Supporting Information

Homology modeling of human UDP-glucuronosyltransferase 1A6 reveals insights into factors influencing substrate and co-substrate binding

*Alexander D. Smith, Brent D.G. Page, Abby C. Collier, and Michael W.H. Coughtrie**

Faculty of Pharmaceutical Sciences, The University of British Columbia, 2405 Wesbrook Mall, Vancouver, BC V6T 1Z3, CANADA

*E-mail: michael.coughtrie@ubc.ca.

Table of Contents

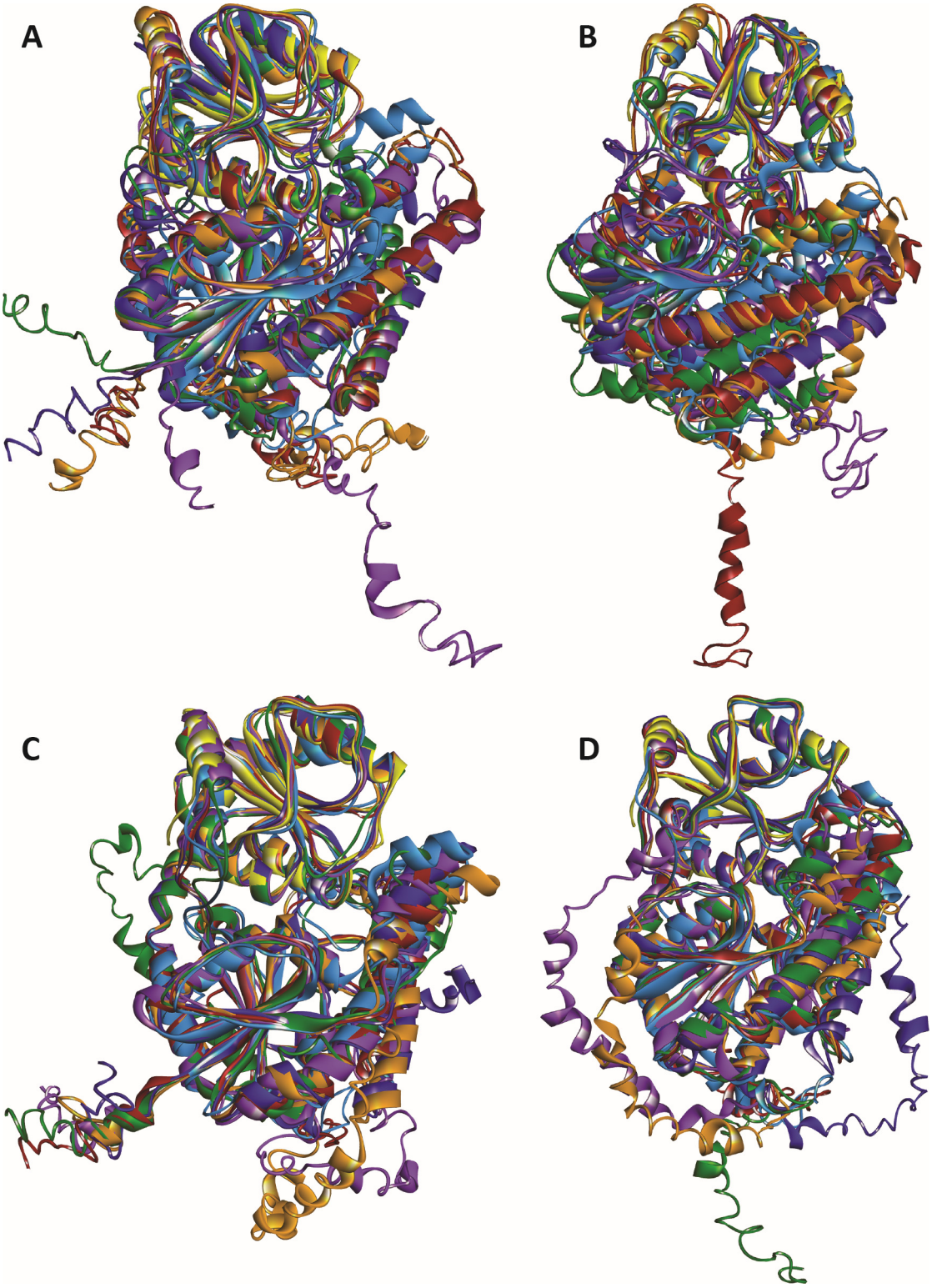


Figure S1. The 5 models produced by I-TASSER for each input condition are shown aligned with the UGT2B7 crystal structure (PDB 2O6L) in yellow, and the final UGT1A6 model in light blue. Model 1 is shown in red, model 2 in orange, model 3 in green, model 4 in blue, and model 5 in purple. A. Models generated with the signal peptide and without the UGT2B7 crystal structure as a guiding template (WSNT), B. Without the signal peptide, and without the UGT2B7 crystal structure (NSNT), C. With the signal peptide, and with the UGT2B7 crystal structure as a guide (WSWT), D. Without the signal peptide and with the UGT2B7 crystal structure as a guide (NSWT). Structural alignments in .mol2 format are in the "Supplemental structures" zip file (I-TASSER aligned models folder).

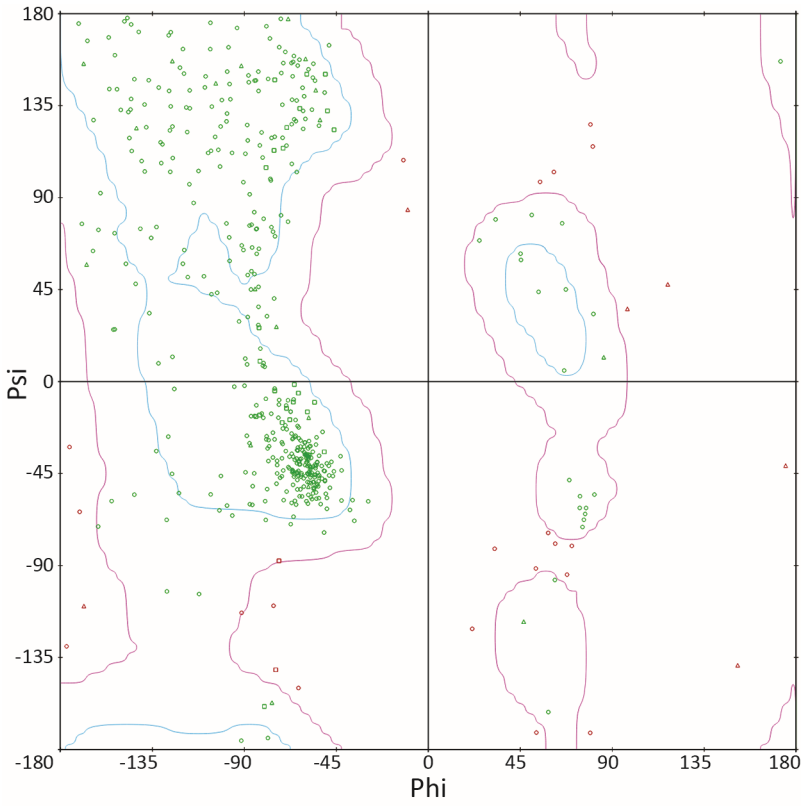


Figure S2. Ramachandran plot of the final UGT1A6 homology model.

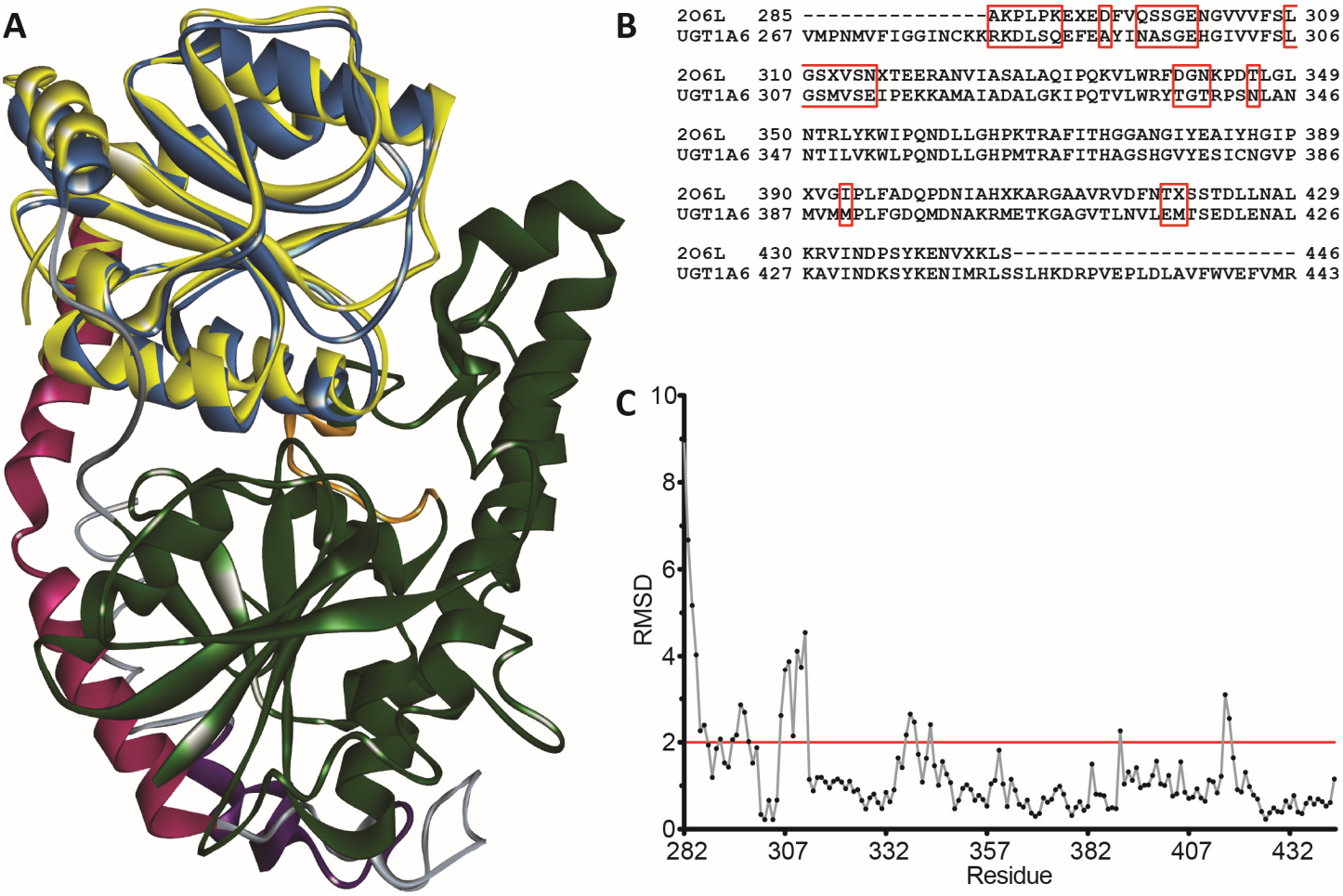
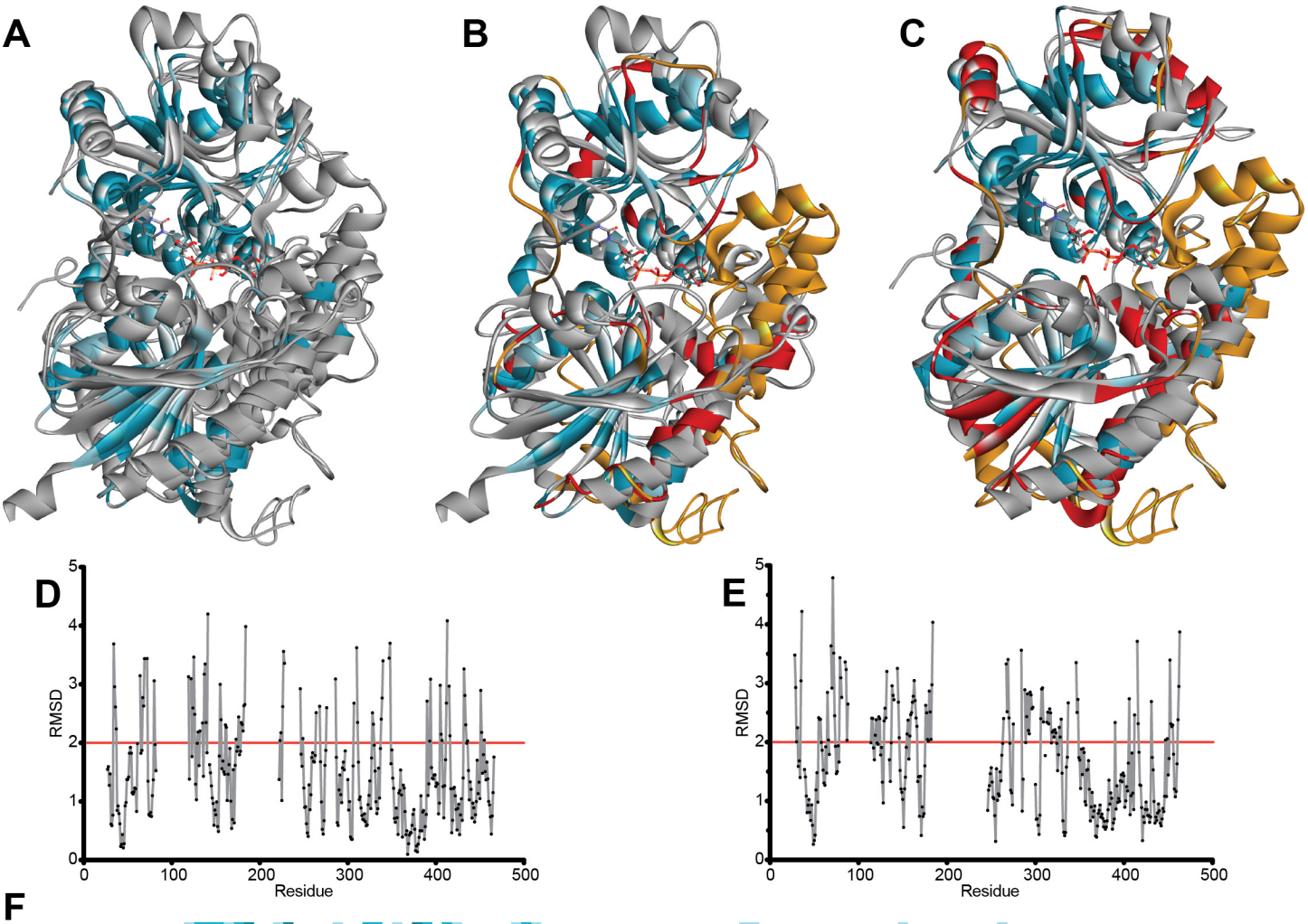


Figure S3. Alignment of the final UGT1A6 model with the template UGT2B7 crystal structure (2O6L). A. Final refined homology model of UGT1A6 with overlaid UGT2B7 crystal structure. Helices are shown as spirals, and β-sheets are shown as broad ribbons. The C-terminal domain (blue) is shown at the top, with the N-terminal domain (green) shown at the bottom of the image. The envelope helices (magenta) are shown to the left, with the transmembrane domain (purple) at the bottom. The putative dimerization domain (PDPVSYIPRCY) of UGT1A6 is highlighted in orange. The UGT2B7 crystal structure is shown in yellow. B. Protein sequence alignment of the aligned regions of the UGT1A6 model with UGT2B7 (2O6L) crystal structure. Residues with RMSDs differing by >2 Å are highlighted in red boxes. Amino acid 'X' in the 2O6L sequence correspond to selenomethionine. C. Graph of RMSD values over the length of the aligned proteins. The structural alignment can be found in the "supplemental structures" zip file, file UGT1A6_2B7_alignment.mol2.



F

```
UGT1A6 1   -------DKLLVVP-QDGSHWLSMKDIVEVLSDRG--HEIVVVVPEV-------NL---LLKESK-------YYTRKIYPVPYDQ---EELKNRYQSFGNNHFAERSFLTAPQTEYR---
2ACV   1   MSDINKNSELIFIPAPGIGHLASALEFAKLLTNHDKNLYITVFCI--KFPGMPFAD---SYIK--SVLASQPQIQLIDLPEVE----------------------------------
5TME   1   --------QVLLLP-FPA--TNPMLQFGRRLAYHG--LRPTLVTTRY-------VLSTTPP-PGD-------PFRVAAISDG---FDDASGM-------------------------AAL

UGT1A6 87  -NNMIV--------------IGLYFINCQSLLQDRDTLNFFKES-K-----FDALFTDPALPCGVILAEYLGLPSVYLFRGFPCSLEHTFSRSPDPVSYIPRCYTKFSDHMTFSQRVANF
2ACV   76  ------PPPQELLKSPEFYIIT-FLESLIPHVKA--TIKTILS-NK-----VVGLVLDFFCVSMIDVGNEFGIPSYLFLTSNVGFLSLMLSL----------------------------
5TME   65  PIPGEY--------------LRTLEAHGARTLAE--LLLSEA----RAGRPARVLVYDPHLPWARRVARAAGVATAAFLSQPCAVDLIYGEV----------------------------

UGT1A6 187 LVNLLEPYL-------------------------------------FY-CLFSKYEELASAVLKRDVDIIT-----------------------------------------------
2ACV   153 ---------KNRQIEEVFDDSDRDHQLLNIPGISNQVPSNVLPDACFNKDGGYIA-----------------------------------------------------------
5TME   136 ------------------------------------------------------------------CARRLALPVTPTDARGLYARGVLGVELGPDDVPPFVAAPELTPAFCEQ

UGT1A6 219 ------------LYQKVSVWLLRYDFVLEYPRP------V----------MPNMVFIGGINCKKRK-------------DL---------------SQEFEAYINASGEH----GIVV
2ACV   199 ------YYKLAERFRDTKGIIVNTFSDLEQS---------SIDALYDHDEKIPPIYAVGP---------------------LLDLKGQPNPKLDQAQHDLILKWLDEQP----DKSVVF
5TME   185 SIEQFA------GLEDDDDVLVNSFSDLLEPK--EAAYMES----------TWRAKTIG--------PSLPSFYLDDGRRST---------------V-PCMEWLDKQP--PR--SVVL

UGT1A6 278 FSLGSM--V-SEIPEKKAMAIADALGKIPQTVLWRYTGTRPSNLA---------------N--------------NTILVKWLPQNDLLGHPMTRAFITHAGSHGVYESICNGVPMVM
2ACV   279 LCFG--SMG-VSFGPSQIREIALGLKHSGVRFLWSNSAEK---------------------KVFPEGFLEWMELEGKGMICGWAPQVEVLAHKAIGGFVSHCGWNSILESMWFGVPILT
5TME   257 VSYGTV---ST-FDVAKLEEIGNGLCNSGKPFLWVVR--------SNEEHKLSVQLRKKCEK---------------RGLIVPFCPQLEVLAHKATGCFLSHCGWNSTLEAIVNGVPLVA

UGT1A6 364 MPLFGDQMDNAKRMET-K-GAGVTLNV----------LEMTSEDLENALKAVIN----DKSYKENIMRLSSLHKDRPV------EPLDLAVFWVEFVMRHKGAPHLRPAAHDLTWYQY
2ACV   374 WPIYAEQQLNAFRLVKEW-GVGLGLR------VDYRKGSDVVAAEEIEKGLKDLMD---KDSIVHKKVQEMKEMSRNAVV---DGGSSLISVGKLIDDITG----------------
5TME   350 MPHWADQPTISKYVES-LWGMGVRV--QLDKS-------GILQREEVERCIREVMDGDR-KEDYRRNATRLMKKAKESMQEGG---SSDKNIAEFAAKYS----------------

UGT1A6 460 HSLDVIGFLLAVVLTVAFITFKCCAYGYRKCLGKKGRVKKAHKSKTH
2ACV   461 ------------------------------------------------
5TME   435 ------------------------------------------------
```

Figure S4. Structural alignments of the final UGT1A6 model with A. both 2ACV and 5TME GT structures, B. 2ACV alone, C. 5TME alone. Structures are coloured according to amino acid similarity: dark teal, identical; blue, strong similarity; cyan, weak similarity. Residues in red have RMSDs >2 Å, and residues in orange failed to align. UDPGA is shown in its binding site. D. Graph of RMSD values over the length of the aligned UGT1A6 and 2ACV amino acids. E. Graph of RMSD values over the length of the aligned UGT1A6 and 5TME amino acids. F. Sequence alignments generated from the superimposed protein structures. Amino acid similarity shading is as in A, strongly similar amino acids: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, and FYW; weakly similar amino acids: CSA, ATV, SAG, STNK, STPA, SGND, SNDEQK, NDEQHK, NEQHRK, FVLIM, and HFY. The structural alignments between our UGT1A6 model and all GTs can be found in the "Supplemental structures" zip file, file UGT1A6_all_alignments.mol2.

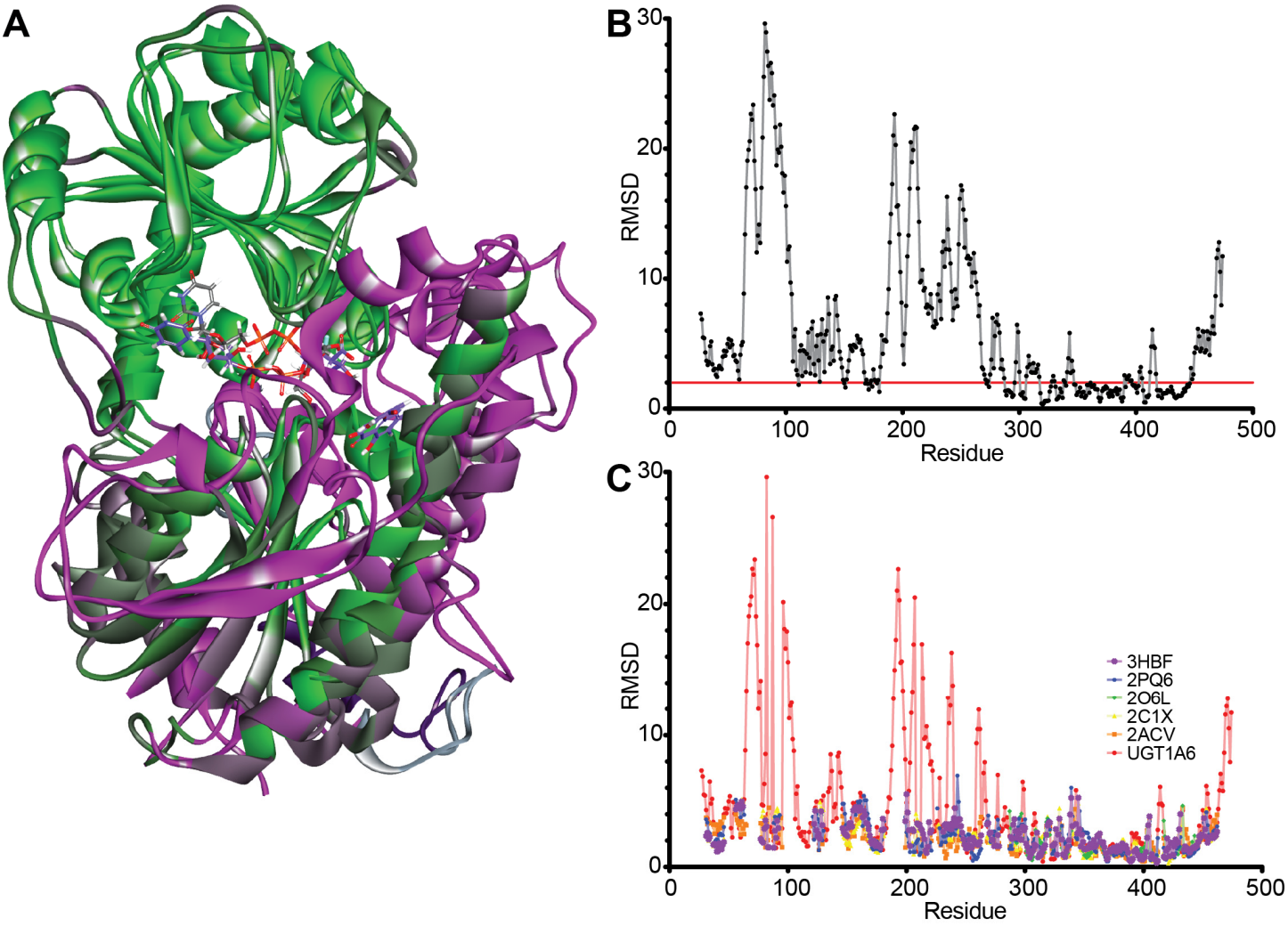


Figure S5. Comparison of our UGT1A6 model with that of Ghemtio *et al* (2014), which has been recapitulated from the coordinates supplied in that paper's supplementary information[1]. A. Structural overlay of both models, coloured by RMSD, with closely aligning regions in green, and more widely deviating regions in magenta. The co-substrates, UDPGA (our model) and U2F (Ghemtio *et al* model) are overlaid in the co-substrate binding region. Quercetin is shown in the substrate binding area in our model. B. Graph of RMSD values over the length of the aligned sequences. C. Graph of RMSD values for our UGT1A6 model, the Ghemtio *et al* model, and the 5 template structures used to create the Ghemtio *et al* model, over the length of alignment generated from the superimposed proteins.

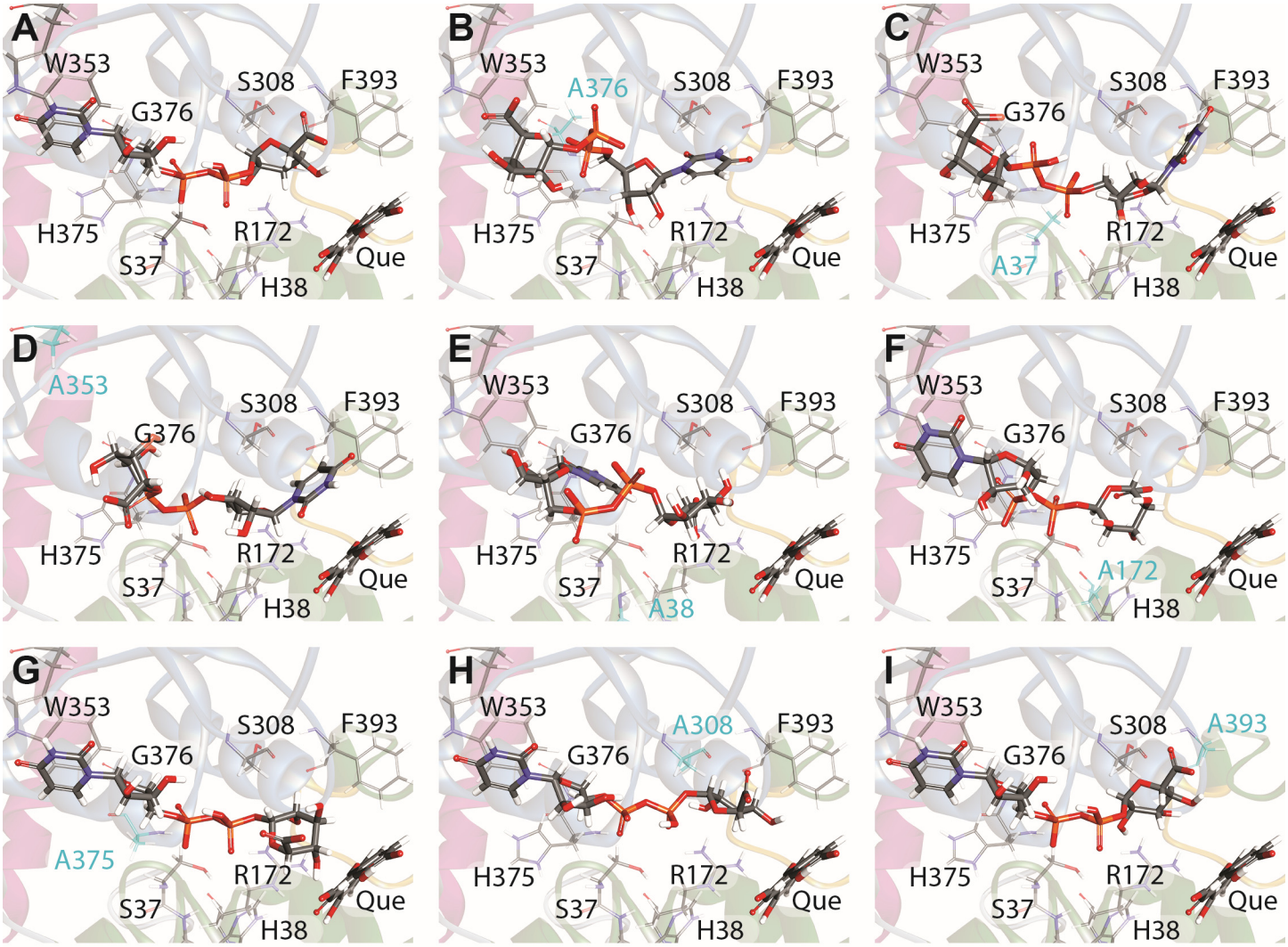


Figure S6. Comparison of UDPGA top docking positions into various mutated UGT1A6 models. Mutations are shown in order from most to least severe effect on UDPGA binding compared to the wild-type enzyme. A. Wild-type UGT1A6, B. G376A mutant, C. S37A mutant, D. W353A mutant, E. H38A mutant, F. R172A mutant, G. H375A mutant, H. S308A mutant, I. F393A mutant. Mutated residues are highlighted in cyan, quercetin (Que) is shown docked into the substrate binding site.

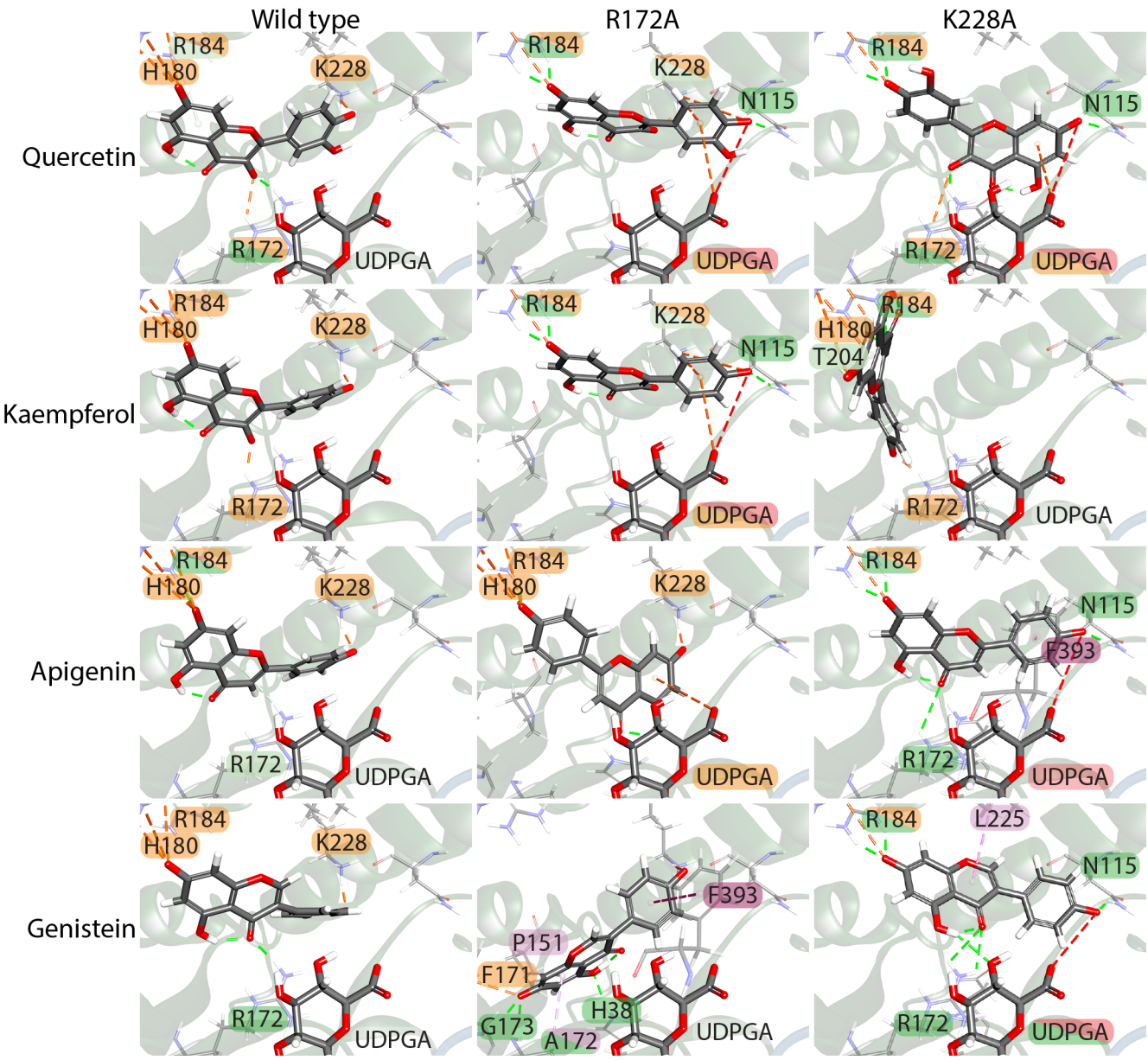


Figure S7. Comparison of the docking of selected substrates into the wildtype and mutated UGT1A6 model. The observed interactions for each docked compound are indicated. Interactions are coloured according to type with electrostatic interactions shown in orange, classical hydrogen bonds in dark green, non-classical hydrogen bonds in pale green, π-π stacking interactions in dark pink, π-alkyl interactions in light pink, and unfavourable interactions in red. Residues with multiple colours had multiple different interactions with the substrate.

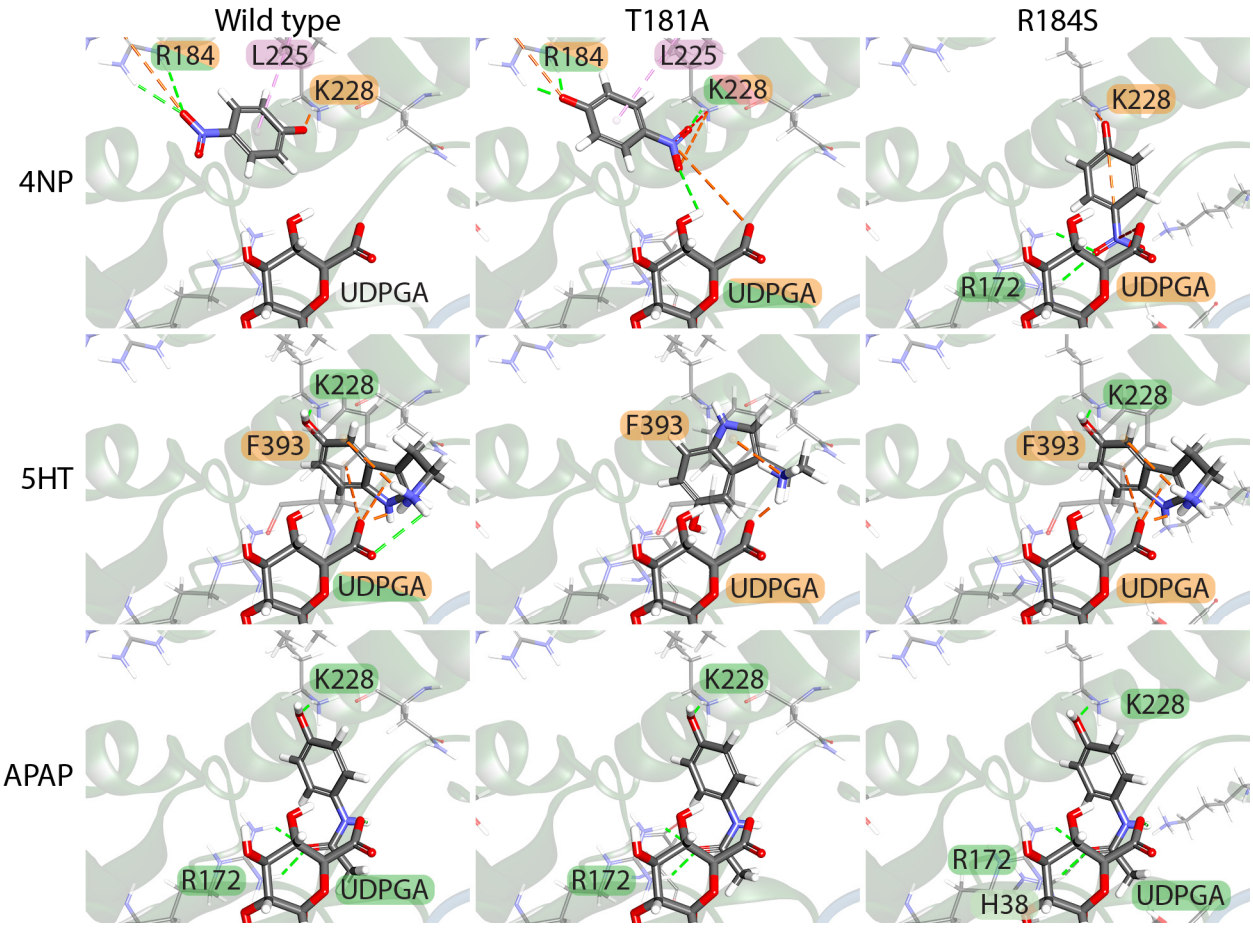


Figure S8. Comparison of the docking of three small planar phenolic compounds into two mutations that contribute to several UGT1A6 polymorphisms. The observed interactions for each docked compound are indicated, and coloured as in Figure S7. 4NP, 4-Nitrophenol; 5HT, serotonin: APAP, acetaminophen.

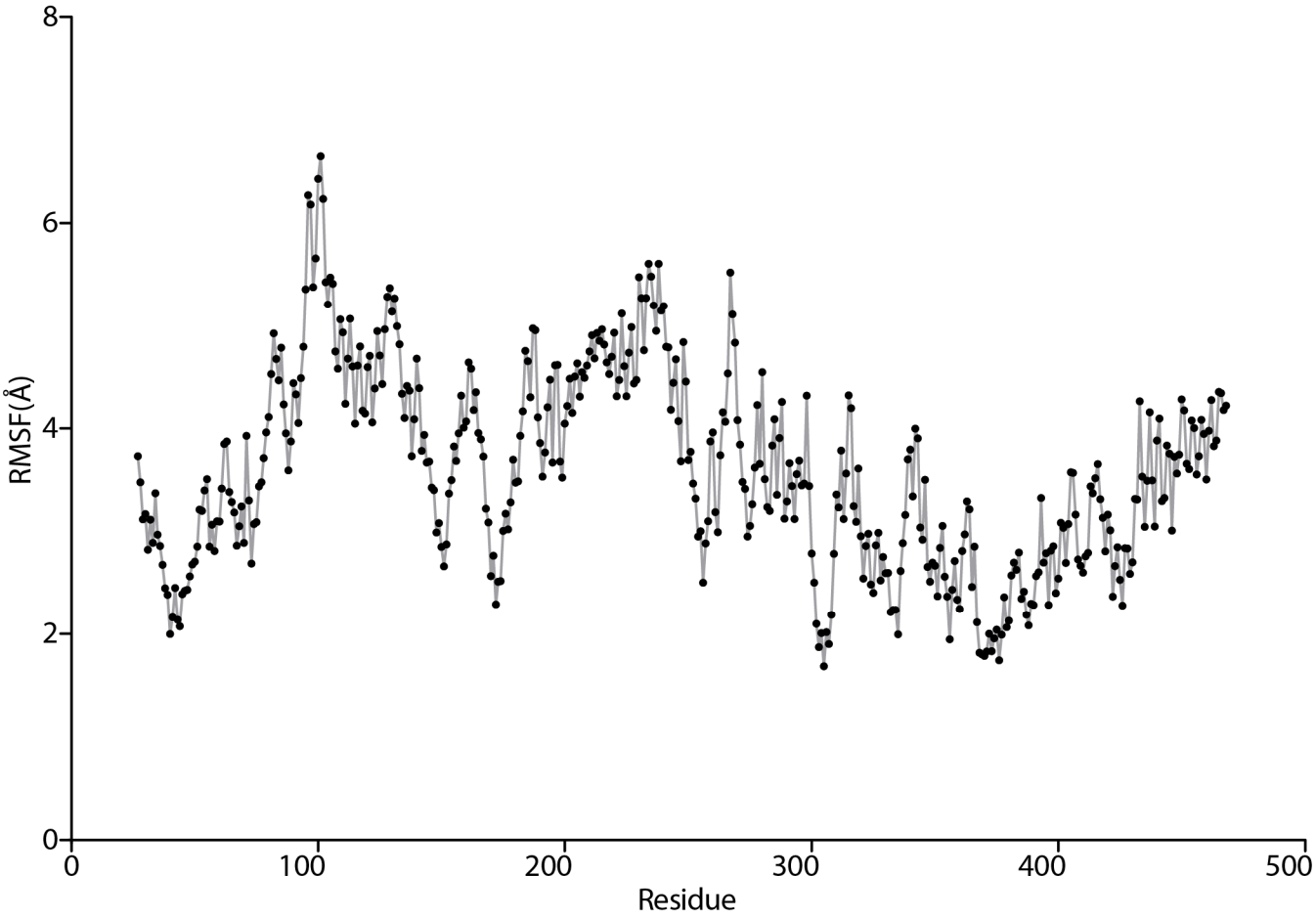


Figure S9. Root mean square fluctuation (RMSF) for each residue over the course of the molecular dynamics simulation.

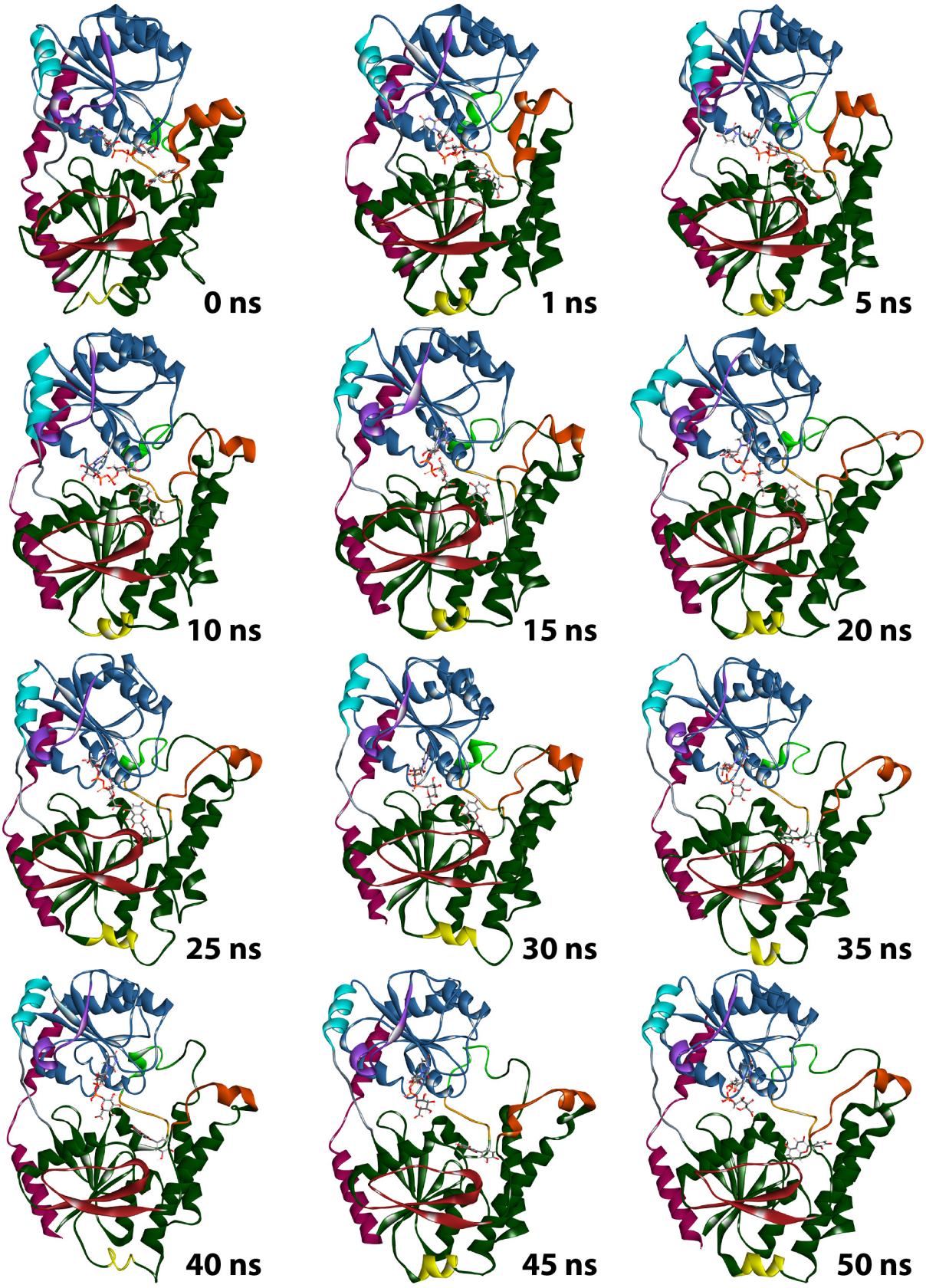


Figure S10. Snapshots from the molecular dynamics simulation trajectory. The cartoon diagram is coloured as in Figure S3, with the helices and sheets that showed substantial differences coloured separately. The bound UDPGA co-substrate, and bound substrate quercetin are shown as sticks. Nβ2, Nα2, and Nβ3 are shown in red, Nα3-1 in bright orange, Nα3-3 in yellow, Nα5-2 in bright green, Cα0 in teal, and Cβ3 – Cα3 in purple.

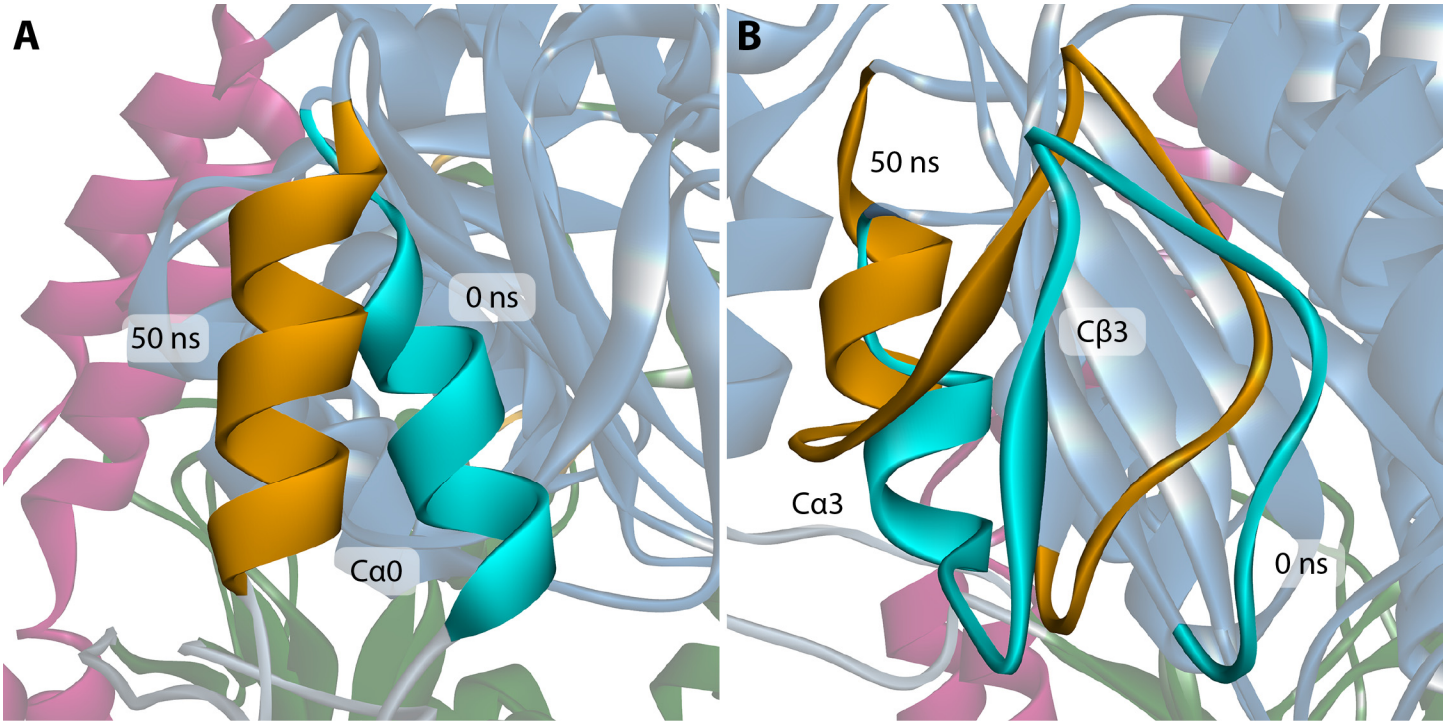


Figure S11. Shifts in the CTD following 20 ns of simulation compared to the initial structure. The initial (0 ns) structure is overlaid with the structure after 50 ns of simulation time. A. Shifts in the Cα0 helix. B. Shifts in the Cβ3 sheet and Cα3 helix. The shifted residues are highlighted in teal (0 ns) and orange (50 ns).

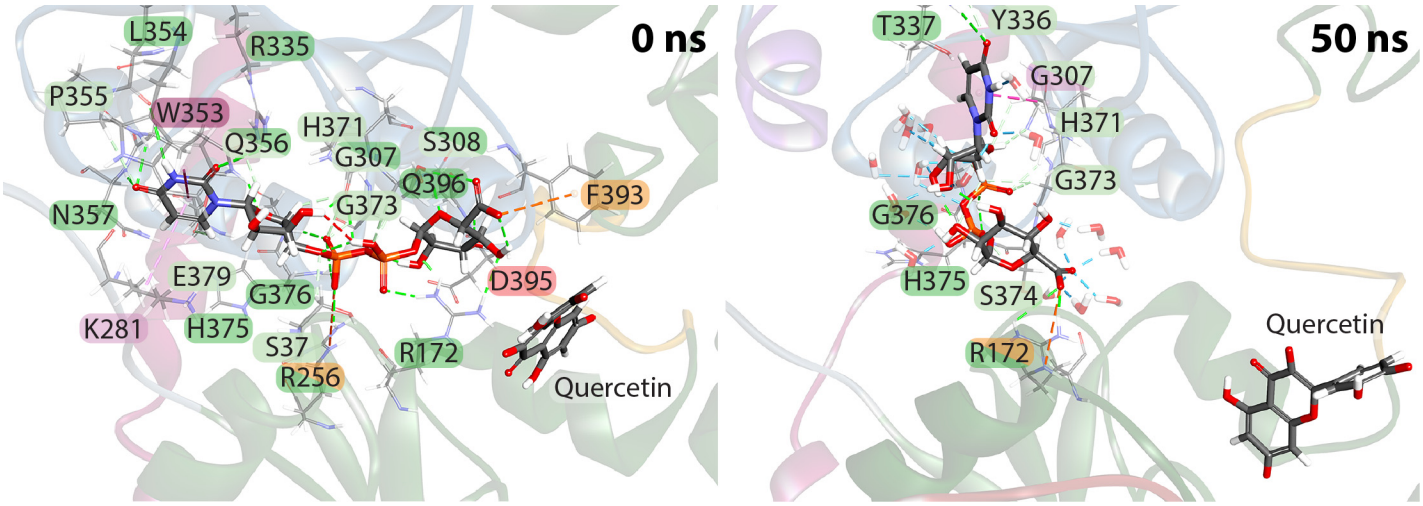


Figure S12. Changes in the UDPGA-UGT1A6 interacting residues following 50 ns of molecular dynamics simulations. Interacting residues are coloured as in Figure S7.

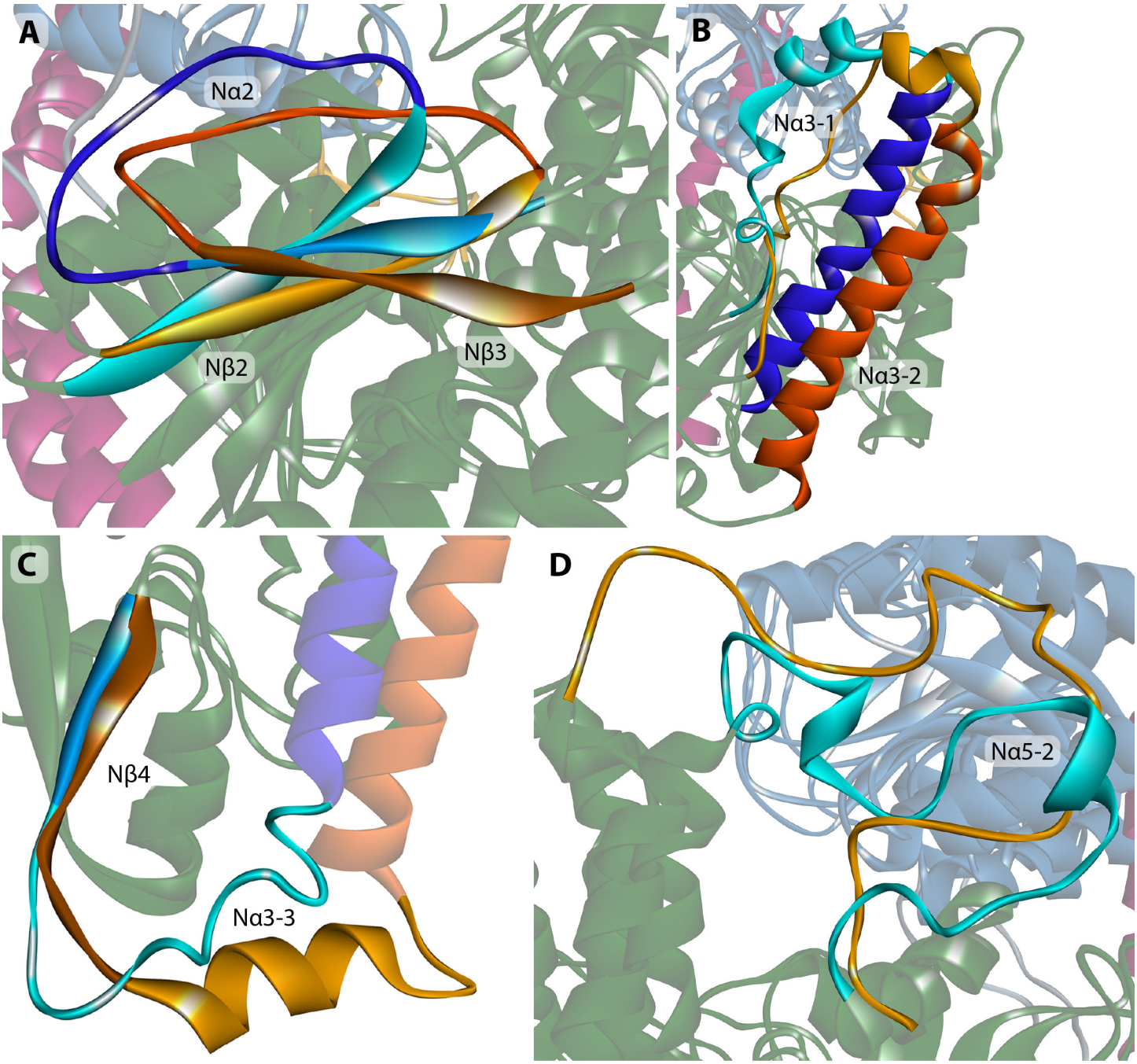


Figure S13. Shifts in the NTD following 50 ns of simulation. The initial (0 ns) structure is overlaid with the structure after 50 ns of simulation time. Secondary structures are highlighted using blues for the 0 ns time point, and oranges for the 50 ns snapshot. A. Shift in the Nβ2 (teal or orange), Nα2 (royal blue and bright orange), and Nβ3 region (cornflower blue and dark orange). B. Shifts in the Nα3-1 and Nα3-2 regions, C. Shifts in the Nα3-3 (teal or orange) and Nβ4 region (cornflower blue or dark orange). D. Shifts in the Nα5-2 region.

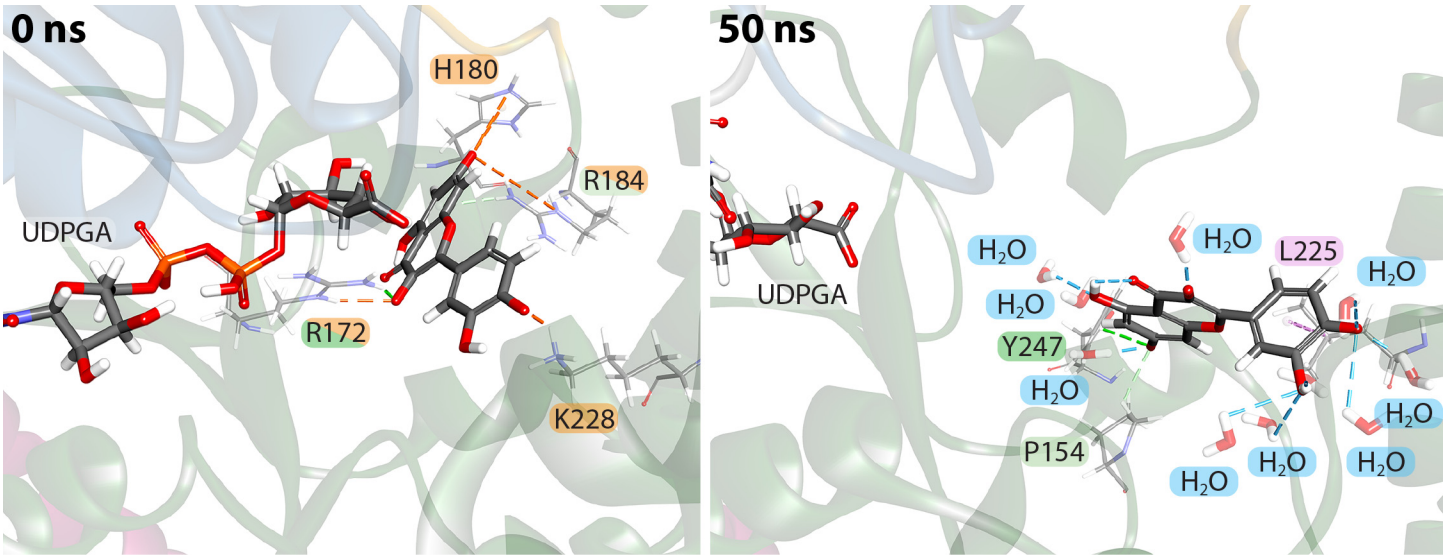


Figure S14. Changes in the quercetin-UGT1A6 interacting residues following 50 ns of molecular dynamics simulations. Interacting residues are coloured as in Figure S7, water-hydrogen bonds are shown in blue.

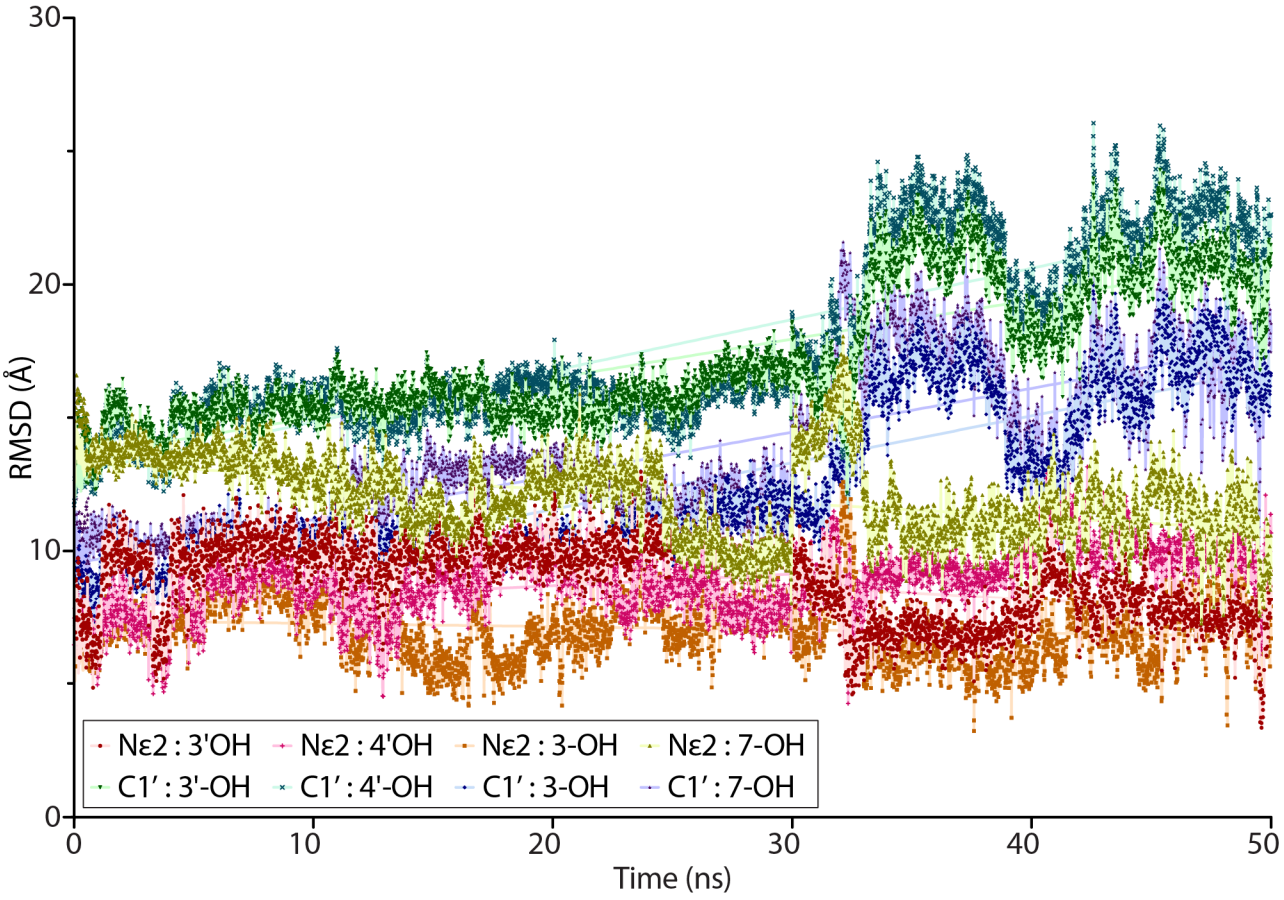


Figure S15. Distances between the Nε2 group of the catalytic H38, the glucuronic acid C1', and the potentially glucuronidated hydroxyl groups of the bound quercetin ligand over 50 ns of molecular dynamics simulations. Distances between the 3'-OH and Nε2 are shown in red, and with the C1' in green; distances between the 4'-OH and Nε2 in pink, and C1' in cyan; distance between the 3-OH and Nε2 in orange, and C1' in blue; and the distance between the 7-OH and Nε2 in yellow, and C1' in purple.

Table S1. Confidence scores for each model produced by I-TASSER

Input	Model #	C-Score[a]	Est. TM-Score[b]	Est RMSD (Å)[c]	No. of decoys[e]	Cluster density[f]
	1	-1.65	0.51 ± 0.15	11.4 ± 4.5	910	0.0591
	2	-2.53			456	0.0247
WSNT	3	-2.56			422	0.0239
	4	-2.6			340	0.0229
	5	-3.07			194	0.0144
	1	-1.67	0.51 ± 0.15	11.3 ± 4.5	805	0.0575
	2	-1.49[g]			630	0.0691
NSNT	3	-1.95			600	0.0436
	4	-2.64			291	0.0218
	5	-3.32			180	0.011
	1	-2.29	0.45 ± 0.14	13.0 ± 4.2	1169	0.0277
	2	-3.06			494	0.0128
WSWT	3	-3.09			493	0.0124
	4	-3.05			470	0.013
	5	-3.73			268	0.0065
	1	-1.65	0.51 ± 0.15	11.2 ± 4.6	2262	0.0518
	2	-3.04			597	0.0129
NSWT	3	-3.22			483	0.0108
	4	-3.13			463	0.0118
	5	-3.33			450	0.0097

[a] C-scores can range from -5 to 2, with higher values indicating higher confidence; [b] I-TASSER only provides TM-scores for the top ranked model, scores >0.5 indicating a model of correct topology and scores <0.17 meaning random similarity; [c] RMSD scores are only provided for the top ranked model, with local errors; [d] the number of low-temperature model replicas generated by excising continuous fragments from the threading alignments and reassembly through replica-exchanged Monte Carlo simulations; [e] the number of structure decoys at a unit of space in the SPICKER cluster, higher cluster density signifies a higher quality model with the structure occurring more frequently in the simulation; [g] models are ranked by cluster size, and although lower ranked models occasionally have higher C-scores, the top ranked model is usually better.

Table S2. DOPE and verify protein scores for I-TASSER produced models

Model	DOPE score	Normalized DOPE score*	Verify score	Verify expected high score	Verify expected low score
WSNT	-60317.38	-0.00	158.84	243.19	109.44
NSNT	-58504.77	-0.33	144.49	231.22	104.05
WSWT	-63160.52	-0.38	160.52	243.19	109.44
NSWT	-60440.23	-0.60	173.76	231.22	104.05
2O6L	-37114.25	-1.99	161.43	149.36	67.21
2ACW	-63436.67	-2.12	256.69	210.496	94.7231

*This is a Z-score derived from the DOPE score that accounts for protein size.

Table S3. Root Mean Square Distances (RMSD) between UGT1A6 models and GT crystal structures

Structure	WSNT (Å)	NSNT (Å)	WSWT (Å)	NSWT (Å)
1IIR	1.944	1.739	1.588	1.59
1PN3	2.331	2.164	1.807	1.782
1PNV	2.304	2.135	1.822	1.799
1RRV	2.102	1.942	1.477	1.491
2ACV	1.487	1.504	1.36	1.497
2ACW	1.463	1.484	1.365	1.457
2C9Z	1.605	1.625	1.608	1.65
2IYA	2.061	1.901	1.743	1.763
2O6L	1.283	1.318	0.841	0.793
2P6P	2.014	1.834	1.705	1.68
2PQ6	1.426	1.42	1.991	1.904
2VCE	1.601	1.617	1.904	1.828
2VCH	1.625	1.641	2.017	1.907
2VG8	1.605	1.622	2.027	1.919
3D0R	2.075	1.867	2.14	2.133
3H4I	2.412	2.226	1.839	1.86
3H4T	2.421	2.236	1.844	1.869
3HBF	1.446	1.467	1.868	2.03
3HBJ	1.452	1.464	1.915	2.103
3IA7	2.311	2.233	2.006	2.01
3OTG	1.984	1.866	1.489	1.452
3OTH	2.007	1.877	1.508	1.47
3RSC	1.95	1.786	1.713	1.695
3WC4	1.325	1.316	1.783	1.981
4REL	1.323	1.316	1.747	1.945
4REM	1.341	1.332	1.685	1.831
4WHM	1.313	1.306	1.628	1.737
5TMB	1.368	1.401	1.798	1.726
5TME	1.41	1.431	1.858	1.756
5U6M	1.415	1.436	1.586	1.549
WSNT		1.283	1.182	1.111
NSNT			1.247	1.168
WSWT				1.338

Table S4. Residues important for co-substrate binding in multiple GT structures

		Crystal structure (bound co-substrate)																						
UDPGA moiety	**UGT1A6 Residue**	**2C1Z**	**2VCE**	**3H4I**	**2ACW**	**5GL5**	**2ACV**	**2C1X**	**2C9Z**	**2IYA**	**2VCH**	**2VG8**	**3H4T**	**3HBF**	**3HBJ**	**4WHM**	**5TMB**	**5TME**	**5U6M**	**1PN3**	**1PNV**	**1RRV**	**3OTG**	**3OTH**	**3RSC**
		(U2F)	**(U2F)**	**(U2F)**	**(UDPG)**	**(UDPG)**	**(UDP)**	**(UDP)**	**(UDP)**	**(UDP)**	**(UDP)**	**(UDP)**	**(UDP)**	**(UDP)**	**(UDP)**	**(UDP)**	**(UDP)**	**(UDP)**	**(UDP)**	**(TYD)**	**(TYD)**	**(TYD)**	**(TYD)**	**(TYD)**	**(TYD)**
Glucose	G36*			S10																					
Glucose	H38*	H20		D13		D752																			
Glucose	R172*	T141																							
Glucose	S374	W353	W367	A294	W360	S1072																			
Glucose	D395	D374	E388	D315	E381	D1093																			
Glucose	Q396	Q375	Q389	Q316	Q382	Q1094																			
Beta phosphate	G36*			S10		S749		S18					S10,R11									T10			
Beta phosphate	S37*	T19		G12		G751	G21	T19	T19	G24	G18	G18	R11,G12	T25	T25	S16			G17	R11,G12	G12	G12	G12	G12	
Beta phosphate	G307*	G279					G284					S277													G236
Beta phosphate	S308*	T280		S230	S285		S285	T280	T280		S277	S277	S230	S282	S282	T273	T291	T291	S273	S230	S230	S246	T231	T231	T237
Beta phosphate	M309*																						S232		T238
Beta phosphate	G373												G293									S311			G303
Beta phosphate	H375									G332															
Beta phosphate	G376									S333															
Beta phosphate	F393		Y386		Y379						Y386	Y386													
Alpha phosphate	S37*						G21			G24	G18	G18		T25	T25		G26						G12	G12	
Alpha phosphate	G307*																								G236
Alpha phosphate	S308*									S264															
Alpha phosphate	H371	H350	H364		H357		H357	H350	H350	H328	H364	H364		H352	H352	H343		H361	H342						H301
Alpha phosphate	G373							G352		G332										S295	S295				
Alpha phosphate	S374					S1072					W367	W367										G313	S298	S298	
Alpha phosphate	H375	N354	N368	G295	N361	G1073	N361	N354	N354		N368	N368	G295	N356	N356	N347	N365	N365	N346			G313	G299	G299	
Alpha phosphate	G376	S355	S369	T296	S362	T1074	S362	S355	S355		S369	S369	T296	S357	S357	S348	S366	S366	S347	T298	T298	T314	T300	T300	T306
(deoxy)Ribose	S37*					R750		T19																	
(deoxy)Ribose	L40*									N27															
(deoxy)Ribose	N357*			H278		H1056							H278												
(deoxy)Ribose	E379	E358	E372		E365		E365	E358	E358	E336	E373	E372		E360	E360	E351	E369	E369	E350						E309
Pyrimidine ring	S37*		G18	R11		R750														R11	R11				
Pyrimidine ring	R282*									Q245										R207					
Pyrimidine ring	K283*																								L218
Pyrimidine ring	S305*				C282		C282									C270			A270						
Pyrimidine ring	R335	S306				N1025			S306	S290				S308	S308								A257	A257	
Pyrimidine ring	W353	W332	W346		W339		W339	W332	W332	W312	W346	W346		W334	W334	W325	F343		W324				W279	W279	W285
Pyrimidine ring	L354	A333	A347	E275	A340		A340	A333	A333	V313	A347	A347	E275	A335	A335	V326	C344	C344	S325	E277	E277	E293	V280	V280	V286
Pyrimidine ring	P355*			V276		V1054							V276							V278	V278	V294			
Pyrimidine ring	Q356							Q335										Q346		L280	L280	F296			H288
Pyrimidine ring	N357			H278		H1056							H278												

*: residue present within a loop region, TYD: Thymidine-5'-diphosphate, UDP: Uridine-5'-diphosphate, UDPG: Uridine-5'-diphosphate-alpha-D-glucose, U2F: Uridine-5'-diphosphate-2-deoxy-2-fluoro-alpha-D-glucose.

Table S5. UDPGA docking site mutagenesis

Importance	Mutation	-CDOCKER score (kcal mol^{-1})	UDPGA Conformation	RMSD from original pose (Å)
Co-substrate binding	S37A	65.6863	reversed linear	10.3539
Catalytic	H38A	51.9961	curved	6.6236
Catalytic	D150A	60.8324	linear	0.0095
Co-substrate and substrate binding	R172A	69.0987	curved	2.9425
Polymorphism	T181A	59.9712	linear	0.0007
Polymorphism	R184S	59.8607	linear	0.0280
Co-substrate binding	S308A	56.5551	slight curve	1.3158
Co-substrate binding	W353A	64.6337	reversed curved	10.0896
Co-substrate binding	L354A	58.6522	linear	0.2295
Co-substrate binding	H375A	60.6102	curved	2.9239
Co-substrate binding	G376A	42.7748	reversed curved	11.1750
Co-substrate and substrate binding	F393A	63.0931	linear	0.3093

Table S6. Substrates, docking scores, and enzyme kinetic data

Substrate (PubChem ID)	-CDOCKER Energy (kcal mol^{-1}); Molecular weight (g mol^{-1})	Metabolism K_m/S_{50} (μM)	Reference	Rationale
(+)-Catechin (9064)	52.7769; 290.27	ND	2	Catechin, plant substrate
1-Naphthol (7005)	6.26024; 144.173	3.36 ± 0.134[a,c]	3	Naphthol, UGT1A6 substrate
		78[b,g]	4	
		67[a,g]		
		84 ± 7[a,c]	5	
		74 ± 7[b,c]		
		3.1 ± 0.02[b,d]	6	
		8.07 ± 0.48[a,c]	7	
		5.25 ± 0.23[a,c,p]		
2,4,5-Trichlorophenol (7271)	34.6477; 197.439	Unknown	8	Plant UGT72B1 substrate
4-Hydroxytamoxifen (449459)	15.0642; 387.523	ND	9	Tamoxifen metabolite, other UGT substrate
4-Hydroxywarfarin (54682146)	49.4215; 324.332	ND	10	Coumarin, other UGT substrate
4-Methylumbelliferone (5280567)	40.2052; 176.171	120 ± 10[a,f]	11	Coumarin, UGT1A6 substrate
		109 ± 3.7[b,d]	6	
		76.4 ± 15.1[a,c]	7	
		57.5 ± 2[a,c]		
4-Nitrophenol (980)	36.5691; 139.11	770 ± 60[a,d]	12	UGT1A6 substrate
		730 ± 80[a,d,o]		
		644 ± 22[a,c]	3	
		1100 ± 230[a,d]	13	
		730 ± 150[a,d,l]		
		570 ± 290[a,d,m]		
		450 ± 160[a,d,n]		
		595 ± 140[a,g]	14	
Acetaminophen (1983)	22.298; 151.165	2200 ± 50[a,d]	15	Phenolic drug, UGT1A6 substrate
		2200 ± 300[a,c]	16	
		4000 ± 400[a,c]	17	
Alvocidib/Flavopiridol (5287969)	47.7067; 401.843	ND	18	Synthetic flavonoid, other UGT substrate
Apigenin (5280443)	64.9334; 270.24	Minimal (rate: 0.4 nmol/min/mg)	19	Flavone, plant substrate & poor UGT1A6 substrate
		Minimal	20	
Aspirin (2244)	32.5819; 180.159	ND	21	Non-metabolized control
Bilirubin (5280352)	39.7054; 584.673	ND	21	Non-metabolized control
Bilirubin Diglucuronide (5459911)	---; 936.328	---		Non-metabolized control

Bilirubin monoglucuronide (5459858)	10.6641; 760.797	---		Non-metabolized control
Biochanin A (5280373)	45.1212; 284.267	Minimal	22	Isoflavone, plant substrate
Bisphenol A (6623)	22.2207; 228.291	ND	23	Plant substrate
Chrysin (5281607)	50.0525; 254.241	12[a,c] 12.8[a,e] Minimal	24 25 20	Flavone, plant substrate & poor UGT1A6 substrate
cis-Resveratrol (1548910)	42.4312; 228.247	402.1 ± 39.2[a,d] 749.8 ± 61.2[a,d,l] 376 ± 28.8[a,d,m] 945.2 ± 93.5[a,d,n] 989.9 ± 92.8[a,c]	26	Stilbene, UGT1A6 substrate
Codeine (5284371)	-34.2624; 299.37	---	27	Opioid, UGT2B7 substrate
Combretastatin A4 (5351344)	12.5753; 316.353	44.8 ± 3.54[b,c]	28	Stilbene, plant & UGT1A6 substrate
Daidzein (5281708)	57.7173; 254.241	ND Minimal [a,e]	22 29	Isoflavone, plant substrate
Daphnetin (5280569)	48.3861; 178.143	74.4 ± 4.1[a,c] 74.4 ± 8.8[a,c]	30	Coumarin, plant & UGT1A6 substrate
Delphinidin (128853)	42.8066; 303.246	Unknown	31	Anthocyanin, plant substrate & possible poor UGT substrate
Diosmetin (5281612)	53.2596; 300.266	3.02 ± 0.59 [a,c]	32	Flavone, plant substrate & UGT1A6 substrate
Dopamine (681)	58.8765; 153.181	Minimal (rate: ~0.4 pmol/min/mg)	33	Poor UGT1A6 substrate, other UGT substrate
Emodin (3220)	54.8205; 270.24	ND	21	Quinone, plant & other UGT substrate
Endoxifen (10090750)	15.4125; 373.496	ND	9	Tamoxifen metabolite, other UGT substrate
Entacapone (5281081)	47.0507; 305.29	ND	34	Phenolic drug, other UGT substrate
Estradiol (5757)	-8.44673; 272.388	ND	21	Non-metabolized control, other UGT substrate
Ezetimibe (150311)	48.6851; 409.433	ND	35	Phenolic drug, other UGT substrate
Fenretinide (5288209)	-13.6979; 391.555	422 ± 46[a,c]	36	Phenolic drug, UGT1A6 substrate
Formononetin (5280378)	43.1571; 268.268	ND	22	Isoflavone, plant substrate
Galangin (5281616)	60.0664; 270.24	ND Minimal Minimal	21 37 20	Flavonol, plant substrate & poor UGT1A6 substrate
Genistein (5280961)	57.9287; 270.24	Minimal Minimal (rate: 0.05 nmol/min/mg) 220[a,e]	22 19 29	Isoflavone, plant substrate & poor UGT1A6 substrate

Glycitein (5317750)	52.3694; 284.267	Unknown	29	Isoflavone, plant substrate, possible substrate
Hesperetin (72281)	51.0055; 302.282	Minimal (rate: 0.018 nmol/min/mg)	38	Flavanone, plant substrate & other UGT substrate
Irinotecan (60838)	-28.3631; 586.689	ND	39,40	Phenolic drug, non-metabolized control
Kaempferol (5280863)	72.648; 286.239	Minimal Minimal	41 20	Flavonol, plant substrate
Morphine (5288826)	-8.70848; 285.343	ND	21	Opioid, other UGT substrate
Mycophenolic acid (446541)	13.8351; 320.341	Low activity	42	Phenolic drug, other UGT substrate
Naringenin (932)	49.4201; 272.256	Unknown	43	Falavanone, plant substrate & other UGT substrate
Norepinephrine (439260)	58.9037; 169.18	Unknown		Similar structure to serotonin/dopamine
Oleandomycin (5284598)	-5.40868; 687.868	Unknown		Bacterial substrate
Phenylephrine (6041)	22.9368; 167.208	Unknown	44	Phenolic drug, UGT substrate
Piceatannol (667639)	52.4984; 244.246	Minimal[c]	45	Stilbene, other UGT substrate, very poor UGT1A6 substrate
Prunetin (5281804)	61.2897; 284.267	Minimal	22	Isoflavone, plant substrate
Quercetin (5280343)	73.3885; 302.238	ND 3% conversion	21 46	Flavonol, plant substrate
Raloxifene (5035)	31.8046; 473.587	ND	47	Phenolic drug, non-metabolized control
Scopoletin (5280460)	40.7312; 192.17	194 ± 6.42[a,c]	3	Scopoletin, UGT1A6 substrate
Serotonin (5202)	26.9016; 176.21	8800 ± 300[a,k] 5900 ± 200[a,e] 6500 ± 900[a,h] 4900 ± 3300[a,i] 12400 ± 2000[a,j] 5000 ± 400[a,c] 6700 ± 200[a,k] 6800 ± 400[a,k] 5200 ± 300[a,k] 8800 ± 300[a,k] 5700 ± 1300[a,k] 6100 ± 800[a,k] 6200 ± 500[a,k] 5400 ± 200[a,k] 7000 ± 300[a,k,l] 2840 ± 1940[a,c] 9800 ± 500[a,d] 9900 ± 500[a,d,o]	48 49 50 12	UGT1A6 substrate
SN38 (104842)	-15.0766;	97.2 ± 7.9[a,c]	40	Phenolic drug, UGT1A6 substrate

	392.411	11.6 ± 5.2[a,d]	39	
Tamoxifen (2733526)	12.3199; 371.524	---		Phenolic drug, non-metabolized control
Testosterone (6013)	-25.8591; 288.431	ND[c]	51	Non-metabolized control, other UGT substrate
trans-Resveratrol (445154)	48.8684; 228.247	Minimal	52	Stilbene, other UGT substrate
Warfarin (54678486)	48.5637; 308.333	ND	10	Coumarin, non-metabolized control
Xanthohumol (639665)	32.1405; 354.402	3% conversion	53	Chalcone, plant substrate & other UGT substrate

ND: not detected, [a] K_m, [b] S_{50}, [c] recombinant protein produced in insect cells infected with baculovirus, [d] recombinant protein produced in HEK293 cells, [e] recombinant protein produced in lymphoblasts, [f] recombinant protein produced in a *Pichia pastoris* yeast expression system, [g] recombinant protein produced in V79 cells, [h] human kidney microsomes, [i] human lung microsomes, [j] human intestinal microsomes, [k] human liver microsomes, [l] *2 allozyme, [m] *3 allozyme, [n] *4 allozyme, [o] i1+i2 splice isoforms, [p] 0.1% BSA. All recombinants are assumed to be the *1 allozyme unless otherwise indicated.

Table S7. Correlation of ligand properties with the -CDOCKER Energy Score

Ligand property	**All docked compounds (55)**		**UGT1A6 substrates (27)[a]**		**UGT1A6 substrates (13)[b]**	
	Correlation (R^2)	**Slope**	**Correlation (R^2)**	**Slope**	**Correlation (R^2)**	**Slope**
Molecular weight	0.1165*	-1.620 ± 0.6127	0.04786	-0.6719 ± 0.5994	0.2836	-2.064 ± 0.9892
Molecular volume	0.1955**	-1.787 ± 0.4979	0.1708*	-1.080 ± 0.4758	0.399*	-2.158 ± 0.7988
Number rotatable bonds	0.05447	-0.02733 ± 0.01564	0.2409**	-0.03791 ± 0.01346	0.2728	-0.04488 ± 0.02210
Number H acceptors	0.005818	0.006280 ± 0.01128	0.1629*	0.02745 ± 0.01244	0.06638	0.01720 ± 0.01944
Number H donors	0.01593	-0.006573 ± 0.007096	0.0289	-0.006331 ± 0.007340	0.09776	-0.01277 ± 0.01170
Number total atoms	0.2180***	-0.3261 ± 0.08484	0.2696**	-0.2347 ± 0.07726	0.4658*	-0.4146 ± 0.1339
Number polar atoms[c]	0.001236	-0.003518 ± 0.01373	0.05134	0.01367 ± 0.01176	0.00046	-0.001404 ± 0.01974

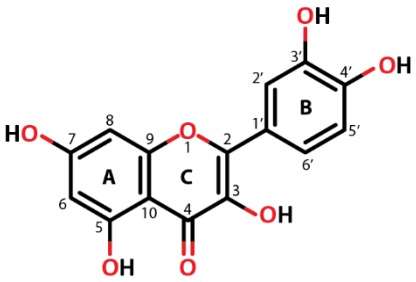
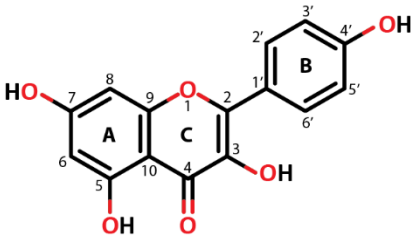
[a] Compounds that have some degree of detectable metabolism by UGT1A6; [b] Only those UGT1A6 substrates that have known Km values; [c] Number of oxygen and nitrogen atoms present in the compound; Slope is significantly non-zero: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

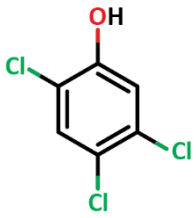
Table S8. Crystal structures used for definition of co-substrate and substrate docking spheres

Organism (PDB)	Enzyme	Bound co-substrate	Bound substrate	Catalytic His[a]	OH[b]	Nε2:OH (Å)[c]	C1':OH (Å)[d]
Medicago truncatula (2ACW)	Triterpene / Flavonoid GT	UDPG	---	H22	---	---	---
Vitis vinifera (2C1Z)	Flavonoid GT	U2F	Kaempferol	H20	3-OH 7-OH	2.696 9.394	4.263 9.028
Vitis vinifera (2C9Z)	Flavonoid GT	UDP*	Quercetin	H20	3'-OH 4'-OH 3-OH 7-OH	8.555 7.921 3.387 10.31	---
Arabidopsis thaliana (2VCE)	Hydroquinone GT	U2F	2,4,5-trichlorophenol	H19	1-OH	2.324	3.779
Actinoplanes teichomyceticus / Amycolatopsis orientalis (3H4I)	Chimeric GT	U2F	---	D13[e]	---	---	---
Medicago truncatula (3HBF)	Flavonoid GT	UDP*	Cannabiscetin	H26	3'-OH 3-OH 7-OH	5.448 4.914 12.201	---
Clitoria ternatea (4REL)	Anthocyanidin GT	---	Kaempferol	H17	3-OH 7-OH	2.492 9.515	---
Clitoria ternatea (4REM)	Anthocyanidin GT	---	Delphinidin	H17	3-OH 7-OH	2.51 9.493	---
Saccharomyces cerevisiae (5GL5)	Sterol 3-beta-GT	UDPG	---	D752[e]	---	---	---
Arabidopsis thaliana (5U6M)	UGT74F2	UDP*	Salicylic acid	H18	2-OH	5.32	---

[a] Position of the catalytic His residue numbered according to the PDB; [b] Position of the commonly glucuronidated OH of the substrate; [c] Distance between the Nε2 of the catalytic His residue and the substrate OH; [d] Distance between the C1' of UDPGA and the substrate OH group to be glucuronidated; [e] These GTs do not have a His residue at this position, instead having an Asp residue *These bound co-substrates were not used in defining the co-substrate docking sphere.

Table S9. Comparison of substrate binding residues between UGT1A6 docked compounds and crystal structures

Compound	Atom	Crystal structure residue	UGT1A6 equivalent	UGT1A6 interacting residue
Quercetin (PDB: 2C9Z) 	3'-OH	H_2O		
		H150	H180	
		S146	S177	
	4'-OH	H150	H180	
		Q188	loop*	
				K228
	B ring	F121	A152	
	3-OH	H20	H38	
				R172
	4-OH	S18	G36	
		H20	H38	
		H_2O		
	7-OH	Q84	N115	
				H180
				R184
Kaempferol (PDBs: 2C1Z | 4REL) 		S146	S177	
		Q188	loop*	
	4'-OH	H150	H180	
			H2O	
			H2O	
				K228
	B ring		F12	Q34
			L82	I118
		F121		A152
			V274	V310*
	3-OH	H20	H17	H38
				R172
	4-OH	S18		G36
			H2O	
			N137	R172
		N84		N115
			D181	loop*
	7-OH	H_2O	H2O	
			H2O	
				H180
				R184
			L196	loop*
	A ring	F200		K228
	C ring	F200		K228
2,4,5-Trichlorophenol (PDB: 2VCE)	1-OH	H19		H38
		U2F		
		H_2O		
				K228



2-Cl	H19	H38	
2-Cl	I86	I119	
	W315	loop*	
	H_2O		
4-Cl	F119	A152	
	L183	loop*	
	P185	loop*	
	A387	F393	
5-Cl	L118	P151	
	V144	S177	
	F148	H180	
	L183	loop*	
			L225
ring	F119	A152	
			L225
	A387	F393	F393

*UGT1A6 model and plant GT crystal structures aligned poorly in these regions. Water interactions are shown in blue, classical hydrogen bonds in dark green, non-classical hydrogen bonds in pale green, electrostatic interactions are shown in orange, π-π stacking interactions in dark pink π-alkyl interactions in light pink, and π–σ interactions in purple.

Table S10. Interactions with selected bound substrates following *in silico* mutagenesis

Mutation	Rank	Substrate	RMSD from WT (Å)	-CDOCKER score (kcal mol^{-1})	-OH[a]	C1':OH (Å)[b]	Nε2:OH (Å)[c]	Interactions[d]
Wild Type	1	Quercetin	---	73.3885	3'-OH	8.34	6.628	R172/R172, H180, R184/R184, K228
					4'-OH	9.412	9.106	
					3-OH	7.513	5.046	
					7-OH	11.743	12.362	
	2	Kaempferol	---	72.648	3-OH	7.455	4.588	R172, H180, R184, K228
					7-OH	11.486	12.003	
	3	Apigenin	---	64.9334	7-OH	11.056	11.184	R172, H180, R184/R184, K228
	4	Prunetin	---	61.2897	4'-OH	11.443	10.525	N115, R184/R184, K228/K228, UDPGA/UDPGA
					5-OH	10.768	10.748	
	5	Galangin	---	60.0664	3-OH	9.677	7.249	R184/R184, K228, UDPGA
					7-OH	11.345	10.577	
	6	Norepinephrine	---	58.9037	5-OH	7.098	6.977	S308, K228, F393, UDPGA
					6-OH	9.33	9.07	
	7	Dopamine	---	58.8765	5-OH	6.996	6.692	R172, K228, F393/F393, UDPGA
					6-OH	9.259	8.861	
	8	Genistein	---	57.9287	7-OH	11.859	12.213	R172, H180, R184, K228
	32	4-Nitrophenol	---	36.5691	1-OH	9.303	8.691	L225, R184/R184, K228
	37	Serotonin	---	26.9016	5-OH	9.332	9.054	F393, K228, UDPGA/UDPGA[e]
	39	Acetaminophen	---	22.298	1-OH	9.535	9.116	R172, K228, UDPGA
H38A	1	Kaempferol	0.1366	71.2639	3-OH	7.439	---	R172, H180, R184, K228
					7-OH	11.531	---	
	2	Quercetin	7.3884	70.1748	3'-OH	10.642	---	N115, R172, R184, L225, K228, UDPGA
					4'-OH	11.381	---	
					3-OH	7.423	---	
					7-OH	9.886	---	
	3	Apigenin	0.1222	64.9835	7-OH	11.08	---	R172, H180, R184, K228
	4	Prunetin	0.1213	61.228	4'-OH	11.407	---	N115, R184, K228/K228, UDPGA/UDPGA
					5-OH	10.802	---	
	5	Galangin	5.7322	60.9539	3-OH	11.342	---	R184, L225, K228/K228
					7-OH	9.211	---	
D150A	1	Quercetin	1.4719	80.8552	3'-OH	8.44	9.083	R172, H180, R184, K228
					4'-OH	9.182	8.015	
					3-OH	7.475	4.497	
					7-OH	11.415	11.896	
	2	Kaempferol	0.7442	72.1385	3-OH	7.538	5.02	R172, H180, R184, K228
					7-OH	11.712	12.304	
	3	Apigenin	0.3409	65.5696	7-OH	11.188	11.491	R172, H180, R184, K228
	4	Galangin	5.7151	62.4565	3-OH	11.336	10.738	R184, L225, K228
					7-OH	9.196	8.186	
	5	Prunetin	0.2342	62.0587	4'-OH	11.414	10.555	

					5-OH	10.883	10.856	N115, R184, K228/K228, UDPGA/UDPGA
R172A	1	Quercetin	3.6115	68.8932	3'-OH 4'-OH 3-OH 7-OH	8.486 9.624 9.79 11.304	8.63 10.956 7.252 10.668	N115, R184/R184, K228/K228, UDPGA/UDPGA
	2	Kaempferol	3.6909	65.3047	3-OH 7-OH	9.756 11.31	7.251 10.622	N115, R184, K228, UDPGA/UDPGA
	3	Apigenin	6.6146	63.289	7-OH	9.087	7.774	H180, R184/R184, K228, UDPGA
	4	Prunetin	2.3509	62.5783	4'-OH 5-OH	10.246 7.459	10.682 8.192	R184/R184, K228/K228/K228, UDPGA
	5	Genistein	4.4617	55.9765	7-OH	9.297	5.787	H38, P151, F171, *A172*/*A172*[f], G173, K228, F393
H180A	1	Quercetin	2.6225	75.2437	3'-OH 4'-OH 3-OH 7-OH	8.933 9.334 7.669 11.316	8.629 7.881 3.105 10.47	Q34, H38, R172/R172, S177, R184, K228
	2	Kaempferol	0.3141	67.5371	3-OH 7-OH	7.497 11.362	4.423 11.813	R172, R184, K228
	3	Apigenin	2.4233	60.8558	7-OH	11.372	10.56	N115, R172/R172, R184, K228
	4	Galangin	5.7445	58.9827	3-OH 7-OH	11.348 9.175	10.726 8.211	R184, L225, K228
	5	Prunetin	0.2389	58.6895	4'-OH 5-OH	11.411 10.849	10.539 10.821	N115, R184, K228/K228, UDPGA
T181A	1	Quercetin	0.6904	72.4485	3'-OH 4'-OH 3-OH 7-OH	9.1456 9.122 7.337 11.717	6.161 8.134 4.948 12.427	R172, H180, R184, K228
	2	Kaempferol	1.9742	70.7189	3-OH 7-OH	7.671 11.467	3.085 10.473	Q34, H38, R172/R172, S177, R184, K228
	3	Apigenin	6.3051	62.139	7-OH	9.194	7.322	Q34, R172, H180, R184, K228
	4	Prunetin	3.0043	61.3194	4'-OH 5-OH	11.354 9.388	10.625 8.508	Q34, R184, K228
	5	Galangin	7.1028	61.1947	3-OH 7-OH	11.61 9.207	10.41 8.256	A181, R184, K228/L225
	32	4-Nitrophenol	3.8724	41.2233	1-OH	11.322	10.797	R184/R184, L225, K228/K228/K228, UDPGA/UDPGA
	38	Serotonin	2.8545	28.0261	5-OH	5.964	4.593	F393, UDPGA
	41	Acetaminophen	0.167	24.6834	1-OH	9.459	9.048	R172, K228
R184A	1	Quercetin	9.7242	71.4492	3'-OH 4'-OH 3-OH	9.684 9.463 4.278	7.016 8.793 4.743	H38, K88, R172, K228, UDPGA

	2	Apigenin	10.0196	62.7969	7-OH 7-OH	10.956 11.212	8.429 8.797	K88, S308, K228
	3	Kaempferol	1.0595	58.3778	3-OH 7-OH	7.269 11.63	5.381 12.91	R172, H180, K228
	4	Dopamine	1.3487	57.2565	5-OH 6-OH	7.016 9.386	8.392 9.169	Y112, K228, F393, UDPGA/UDPGA
	5	Norepinephrine	1.4025	55.6979	5-OH 6-OH	6.983 9.315	6.601 8.773	R172, K228, F393, UDPGA/UDPGA
R184S	1	Quercetin	9.7158	71.4267	3'-OH 4'-OH 3-OH	9.694 9.479 4.287	7.018 8.798 4.749	H38, K88, R172, K228, UDPGA
	2	Apigenin	10.0219	62.7472	7-OH 7-OH	10.953 11.199	8.426 8.747	K88, S308, K228
	3	Kaempferol	0.9558	58.9523	3-OH 7-OH	7.265 11.864	5.311 12.858	R172, H180, K228
	4	Dopamine	0.2299	57.7143	5-OH 6-OH	7.07 7.844	6.743 8.881	K228, F393, UDPGA
	5	Norepinephrine	0.4429	56.8494	5-OH 6-OH	7.198 9.683	7.538 9.592	K228, F393, UDPGA
	32	4-Nitrophenol	6.0794	34.5556	1-OH	9.511	9.373	R172, K228, UDPGA
	37	Serotonin	0.1457	26.737	5-OH	9.313	8.992	K228, F393, UDPGA
	41	Acetaminophen	0.2466	22.1607	1-OH	9.396	8.773	H38, R172, K228, UDPGA
K228A	1	Quercetin	7.3353	63.2987	3'-OH 4'-OH 3-OH 7-OH	10.638 11.355 7.493 9.693	11.426 10.889 5.13 10.643	N115, R172/R172, R184/R184, UDPGA/UDPGA
	2	Kaempferol	5.9363	62.9417	3-OH 7-OH	11.203 11.38	10.78 14.717	R172, H180, R184/R184, T204
	3	Apigenin	2.9312	55.546	7-OH	11.35	10.787	N115, R172, R184, F393, UDPGA
	4	Galangin	3.5476	53.2515	3-OH 7-OH	7.719 11.364	3.111 10.576	Q34, H38, R172/R172, R184
	5	Genistein	3.4702	52.1346	7-OH	11.336	10.686	N115, R172, R184/R184, L225, UDPGA/UDPGA
F393A	1	Quercetin	1.4077	74.2709	3'-OH 4'-OH 3-OH 7-OH	8.996 9.371 7.447 11.729	10.145 9.012 4.969 12.353	R172, H180, R184, K228
	2	Kaempferol	7.746	71.9498	3-OH 7-OH	9.483 8.152	8.59 12.596	R184, L225, K228, *A393*, UDPGA
	3	Apigenin	7.759	66.7926	7-OH	9.233	8.135	H180, R184, K228, *A393*, UDPGA
	4	Galangin	2.1694	65.4804	3-OH 7-OH	9.291 11.327	9.006 10.646	Y112, R184, K225, K228, *A393*
	5	Prunetin	1.5133	64.5641	4'-OH 5-OH	11.367 10.318	10.614 10.133	N115, R184, L225, K228/K228

[a] Position of the commonly glucuronidated –OH of the substrate; [b] Distance between the C1' of UDPGA and the substrate OH group to be glucuronidated; [c] Distance between the Nε2 of the catalytic H38 residue of UGT1A6 and the substrate OH; [d] Interactions are coloured according to type with electrostatic interactions shown in orange, classical hydrogen bonds in dark green, non-classical hydrogen bonds in pale green, π-π stacking interactions in dark pink, π-alkyl interactions in light pink, and unfavourable interactions in red; [e] Residues shown multiple times in different colours had multiple different types of interactions with the substrate involving different atoms of the residue and/or substrate; [f] Interactions with the mutated residue shown in italics.

Table S11. Docking scores and residue distances of top 5 substrates and 3 well-metabolized substrates docked into Ghemtio *et al* (2014) model

Rank	Name	Glucuronidated[a]	-CDOCKER Energy (kcal mol^{-1})	-OH[b]	C1':OH (Å)[c]	Nε2:OH (Å)[d]	Interactions[e]
1	Quercetin	Minimally	64.4719	3'-OH 4'-OH 3-OH 7-OH	6.673 8.08 6.639 8.368	3.75 6.059 7.244 8.254	G36, F171, G173, G394, D395, U2F
2	Apigenin	Minimally	59.4991	7-OH	6.837	5.73	E111, F171, G173, G394, U2F
3	Kaempferol	Minimally	59.3939	3-OH 7-OH	6.225 7.697	7.187 5.132	H180, P219, S234, G394, D395, U2F/U2F
4	Genistein	Minimally	54.2317	7-OH	12.482	11.042	G36, H38, H180, P219, S234, F393, G394, D395, U2F
5	Daidzein	Minimally	52.8188	7-OH	3.114	4.48	G36, H38, H180, P219, S234, U2F
30	4-Nitrophenol	Yes	33.5717	1-OH	6.627	5.507	F171/F171, G173, U2F
36	Acetaminophen	Yes	21.6177	1-OH	3.338	4.337	U2F
37	Serotonin	Yes	19.8037	5-OH	7.844	6.165	G36, H38/H38/H38, A152, R172, U2F

[a] Ability of UGT1A6 to glucuronidate this compound as per table S6. [b] Position of the commonly glucuronidated –OH positions of the substrate; [c] Distance between the C1' of U2F and the substrate OH group to be glucuronidated; [d] Distance between the Nε2 of the catalytic H38 residue of UGT1A6 and the substrate OH; [e] Interactions are coloured as in Table S14.

Table S12. Change in structures over the course of molecular dynamics simulations

Time	RMSD (Å)			
	UGT1A6 (main chain)	UDPGA	Quercetin	2O6L
0 ns	0	0	0	1.712
1 ns	3.44	3.1697	3.1697	3.13
5 ns	3.906	3.7971	3.7971	3.307
10 ns	4.093	4.0673	4.0673	3.712
15 ns	4.345	4.8031	4.8031	3.391
20 ns	5.067	5.1696	5.1696	4.295
25 ns	4.517	5.0822	5.0822	3.548
30 ns	5.277	5.7387	5.7387	4.694
35 ns	5.677	7.2824	7.2824	4.893
40 ns	5.668	6.8596	6.8596	4.259
45 ns	5.930	6.6732	6.6732	4.876
50 ns	5.795	6.895	6.895	4.672

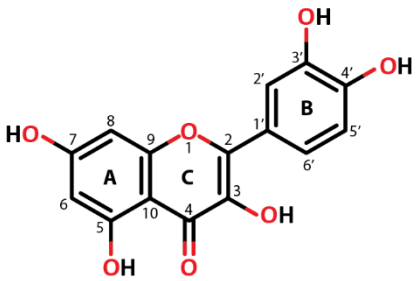
Table S13. Change in UDPGA interacting residues over time

Substructure	0ns	1 ns	5 ns	10 ns	15 ns	20 ns	25 ns	30 ns	35 ns	40 ns	45 ns	50 ns
Glucose	R172[a]	R172 / R172	R172	R172 / R172	H38 / H38	R172 / R172	H38 / H38	S41	R172 / R172	R172 / R172 / R172	R256 / R256	R172 / R172
	S308	R256	R256	R256	R172 / R172	G373	R172 / R172	R172	R256	11x H_2O / H_2O	6x H_2O	7x H_2O
	F393	S374	H371	G373	R256	S374	H371	S374 / S374	S308			
	D395	D395 / D395	G373	S374	G373	D395 / D395	G373	7x H_2O	10x H_2O			
	Q396	Q396	D395 / D395	D395	S374	4x H_2O / 2x H_2O	S374					
		9x H_2O	7x H_2O	Q396	D395		D395					
				8x H_2O	Q396 / Q396		Q396					
					3x H_2O / H_2O		6x H_2O					
β-phosphate	R172	S41	2x H_2O	R172	6x H_2O	H_2O	3x H_2O	S374	H375	S374 / S374	H375	S374 / S374
	G307	R172		R256				H375	G376	H375	G376	G376
	S308	2x H_2O		3x H_2O				2x H_2O	H_2O	H_2O	2x H_2O	2x H_2O
α-phosphate	S37	S41	3x H_2O	S41	H371	G307	H371	H371	G307	H371	H371	H371
	R256	R256		R256	3x H_2O	S308	3x H_2O	G373	H371	G373	G376 / G376	G373
	G307	H_2O		2x H_2O		H371		3x H_2O	G376	G376	3x H_2O	G376 / G376
	H371					H_2O				2x H_2O		2x H_2O
	G373											
	H375											
	G376											
Ribose	G307 / G307	G307 / G307	S308	3x H_2O / H2O	G307	2x H_2O	S308	4x H_2O	S308	G307	S308	G307
	N356 / N356	4x H_2O	4x H_2O		S308		5x H_2O		8x H_2O	S308	4x H_2O / 2x H_2O	H371
					4x H_2O					5x H_2O		6x H_2O

Pyrimidine ring	L281	Q356	K283	W353	G307	R335	R335	G307	G307	S308	Y336	R335
	R335	3x H_2O	3x H_2O	6x H_2O	R335	Y336	Y336	R335	S308 / S308	Y336	T337	Y336
	W353				T337	T337	T337	6x H_2O	Y336	T337	4x H_2O	T337
	L354				6x H_2O	2x H_2O	7x H_2O		T337	2x H_2O		3x H_2O
	P355								4x H_2O			
	Q356											
	N357											
	E379											

[a] Interactions are coloured as in Table S10, with water interactions shown in blue.

Table S14. Change in Quercetin interacting residues over time.

Compound	Atom	0ns	1 ns	5 ns	10 ns	15 ns	20 ns	25 ns	30 ns	35 ns	40 ns	45 ns	50 ns
Quercetin 	3'-OH		A152 H_2O	2x H_2O	Y229 H_2O	H_2O		Y247	Y247 H_2O	K228 H_2O	K228 / K228	K228 H_2O	3x H_2O
	4'-OH	K228	H38 K228 3x H_2O	K228 3x H_2O	K228 Y229	Y229 H_2O	Y229	Y229 H_2O	Y229	C224 K228 / K228 H_2O	K228 H_2O	K228 H_2O	3x H_2O
	B ring		A152	A152	L225	A152 L225		L225		C224** L225	L225	L221 L225	L225
	A ring	R184	F171	F171	F171					P154	A152	A152	
	C ring		F171	F171	F171	H_2O*				A152 / A152*	A152		
	3-OH	R172	A152 R172	R172	R172 A152 H_2O	H38 A152 R172	R172	R172 H2O	A152 H_2O	2x H_2O	H38 2x H_2O	A152 2x H_2O	H_2O
	4-OH		H_2O		H_2O	R172 3x H_2O		R172 H_2O	3x H_2O	2x H_2O	H_2O	A152 H_2O	H_2O
	5-OH		G173	G173		G394 H_2O	G394 D395 H_2O	2x H_2O	D395 H_2O	H_2O/H_2O	R172 H_2O	H_2O	H_2O
	7-OH	H180 R184	S177 F393 G394 H_2O	S177 G394 H_2O	G394 H_2O	H180 2x H_2O	R172 H180 H_2O	H180 2x H_2O	H180 H_2O	Y247	2x H_2O	4x H_2O	P154 Y247 H_2O

*π-lone pair interaction, **π-sulfur interaction, other interactions are coloured according to type as in Table S10, with water interactions shown in blue.

Table S15. Distances between potential quercetin glucuronidation sites and important enzyme and co-substrate atoms over time.

Measure	Slope	Distance (Å)										
		0 ns	5 ns	10 ns	15 ns	20 ns	25 ns	30 ns	35 ns	40 ns	45 ns	50 ns
3'-OH:Nε2	-0.0005236 ± 1.286e-005	6.628	10.590	10.230	9.845	9.541	10.533	10.599	7.0179	8.156	6.900	8.156
3'-OH: C1’	0.001473 ± 1.395e-005	8.34	15.590	15.413	16.034	14.635	15.452	18.611	21.736	18.421	20.506	21.140
4'-OH:Nε2	0.0004419 ± 1.046e-005	6.451	8.338	8.420	8.699	9.999	8.920	9.050	9.261	9.885	9.445	10.680
4'-OH:C1’	0.001922 ± 1.689e-005	11.732	15.086	15.219	15.845	16.206	14.869	18.766	23.312	19.335	21.978	22.626
3-OH:Nε2	-0.0001180 ± 1.155e-005	5.046	7.906	8.159	4.779	7.180	6.996	6.565	6.479	4.993	5.675	6.695
3-OH:C1’	0.001755 ± 1.489e-005	7.513	10.231	10.508	9.794	10.101	10.550	12.856	17.537	13.382	15.648	16.670
7-OH:Nε2	-0.0005080 ± 1.322e-005	12.362	14.183	12.662	11.119	12.457	11.966	10.599	11.614	10.269	13.112	9.142
7-OH:C1’	0.001642 ± 1.767e-005	11.743	11.079	10.588	13.092	12.616	11.428	14.931	19.549	11.811	17.492	18.535

Table S16. Top 10 I-TASSER threading alignments for each set of input conditions

Input	Rank[a]	Algorithm	PDB:Chain	Thread ID (%)[b]	Total ID (%)[c]	Coverage[d]	Z-score[e]
WSNT	1	MUSTER	5gl5:A	18	18	0.78	1.75
	2	FFAS-3D	5gl5:A	17	18	0.78	2.99
	3	SPARKS-X	2iya:A	17	18	0.72	1.74
	4	HHSERARCH2	5gl5:A	18	18	0.78	2.56
	5	HHSEARCH I	5gl5:A	18	18	0.78	2.23
	6	Neff-PPAS	2iyf:B	18	18	0.73	2.46
	7	HHSEARCH	2pq6:A	20	17	0.74	3.22
	8	pGenTHREADER	4m83:A	17	19	0.73	4.39
	9	wdPPAS	5gl5:A	18	18	0.78	1.62
	10	cdPPAS	2iya:A	00	18	0.73	2.83
NSNT	1	MUSTER	5gl5:A	17	20	0.82	1.78
	2	FFAS-3D	2iyf:A	18	19	0.75	3.03
	3	SPARKS-X	2iya:A	18	19	0.74	1.86
	4	HHSEARCH2	5gl5:A	18	20	0.81	2.59
	5	HHSEARCH I	5gl5:A	18	20	0.82	2.20
	6	Neff-PPAS	2iyf:B	18	19	0.76	2.57
	7	HHSEARCH	2pq6:A	20	18	0.79	3.23
	8	pGenTHREADER	2iya:A	18	19	0.76	4.28
	9	wdPPAS	2iyf:B	18	19	0.76	1.66
	10	PROSPECT2	2iyf:A	18	19	0.75	2.93
WSWT	1	USER[f]	2o6l:A	55	17	0.30	10.00
	2	FFAS-3D	2o6l:A	58	16	0.29	2.35
	3	FFAS03	2o6l:A	54	16	0.29	2.95
	4	pGenTHREADER	2o6l:A	59	18	0.30	3.54
	5	HHSEARCH I	2acv:A	16	18	0.75	1.82
	6	HHSEARCH2	2acv:A	16	18	0.75	2.33
	7	SPARKS-X	2acv:A	16	18	0.76	1.30
	8	HHSEARCH	2acv:A	15	18	0.76	2.69
	9	HHSEARCH2	5tme:A	18	18	0.73	2.33
	10	HHSEARCH	5tme:A	18	18	0.73	2.86
NSWT	1	USER	2o6l:A	55	17	0.32	10.00
	2	FFAS-3D	2o6l:A	58	17	0.31	2.35
	3	FFAS03	2o6l:A	53	17	0.30	2.98
	4	pGenTHREADER	2o6l:A	59	19	0.32	3.43
	5	HHSEARCH I	2acv:A	16	19	0.77	1.83
	6	HHSEARCH I	5tme:A	18	19	0.77	1.98
	7	HHSEARCH2	2acv:A	16	19	0.77	2.35
	8	SPARKS-X	2acv:A	18	19	0.79	1.38
	9	HHSEARCH	2acv:A	15	19	0.79	2.76
	10	HHSEARCH2	5tme:A	18	19	0.77	2.42

[a] Top ten threading templates used by I-TASSER, [b] Percent sequence identity of template and query sequence within threading aligned region, [c] percent sequence identity over the entire template and query sequence, [d] coverage of the threading alignment with query (number aligned residues divided by length of query), [e] Normalized Z-score of the threading alignments, alignments with a score >1 indicate a good alignment, [f] user supplied template sequence – UGT2B7 C-terminal end. Structural and sequence alignments are available in the "Supplemental structures" zip file.

Table S17. Top 10 identified structural analogues to I-TASSER constructed models

Model	Rank[a]	PDB	Enzyme	TM-score	RMSD[b]	ID[c]	Coverage[d]
WSNT	1	5gl5:A	Sterol 3-beta-glucosyltransferase	0.777	0.9	0.165	0.786
	2	1rrv:A	TDP-vancosaminyltransferase GtfD	0.663	3.09	0.157	0.737
	3	1iir:A	UDP-glucosyltransferase GtfB	0.651	2.48	0.152	0.703
	4	3oth:A	Calicheamicin Glycosyltransferase	0.632	3.47	0.139	0.714
	5	3rsc:A	Calicheamicin Glycosyltransferase	0.629	3.54	0.178	0.716
	6	3h4t:A	Chimeric glycosyltransferase	0.625	3.18	0.159	0.701
	7	2acw:A	Triterpene/Flavonoid Glycosyltransferase	0.624	4.25	0.155	0.733
	8	2iya:A	Oleandomycin Glycosyltransferase	0.622	3.6	0.179	0.707
	9	3wc4:A	anthocyanidin 3-O-glucosyltransferase	0.62	4.41	0.15	0.733
	10	3ia7:A	Calicheamicin Glycosyltransferase	0.618	3.76	0.161	0.724
NSNT	1	5gl5:A	Sterol 3-beta-glucosyltransferase	0.806	1.12	0.181	0.818
	2	1rrv:A	TDP-vancosaminyltransferase GtfD	0.701	2.87	0.155	0.773
	3	1iir:A	UDP-glucosyltransferase GtfB	0.683	2.51	0.147	0.737
	4	3rsc:A	Calicheamicin Glycosyltransferase	0.666	3.37	0.175	0.751
	5	3h4i:A	Chimeric glycosyltransferase	0.659	3.01	0.161	0.731
	6	2iya:A	Oleandomycin Glycosyltransferase	0.658	3.32	0.177	0.743
	7	3oth:A	Calicheamicin Glycosyltransferase	0.656	3.58	0.14	0.745
	8	5tmb:A	Ox70 UDP-glucosyltransferase	0.649	4.31	0.162	0.767
	9	3ia7:A	Calicheamicin Glycosyltransferase	0.648	3.77	0.161	0.759
	10	1pnv:B	TDP-epi-Vancosaminyltransferase GtfA	0.646	3.19	0.14	0.727
WSWT	1	2acw:A	Triterpene/Flavonoid Glycosyltransferase	0.715	2.05	0.16	0.752
	2	2vce:A	Hydroquinone Glucosyltransferase	0.699	2.49	0.157	0.748
	3	3wc4:A	anthocyanidin 3-O-glucosyltransferase	0.694	2.9	0.16	0.763
	4	5tmb:A	Ox70 UDP-glucosyltransferase	0.69	2.76	0.163	0.75
	5	3hbj:A	Flavonoid 3-O-glucosyltransferase	0.689	2.84	0.181	0.757
	6	2c9z:A	Flavonoid 3-O-glucosyltransferase	0.685	2.88	0.162	0.752
	7	5u6m:A	UDP-glycosyltransferase 74F2	0.684	2.57	0.178	0.739
	8	2pq6:A	UDP-glucosyltransferase	0.67	2.81	0.195	0.729
	9	5gl5:A	Sterol 3-beta-glucosyltransferase	0.66	3.6	0.156	0.748
	10	3oth:A	Calicheamicin Glycosyltransferase	0.629	3.5	0.158	0.72
NSWT	1	2acw:A	Triterpene/Flavonoid Glycosyltransferase	0.73	2.28	0.163	0.777
	2	5tmb:A	Ox70 UDP-glucosyltransferase	0.726	2.51	0.167	0.781
	3	2vce:A	Hydroquinone Glucosyltransferase	0.726	2.6	0.158	0.783
	4	5u6m:A	UDP-glycosyltransferase 74F2	0.72	2.44	0.181	0.775
	5	3wc4:A	anthocyanidin 3-O-glucosyltransferase	0.713	2.95	0.162	0.79
	6	3hbj:A	Flavonoid 3-O-glucosyltransferase	0.711	2.91	0.168	0.787
	7	2c9z:A	Flavonoid 3-O-glucosyltransferase	0.71	2.94	0.156	0.785
	8	2pq6:A	UDP-glucosyltransferase	0.697	2.77	0.193	0.761
	9	5gl5:A	Sterol 3-beta-glucosyltransferase	0.682	3.83	0.15	0.785
	10	2iya:A	Oleandomycin Glycosyltransferase	0.65	3.55	0.192	0.745

[a] Rank is based on TM-score; [b] RMSD is the RMSD in Å between residues that are structurally aligned by TM-align; [c] the percent sequence identity in the structurally aligned region; [d] the number of structurally aligned residues divided by length of the query protein.

Table S18. Glycosyltransferase crystal structures available in the RCSB database

PDB	Kingdom	Organism	Enzyme	Gene	Uniprot
1F0K	Bacteria	*Escherichia coli* (K12)	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	murG	P17443
1IIR	Bacteria	*Amycolatopsis orientalis*	UDP-glucosyltransferase GtfB	GtfB	P96559
1NLM	Bacteria	*Escherichia coli* (K12)	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	murG	P17443
1PN3	Bacteria	*Amycolatopsis orientalis*	TDP-epi-Vancosaminyltransferase GtfA	GtfA	P96558
1PNV	Bacteria	*Amycolatopsis orientalis*	TDP-epi-Vancosaminyltransferase GtfA	GtfA	P96558
1RRV	Bacteria	*Amycolatopsis orientalis*	TDP-vancosaminyltransferase GtfD	GtfD	Q9AFC7
2ACV	Plantae	*Medicago truncatula*	Triterpene/Flavonoid Glycosyltransferase	UGT71G1	Q5IFH7
2ACW	Plantae	*Medicago truncatula*	Triterpene/Flavonoid Glycosyltransferase	UGT71G1	Q5IFH7
2C1X	Plantae	*Vitis vinifera*	Flavonoid 3-O-glucosyltransferase	UFGT	P51094
2C1Z	Plantae	*Vitis vinifera*	Flavonoid 3-O-glucosyltransferase	UFGT	P51094
2C9Z	Plantae	*Vitis vinifera*	Flavonoid 3-O-glucosyltransferase	UFGT	P51094
2HY7	Bacteria	Xanthomonas campestris pv. Campestris	beta-1,2-glucuronosyltransferase	gumK	Q8GCH2
2IYA	Bacteria	*Streptomyces antibioticus*	Oleandomycin Glycosyltransferase	OleI	Q3HTL7
2O6L	Animalia	*Homo sapiens*	UDP-Glucuronosyltransferase 2B7	UGT2B7	P16662
2P6P	Bacteria	*Streptomyces fradiae*	dTDP-D-Olivose-transferase UrdGT2	UrdGT2	Q9RPA7
2PQ6	Plantae	*Medicago truncatula*	UDP-glucosyltransferase	UGT85H2	A6XNC5
2Q6V	Bacteria	*Xanthomonas campestris* pv. Campestris	beta-1,2-glucuronosyltransferase	gumK	Q8GCH2
2VCE	Plantae	*Arabidopsis thaliana*	Hydroquinone Glucosyltransferase	UGT72B1	Q9M156
2VCH	Plantae	*Arabidopsis thaliana*	Hydroquinone Glucosyltransferase	UGT72B1	Q9M156
2VG8	Plantae	*Arabidopsis thaliana*	Hydroquinone Glucosyltransferase	UGT72B1	Q9M156
3CUY	Bacteria	Xanthomonas campestris pv. Campestris	beta-1,2-glucuronosyltransferase	gumK	Q8GCH2
3CV3	Bacteria	Xanthomonas campestris pv. Campestris	beta-1,2-glucuronosyltransferase	gumK	Q8GCH2
3D0R	Bacteria	*Micromonospora echinospora*	Calicheamicin Glycosyltransferase	CalG3	Q8KND7
3H4I	Bacteria	*Actinoplanes teichomyceticus, Amycolatopsis orientalis*	Chimeric glycosyltransferase	tcp8, gtfA	Q6ZZJ7
3H4T	Bacteria	*Actinoplanes teichomyceticus, Amycolatopsis orientalis*	Chimeric glycosyltransferases	tcp8, gtfA	P96558
3HBF	Plantae	*Medicago truncatula*	Flavonoid 3-O-glucosyltransferase	UGT78G1	A6XNC6
3HBJ	Plantae	*Medicago truncatula*	Flavonoid 3-O-glucosyltransferase	UGT78G1	A6XNC6
3IA7	Bacteria	*Micromonospora echinospora*	Calicheamicin Glycosyltransferase	CalG4	Q8KNC3
3OTG	Bacteria	*Micromonospora echinospora*	Calicheamicin Glycosyltransferase	CalG1	Q8KNF2
3OTH	Bacteria	*Micromonospora echinospora*	Calicheamicin Glycosyltransferase	CalG1	Q8KNF2
3RSC	Bacteria	*Micromonospora echinospora*	Calicheamicin Glycosyltransferase	CalG2	Q8KNE0

3S2U	Bacteria	*Pseudomonas aeruginosa*	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	murG	Q9HW01
3WC4	Plantae	*Clitoria ternatea*	anthocyanidin 3-O-glucosyltransferase	Ct3GT-A	A4F1R4
4PQG	Bacteria	*Streptococcus pneumoniae* (strain ATCC BAA-334 / TIGR4)	Glycosyltransferase GtfA	gtfA	A0A0H2URG7
4REL	Plantae	*Clitoria ternatea*	anthocyanidin 3-O-glucosyltransferase	Ct3GT-A	A4F1R4
4REM	Plantae	*Clitoria ternatea*	anthocyanidin 3-O-glucosyltransferase	Ct3GT-A	A4F1R4
4REN	Plantae	*Clitoria ternatea*	anthocyanidin 3-O-glucosyltransferase	Ct3GT-A	A4F1R4
4WHM	Plantae	*Clitoria ternatea*	anthocyanidin 3-O-glucosyltransferase	Ct3GT-A	A4F1R4
4WYI	Plantae	*Arabidopsis thaliana*	Monogalactosyldiacylglycerol synthase 1	MGD1	O81770
4X1T	Plantae	*Arabidopsis thaliana*	Monogalactosyldiacylglycerol synthase 1	MGD1	O81770
5E9T	Bacteria	*Streptococcus gordonii*	Glycosyltransferase GtfA/GtfB complex	GtfA/GtfB	Q9AET5 Q79T00
5E9U	Bacteria	*Streptococcus gordonii*	Glycosyltransferase GtfA/GtfB complex	GtfA/GtfB	Q9AET5 Q79T00
5GL5	Fungi	*Saccharomyces cerevisiae* (strain ATCC 204508 / S288c)	Sterol 3-beta-glucosyltransferase	ugt51	Q06321
5NLM	Plantae	*Polygonum tinctorium*	UDP-glucosyltransferase	PtUGT1	A0A2R2JFJ4
5TMB	Plantae	*Oryza sativa subsp. japonica*	Os70 UDP-glucosyltransferase	UGT70	Q7XT97
5TMD	Plantae	*Oryza sativa subsp. japonica*	UDP-glycosyltransferase 79	UGT79	Q7XT97
5TME	Plantae	*Oryza sativa subsp. Japonica*	UDP-glycosyltransferase 79	UGT79	Q7XT97
5U6M	Plantae	*Arabidopsis thaliana*	UDP-glycosyltransferase 74F2	UGT74F2	O22822
5U6N	Plantae	*Arabidopsis thaliana*	UDP-glycosyltransferase 74F2	UGT74F2	O22822
5U6S	Plantae	*Arabidopsis thaliana*	UDP-glycosyltransferase 74F2	UGT74F2	O22822
5V2J	Plantae	*Arabidopsis thaliana*	UDP-glycosyltransferase 74F2	UGT74F2	O22822
5V2K	Plantae	*Arabidopsis thaliana*	UDP-glycosyltransferase 74F2	UGT74F2	O22822
5XVM	Fungi	*Saccharomyces cerevisiae* (strain ATCC 204508 / S288c)	Sterol 3-beta-glucosyltransferase	ugt51	Q06321
6BK2	Plantae	*Oryza sativa subsp. japonica*	UDP-glycosyltransferase 79	UGT79	Q7XT97
6BK0	Plantae	*Oryza sativa subsp. japonica*	UDP-glycosyltransferase 79	UGT79	Q7XT97
6BK3	Plantae	*Oryza sativa subsp. japonica*	UDP-glycosyltransferase 79	UGT79	Q7XT97
6BK1	Plantae	*Oryza sativa subsp. japonica*	UDP-glycosyltransferase 79	UGT79	Q7XT97
6IJ7	Plantae	*Arabidopsis thaliana*	Flavonol 7-O-rhamnosyltransferase	UGT89C1	Q9LNE6
6IJA	Plantae	*Arabidopsis thaliana*	Flavonol 7-O-rhamnosyltransferase	UGT89C1	Q9LNE6
6IJD	Plantae	*Arabidopsis thaliana*	Flavonol 7-O-rhamnosyltransferase	UGT89C1	Q9LNE6
6IPB	Animalia	*Homo sapiens*	UDP-Glucuronosyltransferase 2B15	UGT2B15	P54855
6O86	Plantae	*Stevia rebaudiana*	UDP-glycosyltransferase 76G1	UGT76G1	Q6VAB4
6O87	Plantae	*Stevia rebaudiana*	UDP-glycosyltransferase 76G1	UGT76G1	Q6VAB4
6O88	Plantae	*Stevia rebaudiana*	UDP-glycosyltransferase 76G1	UGT76G1	Q6VAB4

Table S19. Mutant UGT1A6 structures produced for investigating effects of specific residues on compound binding

Importance	Original amino acid	Position	Changed amino acid	Purpose
Co-substrate binding	Ser	37	Ala	Examine co-substrate binding
Catalytic	His	38	Ala	Examine co-substrate and substrate binding
Catalytic	Asp	150	Ala	Examine co-substrate and substrate binding
Co-substrate and substrate binding	Arg	172	Ala	Examine co-substrate and substrate binding
Substrate binding	His	180	Ala	Examine substrate binding
Polymorphism	Thr	181	Ala	Examine substrate binding
Substrate binding	Arg	184	Ala	Examine substrate binding
Polymorphism	Arg	184	Ser	Examine substrate binding
Substrate binding	Lys	228	Ala	Examine substrate binding
Co-substrate binding	Ser	308	Ala	Examine co-substrate binding
Co-substrate binding	Trp	353	Ala	Examine co-substrate binding
Co-substrate binding	Leu	354	Ala	Examine co-substrate binding
Co-substrate binding	His	375	Ala	Examine co-substrate binding
Co-substrate binding	Gly	376	Ala	Examine co-substrate binding
Co-substrate and substrate binding	Phe	393	Ala	Examine co-substrate and substrate binding

(1) Ghemtio, L.; Soikkeli, A.; Yliperttula, M.; Hirvonen, J.; Finel, M.; Xhaard, H. SVM Classification and CoMSIA Modeling of UGT1A6 Interacting Molecules. *J. Chem. Inf. Model.* **2014**, *54* (4), 1011–1026. https://doi.org/10.1021/ci400577a.

(2) Vaidyanathan, J. B.; Walle, T. Glucuronidation and Sulfation of the Tea Flavonoid (−)-Epicatechin by the Human and Rat Enzymes. *Drug Metab. Dispos.* **2002**, *30* (8), 897–903. https://doi.org/10.1124/dmd.30.8.897.

(3) Luukkanen, L.; Taskinen, J.; Kurkela, M.; Kostiainen, R.; Hirvonen, J.; Finel, M. Kinetic Characterization of the 1a Subfamily of Recombinant Human Udp-Glucuronosyltransferases. *Drug Metab. Dispos.* **2005**, *33* (7), 1017–1026. https://doi.org/10.1124/dmd.105.004093.

(4) Soars, M. G.; Ring, B. J.; Wrighton, S. A. The Effect of Incubation Conditions on the Enzyme Kinetics of Udp-Glucuronosyltransferases. *Drug Metab. Dispos.* **2003**, *31* (6), 762–767. https://doi.org/10.1124/dmd.31.6.762.

(5) Liu, Y.; She, M.; Wu, Z.; Dai, R. The Inhibition Study of Human UDP-Glucuronosyltransferases with Cytochrome P450 Selective Substrates and Inhibitors. *J. Enzyme Inhib. Med. Chem.* **2011**, *26* (3), 386–393. https://doi.org/10.3109/14756366.2010.518965.

(6) Uchaipichat, V.; Mackenzie, P. I.; Guo, X.-H.; Gardner-Stephen, D.; Galetin, A.; Houston, J. B.; Miners, J. O. Human Udp-Glucuronosyltransferases: Isoform Selectivity and Kinetics of 4-Methylumbelliferone and 1-Naphthol Glucuronidation, Effects of Organic Solvents, and Inhibition by Diclofenac and Probenecid. *Drug Metab. Dispos. Biol. Fate Chem.* **2004**, *32* (4), 413–423. https://doi.org/10.1124/dmd.32.4.413.

(7) Manevski, N.; Troberg, J.; Svaluto-Moreolo, P.; Dziedzic, K.; Yli-Kauhaluoma, J.; Finel, M. Albumin Stimulates the Activity of the Human UDP-Glucuronosyltransferases 1A7, 1A8, 1A10, 2A1 and 2B15, but the Effects Are Enzyme and Substrate Dependent. *PLOS ONE* **2013**, *8* (1), e54767. https://doi.org/10.1371/journal.pone.0054767.

(8) Brazier-Hicks, M.; Edwards, R. Functional Importance of the Family 1 Glucosyltransferase UGT72B1 in the Metabolism of Xenobiotics in Arabidopsis Thaliana. *Plant J. Cell Mol. Biol.* **2005**, *42* (4), 556–566. https://doi.org/10.1111/j.1365-313X.2005.02398.x.

(9) Sun, D.; Sharma, A. K.; Dellinger, R. W.; Blevins-Primeau, A. S.; Balliet, R. M.; Chen, G.; Boyiri, T.; Amin, S.; Lazarus, P. Glucuronidation of Active Tamoxifen Metabolites by the Human UDP Glucuronosyltransferases. *Drug Metab. Dispos.* **2007**, *35* (11), 2006–2014. https://doi.org/10.1124/dmd.107.017145.

(10) Zielinska, A.; Lichti, C. F.; Bratton, S.; Mitchell, N. C.; Gallus-Zawada, A.; Le, V.-H.; Finel, M.; Miller, G. P.; Radominska-Pandya, A.; Moran, J. H. Glucuronidation of Monohydroxylated Warfarin Metabolites by Human Liver Microsomes and Human Recombinant UDP-Glucuronosyltransferases. *J. Pharmacol. Exp. Ther.* **2008**, *324* (1), 139–148. https://doi.org/10.1124/jpet.107.129858.

(11) Li, D.; Fournel-Gigleux, S.; Barré, L.; Mulliert, G.; Netter, P.; Magdalou, J.; Ouzzine, M. Identification of Aspartic Acid and Histidine Residues Mediating the Reaction Mechanism and the Substrate Specificity of the Human UDP-Glucuronosyltransferases 1A. *J. Biol. Chem.* **2007**, *282* (50), 36514–36524. https://doi.org/10.1074/jbc.M703107200.

(12) Benoit-Biancamano, M.-O.; Connelly, J.; Villeneuve, L.; Caron, P.; Guillemette, C. Deferiprone Glucuronidation by Human Tissues and Recombinant UDP Glucuronosyltransferase 1A6: An in Vitro Investigation of Genetic and Splice Variants. *Drug Metab. Dispos.* **2009**, *37* (2), 322–329. https://doi.org/10.1124/dmd.108.023101.

(13) Nagar, S.; Zalatoris, J. J.; Blanchard, R. L. Human UGT1A6 Pharmacogenetics: Identification of a Novel SNP, Characterization of Allele Frequencies and Functional Analysis of Recombinant Allozymes in Human Liver Tissue and in Cultured Cells. *Pharmacogenetics* **2004**, *14* (8), 487–499.

(14) Ethell, B. T.; Ekins, S.; Wang, J.; Burchell, B. Quantitative Structure Activity Relationships for the Glucuronidation of Simple Phenols by Expressed Human UGT1A6 and UGT1A9. *Drug Metab. Dispos.* **2002**, *30* (6), 734–738. https://doi.org/10.1124/dmd.30.6.734.

(15) Miners, J. O.; Bowalgaha, K.; Elliot, D. J.; Baranczewski, P.; Knights, K. M. Characterization of Niflumic Acid as a Selective Inhibitor of Human Liver Microsomal UDP-Glucuronosyltransferase 1A9: Application to the Reaction Phenotyping of Acetaminophen Glucuronidation. *Drug Metab. Dispos.* **2011**, *39* (4), 644–652. https://doi.org/10.1124/dmd.110.037036.

(16) Court, M. H.; Duan, S. X.; Moltke, L. L. von; Greenblatt, D. J.; Patten, C. J.; Miners, J. O.; Mackenzie, P. I. Interindividual Variability in Acetaminophen Glucuronidation by Human Liver Microsomes: Identification of

Relevant Acetaminophen UDP-Glucuronosyltransferase Isoforms. *J. Pharmacol. Exp. Ther.* **2001**, *299* (3), 998–1006.
(17) Mutlib, A. E.; Goosen, T. C.; Bauman, J. N.; Williams, J. A.; Kulkarni, S.; Kostrubsky, S. Kinetics of Acetaminophen Glucuronidation by UDP-Glucuronosyltransferases 1A1, 1A6, 1A9 and 2B15. Potential Implications in Acetaminophen−Induced Hepatotoxicity. *Chem. Res. Toxicol.* **2006**, *19* (5), 701–709. https://doi.org/10.1021/tx050317i.
(18) Hagenauer, B.; Salamon, A.; Thalhammer, T.; Kunert, O.; Haslinger, E.; Klingler, P.; Senderowicz, A. M.; Sausville, E. A.; Jäger, W. In Vitro Glucuronidation of the Cyclin-Dependent Kinase Inhibitor Flavopiridol by Rat and Human Liver Microsomes: Involvement of UDP-Glucuronosyltransferases 1A1 and 1A9. *Drug Metab. Dispos.* **2001**, *29* (4), 407–414.
(19) Liu, X.; Tam, V. H.; Hu, M. Disposition of Flavonoids via Enteric Recycling: Determination of the UDP-Glucuronosyltransferase Isoforms Responsible for the Metabolism of Flavonoids in Intact Caco-2 TC7 Cells Using SiRNA. *Mol. Pharm.* **2007**, *4* (6), 873–882. https://doi.org/10.1021/mp0601190.
(20) Singh, R.; Wu, B.; Tang, L.; Hu, M. Uridine Diphosphate Glucuronosyltransferase Isoform-Dependent Regiospecificity of Glucuronidation of Flavonoids. *J. Agric. Food Chem.* **2011**, *59* (13), 7452–7464. https://doi.org/10.1021/jf1041454.
(21) Ebner, T.; Burchell, B. Substrate Specificities of Two Stably Expressed Human Liver UDP-Glucuronosyltransferases of the UGT1 Gene Family. *Drug Metab. Dispos.* **1993**, *21* (1), 50–55.
(22) Tang, L.; Singh, R.; Liu, Z.; Hu, M. Structure and Concentration Changes Affect Characterization of UGT Isoform-Specific Metabolism of Isoflavones. *Mol. Pharm.* **2009**, *6* (5), 1466–1482. https://doi.org/10.1021/mp8002557.
(23) Hanioka, N.; Naito, T.; Narimatsu, S. Human UDP-Glucuronosyltransferase Isoforms Involved in Bisphenol A Glucuronidation. *Chemosphere* **2008**, *74* (1), 33–36. https://doi.org/10.1016/j.chemosphere.2008.09.053.
(24) Galijatovic, A.; Otake, Y.; Walle, U. K.; Walle, T. Extensive Metabolism of the Flavonoid Chrysin by Human Caco-2 and Hep G2 Cells. *Xenobiotica* **1999**, *29* (12), 1241–1256. https://doi.org/10.1080/004982599237912.
(25) Walle, T.; Otake, Y.; Galijatovic, A.; Ritter, J. K.; Walle, U. K. Induction of UDP-Glucuronosyltransferase UGT1A1 by the Flavonoid Chrysin in the Human Hepatoma Cell Line Hep G2. *Drug Metab. Dispos.* **2000**, *28* (9), 1077–1082.
(26) Iwuchukwu, O. F.; Ajetunmobi, J.; Ung, D.; Nagar, S. Characterizing the Effects of Common UDP Glucuronosyltransferase (UGT) 1A6 and UGT1A1 Polymorphisms on Cis- and Trans-Resveratrol Glucuronidation. *Drug Metab. Dispos.* **2009**, *37* (8), 1726–1732. https://doi.org/10.1124/dmd.109.027391.
(27) Holmquist, G. L. Opioid Metabolism and Effects of Cytochrome P450. *Pain Med.* **2009**, *10* (suppl_1), S20–S29. https://doi.org/10.1111/j.1526-4637.2009.00596.x.
(28) Aprile, S.; Grosso, E. D.; Grosa, G. Identification of the Human UDP-Glucuronosyltransferases Involved in the Glucuronidation of Combretastatin A-4. *Drug Metab. Dispos.* **2010**, *38* (7), 1141–1146. https://doi.org/10.1124/dmd.109.031435.
(29) Doerge, D. R.; Chang, H. C.; Churchwell, M. I.; Holder, C. L. Analysis of Soy Isoflavone Conjugation In Vitro and in Human Blood Using Liquid Chromatography-Mass Spectrometry. *Drug Metab. Dispos.* **2000**, *28* (3), 298–307.
(30) Liang, S.-C.; Ge, G.-B.; Liu, H.-X.; Zhang, Y.-Y.; Wang, L.-M.; Zhang, J.-W.; Yin, L.; Li, W.; Fang, Z.-Z.; Wu, J.-J.; et al. Identification and Characterization of Human UDP-Glucuronosyltransferases Responsible for the In Vitro Glucuronidation of Daphnetin. *Drug Metab. Dispos.* **2010**, *38* (6), 973–980. https://doi.org/10.1124/dmd.109.030734.
(31) Fernandes, I.; Marques, C.; Évora, A.; Cruz, L.; Freitas, V. de; Calhau, C.; Faria, A.; Mateus, N. Pharmacokinetics of Table and Port Red Wine Anthocyanins: A Crossover Trial in Healthy Men. *Food Funct.* **2017**, *8* (5), 2030–2037. https://doi.org/10.1039/C7FO00329C.
(32) Zeng, X.; Shi, J.; Zhao, M.; Chen, Q.; Wang, L.; Jiang, H.; Luo, F.; Zhu, L.; Lu, L.; Wang, X.; et al. Regioselective Glucuronidation of Diosmetin and Chrysoeriol by the Interplay of Glucuronidation and Transport in UGT1A9-Overexpressing HeLa Cells. *PLOS ONE* **2016**, *11* (11), e0166239. https://doi.org/10.1371/journal.pone.0166239.
(33) Itäaho, K.; Court, M. H.; Uutela, P.; Kostiainen, R.; Radominska-Pandya, A.; Finel, M. Dopamine Is a Low-Affinity and High-Specificity Substrate for the Human UDP-Glucuronosyltransferase 1A10. *Drug Metab. Dispos.* **2009**, *37* (4), 768–775. https://doi.org/10.1124/dmd.108.025692.
(34) Lautala, P.; Ethell, B. T.; Taskinen, J.; Burchell, B. The Specificity of Glucuronidation of Entacapone and Tolcapone by Recombinant Human Udp-Glucuronosyltransferases. *Drug Metab. Dispos.* **2000**, *28* (11), 1385–1389.

(35) Ghosal, A.; Hapangama, N.; Yuan, Y.; Achanfuo-Yeboah, J.; Iannucci, R.; Chowdhury, S.; Alton, K.; Patrick, J. E.; Zbaida, S. Identification of Human Udp-Glucuronosyltransferase Enzyme(s) Responsible for the Glucuronidation of Ezetimibe (Zetia). *Drug Metab. Dispos.* **2004**, *32* (3), 314–320. https://doi.org/10.1124/dmd.32.3.314.

(36) Illingworth, N. A.; Boddy, A. V.; Daly, A. K.; Veal, G. J. Characterization of the Metabolism of Fenretinide by Human Liver Microsomes, Cytochrome P450 Enzymes and UDP-Glucuronosyltransferases. *Br. J. Pharmacol.* **2011**, *162* (4), 989–999. https://doi.org/10.1111/j.1476-5381.2010.01104.x.

(37) Otake, Y.; Hsieh, F.; Walle, T. Glucuronidation versus Oxidation of the Flavonoid Galangin by Human Liver Microsomes and Hepatocytes. *Drug Metab. Dispos.* **2002**, *30* (5), 576–581. https://doi.org/10.1124/dmd.30.5.576.

(38) Brand, W.; Boersma, M. G.; Bik, H.; Hil, E. F. H. den; Vervoort, J.; Barron, D.; Meinl, W.; Glatt, H.; Williamson, G.; Bladeren, P. J. van; et al. Phase II Metabolism of Hesperetin by Individual UDP-Glucuronosyltransferases and Sulfotransferases and Rat and Human Tissue Samples. *Drug Metab. Dispos.* **2010**, *38* (4), 617–625. https://doi.org/10.1124/dmd.109.031047.

(39) Gagné, J.-F.; Montminy, V.; Belanger, P.; Journault, K.; Gaucher, G.; Guillemette, C. Common Human UGT1A Polymorphisms and the Altered Metabolism of Irinotecan Active Metabolite 7-Ethyl-10-Hydroxycamptothecin (SN-38). *Mol. Pharmacol.* **2002**, *62* (3), 608–617. https://doi.org/10.1124/mol.62.3.608.

(40) Hanioka, N.; Ozawa, S.; Jinno, H.; Ando, M.; Saito, Y.; Sawada, J. Human Liver UDP-Glucuronosyltransferase Isoforms Involved in the Glucuronidation of 7-Ethyl-10-Hydroxycamptothecin. *Xenobiotica* **2001**, *31* (10), 687–699. https://doi.org/10.1080/00498250110057341.

(41) Li, Y.; Lu, L.; Wang, L.; Qu, W.; Liu, W.; Xie, Y.; Zheng, H.; Wang, Y.; Qi, X.; Hu, M.; et al. Interplay of Efflux Transporters with Glucuronidation and Its Impact on Subcellular Aglycone and Glucuronide Disposition: A Case Study with Kaempferol. *Mol. Pharm.* **2018**, *15* (12), 5602–5614. https://doi.org/10.1021/acs.molpharmaceut.8b00782.

(42) Picard, N.; Ratanasavanh, D.; Prémaud, A.; Meur, Y. L.; Marquet, P. Identification of the Udp-Glucuronosyltransferase Isoforms Involved in Mycophenolic Acid Phase Ii Metabolism. *Drug Metab. Dispos.* **2005**, *33* (1), 139–146. https://doi.org/10.1124/dmd.104.001651.

(43) Jiang, W.; Hu, M. Mutual Interactions between Flavonoids and Enzymatic and Transporter Elements Responsible for Flavonoid Disposition via Phase II Metabolic Pathways. *RSC Adv.* **2012**, *2* (21), 7948–7963. https://doi.org/10.1039/C2RA01369J.

(44) Hengstmann, J. H.; Goronzy, J. Pharmacokinetics Of3H-Phenylephrine in Man. *Eur. J. Clin. Pharmacol.* **1982**, *21* (4), 335–341. https://doi.org/10.1007/BF00637623.

(45) Miksits, M.; Maier-Salamon, A.; Vo, T. P. N.; Sulyok, M.; Schuhmacher, R.; Szekeres, T.; Jäger, W. Glucuronidation of Piceatannol by Human Liver Microsomes: Major Role of UGT1A1, UGT1A8 and UGT1A10. *J. Pharm. Pharmacol.* **2010**, *62* (1), 47–54. https://doi.org/10.1211/jpp.62.01.0004.

(46) Boersma, M. G.; van der Woude, H.; Bogaards, J.; Boeren, S.; Vervoort, J.; Cnubben, N. H. P.; van Iersel, M. L. P. S.; van Bladeren, P. J.; Rietjens, I. M. C. M. Regioselectivity of Phase II Metabolism of Luteolin and Quercetin by UDP-Glucuronosyl Transferases. *Chem. Res. Toxicol.* **2002**, *15* (5), 662–670. https://doi.org/10.1021/tx0101705.

(47) Sun, D.; Jones, N. R.; Manni, A.; Lazarus, P. Characterization of Raloxifene Glucuronidation: Potential Role of UGT1A8 Genotype on Raloxifene Metabolism In Vivo. *Cancer Prev. Res. (Phila. Pa.)* **2013**, *6* (7), 719–730. https://doi.org/10.1158/1940-6207.CAPR-12-0448.

(48) Krishnaswamy, S.; Duan, S. X.; Moltke, L. L. V.; Greenblatt, D. J.; Sudmeier, J. L.; Bachovchin, W. W.; Court, M. H. Serotonin (5-Hydroxytryptamine) Glucuronidation in Vitro : Assay Development, Human Liver Microsome Activities and Species Differences. *Xenobiotica* **2003**, *33* (2), 169–180. https://doi.org/10.1080/0049825021000048809.

(49) Krishnaswamy, S.; Hao, Q.; Al-Rohaimi, A.; Hesse, L. M.; Moltke, L. L. von; Greenblatt, D. J.; Court, M. H. UDP Glucuronosyltransferase (UGT) 1A6 Pharmacogenetics: II. Functional Impact of the Three Most Common Nonsynonymous UGT1A6 Polymorphisms (S7A, T181A, and R184S). *J. Pharmacol. Exp. Ther.* **2005**, *313* (3), 1340–1346. https://doi.org/10.1124/jpet.104.081968.

(50) Ishii, Y.; Koba, H.; Kinoshita, K.; Oizaki, T.; Iwamoto, Y.; Takeda, S.; Miyauchi, Y.; Nishimura, Y.; Egoshi, N.; Taura, F.; et al. Alteration of the Function of the UDP-Glucuronosyltransferase 1A Subfamily by Cytochrome P450 3A4: Different Susceptibility for UGT Isoforms and UGT1A1/7 Variants. *Drug Metab. Dispos.* **2014**, *42* (2), 229–238. https://doi.org/10.1124/dmd.113.054833.

(51) Sten, T.; Bichlmaier, I.; Kuuranne, T.; Leinonen, A.; Yli-Kauhaluoma, J.; Finel, M. UDP-Glucuronosyltransferases (UGTs) 2B7 and UGT2B17 Display Converse Specificity in Testosterone and Epitestosterone Glucuronidation, Whereas UGT2A1 Conjugates Both Androgens Similarly. *Drug Metab. Dispos.* **2009**, *37* (2), 417–423. https://doi.org/10.1124/dmd.108.024844.

(52) Brill, S. S.; Furimsky, A. M.; Ho, M. N.; Furniss, M. J.; Li, Y.; Green, A. G.; Green, C. E.; Iyer, L. V.; Bradford, W. W.; Kapetanovic, I. M. Glucuronidation of Trans-Resveratrol by Human Liver and Intestinal Microsomes and UGT Isoforms. *J. Pharm. Pharmacol.* **2006**, *58* (4), 469–479. https://doi.org/10.1211/jpp.58.4.0006.

(53) Ruefer, C. E.; Gerhäuser, C.; Frank, N.; Becker, H.; Kulling, S. E. In Vitro Phase II Metabolism of Xanthohumol by Human UDP-Glucuronosyltransferases and Sulfotransferases. *Mol. Nutr. Food Res.* **2005**, *49* (9), 851–856. https://doi.org/10.1002/mnfr.200500057.