SUPPLEMENTAL METHODS

Expression and purification of hUXS1—For crystallization experiments, a synthetic hUXS1 gene (ORF NM_025076) that was codon-optimized for high-level expression in *E. coli* was inserted into the pET-derived expression vector p11. The expression plasmid encoded a truncated hUXS1 (AA 85-402) fused to a 23 amino acid-long N-terminal peptide that consisted of the His_6 tag and a tobacco etch virus (TEV) protease cleavage site. The Met auxotroph *E. coli* strain B834(DE3) cells were transformed with the hUXS1 plasmid and a single colony was used to start a 5mL overnight culture in LB medium supplemented with 100 μg/mL ampicillin. The overnight culture (1 mL) was used to inoculate 50 mL of LB/AMP medium and grown at 37 °C to an OD_{600} of 1.0. The cells were harvested by centrifugation, washed four times and inoculated in 1 L of SelenoMethionine Medium Base plus Nutrient Mix containing Se-Met at 40 mg/L (Molecular Dimensions, Newmarket, U.K.), and grown at 37 °C to an OD_{600} of 0.6. The culture was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM, cultured overnight at 25 °C, and collected by centrifugation. Pellets were resuspended in 20 mL lysis buffer (5 mM Imidazole, 500 mM NaCl, 50 mM HEPES, pH 7.5, 5% glycerol) including Protease inhibitor (complete; Roche, Basel, Switzerland), lysed by French Press, and the solution was centrifuged (30 min, 4 °C, 20000 x g). The supernatant was loaded onto a 2 mL gravity flow Ni-NTA column (Sigma-Aldrich, St. Louis, MO, U.S.A.), washed with 10 column volumes lysis buffer containing 30 mM imidazole and eluted in lysis buffer plus 250 mM imidazole. The protein was further purified by gel filtration using a Superdex S200 column (GE Healthcare, Little Chalfont, U.K.) developed with 10 mM HEPES, pH 7.5, 500 mM NaCl, 5% glycerol, 1 mM TCEP.

Wild-type hUXS1 for kinetic analysis was prepared in an analogous manner, except that the protein was produced in the *E. coli* strain Rosetta 2(DE3). Cells were grown in LB medium supplemented with 34 μg/mL chloramphenicol and 100 μg/mL ampicillin. For mutants of hUXS1, the pNIC28-Bsa4 expression vector (GenBank ID: EF198106) was used and ampicillin was replaced by kanamycin (50 μg/mL). Bacterial cultures (250 mL) were inoculated from an overnight pre-culture and first grown at 37 °C to an OD_{600} of ~0.8. After cooling to 18 °C, recombinant protein production took place overnight using induction with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Cells were recovered by centrifugation (20 min, 4 °C, 4400 x g), and the pellet was resuspended 1:1 in 50mM Tris/HCl buffer (pH 7.5) containing glycerol (5%, by volume). The cell extract obtained by high-pressure cell disruption (2 passes at 1500 psi) was fractionated on a Cu^{2+} loaded IMAC Sepharose High Performance column (GE Healthcare) using stepwise elution with imidazole. hUXS1 eluted at a concentration of 120 mM imidazole, and, after pooling the fractions containing UXS1 activity, imidazole was removed with a NAP-25 desalting column (GE Healthcare) or Vivaspin-20 centrifugal concentrator (Sartorius Stedim, Göttingen, Germany). Preparations of hUXS1 were brought to a concentration of 15-20 mg/mL and stored at -70 °C in the above Tris/HCl buffer supplemented with 1 mM DTT. The protein concentration was obtained from absorbance at 280 nm,

applying a molar extinction coefficient of 37360 $M^{-1}cm^{-1}$. Purity of wild-type and mutated enzymes was checked by SDS PAGE.

Force Field Test—GAFF is a widely used and well established force field for organic compounds. However, it has not been tested extensively for sugar molecules. Therefore, we performed test calculations comparing structures of the sugar moieties in the substrate and intermediate-1 compounds minimized at the DFT-level of theory to the same compounds modeled with Gromacs and the GAFF force field parameters. Starting from the relaxed chair conformation the substrate sugar was minimized using GAFF and Gromacs with a quasi-Newtonian minimization algorithm (1). For intermediate-1, the average structure from the MD simulation was used as conformation representing the classical force field. Both structures were then optimized using GAMESS-US (2). Here the DFT level of theory was combined with the B3LYP functional and a 6-31+G(d,p) basis set. The structures minimized with DFT and modeled with the classical force field are very close with an RMSD, including all non-hydrogen atoms, of 0.25 Å and 0.30 Å for the substrate and intermediate-1, respectively. The fact that the DFT optimized structure of intermediate-1 is very close to the average structure from the MD simulation suggests that this conformation corresponds to local minimum of the sugar ring. These two calculations can of course not serve as a comprehensive test of GAFFs ability to model sugars, but doing so would be beyond the scope of this work. Given these two results together with the extensive effort put into the development of GAFF as a force field for general organic molecules (3), we conclude that the force fields accuracy is sufficient for our purposes.

HPLC analysis—Samples were analyzed on an Agilent 1200 HPLC system equipped with a 5µm Zorbax SAX Analytical HPLC Column (4.6 × 250 mm; Agilent, Santa Clara, CA, U.S.A.) and a UV detector (λ = 254 nm). After 5 min elution with 5 mM potassium phosphate buffer (pH 3.2), a linear gradient of potassium phosphate buffer (pH 3.2) between 5 mM and 360 mM was used over 20 min. The column was washed (15 min each) with 600 mM and 5 mM potassium phosphate buffer (pH 3.2) after each analysis. The flow rate was 1.5 mL/min. Authentic standards were used for calibration. Note that the applied conditions were carefully optimized for reaction and analysis. Under the conditions used, UDP-GlcUA, UDP-xylose, and UDP-4-keto-pentose were baseline separated (Fig. S5).

In situ NMR analysis—Samples contained 1 mM UDP-GlcUA, 0.5 mM NAD^+ and 7 µM enzyme in case of the wild type. For reactions with mutants, 1 mM UDP-GlcUA, 1 mM NAD^+ and 15 µM enzyme (E120A) or 10 mM UDP-GlcUA, 1 mM NAD^+ and 100 µM enzyme (Y147F, R277Q) were used. All experiments were made in a 20 mM Tris/HCl buffer prepared in D_2O (0.70 mL, 99.9 % deuterium, pD 7.9). The ^1H NMR spectra were recorded at 303.2 K with presaturation (1 s) and acquisition of 32k data points. After zero filling to 64k data points, spectra were performed with a range of 7200 Hz.

Non-denaturing anionic PAGE—Protein samples were mixed with a 1:1 mixture of glycerol and 20mM Tris/HCl buffer (pH 7.4), containing 1% Triton X-100 and 0.9 g/L bromophenol blue. Separation was done on a 10% polyacrylamide gel according to the instructions of the supplier (Bio-

Rad Laboratories, Hercules, CA, U.S.A.), using a Tris-glycine native running buffer (24 mM Tris and 188 mM glycine) and a voltage of 100 V. Buffers and gel were prepared without SDS. After electrophoresis, the proteins were visualized by Coomassie blue staining.

References

1. Shanno, D. F. (1970) Conditioning of quasi-Newton methods for function minimization. *Math. Comp.* **24**, 647-656
2. Gordon, M. S., and Schmidt, M. W. (2005) Advances in electronic structure theory: GAMESS a decade later. in *Theory and Applications of Computational Chemistry: the first forty years* (C. E. Dykstra, Frenking, G., Kim, K. S., and Scuseria, G. E. eds.), Elsevier, Amsterdam. pp 1167-1189
3. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157-1174

SUPPLEMENTAL FIGURE LEGENDS

FIGURE S1. Protein sequence alignment of human UXS and similar proteins from other species. Protein sequences of *Homo sapiens* UDP-glucuronic acid decarboxylase 1 (UXS_human; NP_079352.2), *Oryza sativa subs. japonica* UDP-glucuronic acid decarboxylase (UXS_rize; Q75PK6), *Cryptococcus neoformans var. neoformans JEC21* UDP-glucuronic acid decarboxylase (UXS_fungus; XP_572003.1), *Escherichia coli K12* bifunctional polymyxin resistance protein ArnA (ArnA; P77398.1, AA 151-693 shown), *Solanum tuberosum* UDP-apiose/xylose synthase (UAXS; ABC75032.1), and *Bacillus cytotoxicus NVH 391-98* UDP-glucuronic acid 4-epimerase (UGlcAE; HM581979.1) were aligned with PRALINE using the BLOSUM62 scoring matrix. Secondary structure motifs in UXS_human were obtained with STRIDE from PDB entry 2B69. Residues of the catalytic triad are marked with black boxes in the numbering row; other residues of the active site are marked with grey boxes in the numbering row.

FIGURE S2. Biochemical characterization of hUXS1. A. SDS-PAGE used to determine the molecular mass of the monomeric form of hUXS1. 10 µg of purified hUXS1 were loaded on the gel and visualized by Coomassie blue staining after electrophoresis. The standard used (left lane) is a low molecular weight marker (masses in kDa) from GE Healthcare (Little Chalfont, U.K.). B**.** Native gel electrophoresis. 6 µg aliquots of purified hUXS1 were loaded on the gel. Lane 1 and 2 are wild-type hUXS1; lane 3 is a high molecular weight marker for native gel electrophoresis (GE Healthcare; masses in kDa). Although the exact molecular mass cannot be determined by native PAGE (4), the position of the hUXS1 bands (between the markers for 66 kDa and 140 kDa) indicates the presence of a dimer (77.2 kDa). C. Evidence for partial saturation of hUXS1 with NAD^+ after isolation. NAD^+ is very tightly bound by the enzyme (K_m = 4 µM), therefore hUXS1 shows about 60% of its maximum activity under conditions where no external NAD^+ is added (2 mM UDP-GlcUA; 10 µM hUXS1).

FIGURE S3. MD simulations. A. Convergence of MD simulations, shown as RMSD between all protein non-hydrogen atoms in the simulation and in the initial structure (based on coordinates in PDB 2B69). B. Hydrogen bonds between the substrate sugar moiety and the protein. Donors and acceptors of the protein are drawn on the x-axes, sugar oxygens on the y-axes. For each combination, the relative time during which an H-bond is established between the two atoms is indicated by a square in gray-scale: black 100% simulation time, white 0% simulation time.

FIGURE S4. Surface model of hUXS1 with substrate and NAD^+ bound and comparison with ArnA. A and B. Front view of hUXS1 (A: closed, B: in section), showing the accommodation of UDP-glucuronic acid (green) and NAD^+ (yellow) in the active site. C and D. Two back views of the UXS active site, C: long shot, D: close up. For residues E120 and R277 the protein surface is not shown to

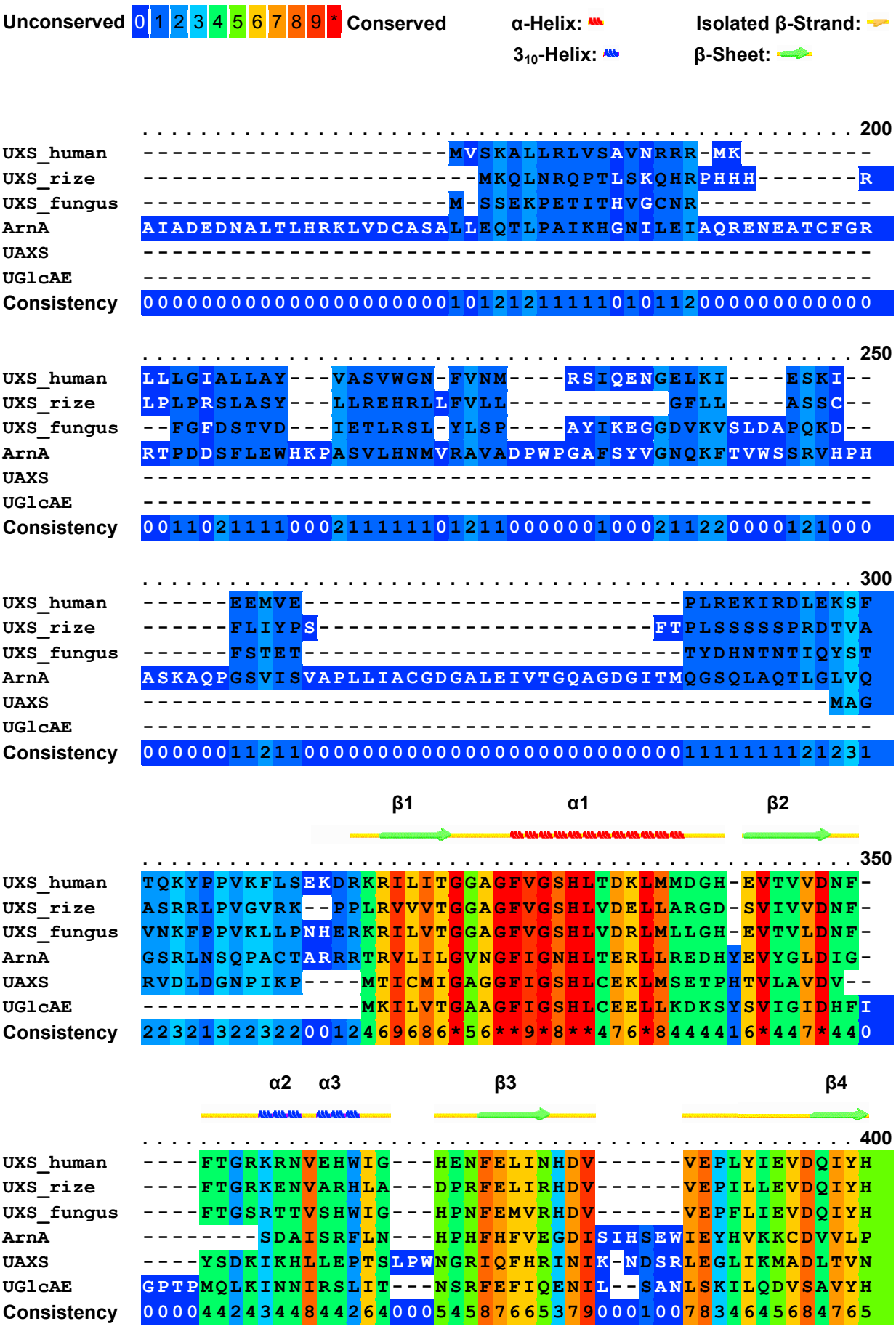
demonstrate that removal of the two side chains can potentially open a channel allowing access to the active site; also shown are the substrate (green) and NAD^+ (yellow). E. Overlay of hUXS1 and ArnA structures, showing the high similarity between the two enzymes. F. Accommodation of UDP-GlcUA in the active site of ArnA. In contrast to hUXS1, the sugar ring is in the undistorted $^4C_1$ *chair* conformation.

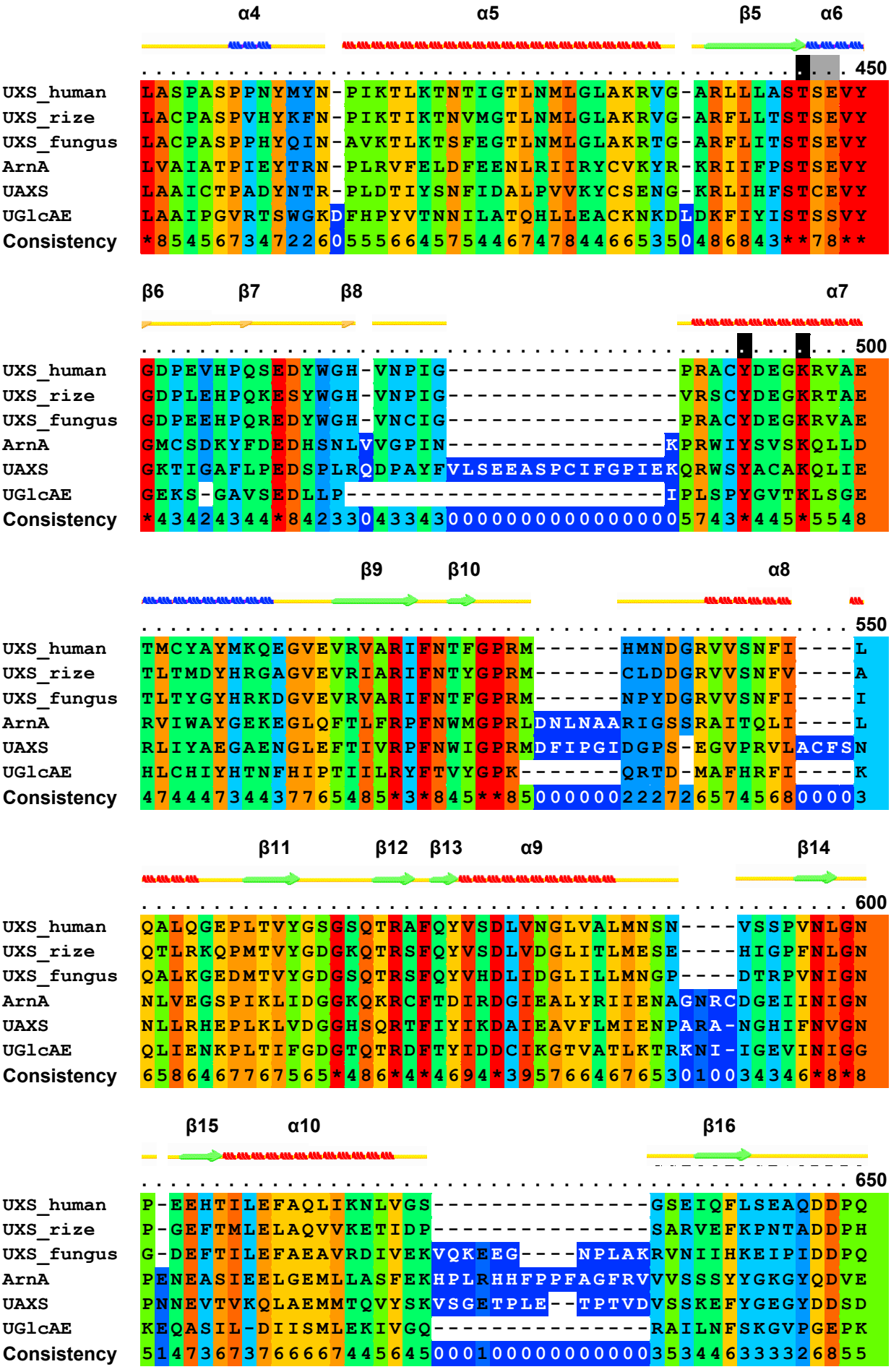
FIGURE S5. HPLC chromatogram showing the separation of UDP-glucuronic acid, UDP-xylose and NAD^+ signals (reaction of wild-type hUXS1). The insert shows an overlay with a chromatogram of Y147F reaction, illustrating separation of UDP-xylose and UDP-4-keto-pentose.

References

4. Razandi, M., Pedram, A., Merchenthaler, I., Greene, G. L., and Levin, E. R. (2004) Plasma membrane estrogen receptors exist and function as dimers. *Mol. Endocrinol.* **18**, 2854-2865

Figure S1


Unconserved 0 1 2 3 4 5 6 7 8 9 * Conserved
α-Helix:
Isolated β-Strand:
3₁₀-Helix:
β-Sheet:
UXS_human
UXS_rize
UXS_fungus
ArnA
UAXS
UGlcAE
Consistency
β1
α1
β2
α2
α3
β3
β4


α4 α5 β5 α6
450
UXS_human LASPASPPNYMYN-PIKTLKTNTIGTLNMLGLAKRVG-ARLLLASTSEVY
UXS_rize LACPASPVHYKFN-PIKTIKTNVMGTLNMLGLAKRVG-ARFLLTSTSEVY
UXS_fungus LACPASPPHYQIN-AVKTLKTSFEGTLNMLGLAKRTG-ARFLITSTSEVY
ArnA LVAIATPIEYTRN-PLRVFELDFEENLRIIRYCVKYR-KRIIFPSTSEVY
UAXS LAAICTPADYNTR-PLDTIYSNFIDALPVVKYCSENG-KRLIHFSTCEVY
UGlcAE LAAIPGVRTSWGKDFHPYVTNNILATQHLLEACKNKDLDKFIYISTSSVY
Consistency *8545673472260555664575446747844665350486843**78**
β6 β7 β8 α7
500
UXS_human GDPEVHPQSEDYWGH-VNPIG----------PRACYDEGKRVAE
UXS_rize GDPLEHPQKESYWGH-VNPIG----------VRSCYDEGKRTAE
UXS_fungus GDPEEHPQREDYWGH-VNCIG----------PRACYDEGKRVAE
ArnA GMCSDKYFDEDHSNLVVGPIN----------KPRWIYSVSKQLLD
UAXS GKTIGAFLPEDSPLRQDPAYFVLSEEASPCIFGPIEKQRWSYACAKQLIE
UGlcAE GEKS-GAVSEDLLP----------IPLSPYGVTKLSGE
Consistency *434243444*842330433430000000000000005743*445*5548
β9 β10 α8
550
UXS_human TMCYAYMKQEGVEVRVARIFNTFGPRM------HMNDGRVVSNFI----L
UXS_rize TLTMDYHRGAGVEVRIARIFNTYGPRM------CLDDGRVVSNFV----A
UXS_fungus TLTYGYHRKDGVEVRVARIFNTFGPRM------NPYDGRVVSNFI----I
ArnA RVIWAYGEKEGLQFTLFRPFNWMGPRLDNLNAARIGSSRAITQLI----L
UAXS RLIYAEGAENGLEFTIVRPFNWIGPRMDFIPGIDGPS-EGVPRVLACFSN
UGlcAE HLCHIYHTNFHIPTIILRYFTVYGPK-------QRTD-MAFHRFI----K
Consistency 474447344377654854*3*8455*85000000222726574568000003
β11 β12 β13 α9 β14
600
UXS_human QALQGEPLTVYGSGSQTRAFQYVSDLVNGLVALMNSN----VSSPVNLGN
UXS_rize QTLRKQPMTVYGDGKQTRSFQYVSDLVDGLITLMESE----HIGPFNLGN
UXS_fungus QALKGEDMTVYGDGSQTRSFQYVHDLIDGLILLMNGP----DTRPVNIGN
ArnA NLVEGSPIKLIDGGKQKRCFTDIRDGIEALYRIIENAGNRCDGEIINIGN
UAXS NLLRHEPLKLVDGGHSQRTFIYIKDAIEAVFLMIENPARA-NGHIFNVGN
UGlcAE QLIENKPLTIFGDGTQTRDFTYIDDCIKGTVATLKTRKNI-IGEVINIGG
Consistency 6586467767565*486*4*46945*39576646765301003434 6*8*8
β15 α10 β16
650
UXS_human P-EEHTILEFAQLIKNLVGS----------GSEIQFLSEAQDDPQ
UXS_rize P-GEFTMLELAQVVKETIDP----------SARVEFKPNTADDPH
UXS_fungus G-DEFTILEFAEAVRDIVEKVQKEEG----NPLAKRVNIIHKEIPIDDPQ
ArnA PENEASIEELGEMLLASFEKHPLRHHFPPFAGFRVVVSSSYYGKGYQDVE
UAXS PNNEVTVKQLAEMMTQVYSKVSGETPLE--TPTVDVSSKEFYGEGYDDSD
UGlcAE KEQASIL-DIISMLEKIVGQ----------RAILNFSKGVPGEPK
Consistency 514736737666674456450001000000000003534463333268 55

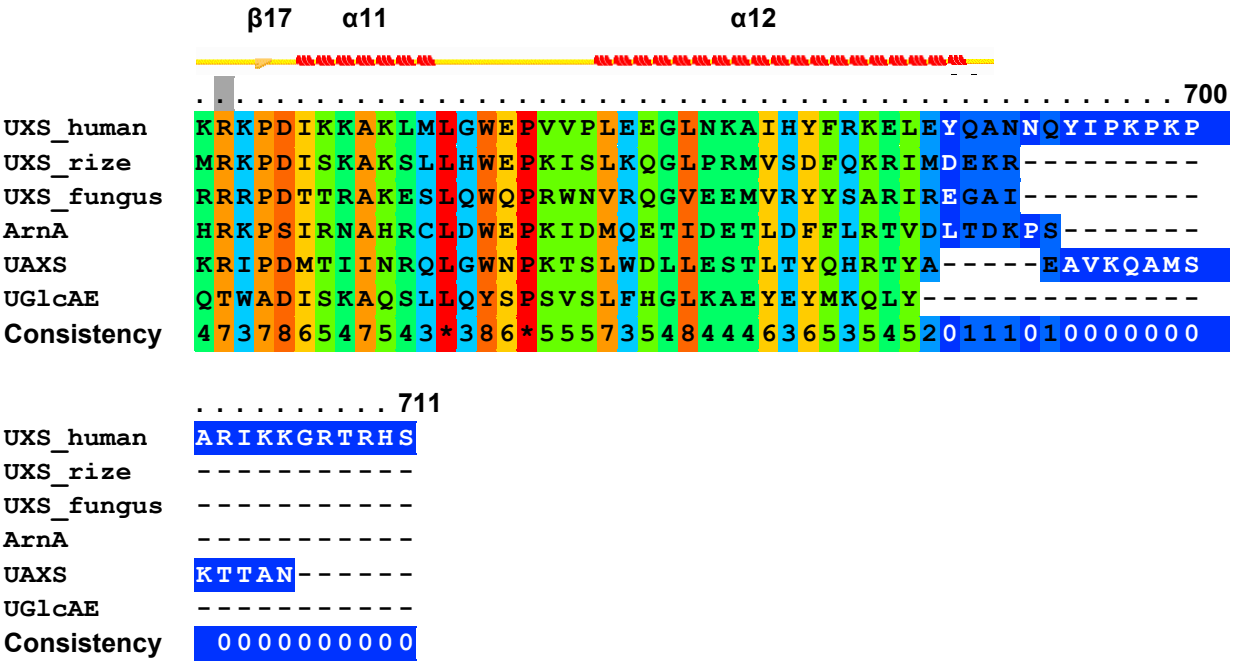

β17
α11
α12
700
UXS_human KRKPDIKKAKLMLGWEPVVPLEEGLNKAIHYFRKELEYQANNQYIPKPKP
UXS_rize MRKPDISKAKSLLHWEPKISLKQGLPRMVSDFQKRIMDEKR----------
UXS_fungus RRRPDTTRAKESLQWQPRWNVRQGVEEMVRYYSARIREGAI----------
ArnA HRKPSIRNAHRCLDWEPKIDMQETIDETLDFFLRTVDLTDKPS--------
UAXS KRIPDMTIINRQLGWNPKTSLWDLLESTLTYQHRTYA-----EAVKQAMS
UGlcAE QTWADISKAQSLLQYSPSVSLFHGLKAEYEYMKQLY----------
Consistency 473786547543*386*555735484446365354520111010000000
711
UXS_human ARIKKGRTRHS
UXS_rize ----------
UXS_fungus ----------
ArnA ----------
UAXS KTTAN------
UGlcAE ----------
Consistency 00000000000

Figure S2

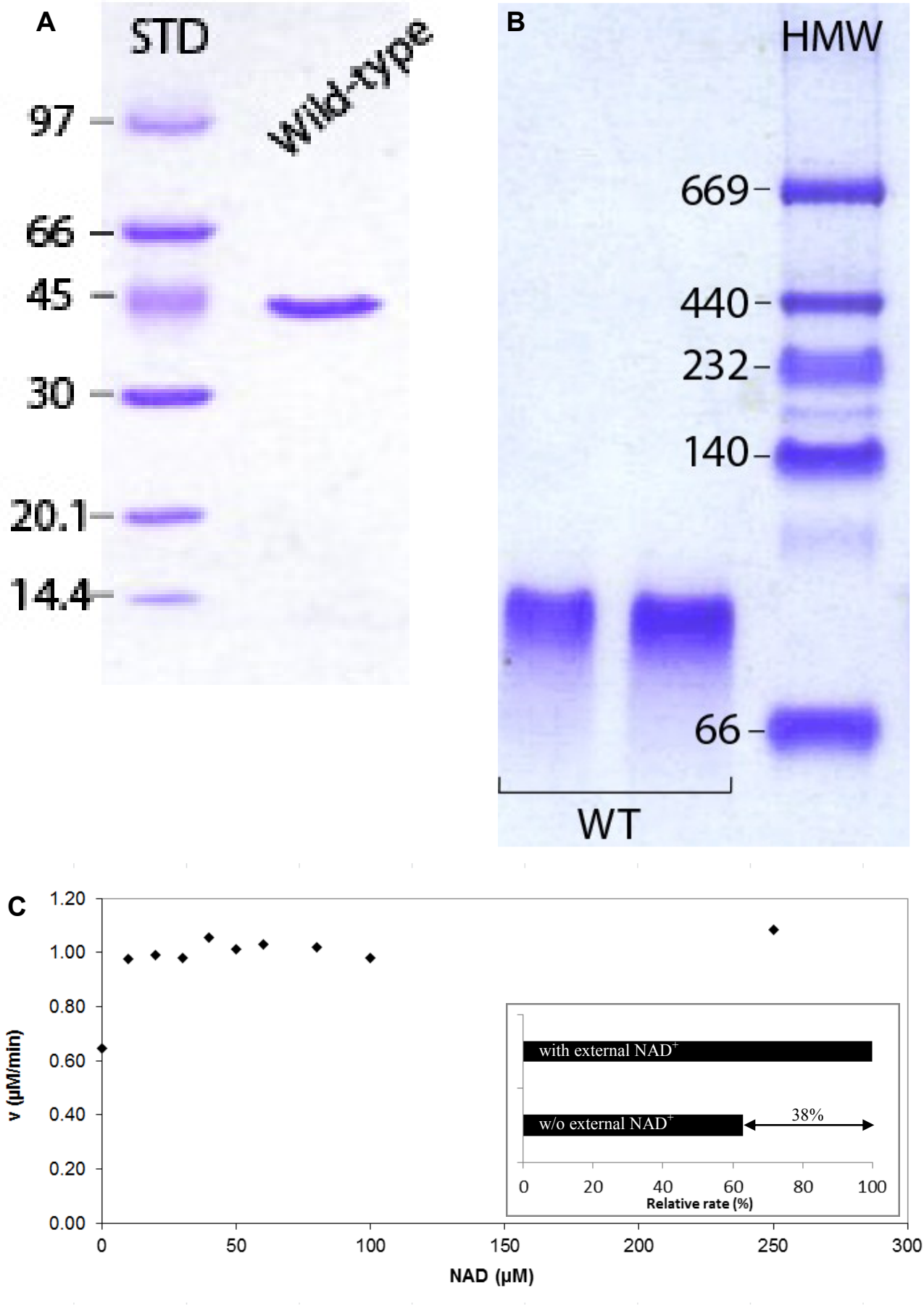
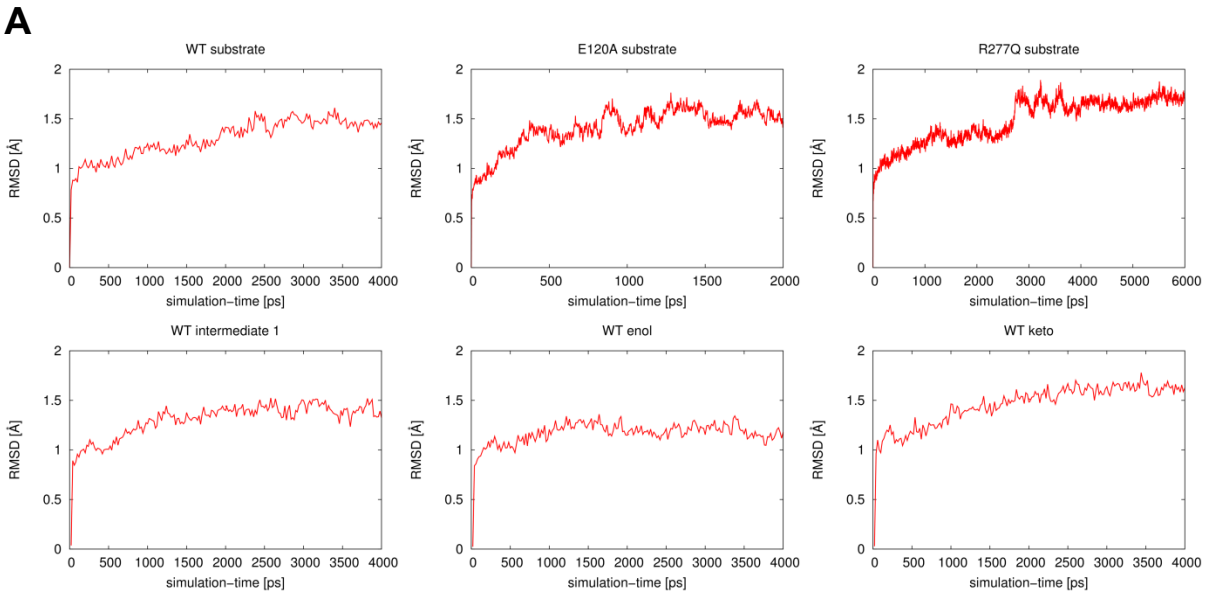

A
STD
Wild-type
97
66
45
30
20.1
14.4
B
HMW
669
440
232
140
66
WT
C
v (μM/min)
NAD (μM)
1.20
1.00
0.80
0.60
0.40
0.20
0.00
0
50
100
150
200
250
300
with external NAD+
w/o external NAD+
38%
0
20
40
60
80
100
Relative rate (%)

Figure S3


A
WT substrate
E120A substrate
R277Q substrate
WT intermediate 1
WT enol
WT keto
RMSD [Å]
simulation–time [ps]

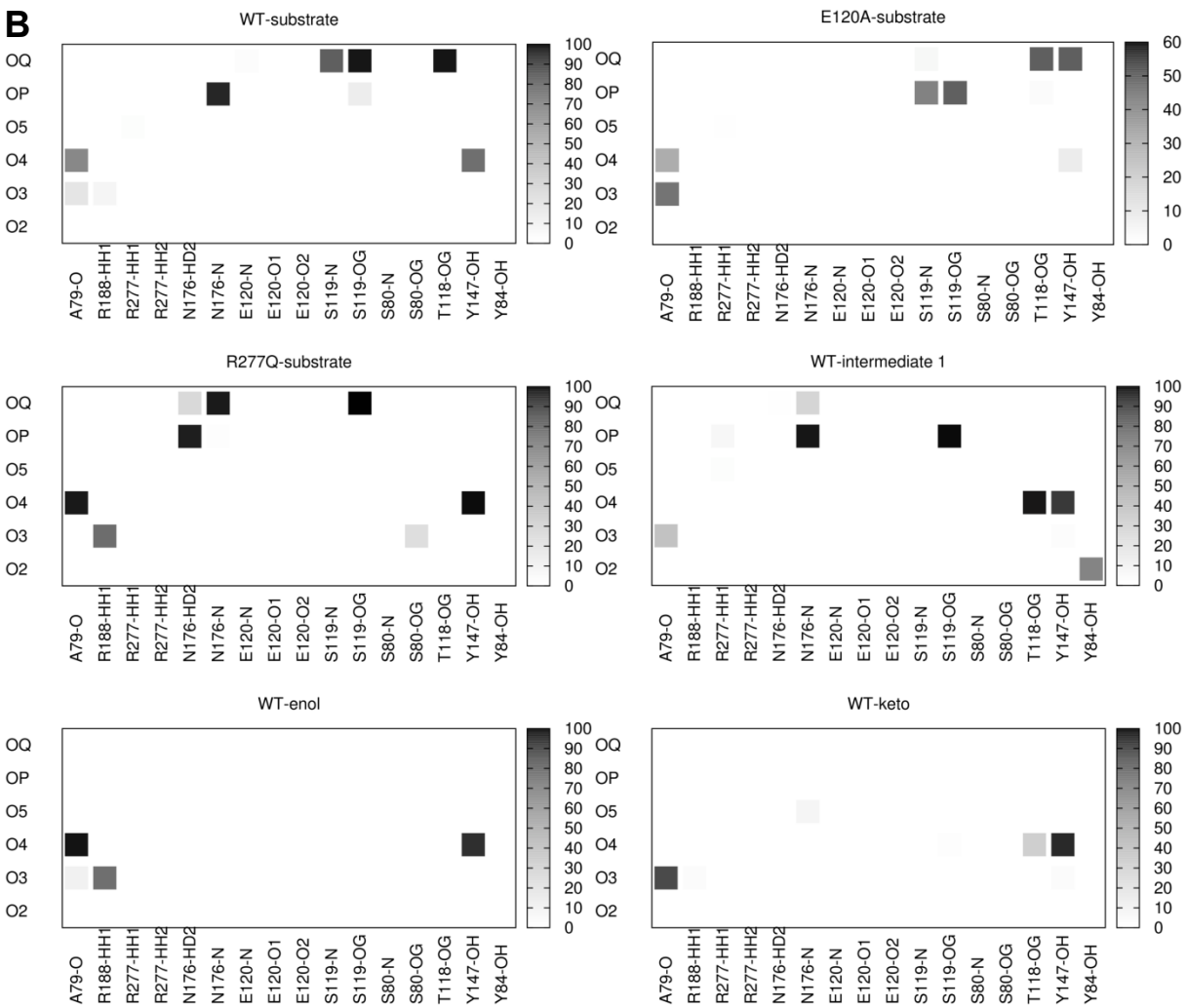

B
WT-substrate
E120A-substrate
R277Q-substrate
WT-intermediate 1
WT-enol
WT-keto
OQ
OP
O5
O4
O3
O2
A79-O
R188-HH1
R277-HH1
R277-HH2
N176-HD2
N176-N
E120-N
E120-O1
E120-O2
S119-N
S119-OG
S80-N
S80-OG
T118-OG
Y147-OH
Y84-OH

Figure S4

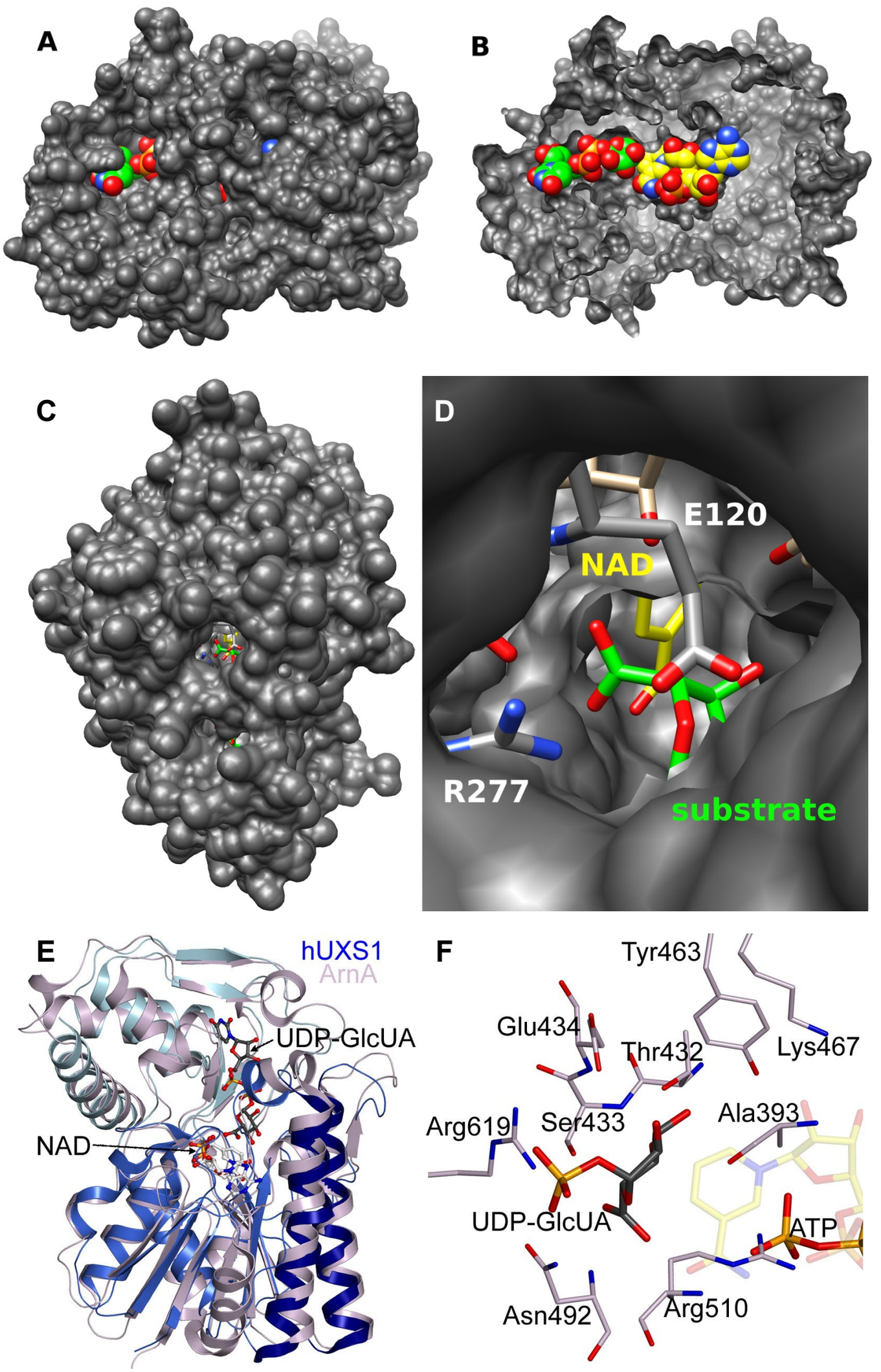
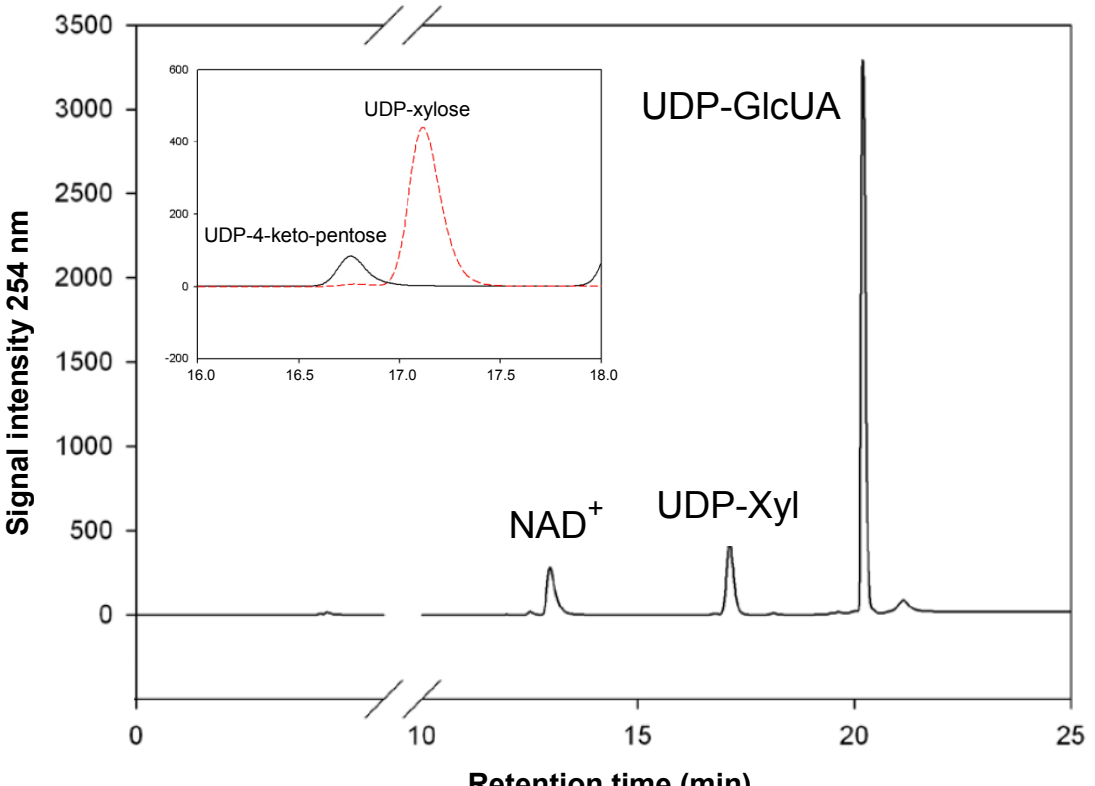

A
B
C
D
E120
NAD
R277
substrate
E
hUXS1
ArnA
UDP-GlcUA
NAD
F
Tyr463
Glu434
Thr432
Lys467
Arg619
Ser433
Ala393
UDP-GlcUA
ATP
Asn492
Arg510

Figure S5


Signal intensity 254 nm
Retention time (min)
UDP-GlcUA
UDP-Xyl
NAD+
UDP-xylose
UDP-4-keto-pentose