SUPPORTING INFORMATION

for

Biosynthesis of Drug Glucuronide Metabolites in the Budding Yeast *Saccharomuces cerevisiae*

By

Shinichi Ikushiro, Miyu Nishikawa, Yuuka Masuyama,Tadashi Shouji, , Miharu Fujii , Masahiro Hamada, Noriyuki Nakajima, Moshe Finel, Kaori Yasuda, Masaki Kamakura, and Toshiyuki Sakaki

Supporting information Figure legends

Figure S1. Chemical structure of the glucuronidation substrates used in this study.

Figure S2. Strategy for construction vectors for co-expressing different mammalian UGTs with the rat UGDH in budding yeast.

Upper part, left column represents the generation of the genome integrating expression vector with a *Not* I site for cloning the rat UGDH (pAUR-N). Upper part, right column represents the generation of rat UGDH. Once the pAUR-UGDH was ready, insertion of the selected UGT gene, through ligation into the Not1 site, results in the generation of co-expression vector of rat UGDH and the selected UGT, generally called pAUR-UGDH/UGT. Halftone boxes in the UGDH or UGT genes in the figure represent the promoter and terminator sequences of glyceraldehydes-3-phosphate dehydrogenase that were derived from *Zygosaccharomyces rouxii* (Ikezawa , *et al*., 2003)[25].

Figure S3. Immunoblot analysis of rat UGDH in yeast cells.

Immunodetection was performed using the anti-rat UGDH antibody. Lanes 1 and 2 represent yeast cell extracts from control, untransfected cells, and from rat UGDH genes-transfected yeast strains, respectively. The arrowhead indicates the mature form of rat UGDH protein.

Figure S4. HPLC chromatograms of UDP-sugars in yeast extracts from rat UGDH expressing strain. (A) UDP-sugars from yeast extracts of a strain that expresses rat UGDH. (B) UDP-sugars from wild type of yeast strain, i.e. a control strain that does not express rat UGDH. (C) UDP-glucose (peak no. 1) and UDPGA (peak no. 2) standards.

Figure S5. Mass spectroscopy measurement of purified diclofenac acyl glucuronide.

LC-MS analysis of diclofenac acyl glucuronide was performed on a Quadrupole 6120 and LC system 1260 Infinity (Agilent Technology, Santa Clara, CA, USA). The nitrogen gas flow rate, spray current, and voltages were adjusted to give maximum sensitivity for the glucuronide. The mass spectrometer was operated in positive ion mode with a scan range from 100 to 500 amu. Choromatographic separation was achieved using a Poroshell 120 EC-C18 (4.6 X 50 mm, 2.7μm, Agilent Technology, Santa Clara, CA, USA) column in conjunction with a gradient solvent system. Elution was performed using a mobile phase containing 0.1% formic acid in water and acetonitrile. The initial conditions consisted of water in 0.1% formic acid. After maintaining the initial conditions for 2 min, the proportion of acetonitrile was increased to 100% from 2 to 10 min and held for additional 1min. The solvent was delivered at a flow rate of 0.5 mL/min. A protonated molecular ion

of the glucuronide (*m/z* 472 $[M+1]^+$) and the aglycone ion (296 $[472\text{-}C_6H_8O_6]^+$) produced from the loss of the glucuronic acid due to in-source fragmentation confirmed the formation of mono-glucuronide of diclofenac.

Figure S6. NMR spectra measurements of the purified diclofenac acyl glucuronide.

^{1}H- and ^{13}C-NMR spectra of diclofenac acyl glucuronide were recorded on a Bruker Biospin AVANCE II 400 (400 and 100 MHz) in $(CD_3)_2CO$. Chemical shifts and coupling constants are reported in δ values (ppm) and hertz, respectively. (A) ^{1}H NMR [400MHz, $(CD_3)_2CO$]: 3.51, 3.60 and 3.67 (3H, 3m, 2-H, 3-H and 4-H), 3.95 (2H, ABq, *J*=15.6 and 9.0 Hzs, $ArCH_2CO$), 4.03 (1H, d, *J*=9.6 Hz, 5-H), 5.66 (1H, d, *J*=8.0 Hz, 1-H, anomeric), 6.47 (1H, d, *J*=8.0 Hz, ArH), 6.82 (1H, br, s, NH), 6.95 (1H, m, ArH), 7.11-7.19 (2H, m, ArH), 7.31 (1H, m, ArH) and 7.47 (2H, m, ArH). (B) ^{13}C NMR [100MHz, $(CD_3)_2CO$]:38.4, 72.5, 73.4, 76.7, 76.9, 95.8 (anomeric), 118.6, 122.9, 125.2, 125.8, 128.9,129.9, 130.7, 132.0, 138.9,143.8, 169.8 and 171.7. These NMR spectra of enzymatically synthesized diclofenac acyl glucuronide were identical to chemically synthesized one, which shows 1β-*O*-acyl glucuronide of diclofenac (Bowkett, *et al.* 2007) [32].

Table S1. List of accession no. of the UGT genes used in this study

Species	UGTs	GenBank accession no.
Human	UGT1A1	M57899
	UGT1A3	M84127
	UGT1A4	M57951
	UGT1A5	NM 019078
	UGT1A6	M39570
	UGT1A7	U89507
	UGT1A8	NM 019076
	UGT1A9	AF056188
	UGT1A10	U89508
	UGT2A1	AJ006054
	UGT2B4	Y00317
	UGT2B7	J05425
	UGT2B10	NM 001075
	UGT2B15	U08854
	UGT2B17	U59209
Rat	UGT1A1	U20551
	UGT1A2	M34007
	UGT1A3	AY435131
	UGT1A5	D38069(E1)
	UGT1A6	J02612
	UGT1A7	D38062(E1)
	UGT2B1	M13506
	UGT2B3	M31109
	UGT2B6	M33746
	UGT2B12	U06273
Mouse	Ugt1a1	L02333
	Ugt1a5	AY227196
	Ugt1a6a	U16818
	Ugt1a6b	AY227198
	Ugt1a9	L27122
	Ugt2b1	BC027200
	Ugt2b5	X06358

	Ugt2b34	AI788959
Porcine	UGT1A3a	AK235866
	UGT1A3b	AK235866*(variant)
	UGT2B18-like	100516628**
	UGT2B31-like	100623255**
	UGT2C1-like	100515394**
Bovine	UGT1A6	AB008677

*Porcine UGT1A3b is a variant of UGT1A3a (AK235866) with replacement of amino acids; K78Q, T82I, R95G, L97F, I180L, K221R, F227S, V228F and V269I.

** The information of porcine UGT gene is from the database resource of Kyoto Encyclopedia of Genes and Genomes (KEGG)

Table S2. Amino acid sequence alignment of the putative N-terminal signal peptide of the human UGTs 1A1, 1A4, 1A7 and 1A9

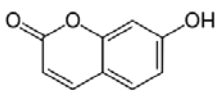
Human UGT	Amino acid sequence alignment of putative signal peptide
UGT1A1	MAVESQGGRPLVLGLLLCVLGPVVSHAG
UGT1A4	MARGLQVPLPRLATGLLLLLSVQPWAESG
UGT1A7	MARAGWTGLLPLYVCLLLTCGFAKAG
UGT1A9	MACTGWTSPLPLCVCLLLTCGFAEAG

Table S3. Glucuronidation rates of budding yeast strains co-expressing rat UGDH and different mammalian UGT enzymes

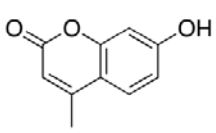
Species	UGT	Specific production rate (μmol/day/g wet weight)		
		7HC	Diclofenac	11α-hydroxyprogesterone
Human	UGT1A1	0.49	ND	ND
	UGT1A3	ND	ND	0.001
	UGT1A4	ND	ND	0.002
	UGT1A5	ND	ND	ND
	UGT1A6	10.26	0.02	ND
	UGT1A7	4.25	0.04	0.008
	UGT1A8	1.17	0.04	0.006
	UGT1A9	2.54	0.09	0.144
	UGT1A10	0.88	ND	ND
	UGT2A1	10.43	0.03	0.006
	UGT2B4	ND	ND	0.001
	UGT2B7	0.03	0.01	0.015
	UGT2B10	ND	ND	0.001
	UGT2B15	0.45	0.08	ND
	UGT2B17	0.09	ND	ND
Rat	UGT1A1	0.02	ND	ND
	UGT1A2	0.62	0.11	0.14
	UGT1A3	0.02	ND	ND
	UGT1A5	0.01	ND	ND
	UGT1A6	0.08	ND	ND
	UGT1A7	1.5	ND	ND
	UGT2B1	7.41	4.19	ND
	UGT2B3	1.67	0.32	0.011
	UGT2B6	0.07	0.08	0.283
	UGT2B12	0.02	ND	ND
Mouse	Ugt1a1	0.02	ND	ND
	Ugt1a5	0.05	0.07	0.001
	Ugt1a6a	2.68	ND	ND
	Ugt1a6b	5.57	ND	ND
	Ugt1a9	1.16	0.02	ND

	Ugt2b1	3.58	1.85	0.002
	Ugt2b5	0.33	ND	0.001
	Ugt2b34	0.02	ND	ND
Porcine	UGT1A3a	1.86	1.36	0.014
	UGT1A3b	0.42	0.89	0.006
	UGT2B18-like	0.16	ND	ND
	UGT2B31-like	0.15	ND	ND
	UGT2C1-like	0.07	0.02	0.022
Bovine	UGT1A6	2.85	ND	ND

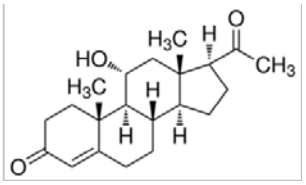
Each value of specific production rate was an average of duplicate experiments



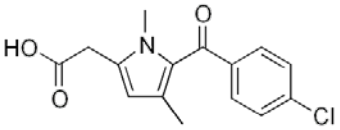
7-Hydroxycoumarin



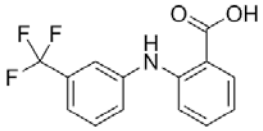
4-Methyl umbelliferone



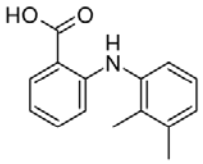
11α-Hydroxyprogesterone



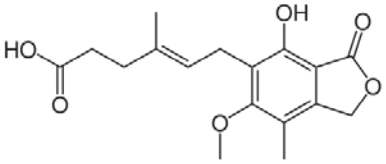
Zomepirac



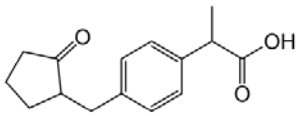
Flufenamic acid



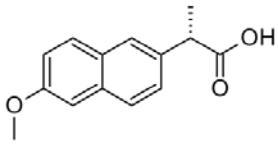
Mefenamic acid



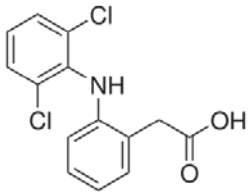
Mycophenolic acid



Loxoprofen



Naproxen



Diclofenac

Figure S1

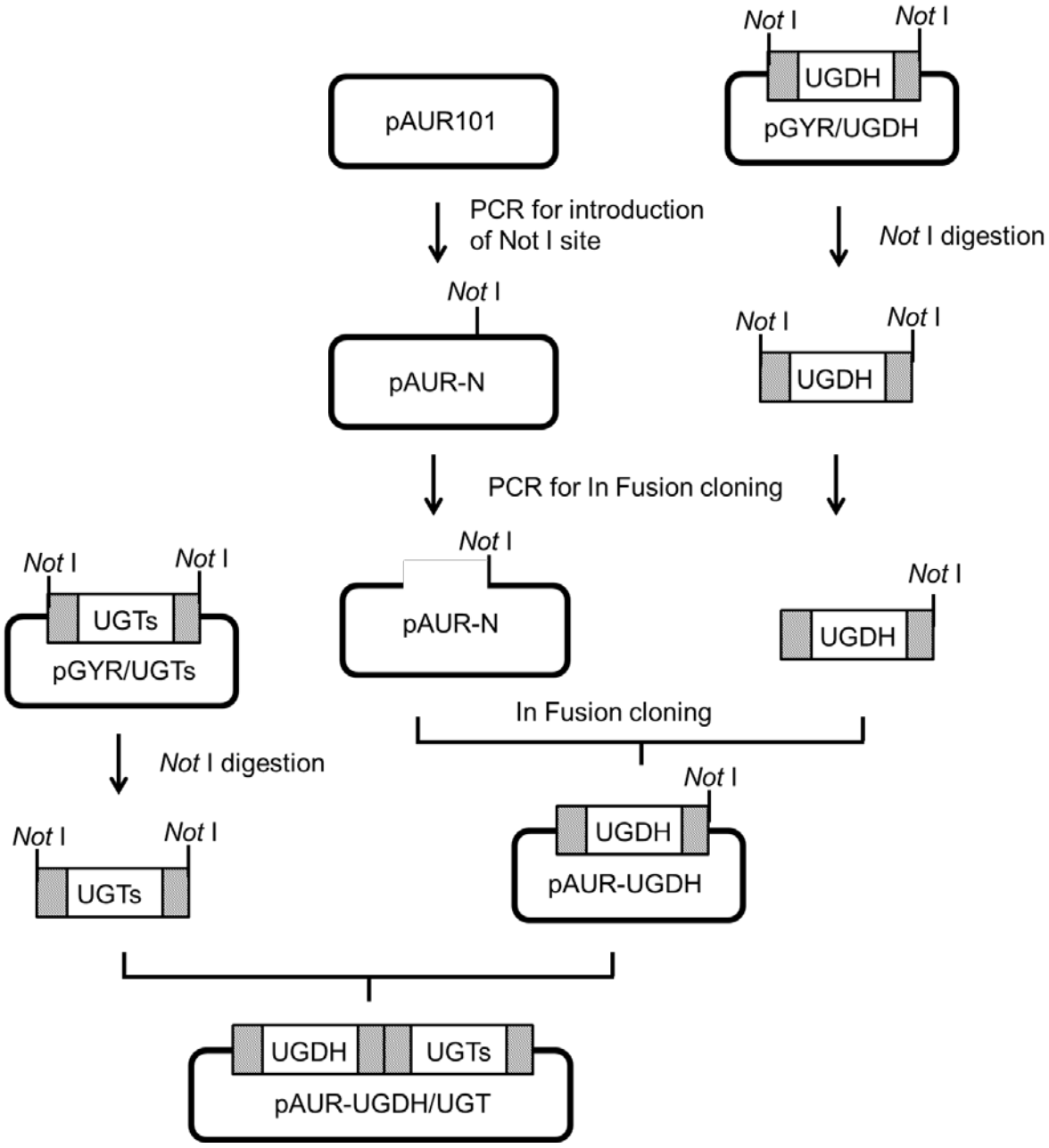

Not I
Not I
UGDH
pGYR/UGDH
pAUR101
PCR for introduction
of Not I site
Not I digestion
Not I
pAUR-N
Not I
Not I
UGDH
PCR for In Fusion cloning
Not I
pAUR-N
Not I
UGDH
Not I
Not I
UGTs
pGYR/UGTs
In Fusion cloning
Not I digestion
Not I
UGDH
pAUR-UGDH
Not I
Not I
UGTs
UGDH
UGTs
pAUR-UGDH/UGT

Figure S2

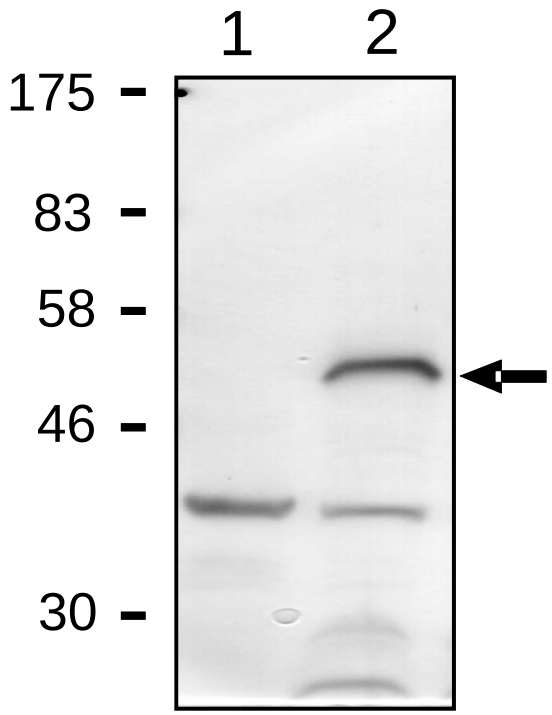


Figure S3

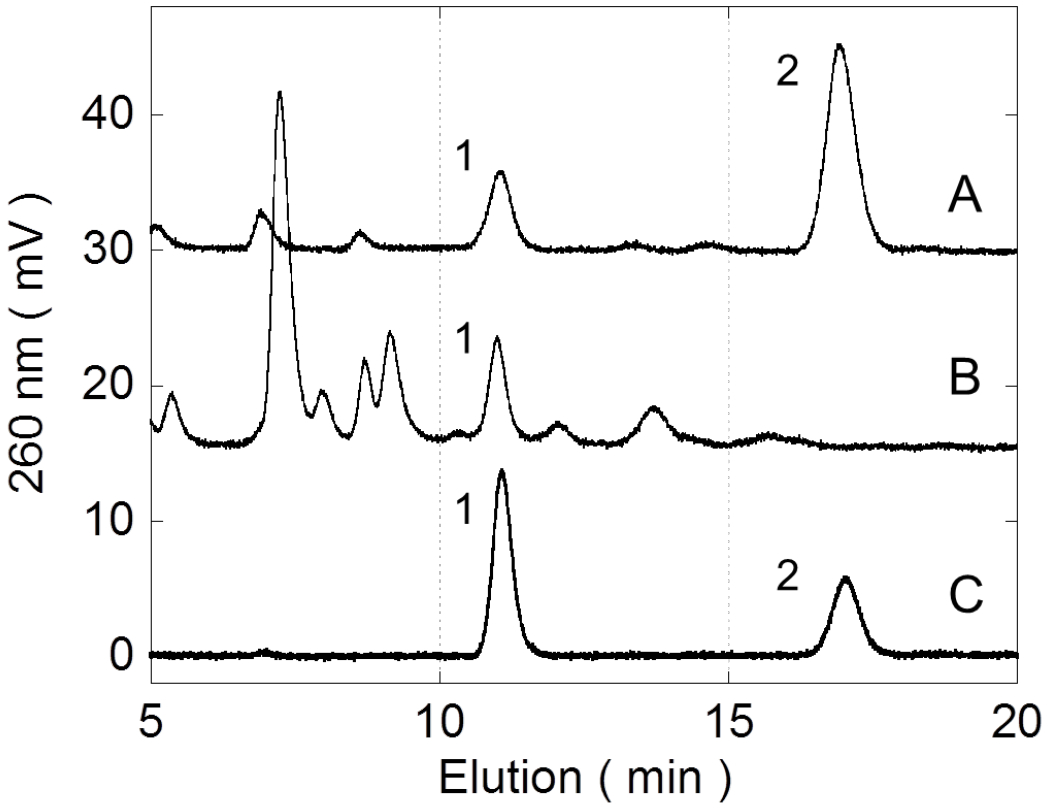


Figure S4.

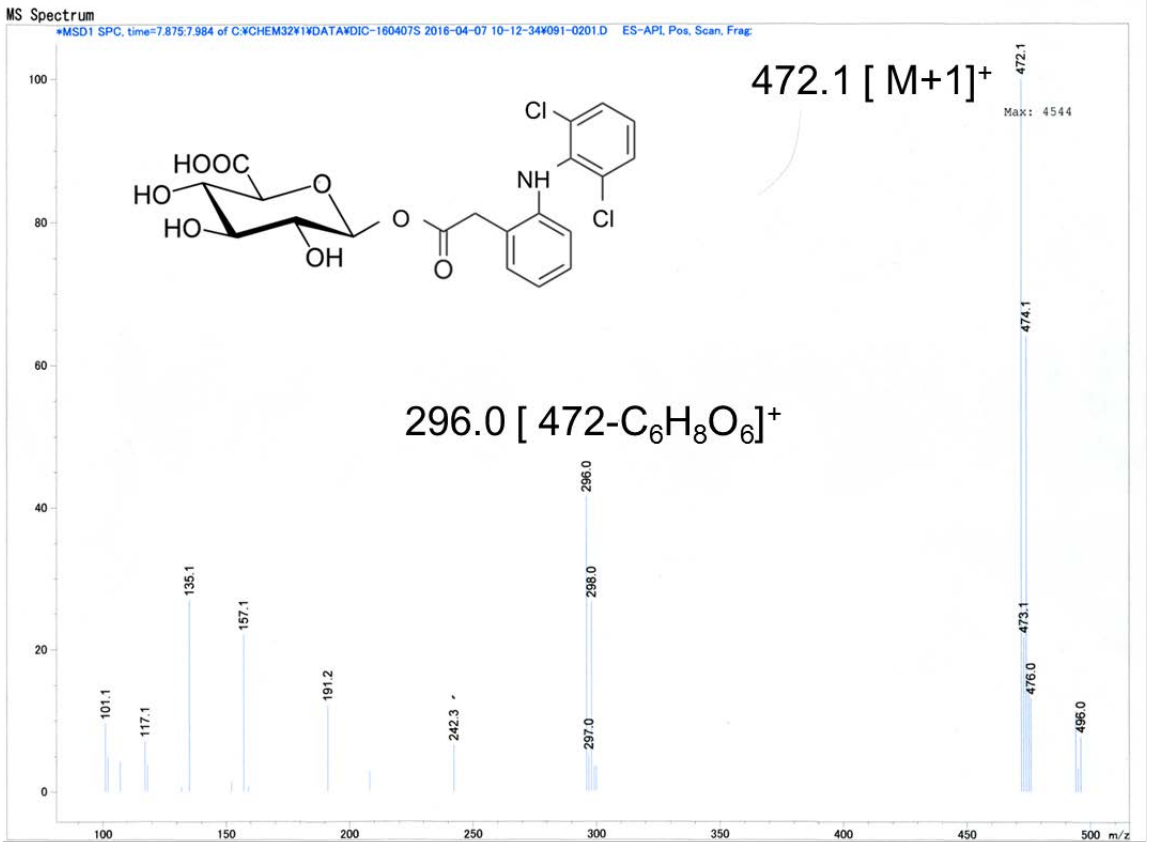

MS Spectrum
*MSD1 SPC, time=7.875:7.984 of C:¥CHEM32¥1¥DATA¥DIC-160407S 2016-04-07 10-12-34¥091-0201.D ES-API, Pos, Scan, Frag:
472.1 [M+1]+
Max: 4544
296.0 [472-C6H8O6]+
HOOC
HO
HO
OH
O
O
NH
Cl
Cl
472.1
474.1
473.1
476.0
496.0
296.0
298.0
297.0
135.1
157.1
191.2
242.3
101.1
117.1
100
80
60
40
20
0
100
150
200
250
300
350
400
450
500 m/z

Figure S5

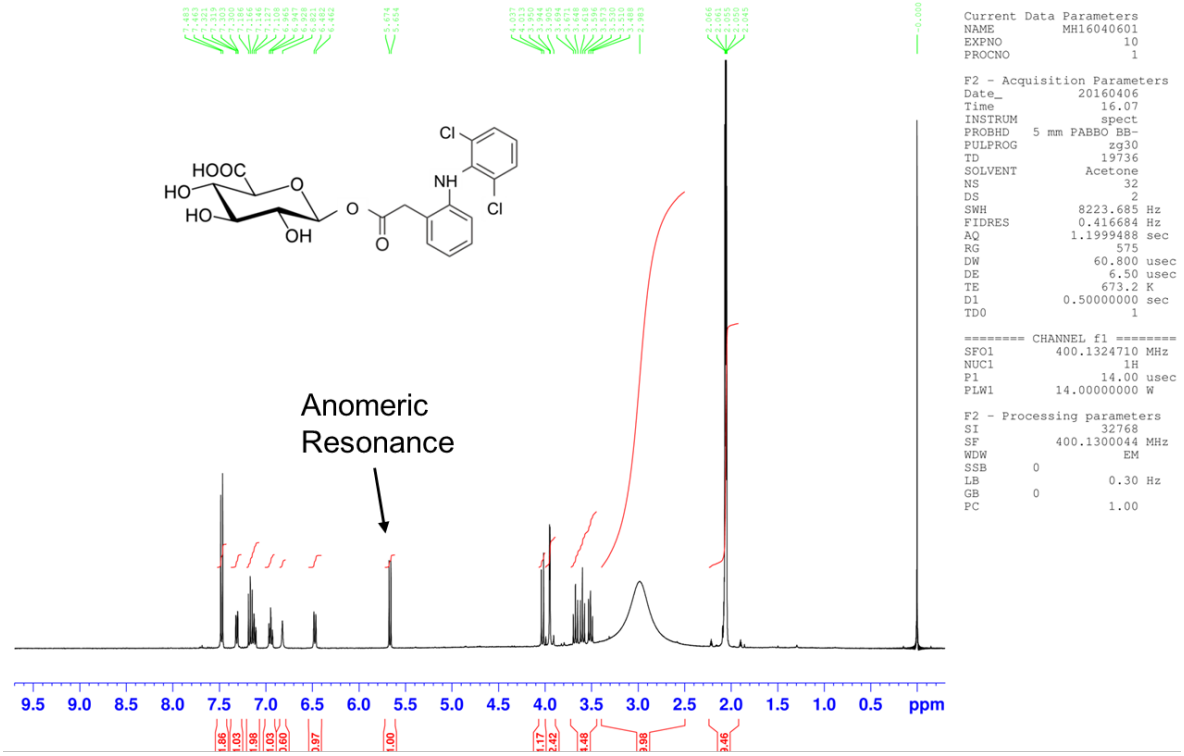

Anomeric
Resonance
HOOC
HO
HO
OH
O
O
NH
Cl
Cl
9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 ppm
1.86 1.03 1.98 1.03 2.60 0.97 1.00 1.17 2.42 4.48 3.98 3.46
Current Data Parameters
NAME MH16040601
EXPNO 10
PROCNO 1
F2 - Acquisition Parameters
Date_ 20160406
Time 16.07
INSTRUM spect
PROBHD 5 mm PABBO BB-
PULPROG zg30
TD 19736
SOLVENT Acetone
NS 32
DS 2
SWH 8223.685 Hz
FIDRES 0.416684 Hz
AQ 1.1999488 sec
RG 575
DW 60.800 usec
DE 6.50 usec
TE 673.2 K
D1 0.50000000 sec
TD0 1
======== CHANNEL f1 ========
SFO1 400.1324710 MHz
NUC1 1H
P1 14.00 usec
PLW1 14.00000000 W
F2 - Processing parameters
SI 32768
SF 400.1300044 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

Figure S6A

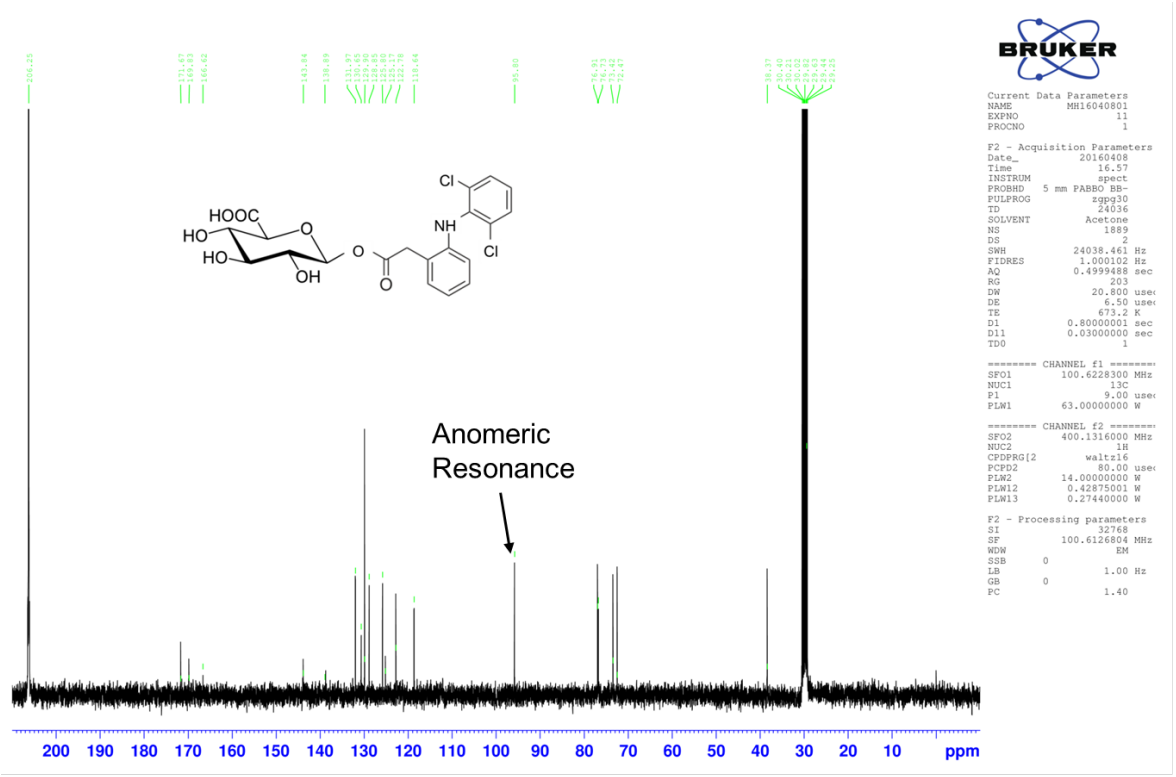

Anomeric
Resonance
HOOC
HO
HO
OH
O
O
NH
Cl
Cl
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 ppm

Figure S6B