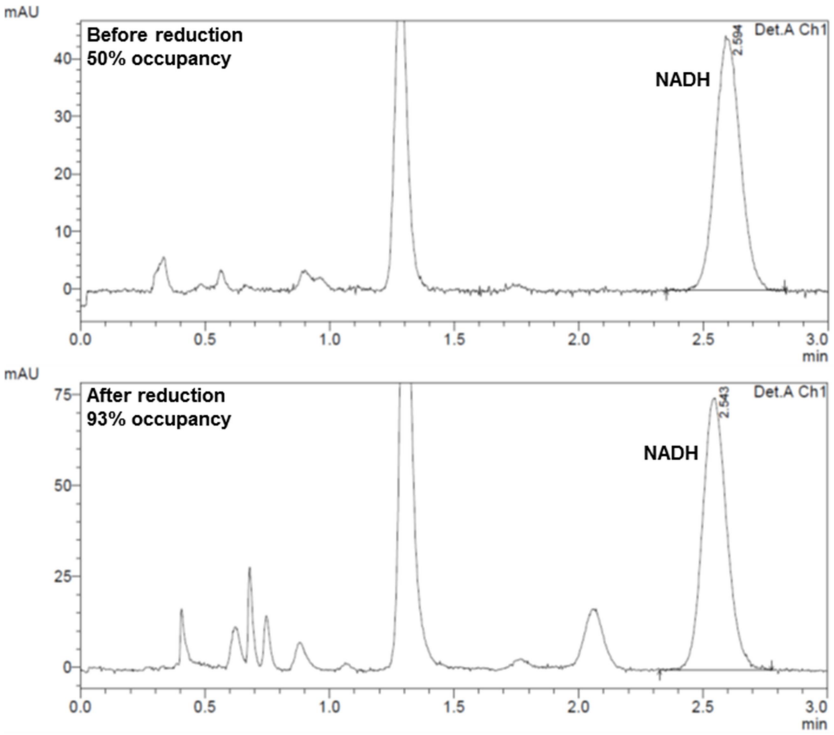
Supplementary Figure 40. HPLC chromatogram of the ArnA reaction mixture after 5 min. The peaks corresponding to the desired product UDP-4-keto-xylose (**5**, here named UDP-4-keto-Xyl) and residual substrate (UDP-GlcA, **2**) are labelled accordingly. The reaction for synthesis of UDP-4-keto-xylose (**5**) contained 2 mM UDP-GlcA (**2**), 2 mM NAD^+, and 5 mg/mL purified ArnA. The sample was incubated for 5 min at 30 °C (no agitation; 70% conversion, >99% selectivity for **5**).



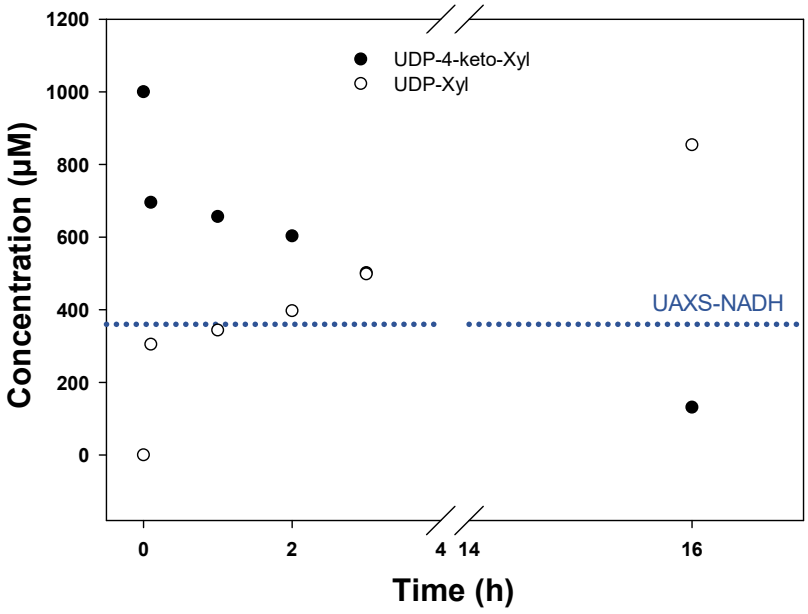
Supplementary Figure 41. ^{1}H-NMR spectrum (300 MHz, deuterium oxide) of isolated UDP-4-keto-xylose (**5**). Signals from the protons attached to respective carbons (C1-C5) in the pyranose ring are labelled accordingly. δ 5.49 (dd, J = 6.8, 3.5 Hz, 1H), 3.83 (d, J = 12.1 Hz, 1H), 3.72 (d, J = 9.9 Hz, 1H), 3.57 (dd, J = 8.1, 5.0 Hz, 1H), 3.46 (d, J = 12.2 Hz, 1H).



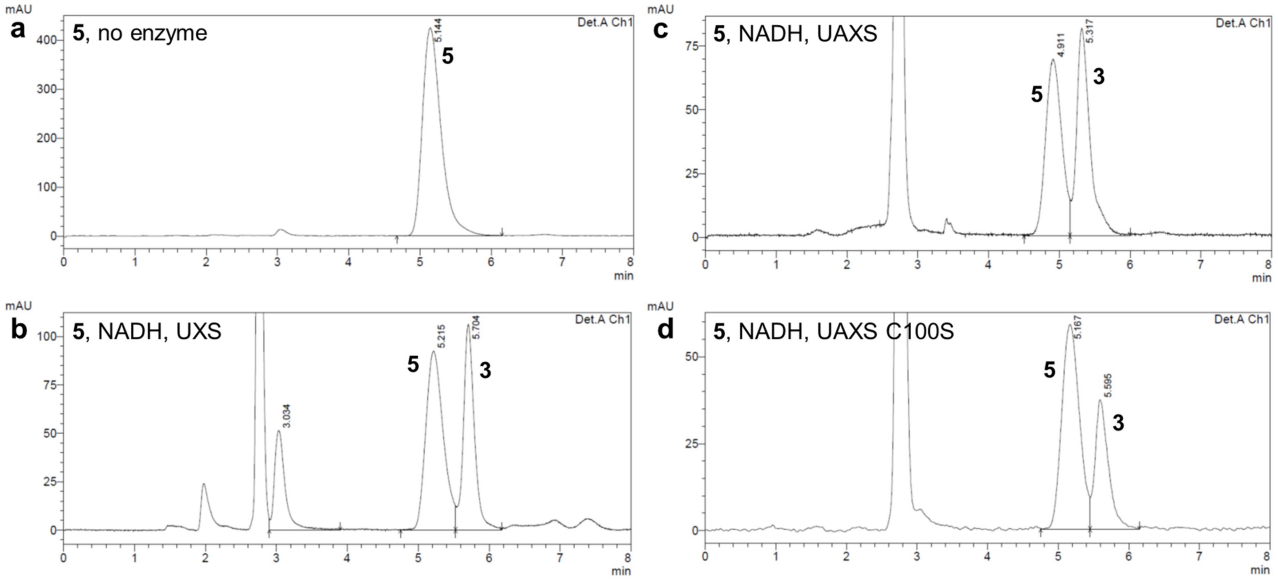
Supplementary Figure 42. Calibration curve for free NADH. Data points were recorded by HPLC-UV.



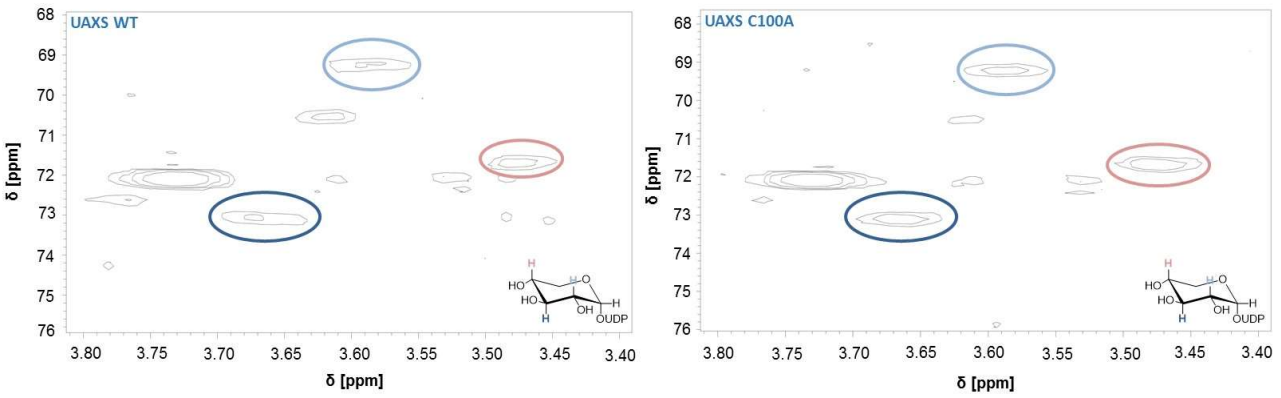
Supplementary Figure 43. HPLC traces from the chemical ($NaBH_4$) reduction of NAD^+ bound to purified wild-type UAXS. The top panel shows a HPLC chromatogram of the as-isolated enzyme with 50% occupancy by NADH. The bottom panel shows NADH released from the active site (93% occupancy with NADH) after the reduction with $NaBH_4$.



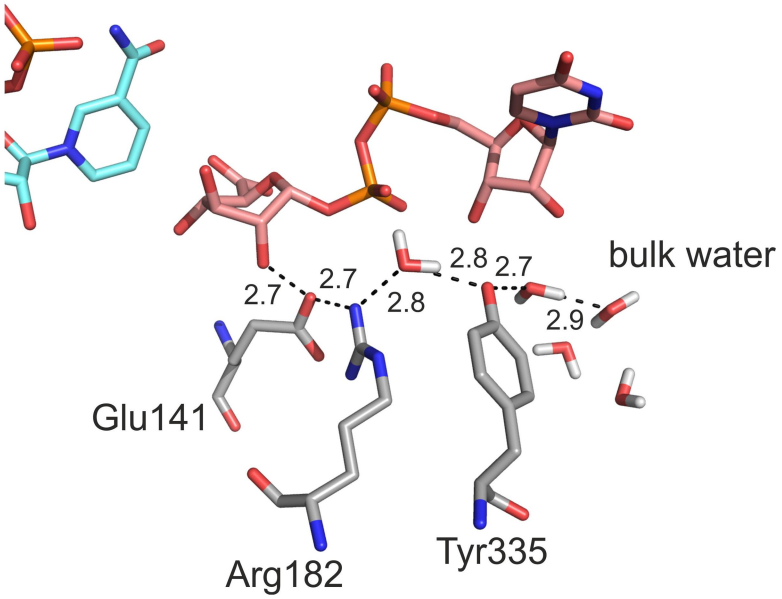
Supplementary Figure 44. Time course for the conversion of 1 mM UDP-4-keto-xylose (**5**) with chemically reduced UAXS (20 mg/mL; NADH content >90%) and excess of NADH (10 mM). The blue dotted line indicates the concentration of enzyme-bound NADH in UAXS.



Supplementary Figure 45. Additional HPLC traces for UDP-4-keto-xylose (**5**) conversions with wild-type UAXS, human UXS, and C100S variant. **a**. Control without enzyme. **b**. Reaction of 1 mM UDP-4-keto-xylose (**5**) with human UXS (5 mg/mL) and 10 mM NADH recorded after 180 min. **c**. Reaction of 1 mM UDP-4-keto-xylose (**5**) with wild-type UAXS (20 mg/mL) and 10 mM NADH recorded after 180 min. **d**. Reaction of 1 mM UDP-4-keto-xylose (**5**) with C100S variant (15 mg/mL) and 10 mM NADH recorded after 30 min.



Supplementary Figure 46. Deuterium wash-in experiment to study potential ring-opening during reduction of UDP-4-keto-xylose (**5**) to UDP-xylose (**3**) using wild-type UAXS (left panel) and C100A variant (right panel). Results show snapshots of heteronuclear single quantum coherence (HSQC) experiments analyzing proton signals from UDP-xylose (**3**). Relevant hydrogens at carbons 2-4 are labelled with colored circles.



Supplementary Figure 47. A possible proton relay in the active site of UAXS is shown. The proton relay connects the C2'-OH of UDP-GlcA (**2**) to bulk solvent via UAXS residues Glu141 and Arg182, and Tyr335. The figure was prepared based on the Michaelis complex shown in Figure 3a,c of the main text.

Supplementary References

1 Eixelsberger, T. *et al.* Isotope probing of the UDP-apiose/UDP-xylose synthase reaction: Evidence of a mechanism via a coupled oxidation and aldol cleavage. *Angew. Chem. Int. Ed. Engl.* **56,** 2503-2507, (2017).

2 Kabsch, W. Xds. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66,** 125-132, (2010).

3 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **67,** 235-242, (2011).

4 McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658-674, (2007).

5 Stein, N. CHAINSAW: A program for mutating pdb files used as templates in molecular replacement. *J. Appl. Crystallogr.* **41,** 641-643, (2008).

6 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66,** 213-221, (2010).

7 Afonine, P. V. *et al.* FEM: feature-enhanced map. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **71,** 646-666, (2015).

8 Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60,** 2126-2132, (2004).

9 Dodson, E. J., Murshudov, G. N. & Vagin, A. A. Description of program using maximum likelihood residual for macromolecular refinement, illustrated by several examples. *Acta Crystallogr., Sect. A: Found. Adv.* **52,** C85-C85, (1996).

10 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66,** 12-21, (2010).

11 Salomon-Ferrer, R. *et al.* Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *J. Chem. Theory Comput.* **9,** 3878-3888, (2013).

12 Jorgensen, W. L. *et al.* Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79,** 926-935, (1983).

13 Kirschner, K. N. *et al.* GLYCAM06: A generalizable Biomolecular force field. Carbohydrates. *J. Comput. Chem.* **29,** 622-655, (2008).

14 Pettersen, E. F. *et al.* UCSF chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605-1612, (2004).

15 Eixelsberger, T. *et al.* Structure and mechanism of human UDP-xylose synthase: Evidence for a promoting role of sugar ring distortion in a three-step catalytic conversion of UDP-glucuronic acid. *J. Biol. Chem.* **287,** 31349-31358, (2012).

16 Fiser, A., Do, R. K. G. & Sali, A. Modeling of loops in protein structures. *Protein Sci.* **9,** 1753-1773, (2000).

17 Izaguirre, J. A., Catarello, D. P., Wozniak, J. M. & Skeel, R. D. Langevin stabilization of molecular dynamics. *J. Chem. Phys.* **114,** 2090-2098, (2001).

18 Berendsen, H. J. C. *et al.* Molecular-dynamics with coupling to an external bath. *J. Chem. Phys.* **81,** 3684-3690, (1984).

19 Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. Numerical-integration of cartesian equations of motion of a system with constraints - Molecular-dynamics of n-alkanes. *J. Comput. Phys.* **23,** 327-341, (1977).

20 Darden, T., York, D. & Pedersen, L. Particle mesh Ewald - an N.Log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98,** 10089-10092, (1993).

21 Warshel, A. & Levitt, M. Theoretical studies of enzymic reactions - Dielectric, electrostatic and steric stabilization of carbonium-ion in reaction of lysozyme. *J. Mol. Biol.* **103,** 227-249, (1976).

22 Sherwood, P. *et al.* QUASI: A general purpose implementation of the QM/MM approach and its application to problems in catalysis. *J. Mol. Struct.: THEOCHEM* **632,** 1-28, (2003).

23 Metz, S. *et al.* ChemShell-a modular software package for QM/MM simulations. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **4,** 101-110, (2014).

24 Ahlrichs, R. *et al.* Electronic-structure calculations on workstation computers - the program system turbomole. *Chem. Phys. Lett.* **162,** 165-169, (1989).

25 Smith, W. & Forester, T. R. DL_POLY_2.0: A general-purpose parallel molecular dynamics simulation package. *J. Mol. Graphics* **14,** 136-141, (1996).

26 Ma, N., Digman, M. A., Malacrida, L. & Gratton, E. Measurements of absolute concentrations of NADH in cells using the phasor FLIM method. *Biomed. Opt. Express* **7,** 2441-2452, (2016).

27 Guyett, P., Glushka, J., Gu, X. G. & Bar-Peled, M. Real-time NMR monitoring of intermediates and labile products of the bifunctional enzyme UDP-apiose/UDP-xylose synthase. *Carbohyd. Res.* **344,** 1072-1078, (2009).

28 Fujimori, T. *et al.* Practical preparation of UDP-apiose and its applications for studying apiosyltransferase. *Carbohydr. Res.* **477,** 20-25, (2019).

29 Eixelsberger, T. & Nidetzky, B. Enzymatic redox cascade for one-pot synthesis of uridine 5'-diphosphate xylose from uridine 5'-diphosphate glucose. *Adv. Synth. Catal.* **356,** 3575-3584, (2014).

30 Gatzeva-Topalova, P. Z., May, A. P. & Sousa, M. C. Crystal structure of *Escherichia coli* ArnA (PmrI) decarboxylase domain. A key enzyme for lipid A modification with 4-amino-4-deoxy-L-arabinose and polymyxin resistance. *Biochemistry* **43,** 13370-13379, (2004).
31 Polizzi, S. J. *et al.* Human UDP-alpha-D-xylose synthase and *Escherichia coli* ArnA conserve a conformational shunt that controls whether xylose or 4-keto-xylose is produced. *Biochemistry* **51,** 8844-8855, (2012).
32 Gatzeva-Topalova, P. Z., May, A. P. & Sousa, M. C. Structure and mechanism of ArnA: Conformational change implies ordered dehydrogenase mechanism in key enzyme for polymyxin resistance. *Structure* **13,** 929-942, (2005).

Supplementary Videos and Data Sets

Supporting Video S1: Motions of the UAXS active site during MD simulations (Replica 1). Representative fragment (1 ns; 100 frames of 1 ps/frame) taken from the trajectory of the MD simulation (total MD simulation for 300 ns) of the proposed Michaelis complex (*Glu141 interacting with the C2'OH for ring-opening aldol cleavage*). The simulation started from the UAXS complex containing NAD^+ and UDP-GlcA. Tyr185 was deprotonated. In the shown video, hydrogen atoms were omitted for clarity, except those involved in relevant hydrogen bonding interactions of UAXS and UDP-GlcA (here: Glu141 and C2'OH; Tyr185 and C4'OH; Thr139 and C4'OH). Interactions of the aforementioned amino acid residues with UDP-GlcA are drawn as dashed lines from the H atom to the corresponding heteroatom. Dashed lines are only shown when the distance is <2.5 Å. The distance (in Å) between the C4'H of UDP-GlcA and C4 of NAD^+ is constantly measured and visualized (dashed line and distance in number) throughout the simulation.

Supporting Video S2: Motions of the UAXS active site during MD simulations (Replica 2). Representative fragment (1 ns; 100 frames of 1 ps/frame) taken from the trajectory of the MD simulation (total MD simulation for 300 ns) showing an alternative (UXS-like) UAXS complex in which the interaction of Glu141 with the C2'OH is lacking. The second replica was chosen according to the heat and density relaxations from the first simulation. In the shown video, hydrogen atoms were omitted for clarity, except those involved in relevant hydrogen bonding interactions of UAXS and UDP-GlcA (here: Tyr105 and C2'OH; Tyr185 and C4'OH; Thr139 and C4'OH). Interactions of the aforementioned amino acid residues with UDP-GlcA are drawn as dashed lines from the H atom to the corresponding heteroatom. Dashed lines are only shown when the distance is <2.5 Å. The distance (in Å) between the C4'H of UDP-GlcA and C4 of NAD^+ is constantly measured and visualized (dashed line and distance in number) throughout the simulation.

Supplementary data for computational work. The supplementary data from computational work contains ten coordinate files. The details of these files are as below:

1. The file ‘initial-coordinate-rep1.pdb’ represents the initial coordinate for the Replica 1. Since we have performed two different replicas with different initial velocities, initial coordinates for replica 1 and replica 2 are the same.

2. The file ‘initial-coordinate-rep2.pdb’ represents the initial coordinate for the Replica 2. Initial coordinates for replica 2 is similar to replica 1.

3. The file ‘final-coordinate-rep1.pdb’ represents the final coordinates after 300 ns of MD simulations for replica 1.

4. The file ‘final-coordinate-rep2.pdb’ represents the final coordinates after 300 ns of MD simulations for replica 2.

5. The file ‘RC-OPT-QM-region-rep1.xyz’ is the QM/MM optimized geometry of the QM-region of the QM/MM calculation in the XYZ file format for a snapshot taken from the MD of replica 1.

6. The file ‘RC-OPT-QM-region-rep2.xyz’ is the QM/MM optimized geometry of the QM-region of the QM/MM calculation in the XYZ file format, the snapshot is taken from the MD for replica 2.

7. The file ‘RC-OPT-rep1-QMMM.pdb’ is the QM/MM optimized geometry of the entire complex including the QM and MM region from the QM/MM calculation in the XYZ file format, the snapshot was taken from the MD of replica 1.

8. The file ‘RC-OPT-rep2-QMMM.pdb’ is the QM/MM optimized geometry of the entire complex including the QM and MM region of the QM/MM calculation in the XYZ file format, the snapshot was taken from the MD of replica 2.

9. The file ‘snap-populated-replica1.pdb’ is a snapshot from replica 1 used for QM/MM calculations, it represents the Michaelis complex before QM/MM optimization.

10. The file ‘snap-populated-replica2.pdb’ is a snapshot from replica 2 used for QM/MM calculations, it represents the complex before QM/MM optimization.