Supporting Information for: Probing the Regiospecificity of Enzyme-Catalyzed Steroid Glycosylation

Maoquan Zhou[†], Yanpeng Hou[†], Adel Hamza[‡], Chang-Guo Zhan[‡], Tim S. Bugni[†,*] and Jon S. Thorson[‡,*]

[†]Pharmaceutical Sciences Division, School of Pharmacy, University of Wisconsin, Madison , Wisconsin 53705, United States; [‡]Center for Pharmaceutical Research and Innovation, University of Kentucky, College of Pharmacy , 789 South Limestone Street, Lexington, Kentucky 40536-0596, United States.

tbugni@pharmacy.wisc.edu; jsthorson@uky.edu

Table of Contents

General methods. All chemicals and reagents were purchased from Sigma-Aldrich, unless otherwise stated. Mass spectra were acquired on a Bruker MaXis high resolution quadrupole time of flight mass spectrometer. Compounds were characterized by NMR with a Bruker Avance III 600 MHz spectrometer with a 1.7 mm ^{1}H {^{13}C/^{15}N} cryogenic probe and Varian ^{UNITY}INOVA 500 MHz spectrometer. ^{1}H and ^{13}C chemical shifts were referenced to internal solvent resonances. Multiplicities are indicated by s (singlet), d (doublet), t (triplet), q (quartet), qn (quintet), m (multiplet) and br (broad). Chemical shifts are reported in parts per million (ppm) and coupling constants *J* are given in Hz.

Protein expression and purification. A single colony of *E. coli* BL21(DE3)pLysS (Stratagene, La Jolla, CA, USA) transformed with the pET28a-based *oleD-ASP* expression vector[1] was used to inoculate 3 mL LB medium supplemented with 50 μg/mL kanamycin and cultured overnight at 37 °C with shaking (250 rpm). The starter culture was then transferred to 1 L LB medium supplemented with 50 μg/mL kanamycin and grown at 37 °C with shaking (250 rpm) until the OD_{600} reached ~0.7. Isopropyl β-D-thiogalactoside (IPTG) was subsequently added to a final concentration of 0.4 mM and the culture was incubated at 28 ºC for approximately 18 hours with shaking at 250 rpm. Cell pellets were collected by centrifugation at 10,000 *g* and 4 °C for 20 min and the supernatant discarded. Cell pellets were resuspended in 10 mL chilled lysis buffer (20 mM phosphate buffer, 0.5 M NaCl, 10 mM imidazole, pH 7.4) and were lysed by sonication (8 pulses of 40 seconds each) in an ice bath. The cell debris was removed by centrifugation at 10,000 *g* and 4 °C for 20 min. The cleared supernatant was immediately applied to 3 mL of nickel nitrilotriacetic acid (Ni-NTA) resin (QIAgen Valencia, CA, USA) pre-equilibrated with wash buffer (20 mM phosphate buffer, pH 7.4, 0.3 M NaCl, 10 mM imidazole). Protein was allowed to bind for 30 min at 4 °C with gentle agitation and the resin was subsequently

(1) (a) Williams, G. J.; Thorson, J. S. *Nat. Protoc*. **2008**, *3*, 357–362. (b) Williams, G. J.; Goff, R. D.; Zhang, C.; Thorson, J. S. *Chem. Biol.* **2008**, *15*, 393–401. (c) Williams, G. J.; Zhang, C.; Thorson, J. S. *Nat. Chem. Biol*. **2007**, *3*, 657–662.

washed with 4 x 50 mL wash buffer. Finally, the enzyme was eluted from the resin via incubation with 2 mL wash buffer containing 250 mM imidazole for 15 min at 4 °C with gentle agitation. The purified protein was applied to a PD-10 desalting column (Amersham Biosciences, Piscataway, NJ, USA) equilibrated with 50 mM Tris-HCl (pH 8.0) and eluted as described by the manufacturer. Protein aliquots were immediately flash frozen in liquid nitrogen and stored at -80 °C. Protein purity was confirmed by SDS-PAGE to be >95% and protein concentration for all studies was determined using the Bradford Protein Assay Kit from Bio-Rad (Hercules, CA, USA).

Enzymatic reaction. Reactions were conducted in a final volume of 1 mL and contained 600 μg of purified OleD ASP, 2.5 mM UDP-glucose, 50 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, and 1 mM of aglycon. Two separate control reactions that withheld either enzyme or UDP-glucose were performed in parallel. Reactions were allowed to proceed at 25 ^{o}C for ~16 hours, subsequently frozen and lyophilized, the debris was resuspended in 2 mL of ice cold MeOH, and filtered. One portion of each clarified reaction mixture was analyzed by analytical reverse-phase HPLC [Phenomenex 250 mm x 4.6 mm Gemini 5μ C18 column (Torrance, CA, USA); flow rate: 1 mL/min; gradient of solvents A (0.1% trifluoroacetic acid/H_2O) and B (100% acetonitrile): (a) 0–20 min, 10–75 % B; (b) 20–21 min, 75-100% B; (c) 21–26 min, 100% B; (d) 26–29 min, 100–10% B; and (e) 29–35 min, 10 % B; A_{220} detection]. HPLC peak areas were integrated using the Star Chromatography Workstation software (Varian, Palo Alto, CA, USA) and the total percent conversion calculated as a percent of the total peak area of substrate and product (Figure S1). The remaining portion of each sample was concentrated to 150 μL for LC-UV/MS-SPE-NMR analysis.

LC-UV/MS-SPE-NMR Analysis. The LC-UV/MS-SPE-NMR system consisted of an Agilent 1200 chromatograph composed of a quaternary pump, a photodiode array detector, and an autosampler, a Bruker micrOTOF-Q II mass spectrometer equipped with an electrospray ionization source and operated via a 5 : 95 flow splitter, a Knauer Smartline 100 pump for postcolumn flow dilution, two Spark Holland Prospect 2 SPE units, a Gilson 215 liquid handler for automated filling of 1.7 mm NMR tubes from the Prospect 2 device configured for SPE cartridge elution, and the Bruker Avance III 600 MHz NMR spectrometer equipped with a 1.7 mm ^{1}H {^{13}C/^{15}N} cryogenic probe. The separation was performed using a Phenomenex 250 mm x 4.6 mm Gemini 5μ C18 column (Torrance, CA, USA) under the following conditions: flow rate: 1 mL/min; gradient of solvents A (H_2O) and B (100% acetonitrile): (a) 0–20 min, 10–75% B; (b) 20–21 min, 75-100% B; (c) 21–26 min, 100% B; (d) 26–29 min, 100–10% B; and (e) 29–35 min, 10% B. The injection volume was 50 μL. The chromatography was monitored by MS and PDA detector simultaneously with thresholds set to trigger SPE trapping (Table S1). The HPLC eluent was diluted with H_2O (3.0 mL/min) prior to trapping on pre-conditioned and equilibrated Spark Holland 2 × 10 mm GP-resin SPE cartridges. The same analyte eluted from the HPLC column was trapped into the same SPE cartridge from three chromatographic runs. The cartridges were dried with pressurized N_2 for 30 min, and the desired products were eluted with 30 μL of CD_3CN into 1.7 mm NMR tubes. The NMR experiments were performed on a Bruker Avance III 600 MHz spectrometer with a 1.7mm ^{1}H {^{13}C/^{15}N} cryogenic probe. Digitoxigenin Glc (**2**): ^{1}H NMR (600 MHz, CD_3CN) δ 5.85 (m, 1H), 4.99 (dd, *J* = 18.3, 1.8 Hz, 1H), 4.85 (dd, *J* = 18.3, 1.8 Hz, 1H), 4.32 (d, *J* = 7.8 Hz, 1H, sugar H1), 4.02 (m, 1H, C3-H), 3.74 (ddd, *J* = 11.4, 6.6, 1.8 Hz, 1H), 3.58 (m, 1H), 3.40 (m, 1H), 3.36 – 3.19 (m, 4H), 3.08 (ddd, *J* = 9.0, 7.8, 3.6 Hz, 1H), 2.81 (dd, *J* = 9.6, 6.0 Hz, 1H), 2.78 – 2.72 (m, 1H), 2.34 (m, 1H), 2.19 – 2.06 (m, 2H), 1.94 – 1.77 (m, 3H), 1.78 – 1.54 (m, 6H), 1.53 – 1.40 (m, 3H), 1.36 – 1.16 (m, 3H), 0.96 (s, 3H), 0.86 (s, 3H). Gitoxigenin Glc (**11**): ^{1}H NMR (600 MHz, CD_3CN) δ 5.90 (m, 1H), 5.03 (dd, *J* = 18.1, 1.8 Hz, 1H), 4.97 (dd, *J* = 18.1, 1.8 Hz, 1H), 4.55 – 4.49 (m, 1H, C16-H), 4.32 (d, *J* = 7.7 Hz, 1H, sugar H1), 4.03 (m, 1H, C3-H), 3.73 (m, 1H), 3.58 (m, 1H), 3.43 – 3.27 (m, 3H), 3.24 (m, 2H), 3.11 – 3.05 (m, 1H), 3.03 (d, *J* = 7.5 Hz, 1H), 2.86 (m, 1H), 2.76 (m, 1H), 2.55 (dd, *J* = 14.7, 7.6 Hz, 1H), 2.10 – 2.06 (m, 1H), 1.91 – 1.81 (m, 1H), 1.81 – 1.14 (m, 12H), 0.95 (s, 3H), 0.92 (s, 3H). Scillarenin Glc (**12**): ^{1}H NMR (600 MHz, CD_3CN) δ 7.91 (dd, *J* = 9.8, 2.5 Hz, 1H), 7.33 (m, 1H), 6.22 (d, *J* = 9.8 Hz, 1H), 5.42 (m, 1H), 4.41 (d, *J* = 7.8 Hz, 1H, sugar H1), 4.19 (m, 1H, C3-H), 3.73 (m, 1H), 3.60 (m, 1H), 3.44 (m, 1H), 3.60 (m, 1H), 3.30 (m, 2H), 3.24 (m, 1H), 3.07 (m, 1H), 2.51 (m, 1H), 2.46 (m, 1H), 2.15 – 2.07 (m, 5H), 1.94 – 1.26 (m, 10H), 1.06 (s, 3H), 0.72 (s, 3H). Bufotalin mono-Glc (**13**): ^{1}H NMR (600 MHz, CD_3CN) δ 8.14 (dd, *J* = 9.7, 2.6 Hz, 1H), 7.31 (d, *J* = 2.4 Hz, 1H), 6.14 (d, *J* = 9.7 Hz, 1H), 5.46 (m, 1H), 4.32 (d, *J* = 7.8 Hz, 1H, sugar H1), 4.03 (m, 1H, C3-H), 3.74 (m, 1H), 3.59 (m, 1H), 3.43 (m, 1H), 3.38 (m, 1H), 3.30 (m, 2H), 3.24 (m, 2H), 3.08 (m, 1H), 2.91 (d, *J* = 8.8 Hz, 1H), 2.79 (m, 1H), 2.69 (dd, *J* = 15.5, 8.9 Hz, 1H), 2.63 (s, 1H), 1.92 – 1.84 (m, 1H), 1.83 (s, 3H), 1.81 – 1.35 (m, 10H), 1.30 – 1.15 (m, 2H), 0.95 (s, 3H), 0.76 (s, 3H). Bufotalin di-Glc (**14**): ^{1}H NMR (600 MHz, CD_3CN) δ 8.14 (dd, *J* = 9.6, 2.4 Hz, 1H), 7.31 (m, 1H), 6.14 (d, *J* =

9.6 Hz, 1H), 5.46 (m, 1H), 4.57 (d, *J* = 7.8 Hz, 1H, sugar 2 H1), 4.46 (d, *J* = 7.8 Hz, 1H, sugar 1 H1), 4.39 (m, 1H), 4.21 (m, 1H), 4.03 (m, 1H, C3-H), 3.75 (m, 1H), 3.66 (m, 1H), 3.59 (m, 1H), 3.49 (m, 2H), 3.40 (m, 1H), 3.37 – 3.15 (m, 5H), 2.91 (d, *J* = 9.0 Hz, 1H), 2.82 (m, 1H), 2.68 (dd, *J* = 15.6, 9.0 Hz, 1H), 2.64 (m, 1H), 1.91 – 1.84 (m, 1H), 1.83 (s, 3H), 1.81 – 1.73 (m, 2H), 1.69 – 1.14 (m, 11H), 0.94 (s, 3H), 0.76 (s, 3H). Cinobufagin mono-Glc (**15**): ^{1}H NMR (600 MHz, CD_3CN) δ 7.91 (m, 1H), 7.26 (m, 1H), 6.17 (d, *J* = 9.8 Hz, 1H), 5.47 (dd, *J* = 9.4, 1.5 Hz, 1H), 4.32 (d, *J* = 7.8 Hz, 1H, sugar H1), 4.04 (m, 1H, C3-H), 3.74 (ddd, *J* = 11.8, 6.4, 2.0 Hz, 1H), 3.71 (m, 1H), 3.59 (m, 1H), 3.43 (m, 1H), 3.37 (m, 1H), 3.33 – 3.27 (m, 2H), 3.24 (m, 2H), 3.08 (ddd, *J* = 8.4, 8.4, 3.5 Hz, 1H), 2.87 (d, *J* = 9.3 Hz, 1H), 2.78 (m, 1H), 2.07 (m, , 1H), 1.93 – 1.71 (m, 5H), 1.85 (s, 1H), 1.63 (m, 1H), 1.59 – 1.40 (m, 5H), 1.37 – 1.27 (m, 1H), 1.20 (m, 1H), 1.05 (m,1H), 0.99 (s, 3H), 0.81 (s, 3H). Cinobufagin di-Glc (**16**): ^{1}H NMR (600 MHz, CD_3CN) δ 7.97 (m, 1H), 7.27 (m, 1H), 6.16 (m, 1H), 5.47 (dd, *J* = 9.2, 1.4 Hz, 1H), 4.57 (d, *J* = 7.8 Hz, 1H, sugar 2 H1), 4.46 (d, *J* = 7.7 Hz, 1H, sugar 1 H1), 4.38 (m, 1H), 4.20 (m, 1H), 4.03 (m, 1H, C3-H), 3.75 (m, 1H), 3.70 (m, 1H), 3.68 – 3.55 (m, 3H), 3.54 – 3.04 (m, 6H), 2.86 (d, *J* = 9.3 Hz, 1H), 2.84 – 2.77 (m, 1H), 2.65 (d, *J* = 9.3 Hz, 1H), 2.11 – 2.00 (m, 1H), 1.91 – 1.40 (m, 11H), 1.85 (s, 3H), 1.36 – 1.26 (m, 1H), 1.19 (m, 1H), 1.06 (m, 1H), 0.97 (s, 3H), 0.80 (s, 3H).

Preparative scale reaction of digoxigenin. Digoxigenin (**3**) (10 mg, 25.6 μmol) was dissolved in DMSO (0.65 mL) and diluted with buffer solution (50 mM Tris HCl, 5 mM $MgCl_2$, pH 8.0, 26 mL total volume). UDP-Glc (62 mg, 102.4 μmol) was added along with OleD ASP (16 mg). After 24 hours of gentle agitation at room temperature, the reaction was diluted with 26 mL of ice cold water and loaded onto a 12 mL HLB column. After drying with N_2 gas, the HLB column was eluted with 20 mL methanol. The methanolic solution was concentrated and subjected to silica gel flash chromatography using 9% CH_2Cl_2/MeOH to afford digoxigenin 12-*O*-*β*-D-Glc **10** (13 mg, 23.5 μmol, 92%). Digoxigenin 12-*O*-*β*-D-Glc (**10**): ^{1}H NMR (1:4 $CD_3OD/CDCl_3$, 500 MHz): δ 5.89 (m, 1H), 4.94 (dd, *J* = 18.0, 1.5 Hz, 1H), 4.87 (dd, *J* = 18.0, 1.5 Hz, 1H), 4.26 (d, *J* = 8.0 Hz, 1H, sugar H1), 4.02 (m, 1H, C3-H), 3.88 (dd, *J* = 11.5, 2.0 Hz, 1H, sugar H6), 3.75 (dd, *J* = 11.5, 5.5 Hz, 1H, sugar H6), 3.61 (dd, *J* = 9.0, 6.5 Hz, 1H, C17-H), 3.47 (dd, *J* = 12.0, 4.0 Hz, 1H, C12-H), 3.39 (dd, *J* = 9.0, 8.5 Hz, 1H, sugar H3), 3.32 (m, 1H, sugar H4), 3.28 (ddd, *J* = 11.5, 5.5, 2.0 Hz, 1H, sugar H5), 3.18 (dd, *J* = 9.0, 8.0 Hz, 1H, sugar H2), 2.15 (m, 1H), 1.92 (m, 2H), 1.78 (m, 4H), 1.63 (m, 3H), 1.52 (m, 3H), 1.30 – 1.20 (m, 6H), 0.98 (s, 3H), 0.83 (s, 3H); ^{13}C NMR (1:4 $CD_3OD/CDCl_3$, 125 MHz): 176.7, 176.1, 116.1, 100.3 (sugar C1), 85.3, 82.1(C12), 76.1, 75.4, 74.3, 73.3, 70.4, 66.0 (C3), 61.8, 54.5, 49.4, 44.7, 40.9, 35.7, 35.0, 32.7, 32.3, 31.7, 29.2, 27.3, 27.1, 26.2, 25.9, 23.2, 21.3; ESIHRMS Calcd for $[C_{29}H_{44}O_{10}+Na]^+$: 575.2827, Found: 575.2823.

Cytotoxicity assay. Testing was performed by the Keck-UWCCC Small Molecule Screening Facility (Madison, WI). Cell lines were maintained and harvested as previously reported, along with compound handling and assay set up.[2] Cells were plated in 50 μL volumes in 384-well clear bottom tissue culture plates. Serial dilutions of 30 mM DMSO compound stock solutions were done in 96-well plates using a BioTek Precision XS liquid handler (Winooski, VT) to a concentration 100 x greater than that of the most dilute assay. Final dilutions were performed in a 384-well plate in quadruplicate using a Beckman-Coulter Biomek FX liquid handler with a 384 channel pipetting head (Fullerton, CA) and were stored at -20 °C when not in use. Compounds were then added to the culture plates by the Biomek FX handler and were incubated at 37 °C for 72 hours. The calcein AM reagent (acetoxymethyl ester; 30 μL, 10 μM) was then added, the cells were incubated for 30 min at 37 °C, and plates were read for fluorescent emission (535 nm). Cell titer-glo reagent (15 μL; Promega Corp., Madison, WI) was added and the plates incubated for 10 min at room temperature with gentle agitation to lyse the cells. Each plate was reexamined for luminescence to verify inhibition. IC_{50} values for cytotoxicity were determined using XLfit 4.2 as previously reported.[2]

Molecular modeling. Various molecular orientations and multiple conformations of each ligand in the OleD (PDBID 2IYF)[3] active site were assessed by using OMEGA (Open Eye Scientific Software). The conformations of each ligand were first filtered using Sabre program and the docking poses of the ligands were performed

(2) (a) Ahmed, A.; Peters, N. R.; Fitzgerald, M. K.; Watson Jr., J. A.; Hoffmann, F. M.; Thorson, J. S. *J. Am. Chem. Soc.* **2006**, *128*, 14224–14225. (b) Langenhan, J. M.; Peters, N. R.; Guzei, I. A.; Hoffmann, F. M.; Thorson, J. S. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 12305–12310.

(3) Bolam, D. N.; Roberts, S.; Proctor, M. R.; Turkenburg, J. P.; Dodson, E. J.; Martinez-Fleites, C.; Yang, M.; Davis, B. G.; Davies, G. J.; Gilbert, H. J. *Proc. Natl. Acad. Sci. USA* **2007**,*104*, 5336–5341.

using Fred software.[4] The docking strategy exhaustively docked/scored all possible positions of each ligand in the binding site. The rigid docking roughly consisted of two steps - shape and optimization. During the shape fitting, the ligand was placed into a 0.5 Å resolution grid box encompassing all active-site atoms (including hydrogen atoms) using smooth Gaussian potential.[5] Two optimization filters were subsequently processed - rigid-body optimization and optimization of the ligand pose in the dihedral angle space. The pose ensemble was then filtered to first reject poses that did not have sufficient shape complimentarily with the active site of the protein followed by rejection of those lacking at least one heavy atom hydrogen bond with the His19 imidazole. In separate docking runs, the binding poses of the ligand structure were refined by MD simulations followed by MM-GBSA calculations using Sander module from Amber11 package[6] as previously described.[7]

Specifically, the OleD-ligand binding complex was neutralized by adding appropriate counter ions and was solvated in a rectangular box of TIP3P water molecules with a minimum solute-wall distance of 10 Å. The solvated systems were energy-minimized and carefully equilibrated. These systems were gradually heated from T = 10 K to T = 298.15 K in 50 ps before running a MD simulation. The MD simulations were performed with a periodic boundary condition in the NPT ensemble at T = 298.15 K with Berendsen temperature coupling[8] and constant pressure (P=1 atm) with isotropic molecule-based scaling. A time step of 2.0 fs was used, with a cutoff of 12 Å for the non-bonded interactions, and the SHAKE algorithm was employed to keep all bonds involving hydrogen atoms rigid.[9] Long-range interactions were handled using the particle mesh Ewald (PME) algorithm.[10] During the energy minimization and MD simulation, only the ligand and residue side chains in the binding pocket were permitted to move to prevent any changes in the OleD structure due to the presence of residues in the loops on the top of the protein active site. A residue-based cutoff of 12 Å was utilized for non-covalent interactions. MD simulations were then carried out for ~4.0 ns. During the simulations, the coordinates of the system were collected every 1 ps. The last 100 snapshots of the simulated structure of the MD trajectory were used to perform the MM-GBSA calculations.

(4) (a) Hamza, A.; Wei, N. N.; Zhan, C. G. *J. Chem. Inf. Model.* **2012**, *52*, 963–974. (b) Fred, V. OpenEye Scientific Software, Inc., Santa Fe, NM, USA. *www.eyesopen.com* **2007**.
(5) McGann, M. R.; Almond, H. R.; Nicholls, A.; Grant, J. A.; Brown, F. K. *Biopolymers* **2003,** *68*, 76–90.
(6) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A., AMBER 10. *San Francisco, CA: University of California* **2008.**
(7) (a) Hamza, A.; Zhao, X.; Tong, M.; Tai, H. H.; Zhan, C. G. *Bioorg. Med. Chem.* **2011**, *19*, 6077–6086. (b) Bargagna-Mohan, P.; Paranthan, R. R.; Hamza, A.; Dimova, N.; Trucchi, B.; Srinivasan, C.; Elliott, G. I.; Zhan, C. G.; Lau, D. L.; Zhu, H. Y.; Kasahara, K.; Inagaki, M.; Cambi, F.; Mohan, R. *J. Biol. Chem.* **2010**, *285*, 7657–7669.
(8) Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
(9) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
(10) Darden, T.; York, D.; Pedersen, l. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

Table S1. LC-UV/MS-SPE-NMR analysis results

compounds*	product retention time (min) †	conversion rate (%)‡	calculated mass $[M+H]^+$	detected mass $[M+H]^+$
2	12.8	20	537.3	537.3
10	10.1	95	553.3	553.3
11	11.4	3	553.3	553.3
12	13.3	7	547.3	547.3
13	12.7	28	607.3	607.3
14	11.5	19	769.3	769.3
15	14.2	39	605.3	605.3
16	12.4	5	767.3	767.3

*See **Figure 1** for structures.
†See **LC-UV/MS-SPE-NMR Analysis** for HPLC parameters.
‡Percent conversions were determined by HPLC and calculated by dividing the integrated area of the glycosylated product by the sum of the integrated area of the product plus the integrated area of the remaining acceptor.

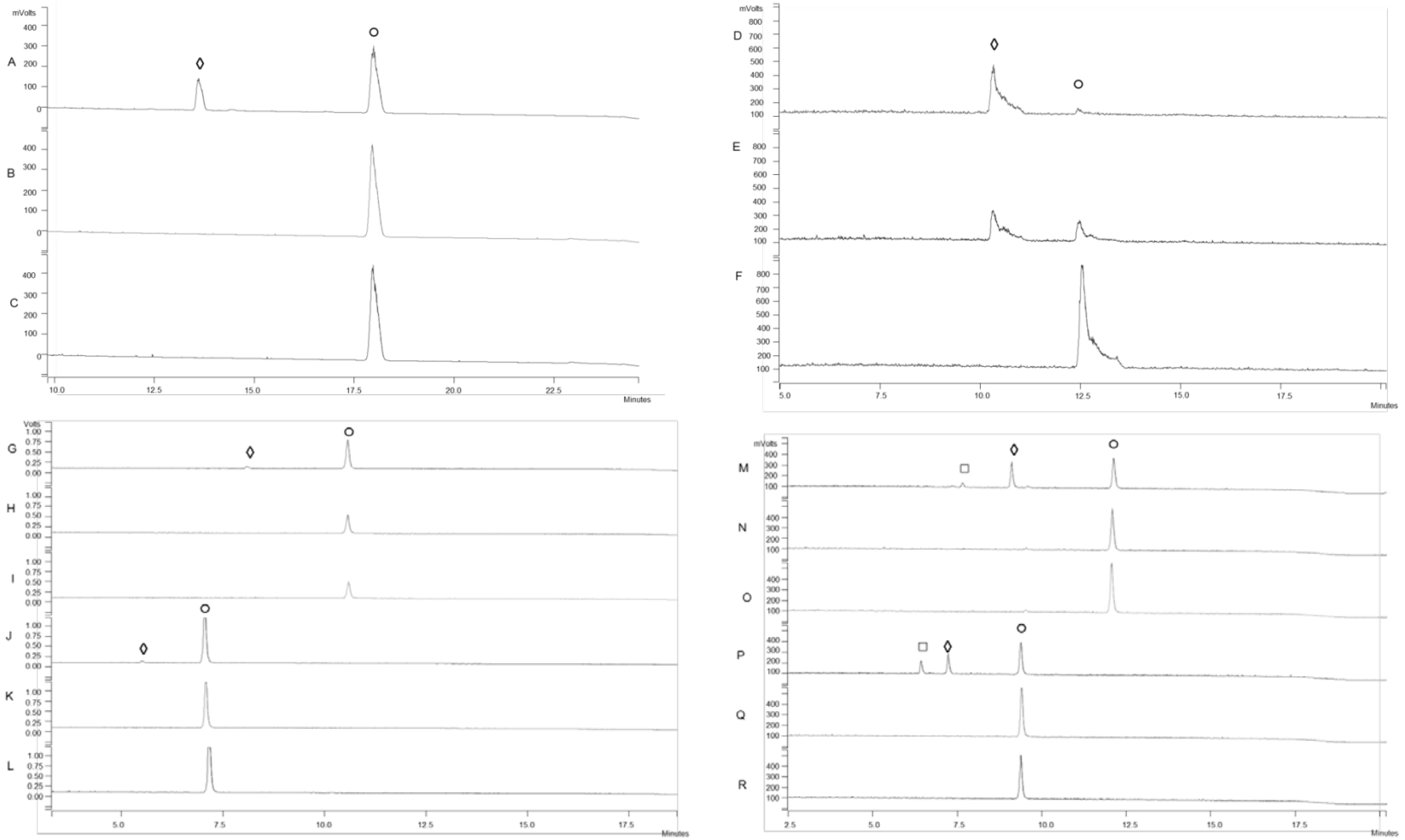


Figure S1. HPLC chromatograms of the enzymatic reactions. (A) digitoxigenin (**1**) real reaction; (B) digitoxigenin (**1**) control reaction without ASP; (C) digitoxigenin (**1**) control reaction without UDP-Glc; (D) digoxigenin (**3**) real reaction; (E) digoxigenin (**3**) real reaction plus digoxigenin; (F) digoxigenin (**3**) control; (G) gitoxigenin (**4**) real reaction; (H) gitoxigenin (**4**) control reaction without ASP; (I) gitoxigenin (**4**) control reaction without UDP-Glc; (J) scillarenin (**7**) real reaction; (K) scillarenin (**7**) control reaction without ASP; (L) scillarenin (**7**)control reaction without UDP-Glc; ; (M) cinobufagin (**9**) real reaction; (N) cinobufagin (**9**) control reaction without ASP; (O) cinobufagin (**9**) control reaction without UDP-Glc. (P) bufotalin (**8**) real reaction; (Q) bufotalin (**8**) control reaction without ASP; (R) bufotalin (**8**) control reaction without UDP-Glc. Open circles (○) denote aglycon, open diamonds (◊) denote the respective mono-Glc, open square (□) denote di-Glc.

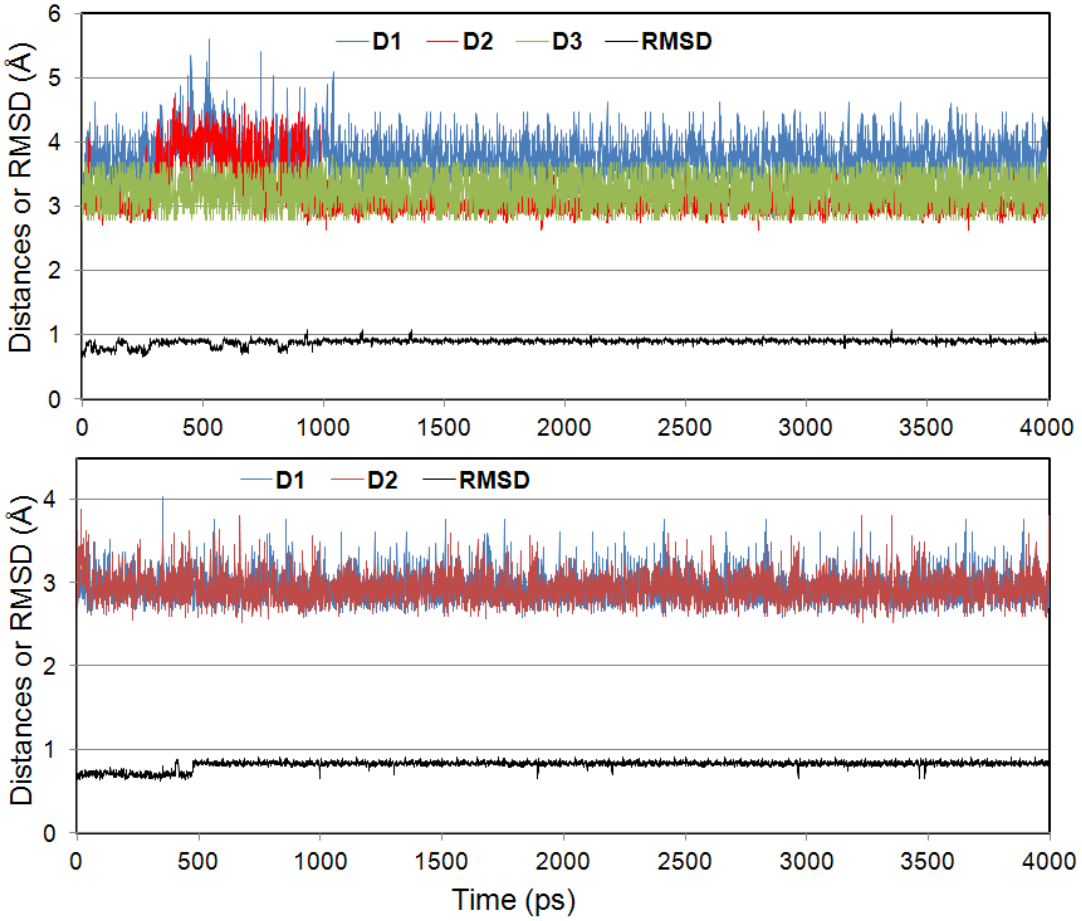


Figure S2. Plots of the MD-simulated internuclear distances and RMSD for atomic positions of the ligand *versus* the simulation time for OleD binding with compounds **1** and **3**. **Top:** Trace D1 represents the internuclear distance between the oxygen of the hydroxyl group at position C3 and the NE2 atom of the His19 side chain. Traces D2 and D3 represent the internuclear distance between the oxygen atoms of the furane group and the side chains of Trp149 and Tyr161. **Bottom:** Trace D1 represents the internuclear distance between the oxygen of the hydroxyl group at position C12 and the NE2 atom of His19 side chain. Trace D2 represents the internuclear distance between the oxygen atom of the hydroxyl group at position C3 and the hydroxyl group of Tyr114 side chain.

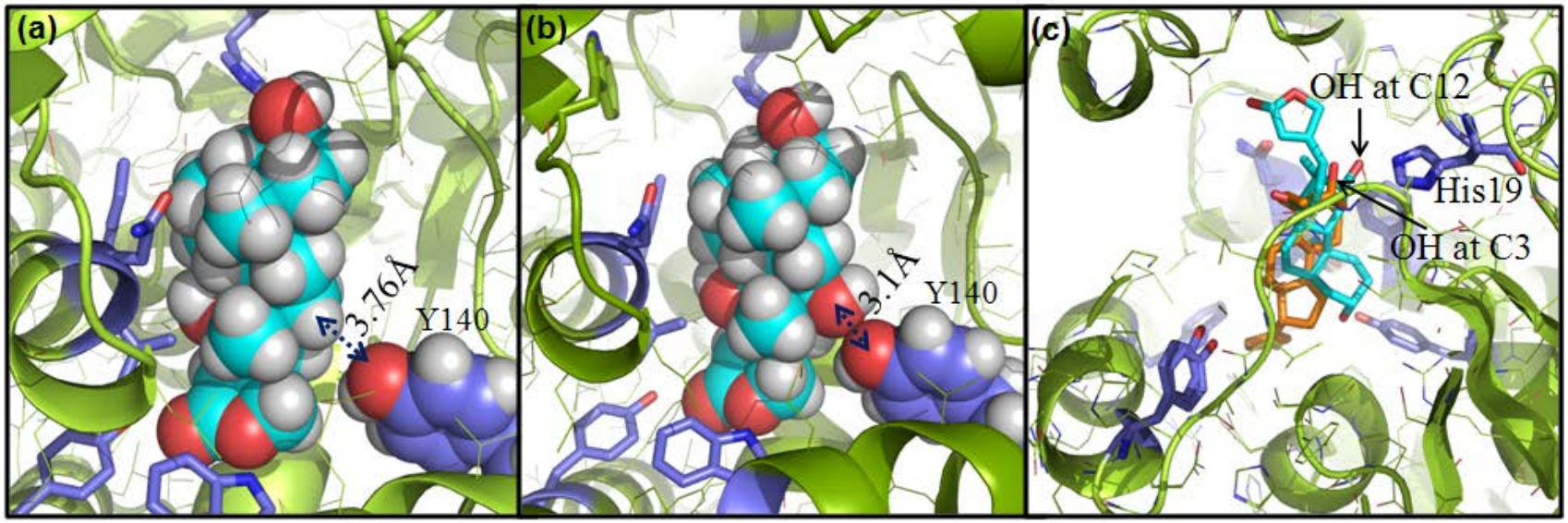


Figure S3. (a) Ball and stick view of the OleD-**1** complex. **(b)** Ball and stick view of a putative OleD-**3** complex wherein 3 adopts the same orientation as 1 in panel **(a)**. The short interatomic distance (~3.1Å) between the **3** C12 OH oxygen atom and the Tyr140 side chain OH prohibits this orientation due to the steric and electrostatic repulsions. **(c)** Ribbon view of the superposed OleD-**1** and OleD-**3** complexes. The ligands **1** and **3** are displayed in brown and light blue, respectively.

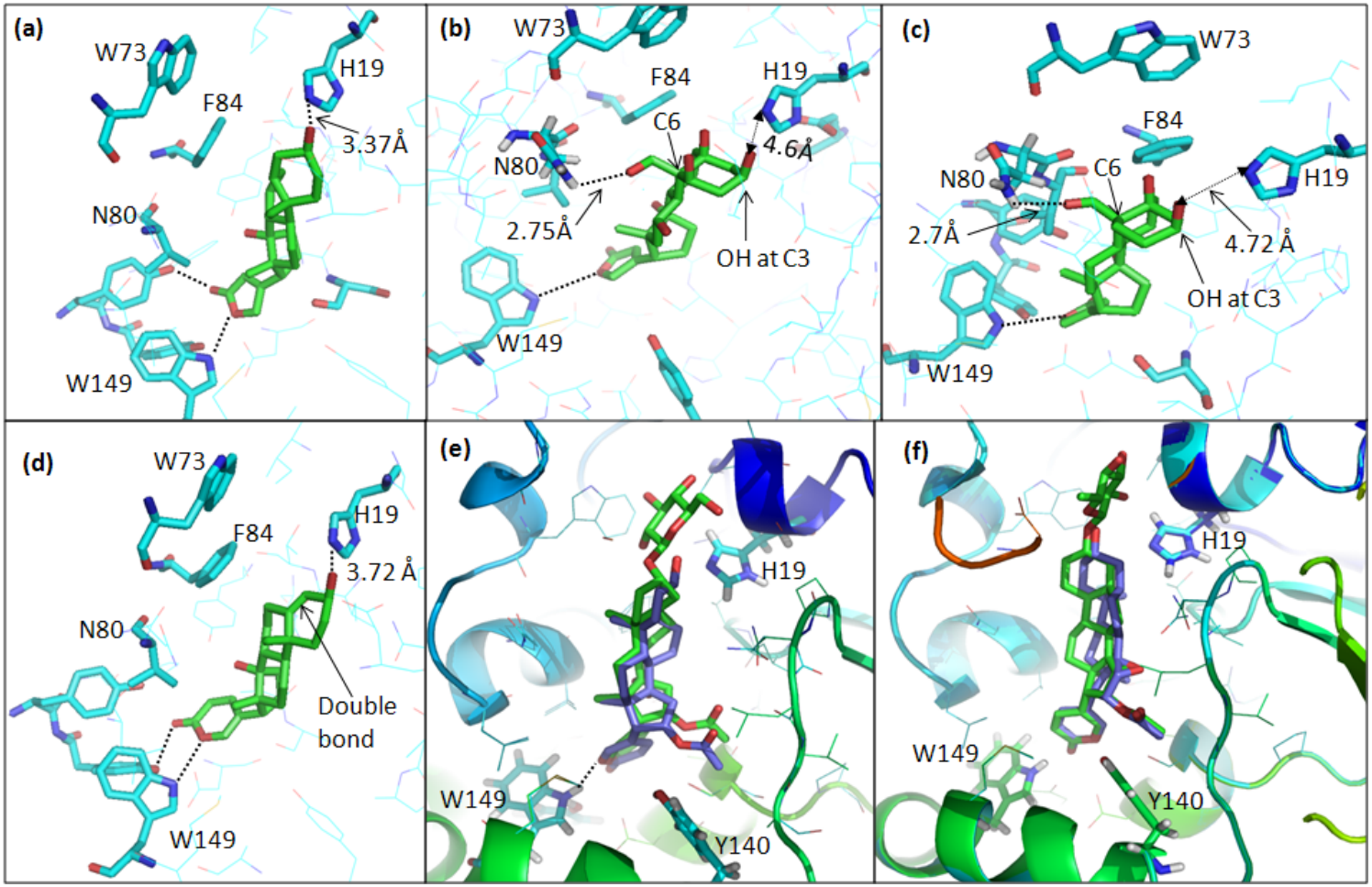


Figure S4. Models of OleD-ligand complexes for **4 (a)**, **5 (b)**, **6 (c)**, **7 (d)**, **8/13 (e)**, **9/15 (f)**. In this model, the C3-OH of **5** and **6** is ~4.5 Å from the His19 (consistent with a lack of turnover). Superimposition of aglycon and monoglucosides **(e)** and **(f)**, **8/13** and **9/15**, respectively, highlight role of Tyr140 in stabilizing the corresponding monoglucoside C2'-OH for diglucosylation.

Figure S5. ^{1}H (CD_3CN, 600 MHz)

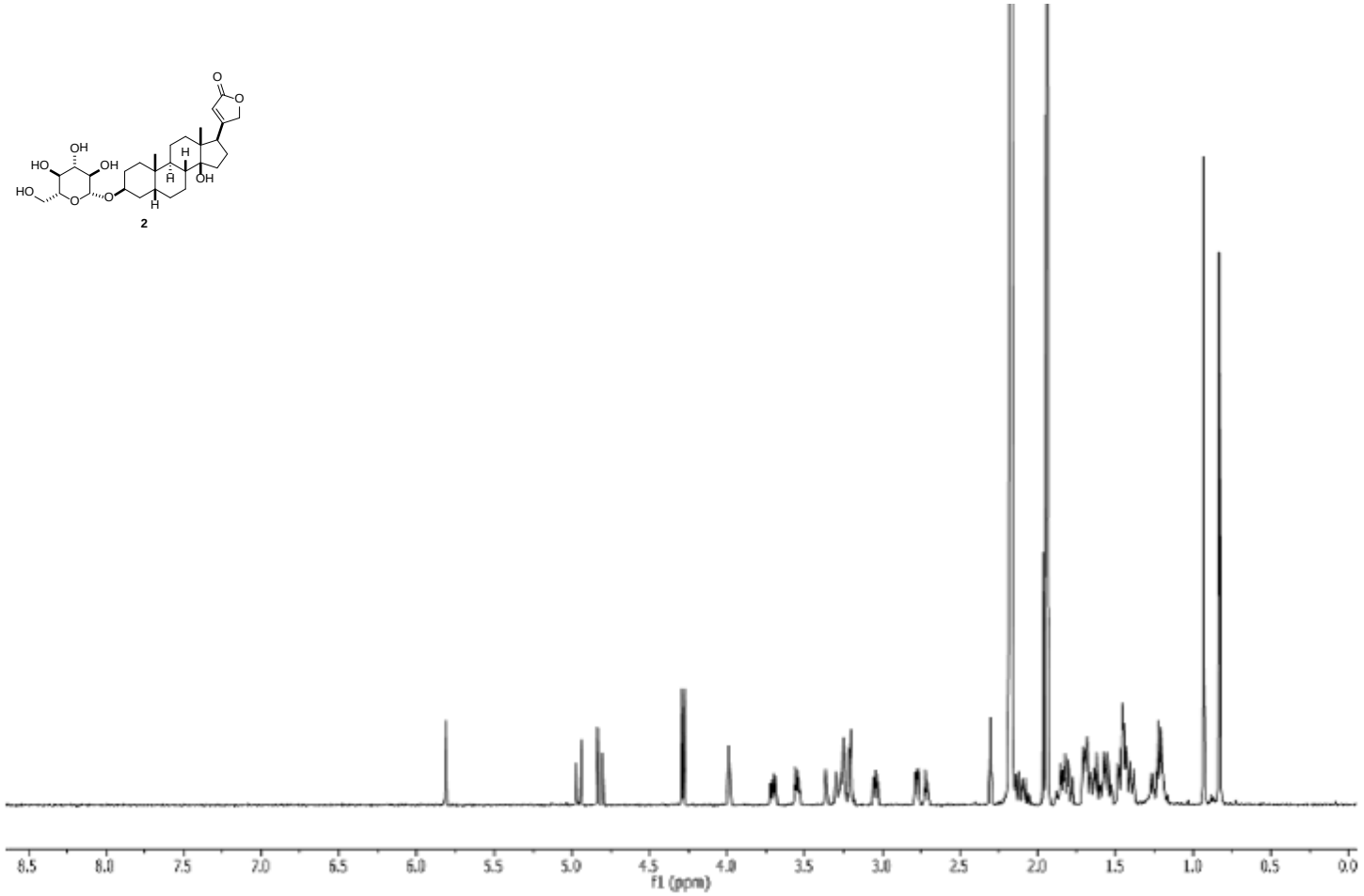


Figure S6. COSY (CD_3CN, 600 MHz)

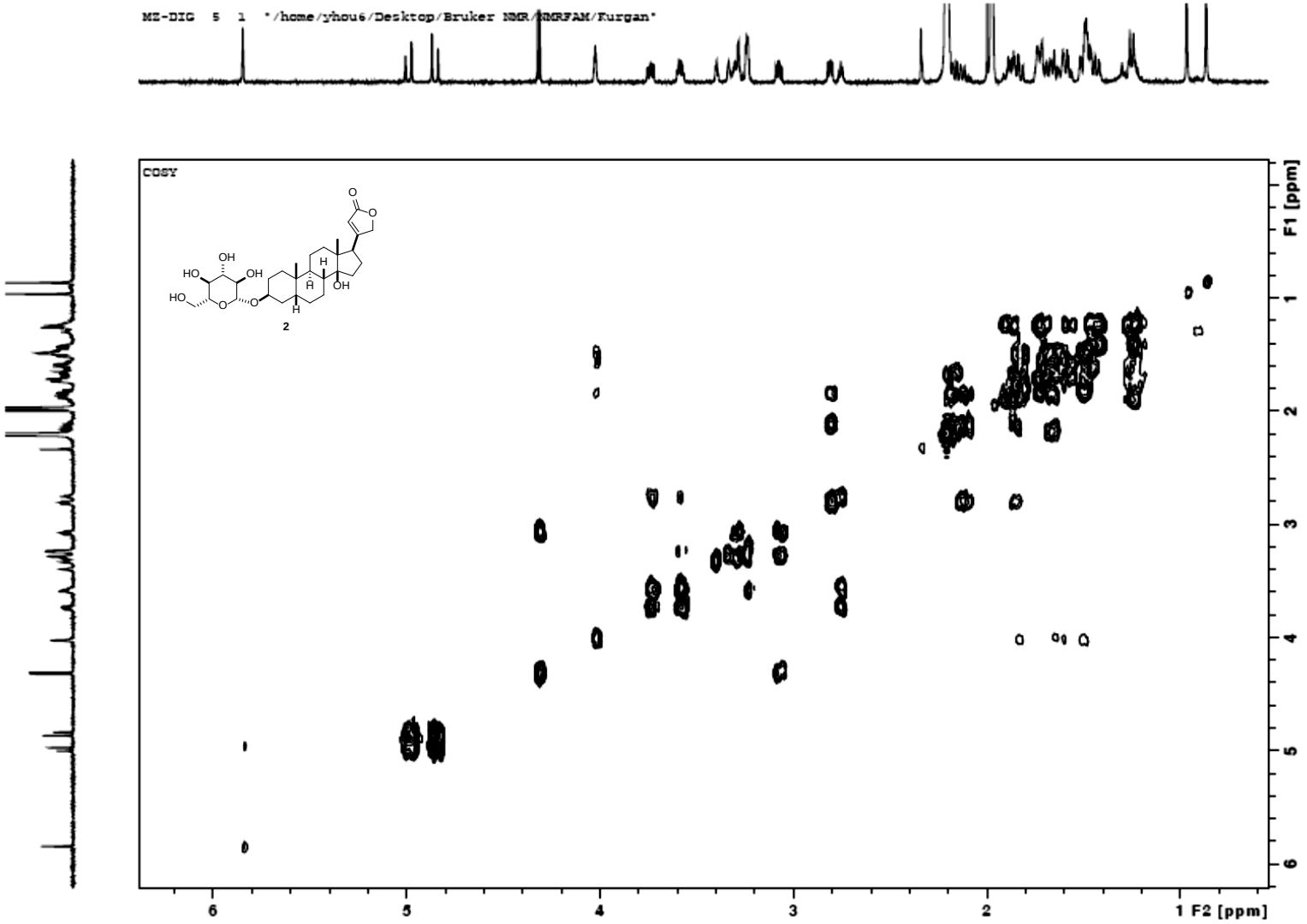


Figure S7. HMBC (CD_3CN, 600 MHz)

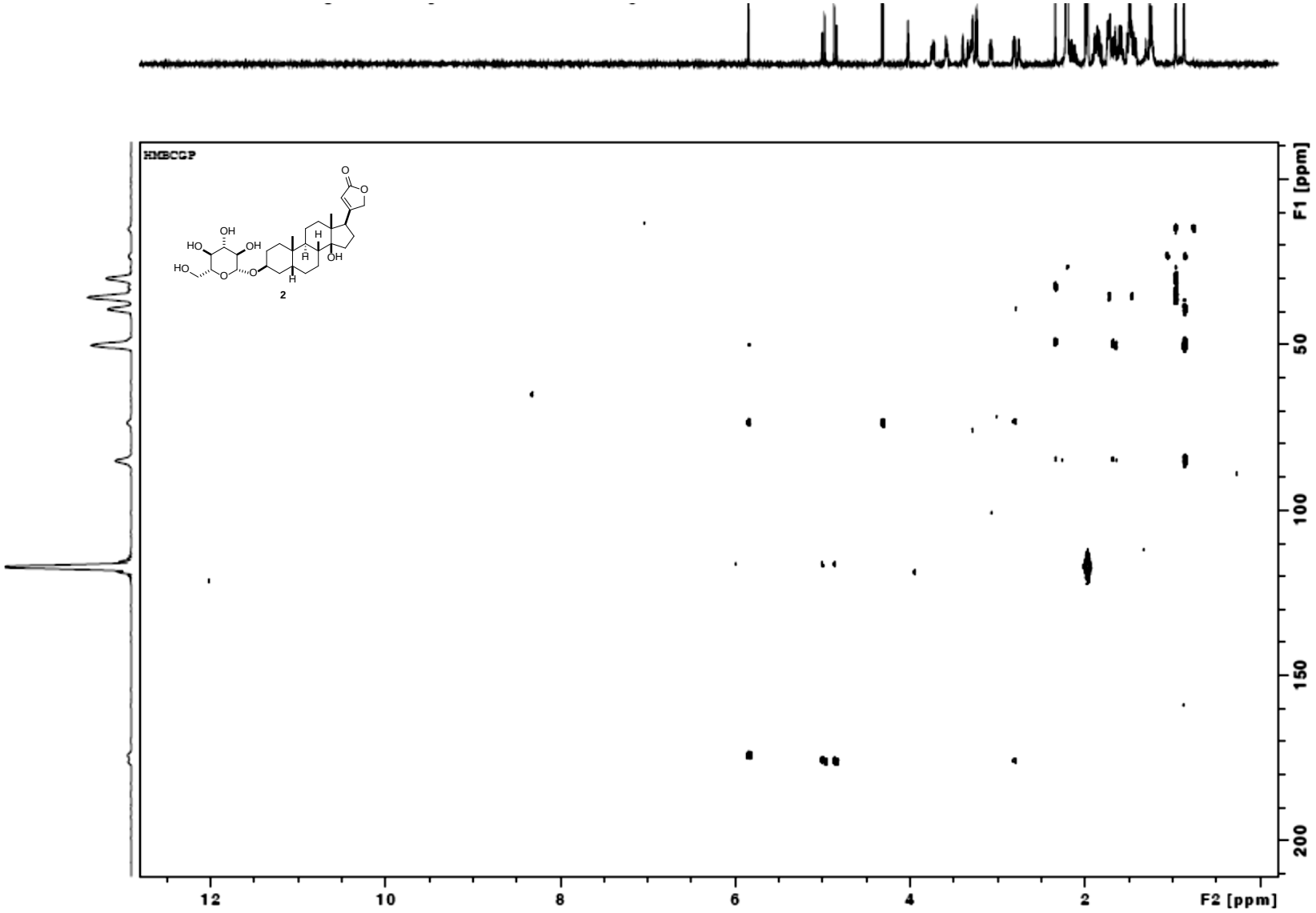


Figure S8. HSQC (CD_3CN, 600 MHz)

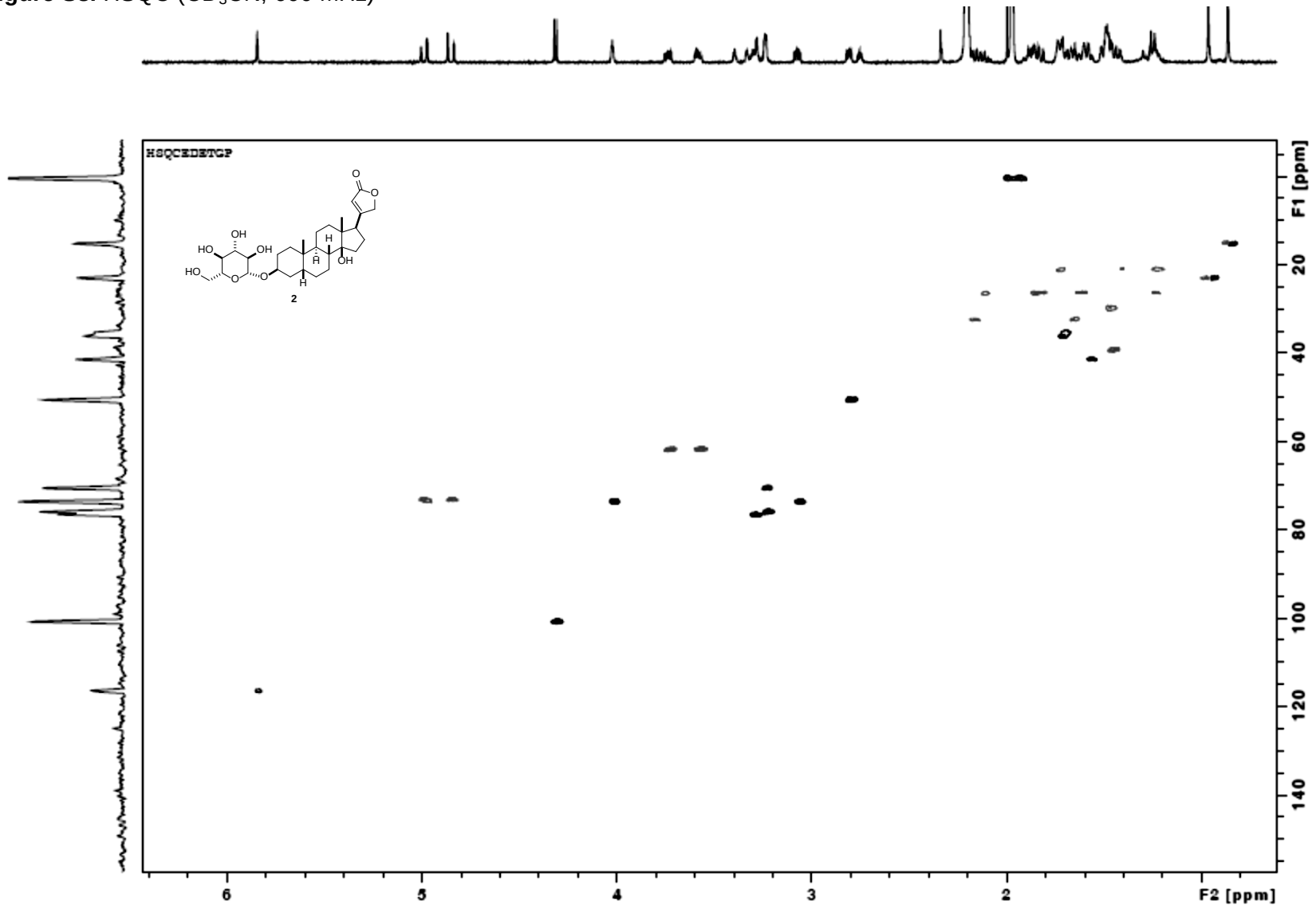
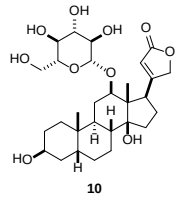


Figure S9. ^{1}H ($CD_3OD/CDCl_3$, 500 MHz)



10

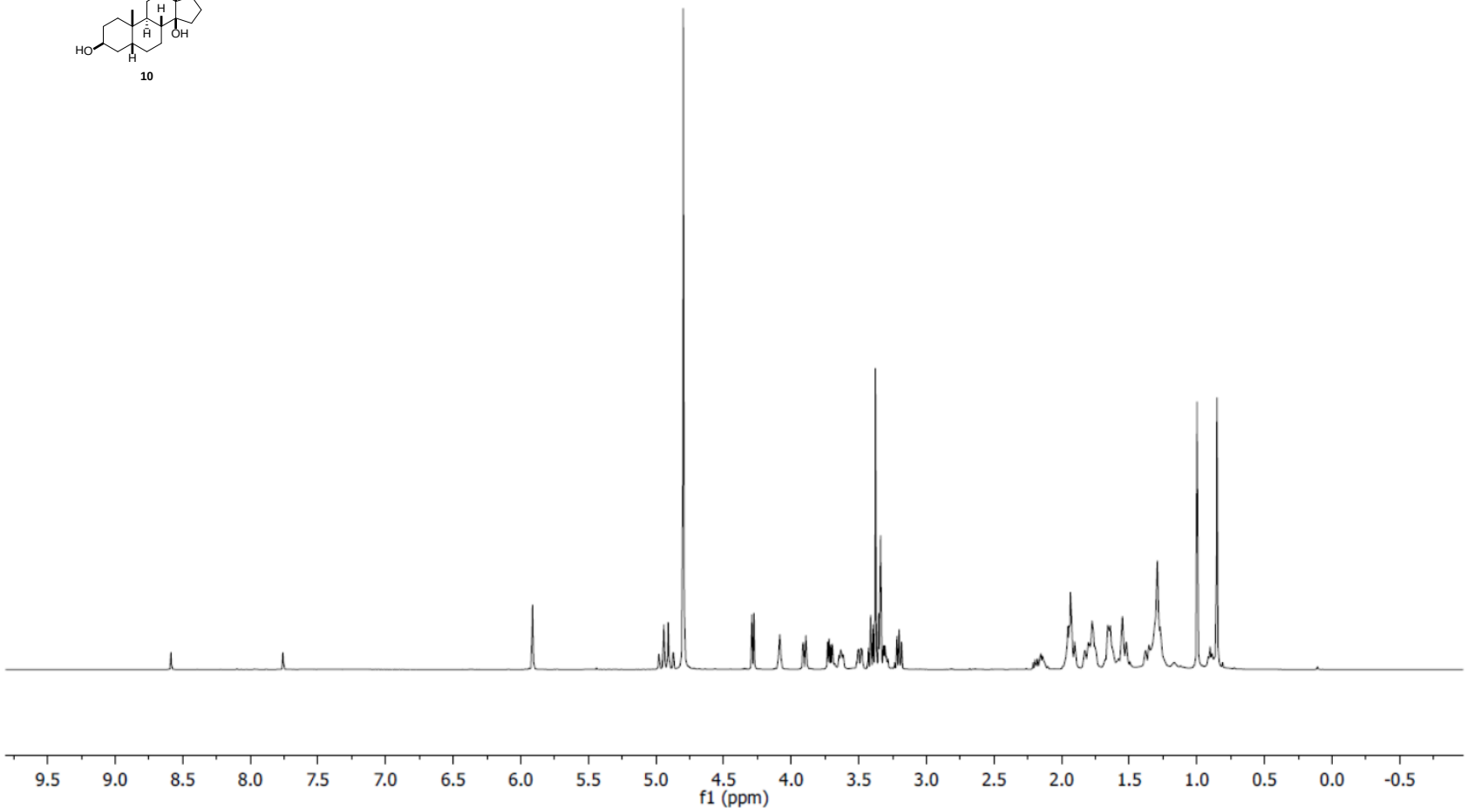
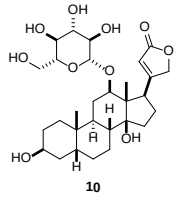


Figure S10. ^{1}H (CD_3CN, 600 MHz)



10

\

Figure S11. HSQC **(**$CD_3OD/CDCl_3$, 500 MHz**)**

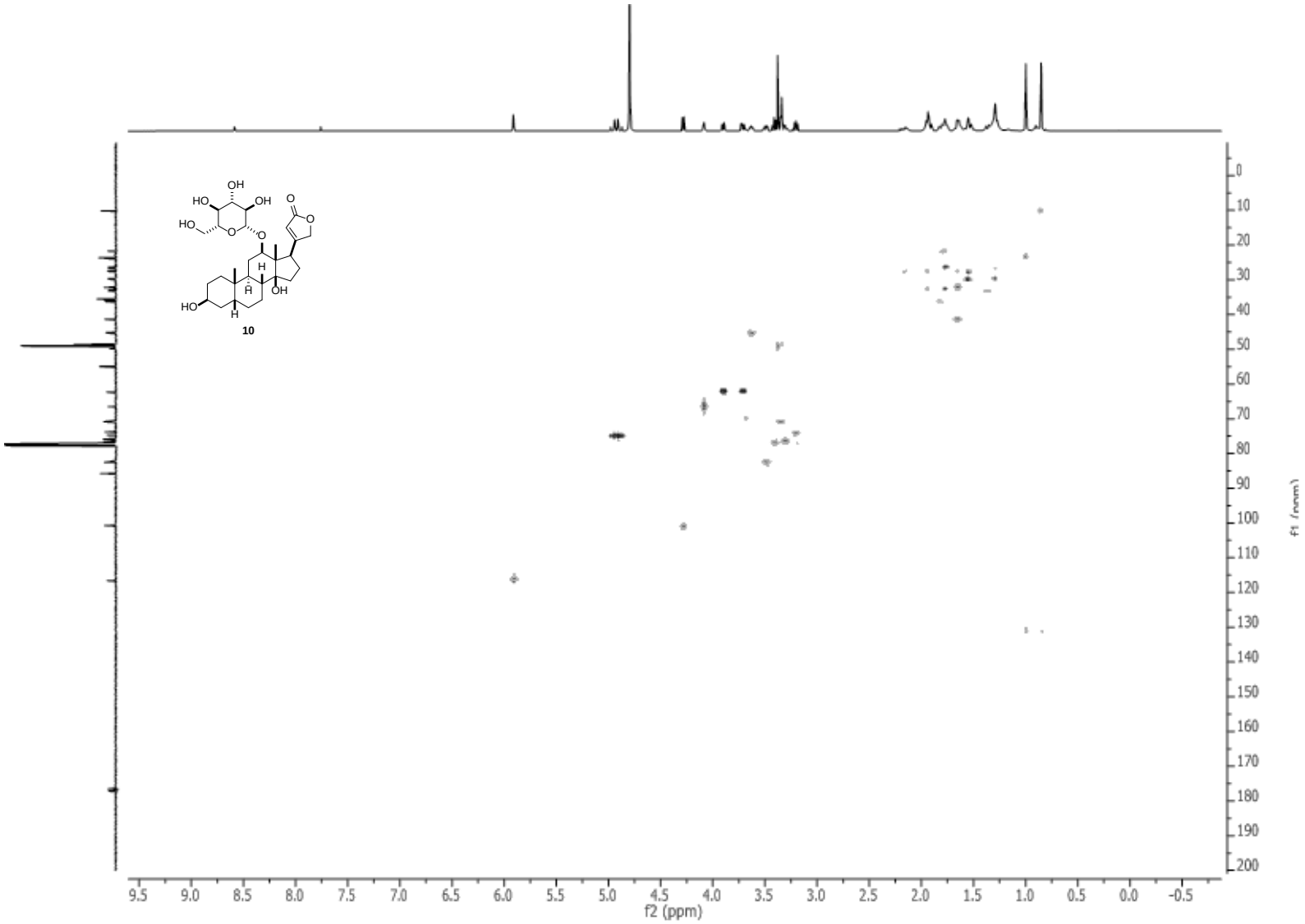


Figure S12. ^{13}C ($CD_3OD/CDCl_3$, 125 MHz)

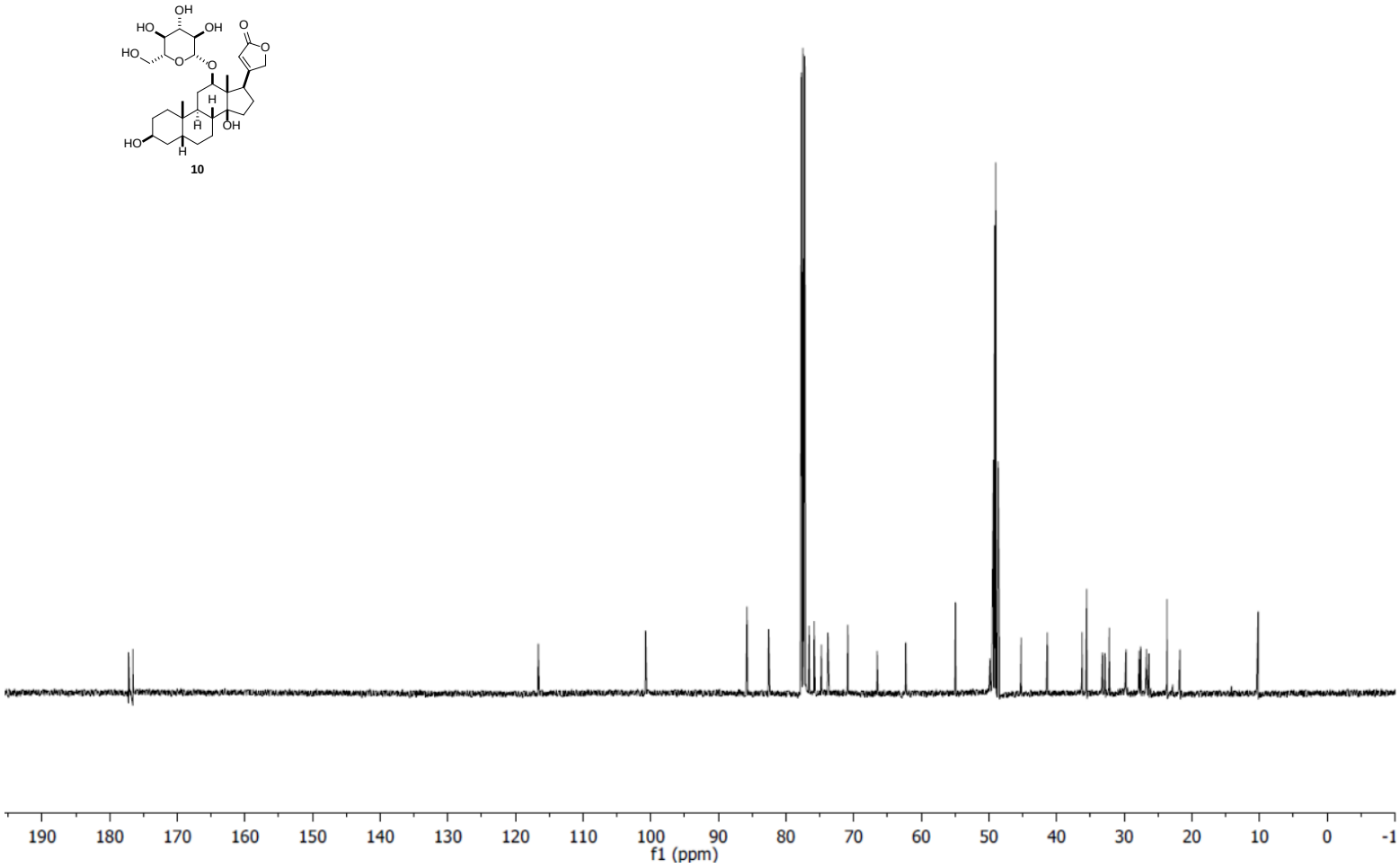
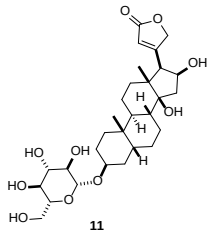


Figure S13. ^{1}H (CD_3CN, 600 MHz)



11

Figure S14. HSQC (CD_3CN, 600 MHz)

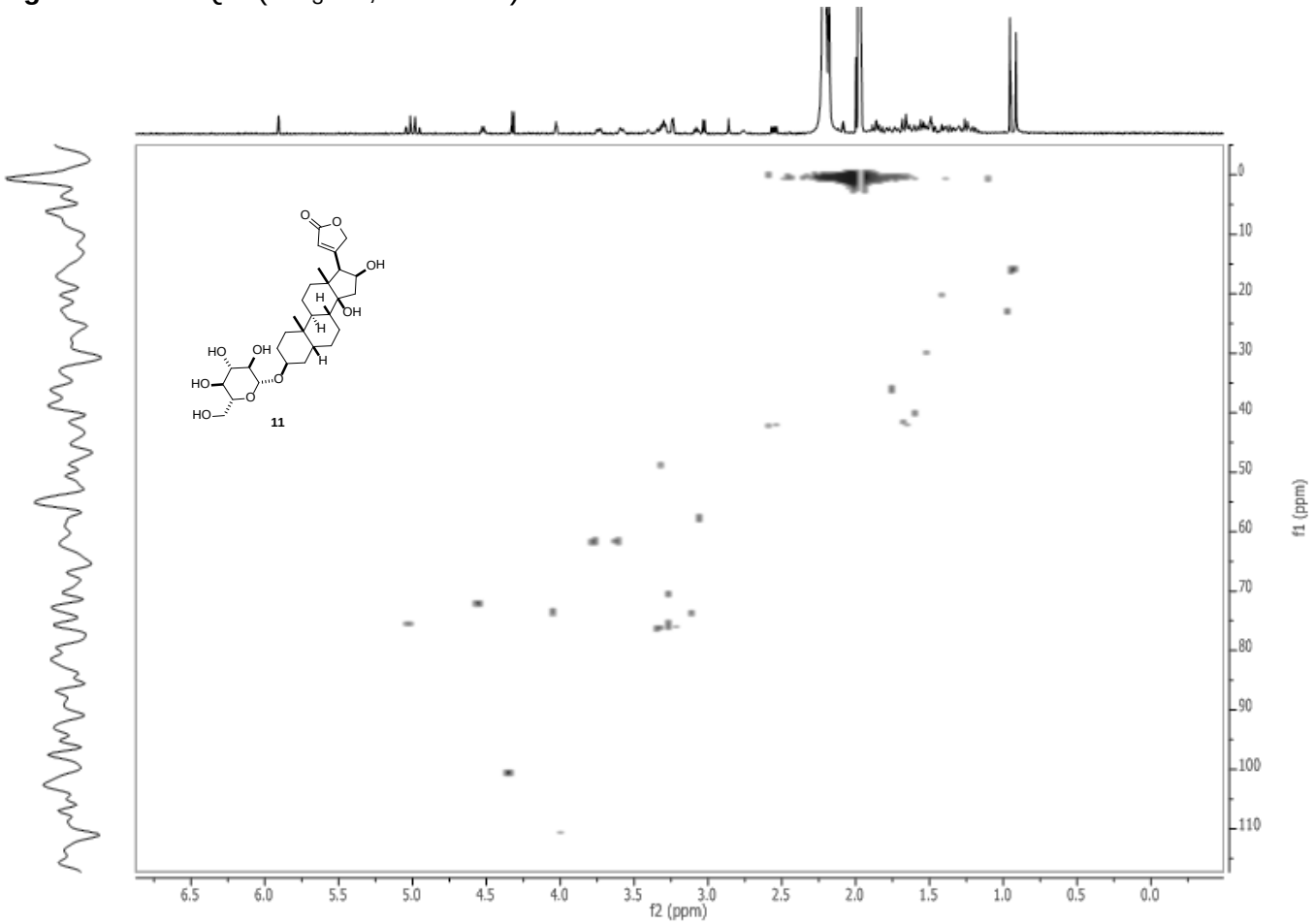
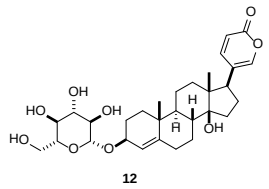


Figure S15. ^{1}H (CD_3CN, 600 MHz)



12

Figure S16. HSQC (CD_3CN, 600 MHz)

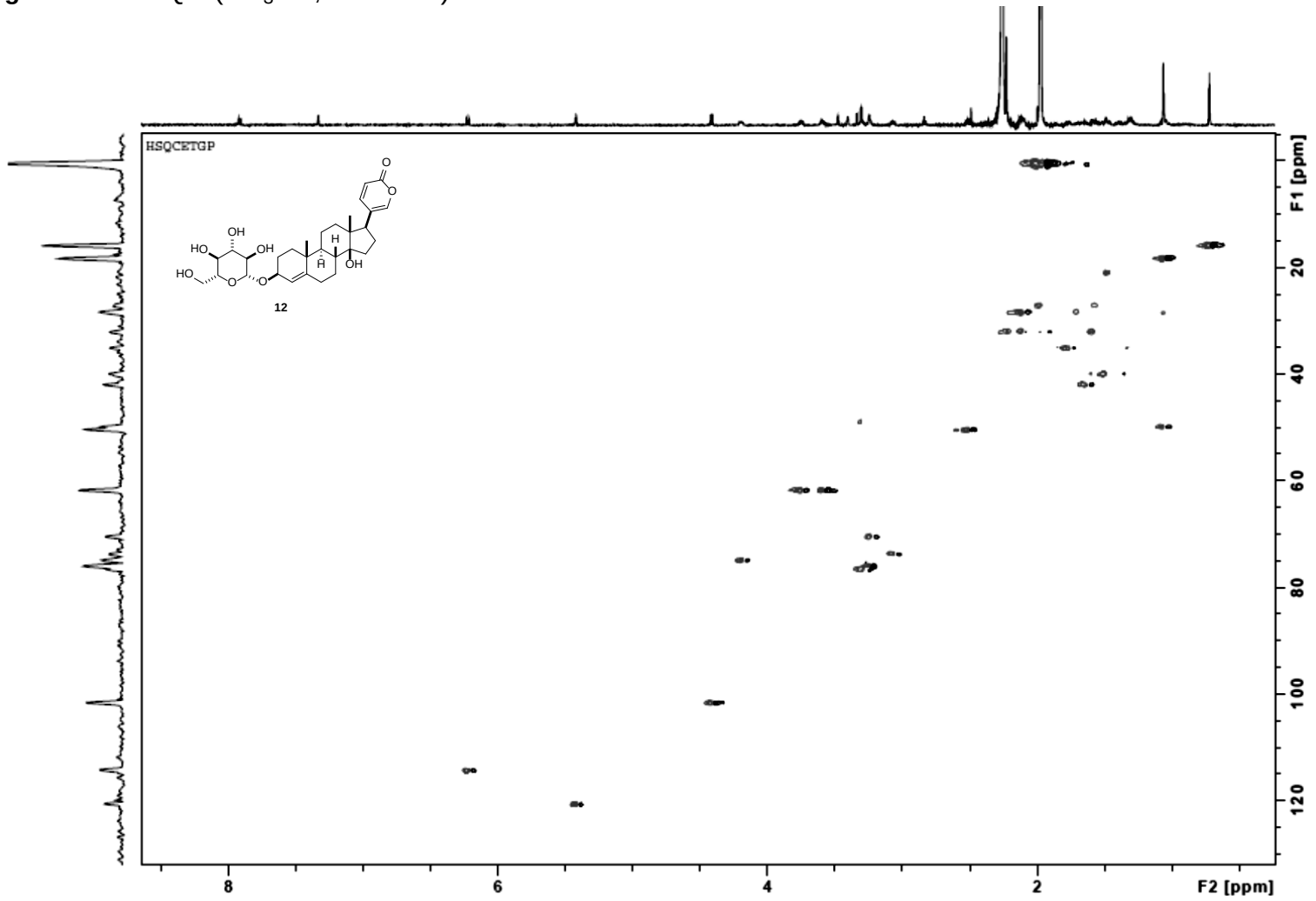
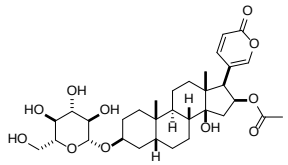


Figure S17. ^{1}H (CD_3CN, 600 MHz)



13

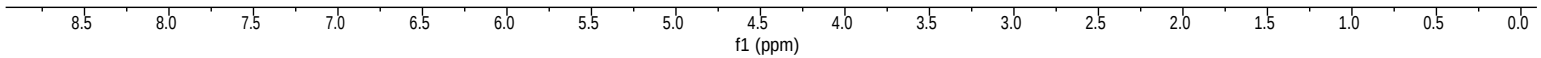


Figure S18. HSQC (CD_3CN, 600 MHz)

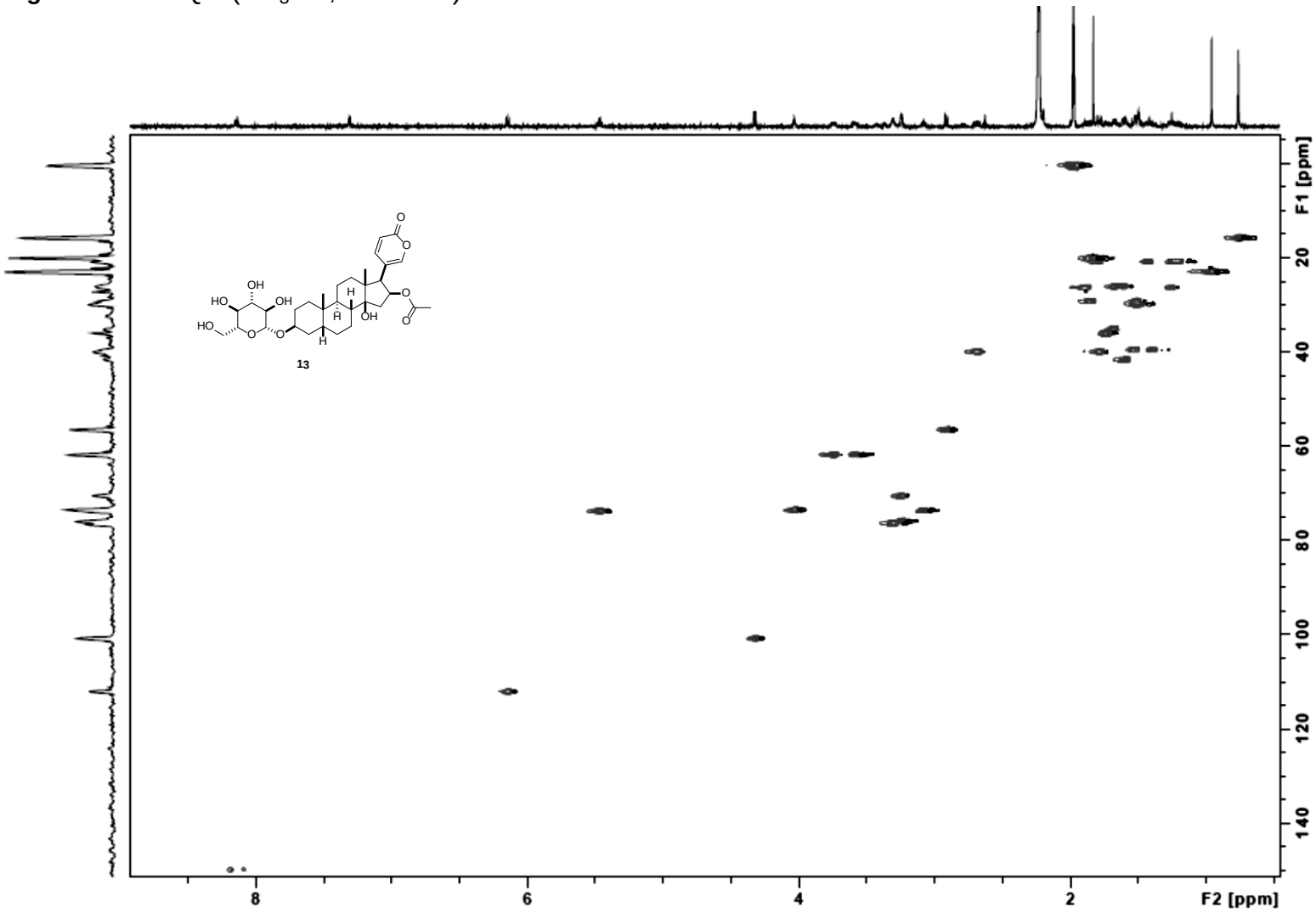
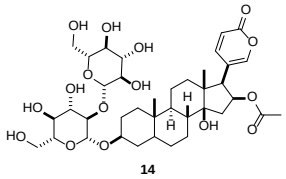


Figure S19. ^{1}H (CD_3CN, 600 MHz)



14

Figure S20. HSQC (CD_3CN, 600 MHz)

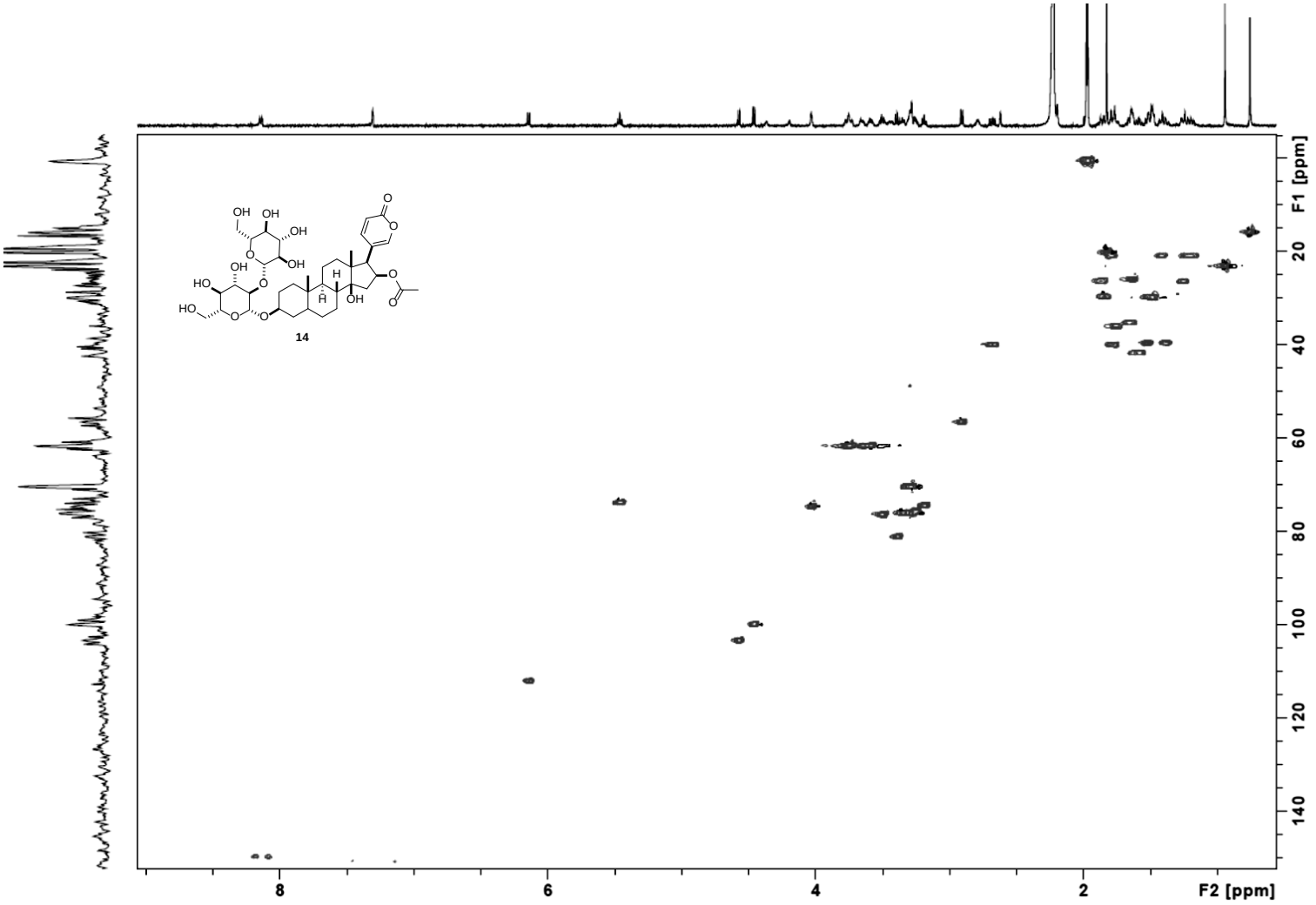
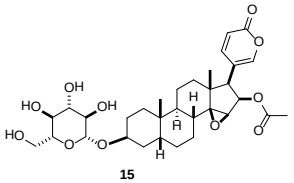


Figure S21. ^{1}H (CD_3CN, 600 MHz)



15

Figure S22. HSQC (CD_3CN, 600 MHz)

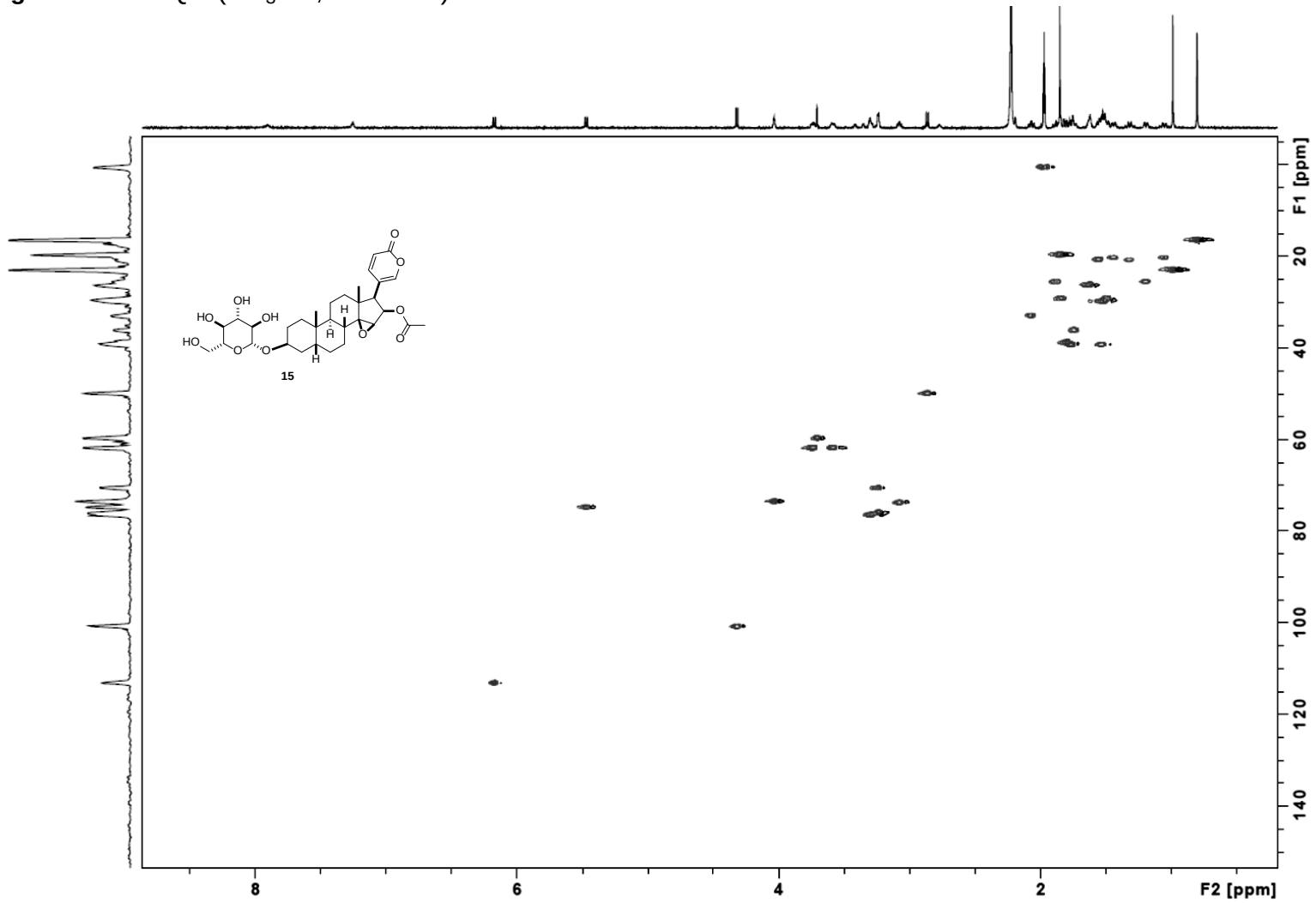
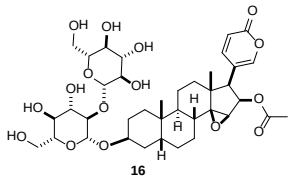


Figure S23. ^{1}H (CD_3CN, 600 MHz)



16

Figure S24. HSQC (CD_3CN, 600 MHz)