Isolation and Identification of 3,4-Seco-Solanidine-3,4-dioic Acid as a Urinary Biomarker of CYP2D6 Activity[S]

Andrew C Behrle[1], Justin Douglas, [iD]J. Steven Leeder and Leon van Haandel[2]

Division of Clinical Pharmacology, Toxicology and Therapeutic Innovation, Department of Pediatrics, and Children's Mercy Research Institute, Children's Mercy Kansas City, Kansas City Missouri (A.C.B., J.S.L., L.V.H.); Nuclear Magnetic Resonance Core Laboratory, University of Kansas, Lawrence, Kansas (J.D.); and Schools of Medicine and Pharmacy, University of Missouri-Kansas City, Kansas City, Missouri (J.S.L.)

ABSTRACT

Cytochrome P450 2D6 (CYP2D6) is responsible for the metabolism and elimination of approximately 25% of clinically used drugs, including antidepressants and antipsychotics, and its activity varies considerably on a population basis, primarily because of genetic variation. CYP2D6 phenotype can be assessed in vivo following administration of an exogenous probe compound, such as dextromethorphan or debrisoquine, but use of a biomarker that does not require administration of an exogenous compound (i.e., drug) has considerable appeal for assessing CYP2D6 activity in vulnerable populations, such as children. The goal of this study was to isolate, purify, and identify an "endogenous" urinary biomarker (M1; m/z 444.3102) of CYP2D6 activity reported previously. Several chromatographic separation techniques (reverse phase HPLC, cation exchange, and analytical reverse phase UPLC) were used to isolate and purify 96 μg of M1 from 40 l of urine. Subsequently, 1D and 2D NMR, and functional group modification reactions were used to elucidate its structure. Analysis of mass spectrometry and NMR data revealed M1 to have similar spectroscopic features to the nitrogen-containing steroidal alkaloid, solanidine. 2D NMR characterization by HMBC, COSY, TOCSY, and HSQC-TOCSY proved to be invaluable in the structural elucidation of M1; derivatization of M1 revealed the presence of two carboxylic acid moieties. M1 was determined to be a steroidal alkaloid with a solanidine backbone that had undergone C-C bond scission to yield 3,4-seco-solanidine-3,4-dioic acid (SSDA). SSDA may have value as a dietary biomarker of CYP2D6 activity in populations where potato consumption is common.

SIGNIFICANCE STATEMENT

Endogenous or dietary biomarkers of processes involved in drug disposition and response may allow improved individualization of drug treatment, especially in vulnerable populations, such as children. Given that several CYP2D6 substrates are commonly used in pediatrics and the ubiquitous nature of potato consumption in Western diets, SSDA has considerable appeal as a noninvasive biomarker of CYP2D6 activity to guide treatment with CYP2D6 substrates in children and adults.

Introduction

One goal of precision therapeutics is to individualize drug doses for patients based on their unique characteristics, thereby optimizing safety and efficacy. The cytochrome P450 2D6 (CYP2D6) enzyme is responsible for metabolizing up to 25% of clinically used drugs (Zhou, 2009a,b). CYP2D6 is highly polymorphic with more than 120 allelic variants identified and classified by the Pharmacogene Variation Consortium (PharmVar) at https://www.pharmvar.org/gene/CYP2D6 (Gaedigk et al., 2008). Inheritance of two recessive loss-of-function alleles results in a "poor-metabolizer phenotype," whereas combinations of full- or reduced-function alleles result in a wide spectrum of phenotypes, including intermediate, extensive, and ultra-rapid metabolizers. As a result, significant interindividual variability in CYP2D6 phenotype has been observed across multiple patient populations, including within groups of nominally the same genotype (Gaedigk et al., 2008; Jones et al., 2010; Llerena et al., 2012; Brown et al., 2016). Given that phenotype prediction from genotype data can be confounded by environmental factors, such as drug-drug interactions, direct assessment of CYP2D6 phenotype may offer a more accurate assessment of CYP2D6 activity at a given point in time relative to genotyping (Wu, 2011; Kantae et al., 2017).

ABBREVIATIONS: 1D, one dimensional; 2D, two dimensional; COSY, ^{1}H-^{1}H correlation spectroscopy; DEPT90, distortionless enhancement by polarization transfer with a final proton pulse angle of 90°; DEPT135, distortionless enhancement by polarization transfer with a final proton pulse angle of 135°; HMBC, heteronuclear multiple-bond correlation; HSQC, heteronuclear single quantum correlation; HSQC-TOCSY, heteronuclear single quantum correlation with additional total correlation spectroscopy transfer; MS/MS, tandem mass spectroscopy; SSDA, 3,4-seco-solanidine-3,4-dioic acid; TOCSY, total correlation spectroscopy; UPLC, ultra-performance liquid chromatography.

Traditional phenotyping studies require the administration of a probe compound that is a substrate of the drug metabolizing enzyme of interest, collection of urine for a defined period of time, followed by determination of the molar ratio of parent compound and metabolite (Streetman et al., 2000). The over-the-counter cough suppressant, dextromethorphan, is a relatively safe probe for assessment of CYP2D6 phenotype in vivo, and the molar ratio of dextromethorphan to dextrorphan in a 4-hour urine collection has been used to phenotype CYP2D6 activity in infants (Blake et al., 2007), children (Evans et al., 1989), pregnant women (Wadelius et al., 1997; Tracy et al., 2005), and adults (Gaedigk et al., 2008). Although widely used as a phenotyping probe, the dextromethorphan/dextrorphan metabolite ratio can be influenced by factors like urinary pH (Labbe et al., 2000), confounding the genotype-phenotype relationship. Other commonly used phenotyping probes, such as debrisoquine, sparteine, and bufurolol, are logistically more difficult to use because of prescription status or they are not approved for clinical use in the United States. Thus, there is considerable appeal for biomarkers reflecting CYP2D6 activity that do not require administration of a drug or other exogenous compound.

Several endogenous compounds have been identified as CYP2D6 substrates (Yu et al., 2003a,b), but none has been explored as a potential phenotyping probe. In contrast, our collaborators applied an untargeted, global metabolomics approach using urine collections from patients previously genotyped and phenotyped for CYP2D6 activity with dextromethorphan to identify an "endogenous" biomarker of CYP2D6 activity (Tay-Sontheimer et al., 2014). This effort identified a compound with parent ion mass of m/z 444.3102 (designated M1) and product ions m/z 98.0964, 370.2731, 206.1883, 56.0494, 55.0550 and 81.0692 (in order of decreasing abundance) that were absent in urine from CYP2D6 poor-metabolizers. M1 was further validated as a product of CYP2D6 activity by observing a decrease in its formation following inhibition of CYP2D6 activity with fluoxetine, a potent CYP2D6 inhibitor. Queries in the common metabolomic databases for the identity of M1 and product ions resulted in no matches, and no further details about the identity of M1 were provided, precluding any further assessment of its clinical utility as a biomarker of CYP2D6 activity.

The goal of this project was to isolate, purify, and elucidate the structure of M1. Through a series of chromatographic methods and 1D and 2D NMR experiments and derivatization reactions, we describe the purification and structural identification of a metabolic product of the nitrogen-containing steroidal alkaloid, solanidine, 3,4-seco-solanidine-3,4-dioic-acid (SSDA), as a potential biomarker CYP2D6 activity.

Materials and Methods

Materials and Reagents. Liquid chromatography-mass spectrometry Optima grade methanol and acetonitrile were purchased from Thermo Fisher Scientific (Waltham, MA). House water was purified using BARNSTEAD E-pure 4-model system to produce 18.1 MΩ-cm water and used without further purification. Liquid chromatography-mass spectrometry grade formic acid, trimethylamine (NEt$_3$), ethanolamine, ammonium hydroxide solution (28.0%–30.0% NH$_3$), acetic acid, sodium acetate (NaOAc), ammonium acetate (NH$_4$OAc), ammonium bicarbonate [(NH$_4$)(HCO$_3$)], spherical C18 flash silica (45-75 μm, 70 Å), Syringe Loading injection Rheodyne Model 7125 and Supelclean LC-SCX SPE tubes 3 ml were obtained from Millipore-Sigma (St. Louis, MO). 10 cm Omnifit chromatography column was purchased from Diba Industries (msscientific Chromatographie-Handel, Berlin, Germany). Waters SunFire C18 (100 Å, 5.0 μm, 4.6 × 150 mm) and Acquity UPLC BEH C18 (130 Å, 1.7 μm, 2.1 × 100 mm) columns were purchased from Waters (Milford, MA). Methanol-D4 was purchased from Cambridge Isotope Laboratories, Inc (Tewksbury, MA). Imject EDC mcKLH Spin Kit was purchased from Thermo Fisher Scientific. A total of 40 l of urine was collected from a single individual, pooled in 4 l aliquots with the pH adjusted of each pooled aliquot to pH = 4.0–4.5 using acetic acid. The collection of urine for

unknown compound purification was declared nonhuman subjects research by Children's Mercy Kansas City Pediatric Institutional Review Board under 45 CFR 46.102(f) (not involving human subjects).

Chromatographic Purification of M1. A comprehensive series of preparative procedures were conducted to isolate and purify M1 for subsequent identification. Briefly, these steps included an initial purification using a 10 cm Omnifit C18 flash silica column, a 3 ml Supelclean LC-SCX solid phase extraction cartridge, a second 10 cm Omnifit C18 flash silica column purification step, followed by purification using a Waters C18 Sunfire column purification, and a final purification procedure using a Waters Acquity UPLC BEH C18 column. The various steps and procedures are summarized in Table 1, and a complete description can be found in the Supplemental Methods.

Chromatography and Mass Spectrometry. Fractions collected during preparative chromatography were analyzed on a Waters TQ-XS triple quadrupole step-wave mass spectrometer (Waters, Manchester, UK) that was connected to an I-Class Acquity UPLC with a flow-through needle sample manager. The instrument was operated in ESI+ mode. The probe capillary was optimized at 3.0 kV and desolvation temperature was set to 400°C. The source offset was set to 60 V. The following gas flows were used: desolvation gas, 800 l/h; collision gas flow, 0.15 ml/min; and nebulizer, 7.0 bar. Argon was used for collision-induced dissociation (CID). Analytes were separated on a Waters Acquity UPLC C18 reversed-phase column (BEH 1.7 μm, 2.1 × 50 mm) that was preceded by a Waters Acquity UPLC C18 VanGuard Precolumn (1.7 μm, 2.1 × 5 mm). The following gradient was used on the UPLC system: initial conditions of 1% B, increasing linearly to 99% B at 4.0 minutes, and held for 1 minute. At 5.10 minutes, the column was stepped down to initial conditions. Mobile phases A and B were H$_2$O and MeOH, respectively, both containing 0.1% formic acid. The flow rate was 0.4 ml per min and column temperature was set to 25°C.

Nuclear Magnetic Resonance (NMR) Conditions. NMR samples were prepared by dissolving the compound in approximately 500 or 300 μl of CD$_3$OD (Cambridge Isotopes, Tewksbury, MA), then transferring to a 5 mm conventional tube (Wilmad LabGlass, Vineland, NJ) or Shigemi tube (Shigemi Co. LTD., Tokyo, Japan), respectively. All NMR spectra were acquired using a Bruker AVIII 500 MHz spectrometer equipped with a cryogenically cooled broadband observe probe or 800 MHz Bruker AV spectrometer equipped with a cryogenically cooled triple resonance inverse probe. NMR experiments were acquired and processed using Bruker Topspin and MestreNova software, respectively, with conventional parameters, unless otherwise noted. ^{1}H NMR shifts given were referenced internally to residual protonated solvent resonance at δ 3.31 ppm (CD$_2$HOD). ^{13}C NMR shifts provided were referenced internally to solvent resonance at δ 49.0 ppm (CD$_3$OD). All values in parentheses are coupling constants (Hz) where ^2J and ^3J denote germinal and vicinal couplings, respectively.

Derivatization of SSDA. A 2 ml glass auto sample vial was charged with 10 μl of a 1:1000 diluted solution of purified unknown M1 and evaporated to dryness and redissolved with 1 ml of Imject EDC conjugation buffer. Then 10 μl of a 1M ethanolamine solution was added to the solution and [H$^+$] was checked to ensure pH = 5. Next, 1-ethyl-3-(2-dimethylaminopropyl) carbodiimide HCl (EDC, 10 mg, 0.522 mmol) was added to the solution and vortexed. The mixture was placed in a 60°C oven. Reaction progress was analyzed by pulling 20 μl samples at time intervals of 30, 60, 90, and 240 minutes and diluted to final volume of 200 μl of using 90:10 10 mM NH$_4$OAc buffer, pH = 7.0/MeOH solution. Samples were analyzed on Xevo G2−XS quadropole time-of-flight mass spectrometer that was connected to a M-Class Acquity UPLC with a flow-through needle sample manager (complete details of the chromatography conditions are provided in the Supplemental Methods).

Results

Elemental Composition By High Resolution Mass Spectrometry. Using high resolution mass spectrometry, the monoisotopic mass (M+H)$^+$ of M1 was determined to be 444.3115 Da, and based on this information, the isotopic distribution and the nitrogen rule, an elemental composition of C$_{27}$H$_{41}$NO$_4$ was proposed. This elemental composition deviated from the theoretical value of 444.3114 by 0.1 mDa (0.2 ppm) and matched the predicted isotopic pattern (I-Fit confidence of 99.52%). The double bond equivalent (DBE) of this elemental composition was 8. Querying the elemental composition against major databases yielded

TABLE 1

Analytical methods used in the isolation and purification of M1 from human urine

Starting Urine Volume (l)	Analytical Media	Compounds Removed	Elution Conditions	Final Volume (ml)
40	C18 spherical flash (45–75 μm, 70 Å)	Salts and very hydrophobic compounds	50:50 MeOH/H$_2$O w/ 0.1% acetic acid, pH = 4.4	500
0.5	3 ml Supelclean LC-SCX SPE cartridge	Neutrals and acids	50:50 MeOH/NH$_4$OH 2-3% NH$_3$ solution	150
0.150	C18 spherical flash (45–75 μm, 70 Å)	Hydrophobic compounds that accumulated on the Sunfire column	95:5 H$_2$O/MeOH (0.1% NEt$_3$, 0.1% acetic acid); 80:20; 60:40; 45:55	150
0.150	Sunfire C18 column (5.0 μm, 4.6 × 150 mm)	Hydrophobic compounds	H$_2$O w/ 0.1% formic acid MeOH w/ 0.1% formic acid	30
0.030	Acquity UPLC BEH C18 column (1.7 μm, 2.1 × 100 mm)	Hydrophobic compounds	90:10 10 mM (NH$_4$)(HCO$_3$) buffer, pH = 10/MeOH MeOH	1 (444.3115 product 99.8% pure)

no identification, as previously reported (Tay-Sontheimer et al., 2014), and a relatively high collision energy (50 eV) was required to fragment M1 into the previously reported fragments. We performed a search of the fragments against the HMDB database (using an in-house developed tool) and found that the fragment at 98.0964, and the couplet of 204.1743 and 206.1906 matched the in-silico generated fragmentation pattern of the steroidal alkaloid solanidine. A literature search revealed that the low molecular weight CID fragments of M1 have been commonly observed in various steroidal alkaloids (Cahill et al., 2010; Zhou et al., 2010); however, the 444.3→370.3 transition, a loss of 74 Da, is unique for this molecule and could not be adequately explained. A search of the steroidal alkaloidal literature for a better match than a solandine-derived structure was unsuccessful, and because of the limited utility of CID spectra for de novo identification of steroidal alkaloids, the decision was made to isolate the compound.

Isolation of M1 From Human Urine. Comparison of peak areas between a known concentration of solanidine and an unknown concentration of M1 resulted in an estimated M1 concentration of 20 nM or 8.9 μg per l in human urine. Based on NMR sensitivity considerations (e.g., https://www.chem.wisc.edu/~cic/nmr/Guides/Other/sensitivity-NMR.pdf) we estimated that approximately 100–300 μg of M1 would be required to obtain an interpretable ^{13}C{^{1}H} NMR spectrum (Supplemental Fig. 15). Therefore, M1 isolation and purification was initiated using 40 l of human urine, using the series of chromatographic steps described in Methods. The challenges associated with each step and the approaches taken to resolve those issues are described in more detail in Supplemental Results. The final product of the isolation and purification process was a solution that appeared pure by quadropole time-of-flight mass spectrometry and served as the source material for interrogation by NMR and determination of M1 structure.

NMR Analysis of M1. Multiple NMR experiments were performed to characterize purified M1, including ^{1}H, ^{13}C{^{1}H}, DEPT135-^{13}C{^{1}H}, DEPT90-^{13}C{^{1}H}, HSQC, HMBC, COSY, 2D TOCSY, and HSQC-TOCSY; raw spectra are provided in the Supplemental Material (Supplemental Figs. 1–37). The interpreted NMR spectra are summarized in Table 2 and Table 3.

^{13}C{^{1}H}, DEPT135-^{13}C{^{1}H} and DEPT90-^{13}C{^{1}H} (Supplemental Figs. 12–15) experiments confirmed the elemental composition of M1 and documented the presence of twenty-seven unique carbons, including four methyl (CH$_3$), nine methylene (CH$_2$), nine methine (CH), and five quaternary (C) carbons. The ^{1}H NMR experiment showed 39 protons, two less than the proposed elemental composition; however, exchangeable protons are not observed in ^{1}H NMR experiments acquired in protic solvents (e.g., CD$_3$OD).

^{1}H and ^{13}C{^{1}H} NMR spectra of solanidine in CD$_3$OD were prepared so that the NMR spectra of M1 could be compared against that of solanidine (Supplemental Fig. 1 and Table 2). Solanidine contains several structural features that result in diagnostic signals in ^{1}H and ^{13}C{^{1}H} NMR experiments because of 1) the 4 methyl groups, 2) the heterocyclic structure (because of the nitrogen), 3) the unsaturation in the B-ring, and 4) the single hydroxyl group. Analogous to solanidine, four methyl groups were observed for M1, and the splitting patterns of the four methyl groups (C-18, C-19, C-21, and C-27) matched the expected pattern of singlet, singlet, doublet and doublet, for each methyl group respectively. Zooming in on the aliphatic region of the ^{1}H NMR spectrum (Supplemental Fig. 2) of M1 highlights the chemical resonances of four methyl groups (Supplemental Fig. 3). Additionally, ^{13}C{^{1}H} chemical shifts of the carbons adjacent to the nitrogen (C-16, C-17, and C-22) are highly characteristic for solanidine and have values of 71.1, 62.9 and 76.3 ppm for M1 (Table 2). The unsaturation in the B-ring of solanidine on C5 and C6 could also be observed in M1, although C6 resonated 13.6 ppm downfield in M1, relative to solanidine (135.9 versus 122.3 ppm). Interestingly the last diagnostic signal in solanidine, is related to the hydroxyl group on carbon 3, which was not observed in M1.

Stepwise connectivity models used to identify M1 as SSDA are summarized in Fig. 1. Our strategy hinged on using the HSQC, HMBC, 2D TOCSY, and HSQC-TOCSY experiments to build these connectivity models, which were initiated at resonances that were separated from the overlapping aliphatic regions of the spectra, namely methyl groups and the olefin. We started with the resonance at 1.27 ppm in the ^{1}H NMR spectrum. In the HMBC experiment, this proton (C19-H) displayed correlations with four carbon atoms, C-5, C-9, C-10, and ambiguous cross peak with either C-1, C-2, C-7, or C-24. C-10 and C-5 are both quaternary carbons because they had no correlations in the HSQC experiment. Given the chemical shift of C-5 (δ 142.2), it is highly suggestive that it is a sp^2-hybridized carbon. Likewise, in the HMBC experiment C6-H displayed correlations with C-8, C-10, C-4, and ambiguous correlation with either C-1, C-2, C-7, or C-24. The ^{1}H and ^{13}C{^{1}H} chemical shifts of C-6 suggest it is an olefin group (-HC=). The resonance assigned to C4 has a chemical shift of 174.1 ppm, implying the presence of a carbonyl moiety. C6-H in M1 is significantly shifted downfield from that of solanidine (Table 2), suggesting the presence of an electron withdrawing group near C6-H. Finally, the C3-H of solanidine is missing in the ^{1}H spectrum of M1, indicating possible structural modification of the A-ring.

The HSQC-TOCSY experiment revealed that C6-H had correlations with only C8-H, and ambiguous correlations with either C-1, C-2, C-7, or C-24 further supported the assignment of C-6 as adjacent to a quaternary carbon, C-5 (Fig. 1A). Analysis of the 2D COSY spectrum revealed

TABLE 2
^{1}H and ^{13}C{^{1}H} data for M1 and solanidine

Carbon #	M1		Solanidine	
	^{1}H NMR (δ)^a	^{13}C{^{1}H} NMR (δ)^a	^{1}H NMR (δ)^b	^{13}C{^{1}H} NMR (δ)^b
1^d	2.07 (td, $^2J_{H-H}$ = 15.0 Hz, $^3J_{H-H}$ = 5.1 Hz) 2.03–1.99 (m)	33.21	1.87 (dt, $^2J_{H-H}$ = 13.2 Hz, $^3J_{H-H}$ = 3.5 Hz) 1.07 (td, $^2J_{H-H}$ = 13.2 Hz, $^3J_{H-H}$ = 3.5 Hz)	38.5
2^d	2.33 (td, $^2J_{H-H}$ = 15.0 Hz, $^3J_{H-H}$ = 5.1 Hz) 1.79–1.71 (m)	32.10	1.80–1.75 (m) 1.54–1.44 (m)	32.3
3	Quaternary carbon	181.5	3.42–3.36 (m)	72.4
4	Quaternary carbon	174.1	2.27–2.16 (m)	43.0
5	Quaternary carbon	142.3	Quaternary carbon	142.3
6	6.59 (bd, $^3J_{H-H}$ = 4.9 Hz)	135.9	5.35 (d, $^3J_{H-H}$ = 5.6 Hz)	122.3
7^d	2.17 (dd, $^2J_{H-H}$ = 13.0 Hz, $^3J_{H-H}$ = 5.4 Hz) 1.77–1.69 (m)	33.13	2.0 (dtd, $^2J_{H-H}$ = 11 Hz, $^3J_{H-H}$ = 5.0 Hz, $^4J_{H-H}$ = 2.6 Hz) 1.59–1.53 (m)	33.2
8	1.73–1.69 (m)	32.1	1.66–1.61 (m)	32.8
9	1.34–1.28 (m)	44.4	1.00–0.95 (m)	51.7
10	Quaternary carbon	40.6	Quaternary carbon	37.8
11^d	1.66 (dd, $^2J_{H-H}$ = 13.0 Hz, $^3J_{H-H}$ = 3.8 Hz) 1.54 (qd, $^2J_{H-H}$ = 13.0 Hz, $^3J_{H-H}$ = 3.8 Hz)	21.5	1.59–1.53 (m)	22.1
12^d	1.88–1.78 (m) 1.34–1.28 (m)	41.0	1.80–1.75 (m) 1.24–1.18 (m)	41.31
13	Quaternary carbon	41.3	Quaternary carbon	41.27
14	1.34–1.28 (m)	58.7	1.23–1.17 (m)	58.9
15^d	2.03–1.99 (m) 1.32–1.29 (m)	30.5	1.86–1.83 (m) 1.19–1.15 (m)	31.77^c
16	3.25–3.20 (m)	71.1	2.75 (s, br)	70.9
17	1.88–1.78 (m)	62.9	1.64–1.60 (m)	64.0
18	0.90 (s)	16.6	0.879 (s)	17.0
19	1.27 (s)	23.6	1.04 (s)	19.85
20	1.88–1.78 (m)	38.1	1.72–1.63 (m)	38.1
21	1.03 (d, $^3J_{H-H}$ = 6.2 Hz)	17.3	0.96 (d, $^3J_{H-H}$ = 6.0 Hz)	17.9
22	2.29–2.23 (m)	76.3	No resonance found	76.2
23^d	1.93 (d, $^2J_{H-H}$ = 13.0 Hz) 1.41–1.34 (m)	28.6	1.80–1.75 (m) 1.25–1.21 (m)	29.6
24^d	1.88–1.78 (m) 1.08–1.04 (m)	33.05	1.80–1.75 (m) 0.90–0.87 (m)	34.1^c
25	1.83–1.78 (m)	31.2	1.69–1.64 (m)	31.77
26^d	3.26–3.20 (m) 1.98–1.96 (m)	60.7	2.95 (s)	61.7
27	0.94 (d, $^3J_{H-H}$ = 6.6 Hz)	19.3	0.88 (d, $^3J_{H-H}$ = 6.0 Hz)	19.73

bd, broad doublet; d, doublet; m, multiplet; q, quartet; s, singlet; t, triplet.
^{a} ^{1}H (800 MHz) and ^{13}C{^{1}H} (125 MHz) NMR in methanol-d_4.
^{b} ^{1}H (500 MHz) and ^{13}C{^{1}H} (125 MHz) NMR in methanol-d_4.
^{c} Carbon resonances overlapped.
^{d} Carbon atoms with diastereotopic protons.

correlation of C6-H with C7-H' (Supplemental Fig. 30). C-7 contained diastereotopic protons (H', δ 2.17) and (H, δ 1.73–1.69) (Table 2). C7-H' had correlations with C-5, C-6, C-8, C-9, and C-14 in the HMBC experiment. The HSQC-TOCSY C7-H' showed only cross peaks with C-6 and C-8. Unfortunately, C7-H overlapped with C2-H and C8-H proton resonances. The HMBC experiment for C8-H had cross peaks with C-5, C-6, C-9, C-14, and C-19. Incorporation of this information expanded the stepwise connectivity model to include C-8, C-9, and C-14 (Fig. 1B).

HSQC data revealed that C-2 contained diastereotopic protons, C2-H' (δ 2.33) and C2-H (δ 1.79–1.71), and the HMBC experiment demonstrated that C2-H' had correlations with C-3, C-9, C-10, C-19, where C-3 had a chemical shift of 181.5 ppm, suggesting it could be a carbonyl group. The HSQC-TOCSY experiment showed a single correlation that could not be resolved unambiguously, however, the observation of only one correlation in the HSQC-TOCSY implies that C-2 is adjacent to at least one quaternary carbon. Since the C2-H' resonance at 2.33 ppm was well separated, we could use the 2D TOCSY experiment to determine that this resonance is coupled with diastereotopic protons at δ 2.07 and δ 2.03–1.99 ppm, along with geminal coupling with C2-H at δ 1.79–1.71 (Supplemental Fig. 6 and 7). Resonances C1-H and C15-H' (δ 2.03–1.99) are overlapped with each other (Supplemental Fig. 6 and 7). Both C-1 and C-15 have diastereotopic protons, C1-H' (δ 2.07) and C1-H (δ 2.03–1.99) and C15-H' (δ 2.03–1.99) and C15-H (δ 1.32–1.29). Since

C2-H' did not have a correlation with C15-H in the TOCSY experiment, the correlation at δ 2.07 and δ 2.03–1.99 represents C1-H' and C1-H, respectively (Supplemental Fig. 6 and 7). This is further supported by the HMBC experimental results wherein C1-H' had correlations with C-3, C-10 and ambiguous cross peaks with C-2, C-7, or C-24. The TOCSY for C1-H'/C15-H and C1-H also had only two correlations, C2-H' (δ 2.33) and C2-H (δ 1.79), indicating it was also adjacent a quaternary carbon (Supplemental Fig. 7). We also noted there was no correlation between C15-H and C2-H' in the 2D TOCSY (Supplemental Fig. 32). With these results, the stepwise connectivity model was broadened to include C-1 through C-3 (Fig. 1C).

The HSQC experiment revealed C-11 to have diastereotopic protons, C11-H' (δ 1.66, dd, $^2J_{H-H}$ = 13.0 Hz, $^3J_{H-H}$ = 3.8 Hz) and C11-H (δ 1.54, qd, $^2J_{H-H}$ = 13.0 Hz, $^3J_{H-H}$ = 3.8 Hz). The HMBC experiment showed C11-H' correlations to C-9 and C-12. C-12 has diastereotopic protons C12-H' (δ 1.88–1.78) and C12-H (δ 1.34–1.28). Focusing on the other methyl resonance that is a singlet, C18-H (δ 0.90), the HMBC experiment revealed that C18-H correlates to C-12, C-13, C-14, C-17, and C-20. HSQC-TOCSY demonstrated that C18-H had no correlations because it was adjacent to a quaternary carbon, C-13. The HMBC experiment for C14-H revealed correlations to C-8, C-13, C-14, C-16, C-17, and C-18, and C-11, C-12, C-13, C-16, C-17, C-18, and C-20 could now be added to the connectivity model (Fig. 1D).

TABLE 3

NMR data for M1

Carbon Number	^{1}H NMR (δ)	^{13}C{^{1}H} NMR (δ)	HMBC (Carbon Numbers)	HSQC-TOCSY (Carbon Numbers)
1	2.07 (td, $^2J_{\text{H-H}}$ = 15.0 Hz, $^3J_{\text{H-H}}$ = 5.1 Hz), H'	33.21[a]	[a], 10, 3	[a]
	2.03–1.99[b] (m)		[a], 13, 14, 17, 3	15, [a], 20, 14, 17, 16
2	2.33 (td, $^2J_{\text{H-H}}$ = 15.0 Hz, $^3J_{\text{H-H}}$ = 5.1 Hz), H'	32.85[a]	19, [a], 10, 9, 3	[a]
	1.79–1.71[c] (m), H		19, 8, 14, 6, 5	11, 15, 8, [a], 12, 9, 14, 6
3	Quaternary carbon	181.5	No cross peaks	No cross peaks
4	Quaternary carbon	174.1	No cross peaks	No cross peaks
5	Quaternary carbon	142.3	No cross peaks	No cross peaks
6	6.59 (bd, $^3J_{\text{H-H}}$ = 4.9 Hz)	135.9	8, [a], 10, 4	8, [a]
7	2.17 (dd, $^2J_{\text{H-H}}$ = 13.0 Hz, $^3J_{\text{H-H}}$ = 5.4 Hz), H'	33.13[a]	8, 9, 14, 5, 6	8, 6
	1.73–1.69[c] (m), H		19, 8, 9, 14, 5, 6	11, 15, 8, [a], 12, 9, 14, 6
8	1.73–1.69[c] (m)	32.1	19, 8, 9, 14, 5, 6	11, 15, 8, [a], 12, 9, 14, 6
9	1.34–1.28[d] (m)	44.4	18, 8, 13, 14, 16, 17	11, 15, 8, [a], 12, 9, 14, 17, 16
10	Quaternary carbon	40.6	No cross peaks	No cross peaks
11	1.66 (dd, $^2J_{\text{H-H}}$ = 13.0 Hz, $^3J_{\text{H-H}}$ = 3.8 Hz), H'	21.5	No cross peaks	11, 12, 9
	1.54 (qd, $^2J_{\text{H-H}}$ = 13.0 Hz, $^3J_{\text{H-H}}$ = 3.8 Hz), H		12, 9	12
12	1.88–1.78[e] (m), H'	41.0	21, 11, [a], 13, 9, 14, 22	21, 27, 11, 23, 15, 25, 8, [a], 20, 12, 9, 14, 26, 17, 16, 22
	1.34–1.28[d] (m), H		18, 8, 13, 14, 16, 17	11, 15, 8, [a], 12, 9, 14, 17, 16
13	Quaternary carbon	41.3	No cross peaks	No cross peaks
14	1.34–1.28 (m)[d]	58.7	18, 8, 13, 14, 16, 17	11, 15, 8, [a], 12, 9, 14, 17, 16
15	2.03–1.99[b] (m), H'	30.5	[a], 13, 14, 17, 3	15, [a], 20, 14, 17, 16
	1.32–1.29[d] (m), H		18, 8, 13, 14, 16, 17	11, 15, 8, [a], 12, 9, 14, 17, 16
16	3.25–3.20[f] (m)	71.1	No cross peaks	27, 15, 25, [a], 26, 16
17	1.88–1.78[e] (m)	62.9	21, 11, [a], 13, 9, 14, 22,	21, 27, 11, 23, 15, 25, 8, [a], 20, 12, 9, 14, 26, 17, 16, 22
18	0.90 (s)	16.6	20, 12, 13, 14, 17	No cross peaks
19	1.27 (s)	23.6	[a], 10, 9, 5	No cross peaks
20	1.88–1.78[e] (m)	38.1	21, [a], 9, 14, 13, 22	21, 27, 11, 23, 15, 25, 8, [a], 20, 12, 9, 14, 26, 17, 16, 22
21	1.03 (d, $^3J_{\text{H-H}}$ = 6.2 Hz)	17.3	20, 17, 22	23, [a], 20, 17, 16, 22
22	2.29–2.23 (m)	76.3	No cross peaks	No cross peaks
23	1.93 (d, $^2J_{\text{H-H}}$ = 13.0 Hz), H'	28.6	No cross peaks	27, [a], 26
	1.41–1.34 (m), H			[a]
24	1.88–1.78[e] (m), H'	33.05[a]	21, [a], 11, 13, 9, 14, 22	21, 27, 11, 23, 15, 25, 8, [a], 20, 12, 9, 14, 26, 17, 16, 22
	1.08–1.04 (m), H		No cross peaks	No cross peaks
25	1.83–1.78[e] (m)	31.2	21, [a], 11, 13, 9, 14, 22	21, 27, 11, 23, 15, 25, 8, [a], 20, 12, 9, 14, 26, 17, 16, 22
26	3.26–3.20[f] (m), H'	60.7	No cross peaks	27, 15, 25, [a], 26, 16
	1.98–1.96 (m), H		No cross peaks	no cross peaks
27	0.94 (d, $^3J_{\text{H-H}}$ = 6.6 Hz)	19.3	25, [a], 26	23, 25, [a], 26

[a] Ambiguous assignment of carbons 1, 2, 7, and 24.
[b] Overlapped proton resonances from carbons 1 and 15.
[c] Overlapped proton resonances from carbons 2, 7, and 8.
[d] Overlapped proton resonances from carbons 9, 12, 14, and 15.
[e] Overlapped proton resonances from carbons 12, 17, 20, 24, and 25.
[f] Overlapped proton resonances from carbons 16 and 26.

D-ring Assignment. The third methyl group (δ 1.03 ppm, d, $^3J_{\text{H-H}}$ = 6.2 Hz), which is a doublet and hence, must be adjacent to a carbon with only one unique hydrogen, making it a methine (-CH) carbon, is assigned to C21-H. The HMBC for C21-H had correlations to C-17, C-20, and C-22. The HSQC-TOCSY for C21-H showed correlations to C-16, C-17, C-20, C-22, C-23, and ambiguous cross peaks to C-1, C-2, C-7, or C-24 (Table 3). Unfortunately, C22-H had no correlations in the HMBC

and HSQC-TOCSY experiments. We examined the COSY experiment for C21-H to see if any correlations could be found. C21-H correlated to one multiplet (δ 1.88–1.79) that represented overlapped resonances from C12-H', C17-H, C20-H, C24-H', and C25-H.

Given the downfield shift of the C-16 (δ = 71.1 ppm; Table 2), we reasoned that C-16 is adjacent to an electronegative atom. C-16 in solanidine has a resonance at δ 70.9 ppm and can explain the downfield
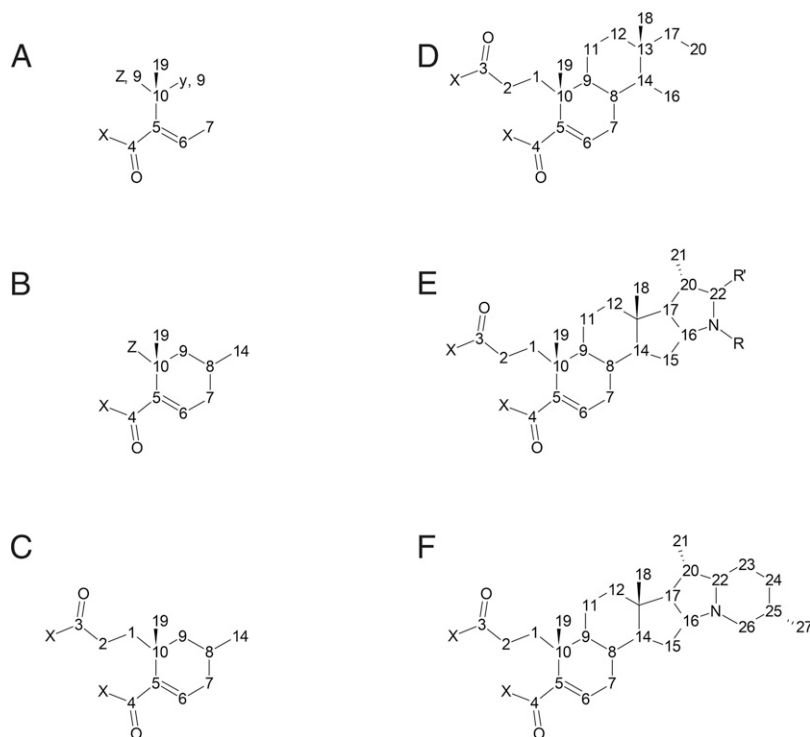
Fig. 1. Stepwise connectivity models of SSDA structure as described in Results with hydrogen atoms omitted for clarity. (A) Z = ambiguous assignment of C-1, C-2, and C-24. X = C or heteroatom; (B) Z = ambiguous assignment of carbons 1, 2, and 24. X = C or heteroatom; (C) X = C or heteroatom; (D) X = C or heteroatom; (E) X = C or heteroatom; R = C or heteroatom; R' = C or heteroatom; (F) X = heteroatom.

shift caused by the presence of the tertiary alkylamine. In the M1 ^{1}H NMR spectrum C16-H and C26-H are overlapped (δ 3.26–3.20). The HMBC experiments for C16-H and C26-H showed no correlations; however, the HSQC-TOCSY displayed correlations with C-15, C-25, C-26, and C-27, and ambiguous cross peaks to either C-1, C-2, C-7, or C-24. Based on the DEPT90 and DEPT135 experiments, C-16 is a methine carbon and C-26 is a methylene. The 2D TOCSY experiment of C16-H had correlations with C15-H' (δ 2.03–1.99) and C21-H. C16-H also had correlations with overlapped two regions δ 1.88–1.80 ppm (C12-H', C17-H, C20-H, C24-H' and C25-H) and δ 1.32–1.29 ppm (C9-H, C12-H, C14-H and C15-H). The COSY experiment for C16-H had two correlations with C15-H' and overlapped resonances of C12-H', C17-H, C20-H, C24-H and C25-H. There are two points of interest here. First, both HSQC-TOCSY experiments for C14-H and C16-H had correlations with C15 and C17. Second, the COSY for C16-H had correlations with C15-H' and C17-H, whereas the 2D TOCSY for C16-H had correlations with C14-H, C15-H', and C17-H. One drawback to an HMBC experiment is that pulse sequences use a single fixed delay based on the average value of $^nJ_{CH}$ (Reynolds and Enríquez, 2002). Complex molecules, such as steroidal-alkaloids, have a range of $^nJ_{CH}$ values (2–15 Hz) and correlations with $^nJ_{CH}$ values that differ significantly from the average will be filtered. Although C-15 did not have correlations with C16-H or C17-H, the HMBC experiments for C15-H' (δ 2.03–1.99) had correlations with C-13, C-14, and C-17. Additionally, the HMBC experiments for C15-H (δ 1.32–1.29) had correlations with C-13, C-14, C-16, and C-17, and C-18 (Table 2). Based on these results, the D-ring assignment was completed and included in the stepwise connectivity model (Fig. 1E).

Proposed M1 Structure. Identification of the remaining carbons was completed by examining the final methyl group, C-27. HMBC experiments for C27-H showed correlations with C-25 and C-26, and ambiguous cross peak to either C-1, C-2, C-7, or C-24. Like C-21, C-27 must be directly adjacent to a methine carbon, in this case C-25. C-26 was significantly downfield shifted compared with C-24 (Δ = 26.7 ppm) and is directly bonded to the nitrogen atom. The HSQC-TOCSY

experiment of C27-H resulted in correlations with C-23, C-25, and C-26, and ambiguous cross peaks to C-1, C-2, C-7, or C-24. Assignment of C-23 was completed using the COSY experiment and was found to have diastereotopic protons, C23-H' (δ 1.93) and C23-H (δ 1.41–1.34). Analysis of C23-H' resulted in correlations with C24-H' (δ 1.83–1.78), the multiplet at δ 1.41–1.34 (geminal coupling), and C24-H (δ 1.08–1.04). The HSQC-TOCSY results for C23-H' had correlations with C27, C24, C25, and C26. HSQC-TOCSY results for C23-H had correlations with only C-24. The completed connectivity model for M1 is presented in Fig. 1F.

Heteroatom Identification and Final Structure of SSDA. For heteroatom identification we reviewed the elemental composition of M1, $C_{27}H_{41}NO_4$ and noted that the proposed structure in Fig. 1F has a molecular formula of $C_{27}H_{39}NO_2$ and is missing two oxygen atoms and two hydrogen atoms. The total sum of the integration values for the resonances in the ^{1}H spectrum had a value of 39. Given that all carbon atoms were identified, the remaining oxygen and hydrogen atoms must be part of the carbonyl moieties at C-3 and C-4. Therefore, we propose that M1 contains two carboxylic acid functional groups and is identified as 3,4-*seco*-solanidine-3,4-dioic acid (SSDA; Fig. 2).

Confirmation of SSDA Dicarboxylic Structure. Orthogonal confirmation of the carboxylic acid moieties was acquired by activating the carboxylate groups with carbodiimide, followed by reaction with ethanolamine (Supplemental Fig. 32). Reaction progress was monitored using mass spectrometry. The first amidation reaction was facile, with the disappearance of the parent ion of m/z 444.31 and appearance of m/z 487.35, indicating the addition of 1 unit of ethanolamine. The second amidation required an increase in temperature to 60°C and was confirmed with the appearance of the m/z 530.39 (Supplemental Fig. 38 and 39), consistent with the proposed SSDA structure.

Determination of Parent and Product Ion Structures. Support for the proposed structure of SSDA is further corroborated by the product ion assignments from the MS/MS fragmentation experiment (Fig. 3). Collision induced dissociation (CID) of the parent ion of m/z 444.3123 resulted in six major product ions: m/z 370.2744, m/z

A

Solanidine

B

3,4-seco-Solanidine-3,4-Dioic Acid
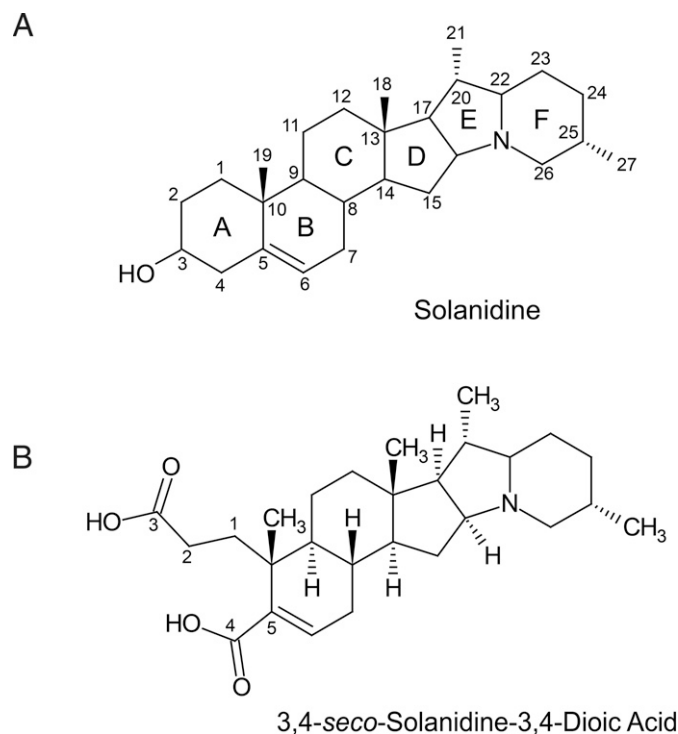
Fig. 2. Structures of solanidine (A) and final proposed structure of M1 as 3,4-seco-solanidine-3,4-dioic acid (SSDA) (B). For solanidine, rings are labeled with letters and carbon atoms are numbered; hydrogen atoms are omitted for clarity. For SSDA, the numbering of carbon atoms has been omitted for clarity, except for A-ring carbons.

356.2582, m/z 326.2851, m/z 206.1908, m/z 150.1270, m/z 136.1116, and m/z 98.0983, and each of the major fragments can now be explained by simple bond breaking. The product ion of m/z 370.2744 had a mass loss of m/z 74.0379 with an elemental composition of $C_3H_6O_2$. This fragmentation represents the loss of the propionic acid chain. The product ion of m/z 356.2582 had an elemental composition of $C_{23}H_{36}NO_2$ and represents a mass loss of m/z 88.0541 from the parent ion, which would be consistent with propionic acid and a methyl group. The product ion 326.2851 m/z had an elemental composition of $C_{23}H_{36}N$. The mass loss difference between product ions m/z 370.2744 and m/z 326.2851 of 43.9893 represents loss of CO_2, or the second carboxylic acid moiety. The product ion of m/z 206.1908 had an elemental composition of $C_{14}H_{24}N$ and represents the loss of B and C rings. The product ion m/z of 150.1270 had an elemental composition of $C_{10}H_{16}N$ and can be described as the fragment composed of a hydro-indolizine ring, a diagnostic feature of solanidine-based alkaloids (Lawson et al., 1997; Shou et al., 2010). The product ion of m/z 98.0983 had an elemental composition of $C_6H_{12}N$ and is described as the six-member amino-alkane ring.

Quantitative NMR was performed using an internal calibrant, and a quantity of 96 μg of SSDA was determined with a purity of 99.8% with the assumption that salts and other inert NMR impurities have been removed from the overall purification methods (Supplemental Tables 1–3).

Discussion

In this investigation, we report the structure of an "endogenous" biomarker of CYP2D6 previously discovered by an untargeted, global metabolomics approach using urine collections from patients previously genotyped and phenotyped for CYP2D6 activity with dextromethorphan (Tay-Sontheimer et al., 2014). The compound, designated "M1" was

characterized by a parent ion mass of m/z 444.3102, was absent in urine of CYP2D6 poor metabolizers, and its formation was reduced in vivo by coadministration of the CYP2D6 inhibitor, fluoxetine. Present at relatively low concentrations in urine, M1 was isolated and purified from ∼40 l of urine and determined to be a metabolite of solanidine, 3,4-seco-solanidine-3,4-dioic acid. As solanidine is a toxic glycoalkaloid produced by members of the *Solanaceae* family and found at particularly high concentrations in the metabolically active parts of plants, such as leaves, flowers, and sprouts of potatoes (Omayio et al., 2016; Saranya et al., 2018), SSDA is dietary in origin and not an "endogenous" biomarker of CYP2D6 activity.

The structure of SSDA was somewhat unusual and unexpected, but the final structure is plausible for several reasons. First, the ^{1}H NMR experiments were conducted in CD_3OD and acidic protons are known to undergo H/D exchange in protic deuterated solvents like methanol. Secondly, the [13]C{[1]H} chemical shifts for C-3 and C-4 are δ 181.5 and 174.1, respectively, and correspond very well to reported chemical shifts of carboxylic acid moieties in C_{27} products. For example, 3,4-seco-12α-hydroxy-5β-cholan-3,4,24-trioic acid, a bile acid produced from cholesterol metabolism, has been synthesized and characterized using NMR techniques (Sekiguchi et al., 2017). The reported [13]C{[1]H} chemical shifts for C-3 and C-4 are δ 176.4 and δ 177.2, respectively. Furthermore, scission of the A-ring is not uncommon. Exploration of the literature reveals that A-ring opening of steroidal compounds is not unprecedented as there are multiple examples that proceed by a mechanism involving oxidation of the alcohol followed by a Baeyer-Villiger oxidation and, subsequently, hydrolysis. For example, *trans*-cyclohexa-1,2-diol has been reported to be oxidized to the dicarboxylic acid adipate by the fungus *Acinetobacter* TD63 via a mechanism that involves an initial dehydrogenation to 2-hydroxycyclohexane-1-one, ring expansion to form a lactone, 1-oxa-2-oxo-7-hydroxycycloheptane, rearrangement to 6-oxohexanoate, and subsequently adipate (Davey and Trudgill, 1977). Similarly, 3β-acetoxy-5α-androstan-3-one is reported to undergo A-ring opening when incubated in the presence of *Myceliophthora thermophile* (Hunter et al., 2009), and the A-ring of testosterone can be cleaved under anaerobic conditions by *Steroidobacter denitrificans* (Wang et al., 2013). Perhaps of most relevance to SSDA formation, the structurally similar steroidal alkaloid, tomatidine, has been shown to undergo degradation via A-ring opening to form 4-hydroxy-3,4-seco-tomatidine-3-oic acid in the presence of the fungus, *Gymnoascus reesii* (Gaberc-Porekar et al., 1983).

There are also several reports describing a role for mammalian oxidative enzymes catalyzing Baeyer-Villiger oxidation of steroidal substrates (Leisch et al., 2011). Verhoeven and coworkers reported that A-ring opening of a progestogen, (17α)-17-hydroxy-11-methylene-19-norpregna-4,15-dien-20-yn-3-one, was the major metabolic route after oral administration of the compound to rats (Verhoeven et al., 1998), and Lai et al. (2011) reported on the oxidation and opening of a keto-piperidine group of a poly(ADP-ribose) polymerase inhibitor; in the latter study, in vitro experiments revealed FMO5 to be responsible for the ring-opening via Baeyer-Villiger oxidation. Finally, a novel secondary bile acid in the common ring tail possum was recently described and mimics the exact A-ring opening that we observed (Sekiguchi et al., 2017); A-ring scission through a Baeyer-Villiger mechanism and oxidation of the resulting alcohol (following hydrolysis) to the carboxylate is a multistep process.

Given that minimal to no formation of SSDA is observed in CYP2D6 poor metabolizers and that formation is reduced in the presence of CYP2D6 inhibitors in vivo, it is likely that CYP2D6 catalyzes the rate-limiting step in SSDA formation. In fact, incubation of solanidine with heterologously expressed CYP2D6 produced a hydroxylated metabolite that has been identified as 4-hydroxysolanidine by an NMR
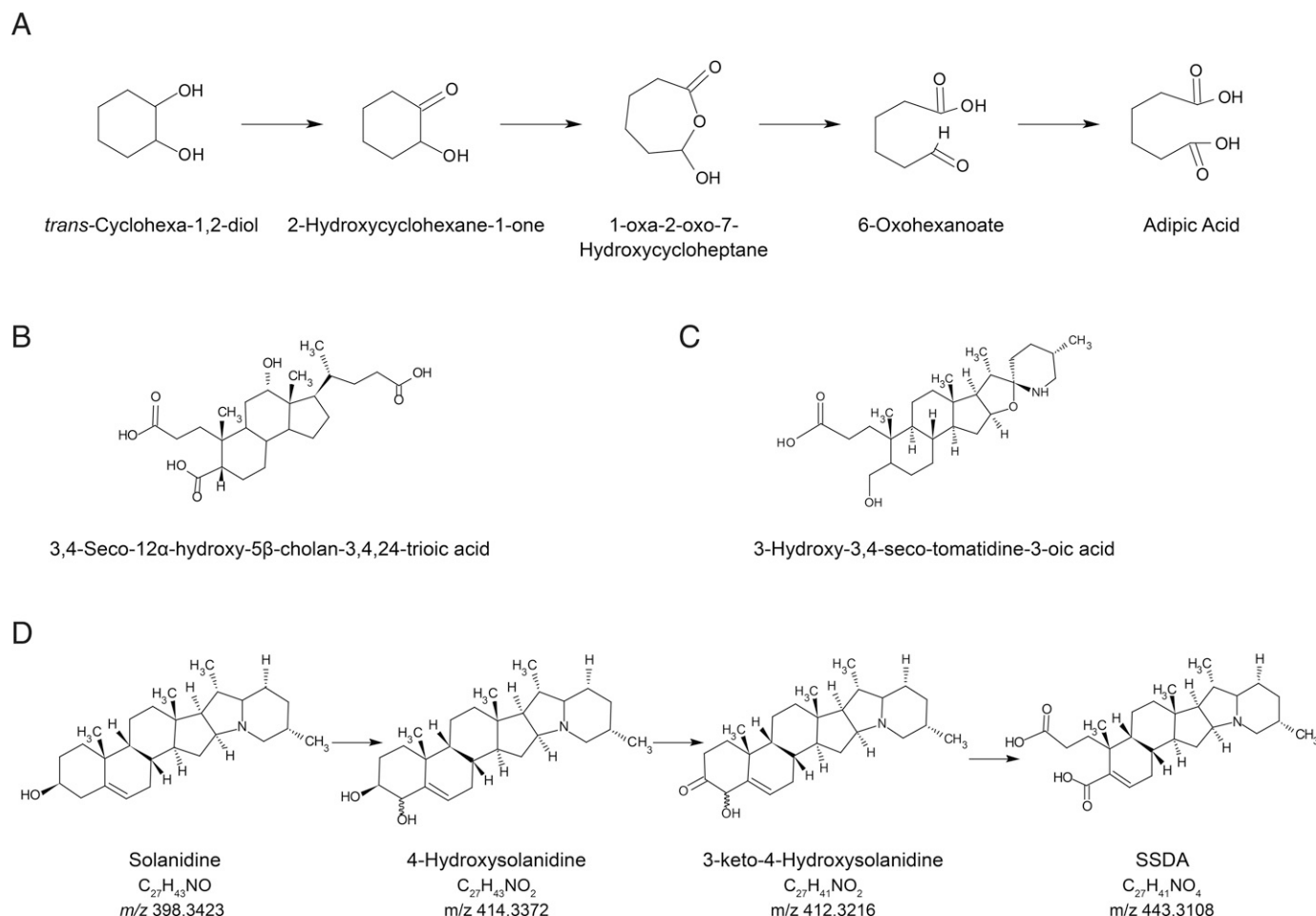
analysis approach analogous to that described for SSDA (manuscript in preparation), a significant finding in that 3-keto-4-hydroxy compounds are potential substrates for Baeyer-Villiger oxidations, as discussed above. Based on the reported *trans*-cyclohexa-1,2-diol to adipate pathway (Davey and Trudgill, 1977) (Fig. 4A) and the reports of 3,4-seco-12α-hydroxy-5β-cholan-3,4,24-trioic acid formed from a secondary bile acid (Sekiguchi et al., 2017) (Fig. 4B) and 4-hydroxy-3,4-seco-tomatidine-3-oic acid from tomatidine (Gaberc-Porekar et al., 1983) (Fig. 4C), we propose a pathway for formation of SSDA from solanidine following CYP2D6-mediated formation of 4-hydroxysolanidine in Fig. 4D. Consistent with this pathway, we have also observed structures consistent with dehydrogenation of 4-hydroxysolanidine to form 4-keto-solanidine and 3-keto-4-hydroxysolanidine in patient urine samples, although the analytes and the enzymes involved in the processes have not yet been identified (manuscript in preparation).

In summary, we have presented the isolation and structural elucidation of SSDA using supporting data from several 1D and 2D NMR experiments, functional group modification reactions, and mass spectrometry data. Collectively, the data all support the proposed structure for SSDA and increase confidence in the final assignments. Derived from solanidine, a steroidal alkaloid naturally

occurring in plants of the *Solanaceae* family, such as potatoes, a urinary metabolite ratio of solanidine/SSDA would be analogous to the dextromethorphan/dextrorphan ratio used to assess CYP2D6 phenotyping following administration of cough suppressant, dextromethorphan. Indeed, during preparation of this manuscript, Magliocco et al. (2021) applied an elegant pharmacometabolomic approach to provide compelling evidence for solanidine-based metabolite ratios in urine and plasma as biomarkers of CYP2D6 activity in vivo. Identification of *m/z* 444.3102 as SSDA and the proposed pathway presented in Fig. 4D provides insight into the identities for the major CYP2D6 metabolites *m/z* 414.3366 and *m/z* 412.3210 reported in that study. Given that potatoes or potato-derived products are relatively common in North American and European diets, and the relatively long half-life of solanidine [$t_{1/2}$ = 5 days; (Harvey et al., 1985)], we concur with Magliocca et al. (2021) that solanidine-based metabolite ratios may have broad utility as a measure of CYP2D6 activity that does not require exogenous drug administration. Clearly, multiple metabolic steps are involved in the biotransformation of solanidine to form SSDA and it is possible that a ratio of solanidine to various combinations of metabolites along the path from 4-hydroxysolanidine
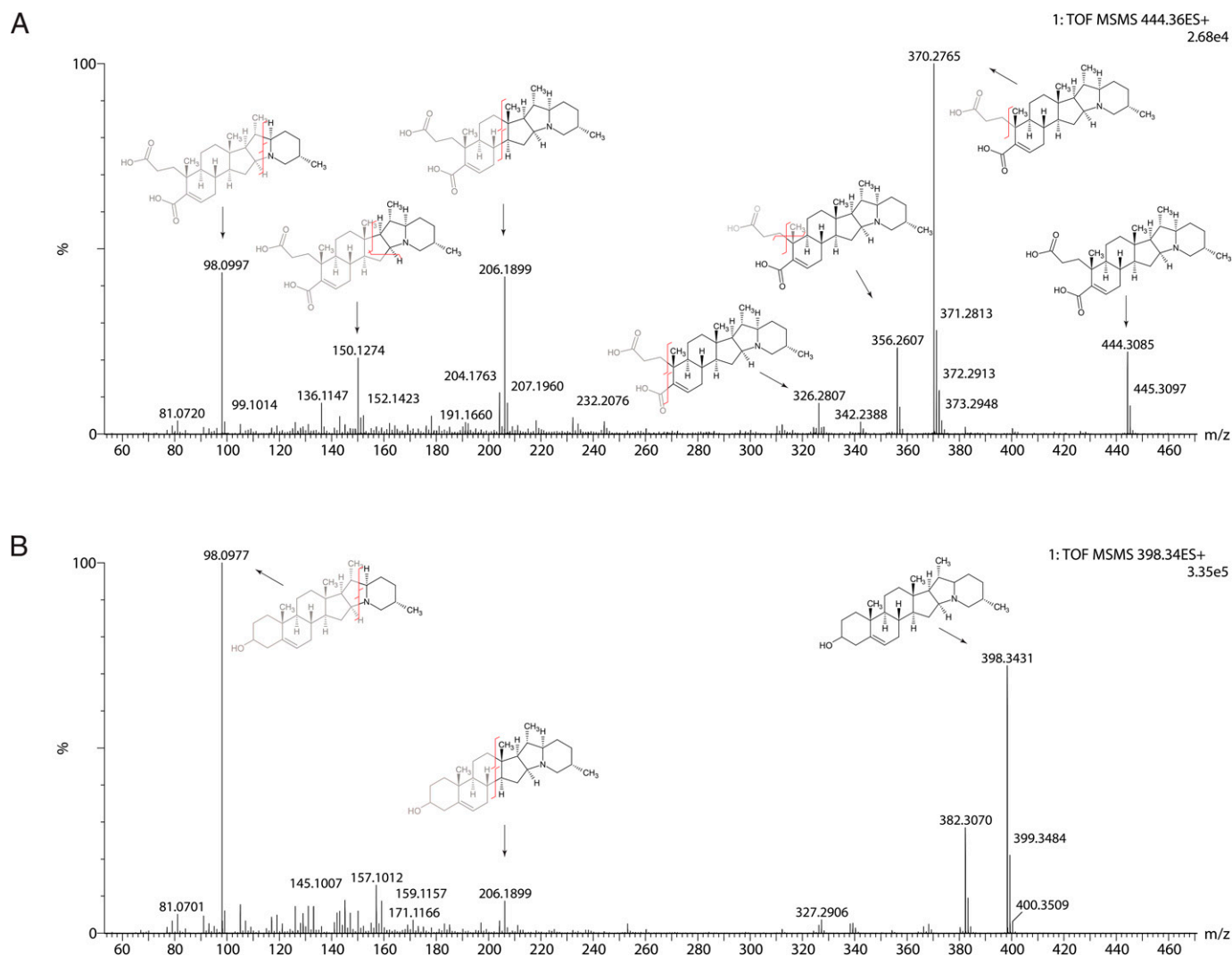
Fig. 3. MS/MS fragmentation chromatogram of SSDA (A), with labeled product ions. The MS/MS fragmentation chromatogram for the parent compound, solanidine, is provided for reference (B).

A

trans-Cyclohexa-1,2-diol   2-Hydroxycyclohexane-1-one   1-oxa-2-oxo-7-Hydroxycycloheptane   6-Oxohexanoate   Adipic Acid

B

3,4-Seco-12α-hydroxy-5β-cholan-3,4,24-trioic acid

C

3-Hydroxy-3,4-seco-tomatidine-3-oic acid

D

Solanidine
$C_{27}H_{43}NO$
m/z 398.3423

4-Hydroxysolanidine
$C_{27}H_{43}NO_2$
m/z 414.3372

3-keto-4-Hydroxysolanidine
$C_{27}H_{41}NO_2$
m/z 412.3216

SSDA
$C_{27}H_{41}NO_4$
m/z 443.3108

Fig. 4. Proposed pathway for SSDA formation involving Baeyer-Villiger chemistry based on literature precedents. (A) Proposed pathway for adipic acid formation from *trans*-cyclohexa-1,2-diol (Davey and Trudgill, 1977); (B) A-ring opening to form 3,4-seco-12α-hydroxy-5β-cholan-3,4,24-trioic acid formed from deoxycholic acid (Sekiguchi et al., 2017); (C) Formation of 4-hydroxy-3,4-seco-tomatidine-3-oic acid from tomatidine (Gaberc-Porekar et al., 1983); (D) Proposed pathway for formation of 3,4-seco-solanidine-3,4-dioic acid following an initial CYP2D6-mediated hydroxylation to form 4-hydroxysolanidine.

to SSDA may ultimately prove superior to administration of an exogenous compound, such as dextromethorphan, as a convenient biomarker of CYP2D6 activity for application in precision therapeutic settings.

Authorship Contributions

Participated in research design: Behrle, Leeder, van Haandel

Conducted experiments: Behrle, Douglas, van Haandel

Performed data analysis and interpretation: Behrle, Douglas, van Haandel

Wrote or contributed to the writing of the manuscript: Behrle, Douglas, Leeder, van Haandel

References

Blake MJ, Gaedigk A, Pearce RE, Bomgaars LR, Christensen ML, Stowe C, James LP, Wilson JT, Kearns GL, and Leeder JS (2007) Ontogeny of dextromethorphan O- and N-demethylation in the first year of life. *Clin Pharmacol Ther* **81**:510–516.

Brown JT, Abdel-Rahman SM, van Haandel L, Gaedigk A, Lin YS, and Leeder JS (2016) Single dose, CYP2D6 genotype-stratified pharmacokinetic study of atomoxetine in children with ADHD. *Clin Pharmacol Ther* **99**:642–650.

Cahill MG, Caprioli G, Vittori S, and James KJ (2010) Elucidation of the mass fragmentation pathways of potato glycoalkaloids and aglycons using Orbitrap mass spectrometry. *J Mass Spectrom* **45**:1019–1025.

Davey JF and Trudgill PW (1977) The metabolism of trans-cyclohexan-1,2-diol by an acinetobacter species. *Eur J Biochem* **74**:115–127.

Evans WE, Relling MV, Petros WP, Meyer WH, Mirro Jr J, and Crom WR (1989) Dextromethorphan and caffeine as probes for simultaneous determination of debrisoquin-oxidation and N-acetylation phenotypes in children. *Clin Pharmacol Ther* **45**:568–573.

Gaberc-Porekar V, Gottlieb HE, and Mervič M (1983) A new microbial degradation pathway of steroid alkaloids. *J Steroid Biochem* **19**:1509–1511.

Gaedigk A, Simon SD, Pearce RE, Bradford LD, Kennedy MJ, and Leeder JS (2008) The CYP2D6 activity score: translating genotype information into a qualitative measure of phenotype. *Clin Pharmacol Ther* **83**:234–242.

Harvey MH, McMillan M, Morgan MR, and Chan HW (1985) Solanidine is present in sera of healthy individuals and in amounts dependent on their dietary potato consumption. *Hum Toxicol* **4**:187–194.

Hunter AC, Watts KR, Dedi C, and Dodd HT (2009) An unusual ring–a opening and other reactions in steroid transformation by the thermophilic fungus *Myceliophthora thermophila*. *J Steroid Biochem Mol Biol* **116**:171–177.

Jones AE, Brown KC, Werner RE, Gotzkowsky K, Gaedigk A, Blake M, Hein DW, van der Horst C, and Kashuba AD (2010) Variability in drug metabolizing enzyme activity in HIV-infected patients. *Eur J Clin Pharmacol* **66**:475–485.

Kantae V, Krekels EHJ, Esdonk MJV, Lindenburg P, Harms AC, Knibbe CAJ, Van der Graaf PH, and Hankemeier T (2017) Integration of pharmacometabolomics with pharmacokinetics and pharmacodynamics: towards personalized drug therapy. *Metabolomics* **13**:9.

Labbé L, Sirois C, Pilote S, Arseneault M, Robitaille NM, Turgeon J, and Hamelin BA (2000) Effect of gender, sex hormones, time variables and physiological urinary pH on apparent CYP2D6 activity as assessed by metabolic ratios of marker substrates. *Pharmacogenetics* **10**:425–438.

Lai WG, Farah N, Moniz GA, and Wong YN (2011) A Baeyer-Villiger oxidation specifically catalyzed by human flavin-containing monooxygenase 5. *Drug Metab Dispos* **39**:61–70.

Lawson DR, Green TP, Haynes LW, and Miller AR (1997) Nuclear magnetic resonance spectroscopy and mass spectrometry of solanidine, leptinidine, and acetylleptinidine. Steroidal alkaloids from *Solanum chacoense* bitter. *J Agric Food Chem* **45**:4122–4126.

Leisch H, Morley K, and Lau PCK (2011) Baeyer-Villiger monooxygenases: more than just green chemistry. *Chem Rev* **111**:4165–4222.

Llerena A, Dorado P, Ramírez R, González I, Alvarez M, Peñas-Lledó EM, Pérez B, and Calzadilla LR (2012) CYP2D6 genotype and debrisoquine hydroxylation phenotype in Cubans and Nicaraguans. *Pharmacogenomics J* **12**:176–183.

Magliocco G, Desmeules J, Matthey A, Quirós-Guerrero LM, Bararpour N, Joye T, Marcourt L, Queiroz EF, Wolfender J-L, Gloor Y, et al. (2021) Metabolomics reveals biomarkers in human urine and plasma to predict cytochrome P450 2D6 (CYP2D6) activity. *Br J Pharmacol* **178**:4708–4725.

Omayio DG, Abong GO, and Okoth MW (2016) A review of occurrence of glycoalkaloids in potato and potato products. *Curr Res Nutr Food Sci* **4**:195–202.

Reynolds WF and Enríquez RG (2002) Choosing the best pulse sequences, acquisition parameters, postacquisition processing strategies, and probes for natural product structure elucidation by NMR spectroscopy. *J Nat Prod* **65**:221–244.

Saranya CV, Madhavankutty A, and Jayachandran K (2018) Solanidine acetate - the modified glyco-alkaloid imparting toxicity in green coloured and sprouted potatoes. *J Chem Edu Res Prac* **2**:1–6.

Sekiguchi S, Namegawa K, Nakane N, Okihara RSN, Omura K, Hofmann AF, and Iida T (2017) 3,4-Seco-12α-hydroxy-5β-cholan-3,4,24-trioic acid, a novel secondary bile acid: isolation from the bile of the common ringtail possum (*Pseudocheirus peregrinus*) and chemical synthesis. *J Pharmacogn Nat Prod* **3**:131.

Shou QY, Tan Q, and Wu Shen Z (2010) Two 22S-solanidine-type steroidal alkaloids from *Fritillaria anhuiensis. Fitoterapia* **81**:81–84.

Streetman DS, Bertino Jr JS, and Nafziger AN (2000) Phenotyping of drug-metabolizing enzymes in adults: a review of in-vivo cytochrome P450 phenotyping probes. *Pharmacogenetics* **10**:187–216.

Tay-Sontheimer J, Shireman LM, Beyer RP, Senn T, Witten D, Pearce RE, Gaedigk A, Gana Fomban CL, Lutz JD, Isoherranen N, et al. (2014) Detection of an endogenous urinary bio-marker associated with CYP2D6 activity using global metabolomics. *Pharmacogenomics* **15**:1947–1962.

Tracy TS, Venkataramanan R, Glover DD, and Caritis SN; National Institute for Child Health and Human Development Network of Maternal-Fetal-Medicine Units (2005) Temporal changes in drug metabolism (CYP1A2, CYP2D6 and CYP3A Activity) during pregnancy. *Am J Obstet Gynecol* **192**:633–639.

Verhoeven CH, Krebbers SF, Wagenaars GN, Booy CJ, Groothuis GM, Olinga P, and Vos RM (1998) In vitro and in vivo metabolism of the progestagen Org 30659 in several species. *Drug Metab Dispos* **26**:1102–1112.

Wadelius M, Darj E, Frenne G, and Rane A (1997) Induction of CYP2D6 in pregnancy. *Clin Pharmacol Ther* **62**:400–407.

Wang PH, Leu YL, Ismail W, Tang SL, Tsai CY, Chen HJ, Kao AT, and Chiang YR (2013) Anaerobic and aerobic cleavage of the steroid core ring structure by Steroidobacter denitrificans. *J Lipid Res* **54**:1493–1504.

Wu AH (2011) Drug metabolizing enzyme activities versus genetic variances for drug of clinical pharmacogenomic relevance. *Clin Proteomics* **8**:12.

Yu AM, Idle JR, Herraiz T, Küpfer A, and Gonzalez FJ (2003a) Screening for endogenous substrates reveals that CYP2D6 is a 5-methoxyindolethylamine O-demethylase. *Pharmacogenetics* **13**:307–319.

Yu AM, Idle JR, Krausz KW, Küpfer A, and Gonzalez FJ (2003b) Contribution of individual cytochrome P450 isozymes to the O-demethylation of the psychotropic beta-carboline alkaloids harmaline and harmine. *J Pharmacol Exp Ther* **305**:315–322.

Zhou J-L, Xin G-Z, Shi Z-Q, Ren M-T, Qi L-W, Li H-J, and Li P (2010) Characterization and identification of steroidal alkaloids in Fritillaria species using liquid chromatography coupled with electrospray ionization quadrupole time-of-flight tandem mass spectrometry. *J Chromatogr A* **1217**:7109–7122.

Zhou S-F (2009a) Polymorphism of human cytochrome P450 2D6 and its clinical significance: Part I. *Clin Pharmacokinet* **48**:689–723.

Zhou S-F (2009b) Polymorphism of human cytochrome P450 2D6 and its clinical significance: part II. *Clin Pharmacokinet* **48**:761–804.

Address correspondence to: Dr. J. Steven Leeder, Division of Clinical Pharmacology, Toxicology and Therapeutic Innovation, Department of Pediatrics, Children's Mercy Kansas City, 2401 Gillham Road, Kansas City, MO 64108. E-mail: sleeder@cmh.edu