Supporting Information

Identification of Poly(ADP-ribose)polymerase 1 and 2 (PARP1/2) as Targets of Andrographolide Using an Integrated Chemical Biology Approach

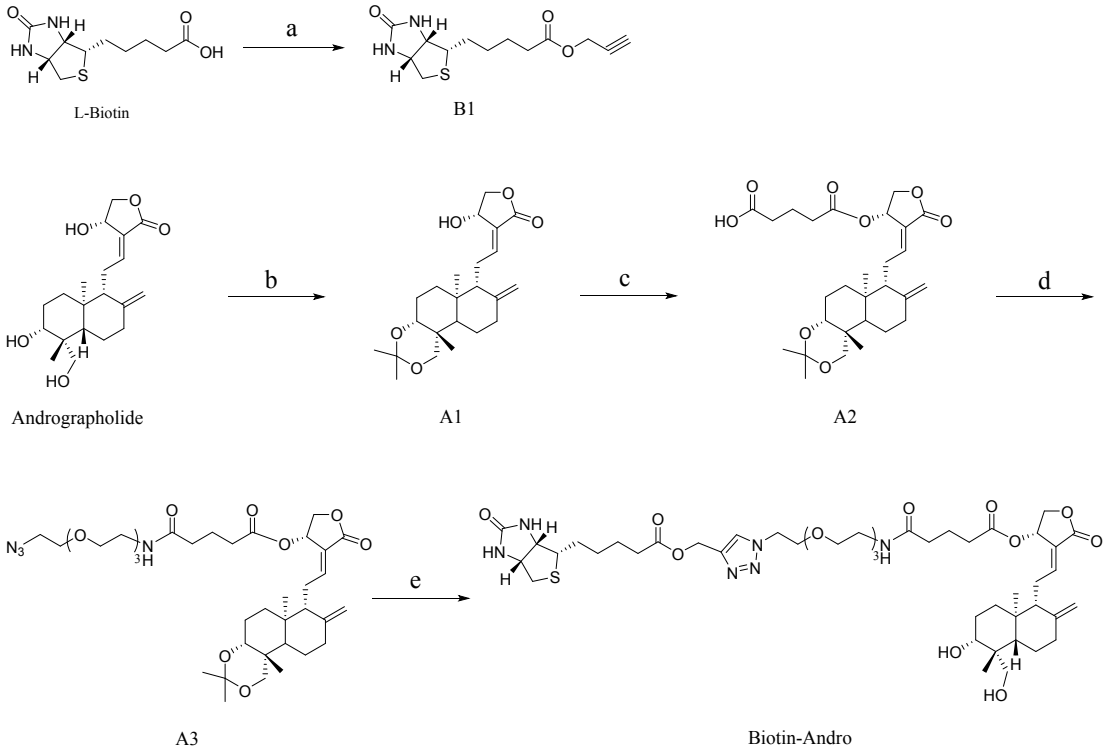
Wenchao Li [a], Bowen Pan [ab], Yang Shi [ab], Meiying Wang [a], Tianjiao Han [a], Qing Wang [a], Guifang Duan [a], and Hongzheng Fu*[a]

[a] State Key Laboratory of Natural and Biomimetic Drugs, School of Pharmaceutical Sciences, Peking University, Beijing 100191, People's Republic of China.

[b] College of Pharmacy, Guizhou University of Traditional Chinese Medicine, Guiyang 550025, People's Republic of China.

Table of Contents

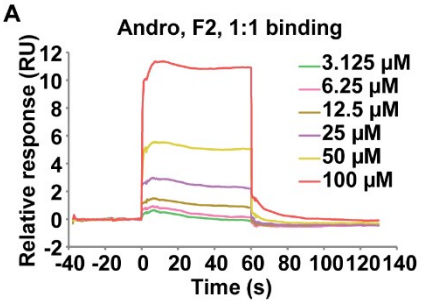
1. Supporting Scheme



Scheme S1: Synthesis of Biotinylated-Andrographolide (**Biotin-Andro**). Reagents and conditions: a) 3-bromoprop-1-yne, K_2CO_3, DMF, 50 °C, 1.5 h, 98%; b) 2,2-dimethoxypropane, P-toluenesulfonic acid, toluene, DMSO, 90 °C, 92%; c) glutaric anhydride, CH_2Cl_2, triethylamine, 0 °C, 3 h, 73%; d) 2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethan-1-amine, CH_2Cl_2, HATU, DIPEA, 0 °C → r.t., overnight, 83%; e) (i) **B1**, $CuSO_4$•$5H_2O$, sodium ascorbate, t-BuOH/H_2O, 5 h, (ii) 1N HCl, THF, 1 h.

2. Supporting Figures

A



B

Target	k_a (M^{-1} s^{-1})	k_d (s^{-1})	$K_D = k_d/k_a$ (M)
F2	1.63×10^2	3.22	1.97×10^{-2}

Figure S1: Sensorgram plots and binding constants of andrographolide with F2. (A) Sensorgram plots of andrographolide with F2. (B) Kinetic constants of andrographolide binding to F2.

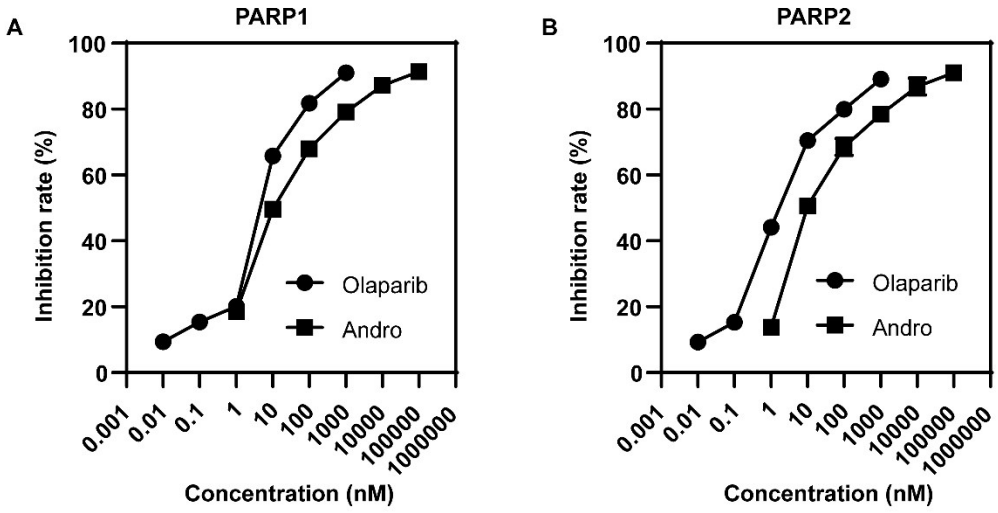


Figure S2: Dose-dependent inhibition assay. (A) PARP1 inhibition of Andro and Olaparib measured by the PARP1 enzyme assay. (B) PARP2 inhibition of Andro and Olaparib measured by the PARP2 enzyme assay.

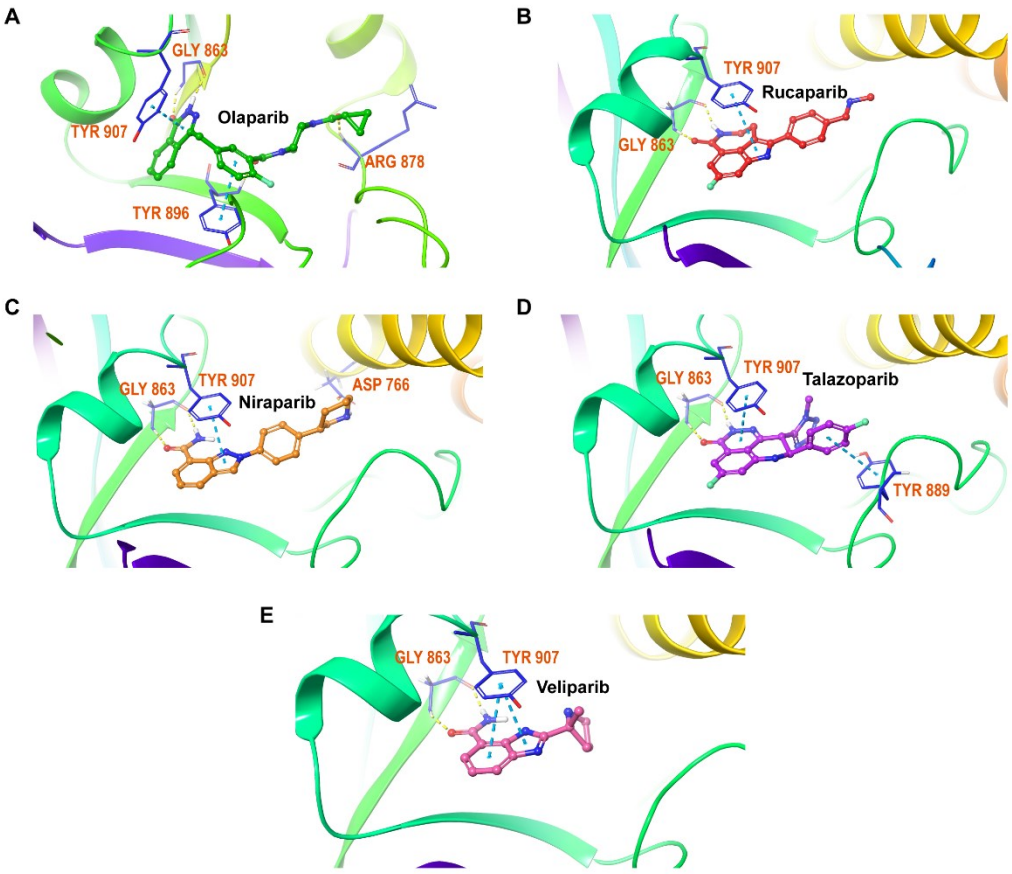


Figure S3: Co-crystal structures of PARP1 in complex with Olaparib (A, PDB code: 5DS3), Rucaparib (B, PDB code: 4RV6), Niraparib (C, PDB code: 7KK5), Talazoparib (D, PDB code: 4PJT) and Veliparib (E, PDB code: 7KK6).

3. Experimental Methods

3.1 Materials and Methods

Chemicals and biochemical reagents. Human coagulation F10 protein (His Tag), Rat Coagulation F2 Protein (His Tag) and Human PARP-1 Protein (His Tag) were purchased from Sino Biological Inc. (Sino Biological Inc., Beijing, China). ^{1}H- and ^{13}C- NMR were recorded on Bruker AVANCE III-400 spectrometers (Bruker, Baden, Switzerland). Chemical shifts (δ) are given in parts per million (ppm) relative to tetramethylsilane as an internal standard or the residual solvent protons. Coupling constants (*J*) are recorded in Hertz (Hz). High-Resolution Mass Spectra (HRMS) were recorded using a Waters Xevo G2 Q-TOF spectrometer (Waters, Milford, MA, USA). Spectral processing was performed using MestReNov.9.0.0 (Mestrelab Research, Santiago de Compostella, Spain). All reagents and anhydrous solvents obtained from commercial sources were used without further purification unless otherwise noted. All compounds were purified by column chromatography using 200-300 mesh silica gel (Qingdao Marine Chemical Factory, Qingdao, China). Thin-layer chromatography (TLC) was performed on silica gel GF254 plates (Merck Darmstadt, Germany) and visualized using UV light (254 or 365 nm) and anisaldehyde/acid spray reagent (ethanol: acetic acid: anisaldehyde: sulfuric acid = 85: 9: 1: 5).

3.2 Biochemical methods

Proteome microarray assays

HuProt™ v4.0 Human Proteome Microarrays (BC Biotechnology Co., Ltd) were blocked with blocking buffer for 1 h at room temperature and then incubated with 10 μM biotinylated-Andrographolide or free-biotin at room temperature for 1 h with gentle agitation. The microarrays were washed three times with TBST (5 min per wash) and were incubated with Cy3-Streptavidin (1:1000, Sigma-Aldrich, St Louis, MO) for 1 h at room temperature. Next, the microarrays were washed with TBST (3 times) and ultrapure water (2 times) at room temperature. The microarrays were spun dry and were scanned with a GenePix 4200A microarray scanner (Axon Instruments, Sunnyvale, CA) to visualize and record the results. Data were analyzed by GenePix Pro 6.0. To eliminate the signal inhomogeneity between different protein spots within the same chip due to inconsistent background values, it is processed by a background correction method. This is implemented as a foreground to background ratio, F/B, for each protein, and on this basis the SNR (signal-to-noise ratio) is defined, which is the mean value of F/B for the two replicate proteins.

Surface Plasmon Resonance (SPR) Studies

The binding of Andro to recombinant his-tagged PARP1, F2, and F10 was performed with a Biacore 8K (GE Healthcare Life Sciences, Uppsala, Sweden). PARP1, F2, and F10 were immobilized on a CM5 sensor chip by a standardized amine coupling method, and the SPR studies were performed in a running buffer containing PBS-P (10 mM phosphate buffer with 137 mM NaCl and 2.7 mM KCl, 0.05% surfactant P20, and pH 7.4). The binding studies were carried out using concentrations in the range of 0.39-100 μM with a flow rate of 30 μL/ min with an association time of 60 s and a dissociation time of 120 s. Data were fit and plotted to both steady state and a 1:1 kinetic model with Biacore evaluation software.

PARP1/2 Enzyme Assay

Three units of human PARP-1 (Enzo Life Sciences, Farmingdale, USA) or human (Abnova, CA, USA) PARP-2 enzyme, exposed to graded concentrations of Andro or Olaparib, were incubated in the presence of 5 mM $MgCl_2$,

2 mM dithiothreitol (DTT), 0.1 mg/mL human thymus DNA, 50 mmol/L Tris-HCl (pH 8.0), and 0.2 μCi [^{3}H]-NAD^+ at 37 °C for 1h. After completion, the reaction was terminated by adding 1 mL of 10% trichloroacetic acid (w/v) and centrifuged. The precipitate was then washed twice with 1 mL H_2O and resuspended with 1 mL NaOH (0.1 mol/L). Levels of radioactivity were determined using scintillation spectroscopy on a Beckman LS6000 liquid scintillation counter (Beckman Instruments Inc, Palo Alto, CA).

FXa Enzyme Assay

Andro and Rivaroxaban were tested for their inhibitory activity using a chromogenic assay. Assays were performed in 96-well microplates with the control (0.1% DMS0 10 μL) or test compounds with both buffer solutions (0.05 M Tris, 0.1 M NaCl, 0.1% BSA, pH 7.4) and human FXa solution (Kordia, Leiden, the Netherlands). After incubating for 15 minutes at 37 °C, 40 μL of chromogenic substrate S-2765 (Boatman Biotech, Shanghai, China) was added. After an hour, each well's absorbance was monitored at 405 nm using a microtiter plate reader (Tecan M200). The inhibition rate was calculated using the formula [(OD/min) control - (OD/min) sample]/(OD/min) control. The IC_{50} value was calculated and analyzed by Graphpad 5.0 software; the inhibition constant Ki= IC_{50}/ (1 + [substrate concentration]/Km).

Molecular Dynamic Simulation

Ligand Preparation. Chemical structure of Andro was prepared with LigPrep module Schrödingers Maestro software package and minimized using OPLS3 (Schrödinger Release 2018-1: Maestro version 11.5, Schrödinger, LLC, New York, NY).[1] The molecule was treated in the ionized form at pH = 7.0 ± 2.0, and the chiralities were preserved.[2]

*Induced Fit Docking***.** Flexible protein-ligand docking was carried out with Schrödinger induced fit docking protocol. The crystal structure of PARP1 and PARP2 (PDB code: 4PJT and 4TVJ) were retrieved from the Protein Data Bank database. The center of the grid was set as protein functional sites (inner box size = 10 Å; the outer box = 20 Å). During Glide docking, the receptor and ligand van der Waals scaling were set to 0.5 in both cases, and a default 20 poses were saved for each ligand. After that, extra precision (XP) was used for the Glide Redocking stage with the default cutoff. IFD score was calculated by default scoring function, and we selected the best-ranked conformation of Andro for the follow-up molecular dynamic simulation.[3]

Molecular Dynamic Simulation. The simulation was carried out with the AMBER 14 molecular simulation package. To get molecular mechanical parameters of Andro, ab initio quantum chemical methods were performed using the Gaussian 09 program. The geometry was fully optimized, and the electrostatic potential around them was measured at the B3LYP/6-31G* theoretical level. The RESP strategy was used to obtain the partial atomic charges. The complexes obtained by induced fit docking were embedded in an octahedral box of TIP3P water molecules, which extended 8 Å away from any solute atom. Sodium ions were added to neutralize the negative charges of simulated molecules. MD simulation was performed using the PMEMD module of AMBER 14. The atoms of the complex were minimized with 500 steps of steepest-descent followed by 500 steps of conjugate-gradient minimization with a large constraint of 500 kcal $mol^{-1}Å^{-2}$. A 1000 step steepest-descent is then performed, followed by a 1500 step conjugate-gradient minimization without any restrictions on the complex atoms. Subsequently, after 20 ps of MD, the temperature was slowly increased from 0 to 300 K with weak (10 kcal $mol^{-1}Å^{-2}$) restraint on the complex, culminating in a final production simulation of 100.0 ns of pressure (1 atm) and temperature (300 K) at constant conditions. In the entire simulation, SHAKE was applied to all bonds involving hydrogen atoms. Application of periodic boundary conditions with minimum image convention to calculate non-

bonded interactions. The cutoff strength of the Lennard-Jones interaction was 10 Å. The final configurations of the complexes were generated by 1000 steps of the minimized average structure of the last 50.0 ns of MD.

Growth inhibition assays

MCF-7, Capan-1, MDA-MB-231, and MDA-MB-436 cell lines were purchased from the Peking Union Medical College Cell Library (Beijing, China). MCF-7 and MDA-MB-231 cells were cultured in DMEM supplemented with 10% fetal bovine serum (FBS) at 37 °C in a 5% CO2, 95% air incubator. MDA-MB-436 cells were maintained in Leibovitzs L-15 medium (L-15) supplemented with 10% FBS at 37 °C in an air incubator, Capan-1 cells were maintained in IMEM supplemented with 20% FBS at 37 °C in a 5% CO_2, 95% air incubator. All the cell lines were supplemented with Penicillin-Streptomycin (0.1 mg/mL). Compounds were dissolved in DMSO to a final concentration of 10 mM and further diluted to a corresponding concentration with a complete culture medium for cell culture experiments. All of the growth inhibition assays were conducted using exponentially growing cells in 96-well plates at a density of 700 cells/well for 7 days**.** Absorbance at 450 nm was measured 2 h after the addition of 10 μL CCK-8 per well. The percent cell viability inhibition was expressed as: Cell viability = [OD (treated) − OD (control)] / OD (control) × 100. Growth curves were plotted using Origin 2018 (Origin Lab Corporation, Northampton, Ma, USA).

Apoptosis Assay

MDA-MB-436 cells were digested with 0.25% trypsin at the logarithmic growth phase and incubated in a 6-well culture plate (3×10^5 cells per well) for 24 h. MDA-MB-436 cells were treated with different concentrations (0, 5, 10, and 20 μM) of Andro for 48 h, after which the culture medium was discarded, and cells were washed 3 times with PBS. After treatment, the cells were washed 3 times with PBS again. Digested cells were collected into a 7 mL centrifuge tube and centrifuged for 5 min. The supernatant was discarded. 300 μL of 1× Binding Buffer was added to each tube. Annexin V-FITC (5 μL) was added, and the mix was incubated for 15 min in the dark. PI (50 μL) and 1×binding buffer (200 μL) were added in order 5 minutes before detection. Fluorescence was analyzed by flow cytometry (MoFlo Legacy; Beckman Coulter). Results were analyzed using FlowJo software (Treestar, Ashland, OR).

Wound Healing Assay

The MDA-MB-436 cells were seeded in a 6-well plate (1×10^6 cells per well) for 48 h and then starved in a serum-free medium for 6 h. The confluent monolayer cells were then scratched with a sterile 10 μL pipette tip. The monolayer was washed several times with media to remove floating cells and added with medium containing either Andro (0, 5, 10, and 20 μM) or Olaparib (0, 5, 10, and 20 μM). The images were captured at 0 and 48 h under an inverted microscope (XDS-1B, COIC, Chongqing, China). The analysis was performed with ImageJ software (NIH, Bethesda, MD).

In Vivo Pharmacology

Female BALB/c nude (nu/nu) mice (weight, 16 - 18 g) were obtained from the Model Animal Research Center of Nanjing University. The mice were housed in an SPF level environment with a controlled temperature (23 ± 2 °C), provided with sterile food and water. Mice were acclimated for seven days before inoculation. The MDA-MB-436 xenograft model was established by inoculating 10^7 cells subcutaneously in the nude mice. Once tumors reached ~100 mm^3, the mice were divided into five groups (n = 5) randomly and administered po at doses of 15, 30, and 60 mg/kg (dissolved in 0.5% CMC-Na); Olaparib (30 mg/kg) served as a positive control. Tumor volumes and body

weight were assessed every three days in two dimensions using a caliper. Tumor volumes were calculated using the following equation: $V = 0.5(XY^2)$, where X and Y are the long and short diameters of the tumor, respectively. Mice were sacrificed 2 h after the last treatment dose, and the tumors were removed. Tumors were washed 2~3 times with D-Hanks solution, drained and stored, weighed, and volume measured.

Whole-cell patch-clamp

The CHO-hERG cells were cultured in F12 medium (Gibco) supplemented with 10% (v/v) fetal bovine serum (FBS) and 0.5 mg/mL Geneticin (Invitrogen) at 37 °C in a humidified environment (5% CO_2/95% air). The cells were seeded out two days before reaching 70% confluency. Before use, the cells were washed in PBS and incubated with 5 mL Detachin (Genlantis) for 4-5 min at 37 °C to detach cells from the culture dish. The harvested cells were resuspended in an F12 medium at a density of 2 million cells/mL. The cells were transferred to a QPatch instrument (Sophion Bioscience, Denmark) and allowed to recover for 20 min in the Qstir cell preparation station on the Qpatch-8 before the experiment.

The tail currents of the hERG channel were evaluated using the Qpatch automated patch clamp platform (Sophion Bioscience, Denmark). The following solutions were used during patch-clamp recording (compositions in mM): internal solution: KCl 120, $CaCl_2$ 5.374, $MgCl_2$ 1.75, KOH 31.25, EGTA 10, HEPES 10, Na_2ATP 4, pH 7.2 (KOH); external solution: NaCl 145, KCl 4, $MgCl_2$ 1, $CaCl_2$ 2, HEPES 10, glucose 10, pH 7.4 (NaOH). All solutions were sterile filtered. Cells were clamped at -80 mV and hyperpolarized to -100 mV to monitor the change of series resistance. The voltage protocol for hERG ion channel started with a short (200 ms) -50 mV step to establish the baseline region. A depolarizing step was applied to the test potential of 20 mV for 2s, and then the cell was depolarized to -50mV to evoke outward tail currents. Currents were filtered using the internal Bessel filter in Qpatch. Recording started in an external solution. Five increasing concentrations of the test compounds were applied after this control period, each for approximately 4 min to record a complete concentration-response curve. The last control period (Saline) was used as a baseline for data normalization. Cisapride (2 μM) was applied as a reference inhibitor at the end of the protocol. The sampling frequency is 2000Hz.

3.3 Synthetic Procedures

Synthesis of compound B1. To a solution of L-biotin (489 mg, 2.0 mmol) in DMF (8 mL) was added potassium carbonate (276 mg, 2.0 mmol) and propargyl bromide (262 mg, 2.2 mmol), and the mixture was stirred at 50 °C for 2 h. The reaction system was added 20 mL H_2O and extracted with ethyl acetate (3 × 20 mL). The combined organic layers were washed with 30 mL brine (1×), dried over anhydrous Na_2SO_4 and concentrated in vacuo. Purification was performed by column chromatography (silica gel, CH_2Cl_2/MeOH, 50: 1) to yield compound **B1** as a white solid (522 mg, 1.85 mmol, 92%).^{1}H NMR (400 MHz, Chloroform-*d*) δ 6.18 (s, 1H), 5.63 (s, 1H), 4.66 (d, *J* = 2.5 Hz, 2H), 4.49 (dd, *J* = 7.7, 4.7 Hz, 1H), 4.30 (dd, *J* = 7.8, 4.5 Hz, 1H), 3.14 (ddd, *J* = 8.4, 6.5, 4.5 Hz, 1H), 2.89 (dd, *J* = 12.8, 4.9 Hz, 1H), 2.73 (d, *J* = 12.8 Hz, 1H), 2.48 (t, *J* = 2.5 Hz, 1H), 2.38 (t, *J* = 7.5 Hz, 2H), 1.78 – 1.59 (m, 4H), 1.44 (tt, *J* = 9.8, 5.9 Hz, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 172.81, 163.85, 77.78, 74.90, 61.98, 60.16, 55.48, 51.86, 40.56, 33.61, 28.30, 28.22, 24.66. ESI-HRMS (m/z): calcd for $C_{13}H_{20}N_2O_3S$ $[M+H]^+$, 283.1116; found, 283.1116.

Synthesis of compound A1. Andrographolide (1.0 g, 2.8 mmol) was dissolved in a mixed solution of DMSO (1.6 mL) and toluene (10 mL). Subsequently, 2,2-dimethoxypropane (1.2 mL, 10 mmol) and TsOH (50 mg, 1.43 mmol) were added and the mixture was stirred at 90 °C for 4 h. The reaction mixture was poured in ethyl acetate (50 mL) and washed with H_2O (2×), and the combined organic portion was washed with 50 mL brine (1×), dried over anhydrous Na_2SO_4. and concentrated in vacuo. Purification was performed by column chromatography (silica gel, petroleum ether/ethyl acetate 6:1 v/v) to yield compound **A1** as a white solid (976 mg, 2.5 mmol, 88%).^{1}H NMR (400 MHz, DMSO-d_6) δ 6.60 (td, *J* = 6.9, 1.8 Hz, 1H), 5.70 (d, *J* = 6.0 Hz, 1H), 4.89 (s, 1H), 4.82 (s, 1H), 4.64 (s, 1H), 4.37 (dd, *J* = 9.9, 6.1 Hz, 1H), 4.04 – 3.98 (m, 1H), 3.85 (d, *J* = 11.6 Hz, 1H), 3.38 (dd, *J* = 9.1, 3.7 Hz, 1H), 3.08 (d, *J* = 11.5 Hz, 1H), 2.47 (dt, *J* = 3.5, 1.8 Hz, 3H), 2.35 – 2.27 (m, 1H), 2.04 – 1.82 (m, 3H), 1.75 – 1.52 (m, 3H), 1.30 (s, 3H), 1.28 – 1.25 (m, 1H), 1.22 (s, 3H), 1.21 – 1.13 (m, 1H), 1.10 (s, 3H), 0.84 (s, 3H).

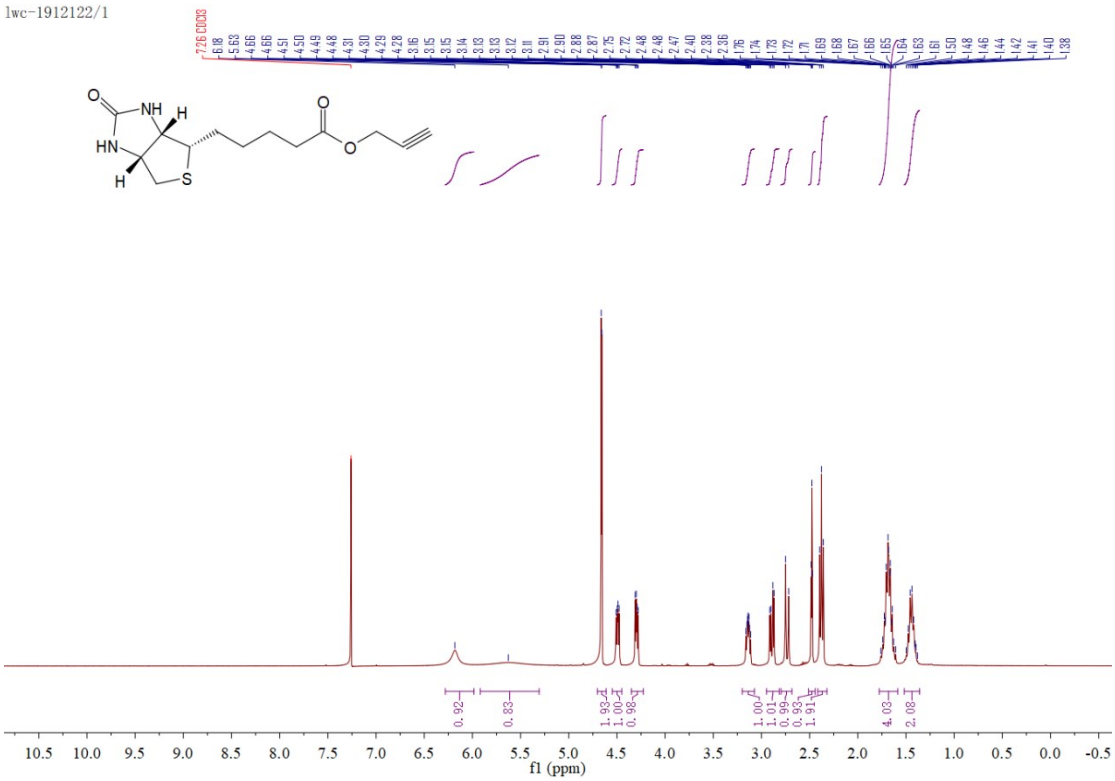
Synthesis of compound A2. To a solution of **A1** (195 mg, 0.5mmol) in CH_2Cl_2 (10 mL) was added glutaric anhydride (68 mg, 0.6 mmol) and DIPEA (87 μL, 0.5 mmol). The reaction was stirred for 3 h in an ice bath. The crude product was purified by column chromatography (silica gel, CH_2Cl_2/MeOH, 80: 1) to yield compound **A2** as a white solid (217 mg, 0.43 mmol, 86%). ^{1}H NMR (400 MHz, Methanol-d_4) δ 6.97 (td, *J* = 6.9, 1.8 Hz, 1H), 6.07 (d, *J* = 6.2 Hz, 1H), 4.94 – 4.92 (m, 1H), 4.64 – 4.58 (m, 2H), 4.31 (dd, *J* = 11.1, 1.9 Hz, 1H), 4.04 (d, *J* = 11.6 Hz, 1H), 3.53 (dd, *J* = 8.9, 3.7 Hz, 1H), 3.21 (d, *J* = 11.6 Hz, 1H), 2.63 – 2.53 (m, 1H), 2.51 – 2.43 (m, 3H), 2.39 (t, *J* = 7.3 Hz, 2H), 2.12 – 1.97 (m, 4H), 1.93 (q, *J* = 7.3 Hz, 2H), 1.87 – 1.73 (m, 3H), 1.42 (s, 3H), 1.40 – 1.36 (m, 2H), 1.34 (s, 3H), 1.32 (d, *J* = 7.1 Hz, 1H), 1.22 (s, 3H), 0.98 (s, 3H). ^{13}C NMR (101 MHz, Methanol-d_4) δ 175.10, 172.62, 169.99, 150.12, 147.78, 124.36, 107.91, 98.99, 76.56, 71.79, 67.94, 63.49, 55.70, 52.04, 38.10, 37.54, 37.29, 34.20, 32.56, 32.46, 26.14, 25.17, 25.10, 24.81, 24.34, 22.91, 19.87, 15.21. ESI-HRMS (m/z): calcd for $C_{28}H_{44}NO_8$ $[M+NH_4]^+$, 522.3067; found, 522.3073.

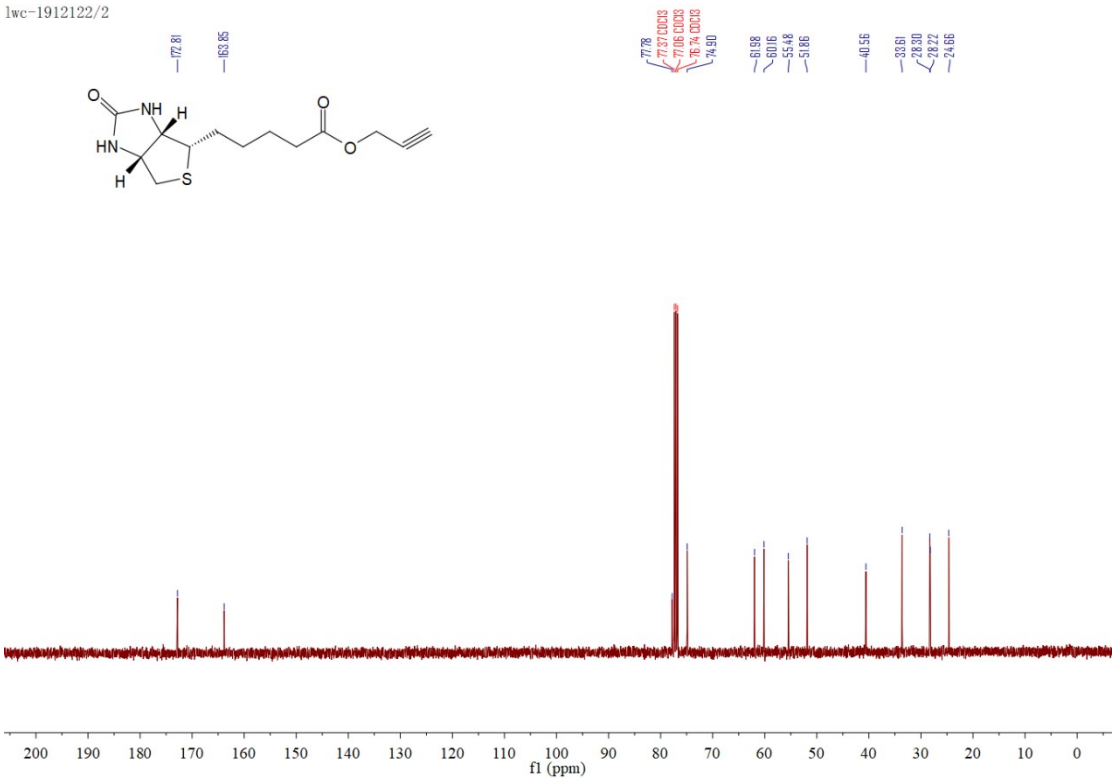
Synthesis of compound A3. A2 (101 mg, 0.2 mmol) and HATU (87 mg, 0.4 mmol) were dissolved in 5 mL of DCM, followed by 50 μL of DIPEA (0.3 mmol). After stirring for 20 min in an ice bath, amine (87 mg, 0.4 mmol) was added and kept mixing for 3 h. The reaction system was added 25 mL H_2O and extracted with ethyl acetate (3 × 20 mL). The combined organic layers were washed with 30 mL brine (1×), dried over anhydrous Na_2SO_4 and concentrated in vacuo. Purification was performed by column chromatography (silica gel, CH_2Cl_2/MeOH, 100: 1→50: 1) to yield compound **A3** as a white solid (105 mg, 1.85 mmol, 85%).^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.01 (td, *J* = 6.9, 1.8 Hz, 1H), 6.11 (t, *J* = 5.6 Hz, 1H), 5.92 (d, *J* = 5.9 Hz, 1H), 4.89 (s, 1H), 4.58 – 4.50 (m, 2H), 4.24 (dd, *J* = 11.2, 1.9 Hz, 1H), 3.95 (d, *J* = 11.7 Hz, 1H), 3.68 (d, *J* = 4.8 Hz, 1H), 3.67 (s, 4H), 3.66 – 3.64 (m,

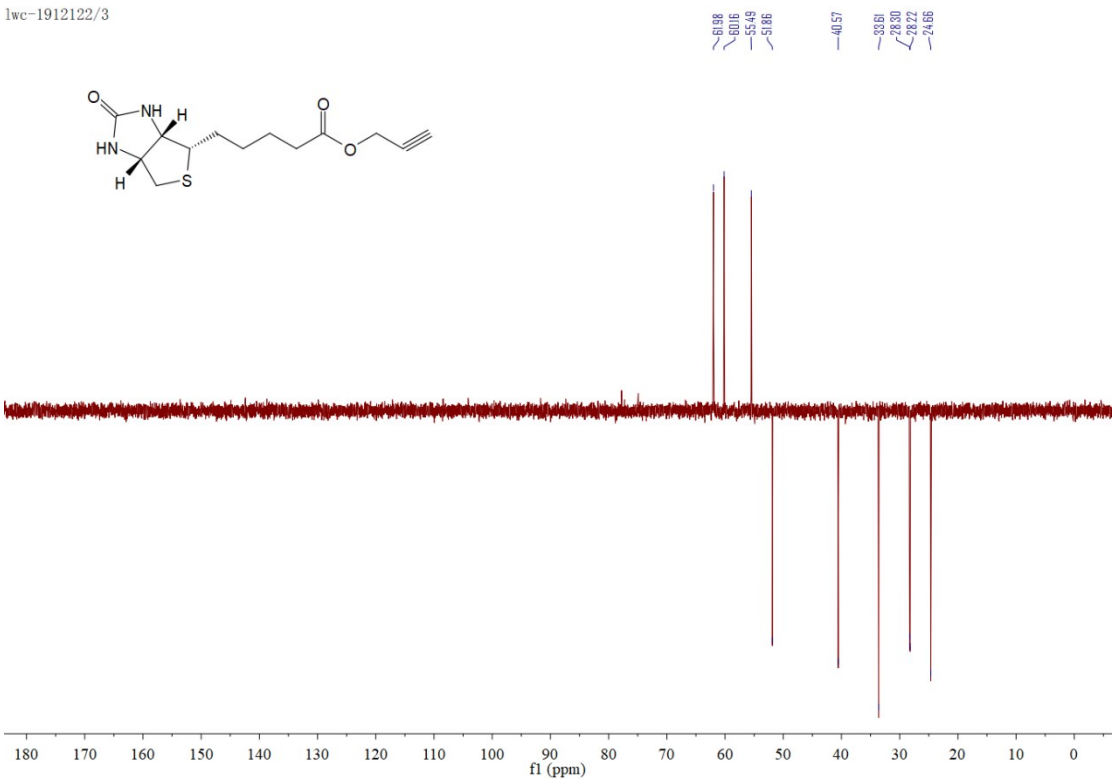
2H), 3.63 (d, *J* = 2.2 Hz, 1H), 3.55 (t, *J* = 5.1 Hz, 2H), 3.50 – 3.42 (m, 3H), 3.39 (t, *J* = 5.0 Hz, 2H), 3.17 (d, *J* = 11.6 Hz, 1H), 2.43 (t, *J* = 7.4 Hz, 4H), 2.39 (q, *J* = 3.3 Hz, 1H), 2.25 (t, *J* = 7.2 Hz, 2H), 2.04 – 1.92 (m, 4H), 1.88 – 1.65 (m, 5H), 1.40 (s, 3H), 1.36 (s, 3H), 1.34 – 1.25 (m, 3H), 1.23 (q, *J* = 5.4, 4.0 Hz, 1H), 1.19 (s, 3H), 0.93 (s, 3H). ^{13}C NMR (101 MHz, Chloroform-*d*) δ 172.70, 171.74, 169.09, 150.61, 147.04, 123.82, 108.88, 99.09, 76.17, 71.63, 70.65, 70.58, 70.51, 70.21, 70.01, 69.80, 67.80, 63.87, 55.88, 52.18, 50.65, 39.20, 38.36, 37.87, 37.57, 35.10, 34.51, 33.14, 27.07, 26.11, 25.46, 25.30, 24.94, 23.11, 20.60, 16.14. ESI-HRMS (m/z): calcd for $C_{36}H_{57}N_4O_{10}$ $[M+H]^+$, 705.4075; found, 705.4077.

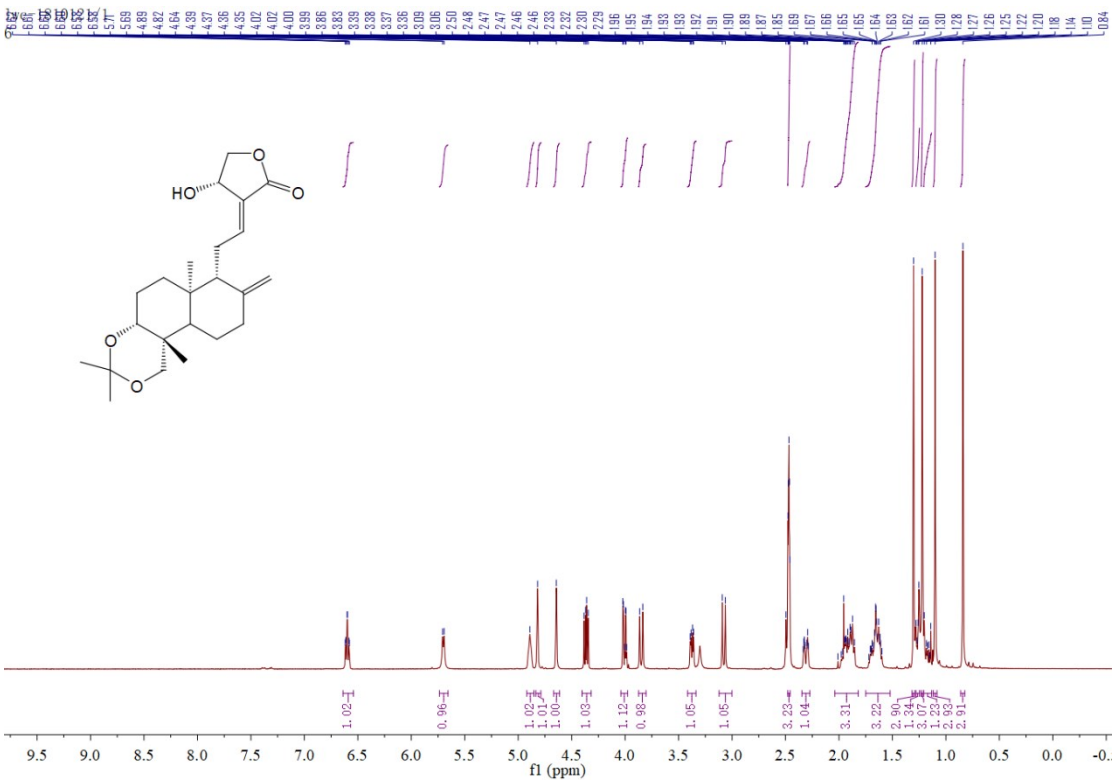
Synthesis of compound Biotin-Andro. A3 (140 mg, 0.2 mmol) and B1 (56 mg, 0.2 mmol) was dissolved in a mixed solution of H_2O (2.5 mL) and t-BuOH (2.5 mL). Subsequently, copper sulfate pentahydrate (98 mg, 0.4 mmol) and sodium ascorbate (158 mg, 0.8 mmol) were added and the mixture was stirred at room temperature for 5 h. The mixture was added 30 mL H_2O and extracted with ethyl acetate (3 × 20 mL). The combined organic layers were washed with 30 mL brine (1×), dried over anhydrous Na_2SO_4 and concentrated in vacuo. A magnetically stirred solution of crude product in THF (3 mL) was added 1N HCl (3 mL) and stirred at room temperature for 1 h. The mixture was added 30 mL saturated aqueous solution of sodium bicarbonate and extracted with ethyl acetate (3 × 15 mL). The combined organic layers were washed with 30 mL brine (1×), dried over anhydrous Na_2SO_4 and concentrated in vacuo. Purification was performed by column chromatography (silica gel, CH_2Cl_2/MeOH, 50: 1→20: 1) to yield compound **Biotin-Andro** as a white solid (172 mg, 0.18 mmol, 91%). ^{1}H NMR (400 MHz, Methanol-d_4) δ 8.08 (s, 1H), 7.97 (t, *J* = 5.7 Hz, 1H), 6.96 (ddd, *J* = 7.7, 6.6, 1.6 Hz, 1H), 6.05 (d, *J* = 6.0 Hz, 1H), 5.21 (s, 2H), 4.89 (s, 1H), 4.60 (q, *J* = 5.7 Hz, 3H), 4.55 (s, 1H), 4.51 (dd, *J* = 7.9, 4.9 Hz, 1H), 4.36 – 4.28 (m, 2H), 4.12 (d, *J* = 10.9 Hz, 1H), 3.91 (t, *J* = 5.0 Hz, 2H), 3.63 (d, *J* = 4.9 Hz, 4H), 3.61 (s, 4H), 3.55 (t, *J* = 5.5 Hz, 2H), 3.44 – 3.38 (m, 2H), 3.37 (s, 4H), 3.21 (dt, *J* = 9.8, 5.3 Hz, 1H), 2.94 (dd, *J* = 12.8, 5.0 Hz, 1H), 2.72 (d, *J* = 12.7 Hz, 1H), 2.59 (ddd, *J* = 16.6, 6.6, 3.3 Hz, 1H), 2.42 (dt, *J* = 22.3, 7.4 Hz, 6H), 2.28 (t, *J* = 7.4 Hz, 2H), 2.04 (dt, *J* = 12.5, 7.2 Hz, 1H), 2.00 – 1.91 (m, 3H), 1.91 – 1.84 (m, 1H), 1.79 (q, *J* = 5.5, 3.6 Hz, 3H), 1.76 – 1.65 (m, 3H), 1.65 – 1.53 (m, 1H), 1.45 (q, *J* = 7.7 Hz, 2H), 1.40 – 1.25 (m, 3H), 1.23 (s, 3H), 0.73 (s, 3H). ^{13}C NMR (101 MHz, Methanol-d_4) δ 173.77, 173.39, 172.61, 170.01, 164.62, 150.16, 147.72, 142.51, 125.22, 124.34, 107.62, 79.47, 71.83, 70.14, 70.07, 70.01, 69.87, 69.18, 68.89, 68.02, 63.56, 61.95, 60.22, 56.79, 55.79, 55.54, 54.85, 50.06, 42.29, 39.69, 39.09, 38.62, 37.51, 36.70, 34.58, 33.26, 32.70, 28.24, 28.04, 27.66, 25.06, 24.48, 23.79, 22.03, 20.78, 14.23. ESI-HRMS (m/z): calcd for $C_{46}H_{71}N_6O_{13}S$ $[M+H]^+$, 947.4800; found, 947.4819.

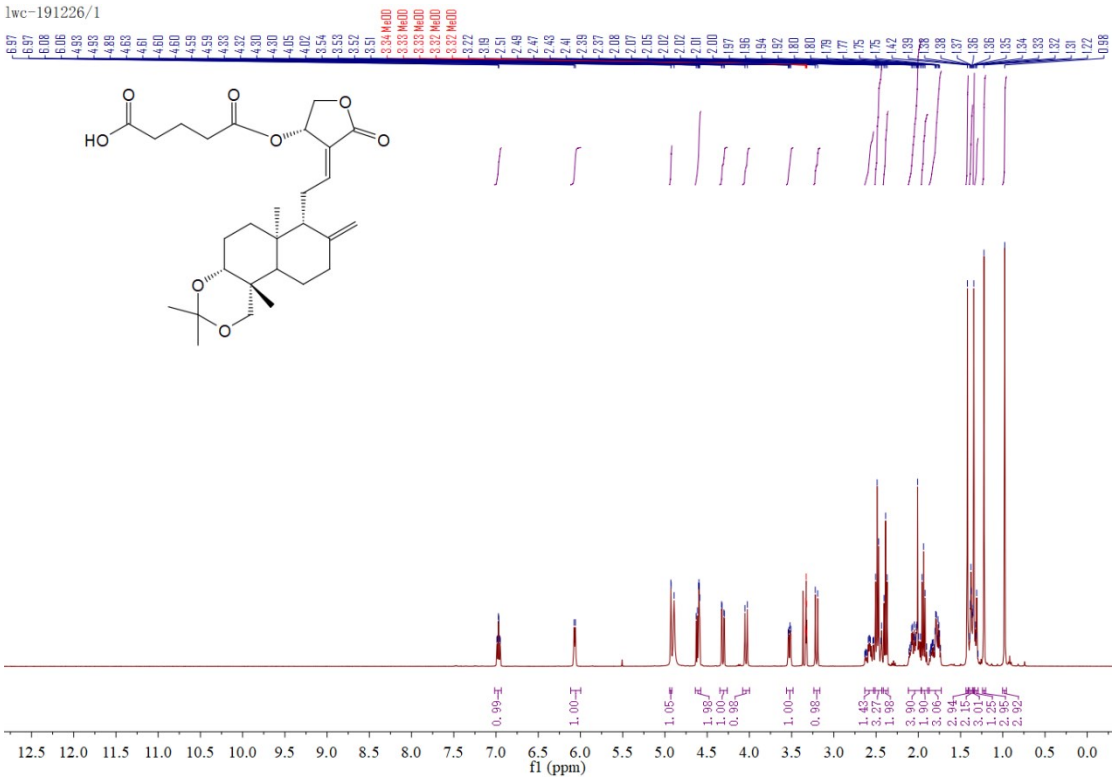
3.4 ^{1}H-NMR, ^{13}C-NMR APT and DEPT-135 spectra

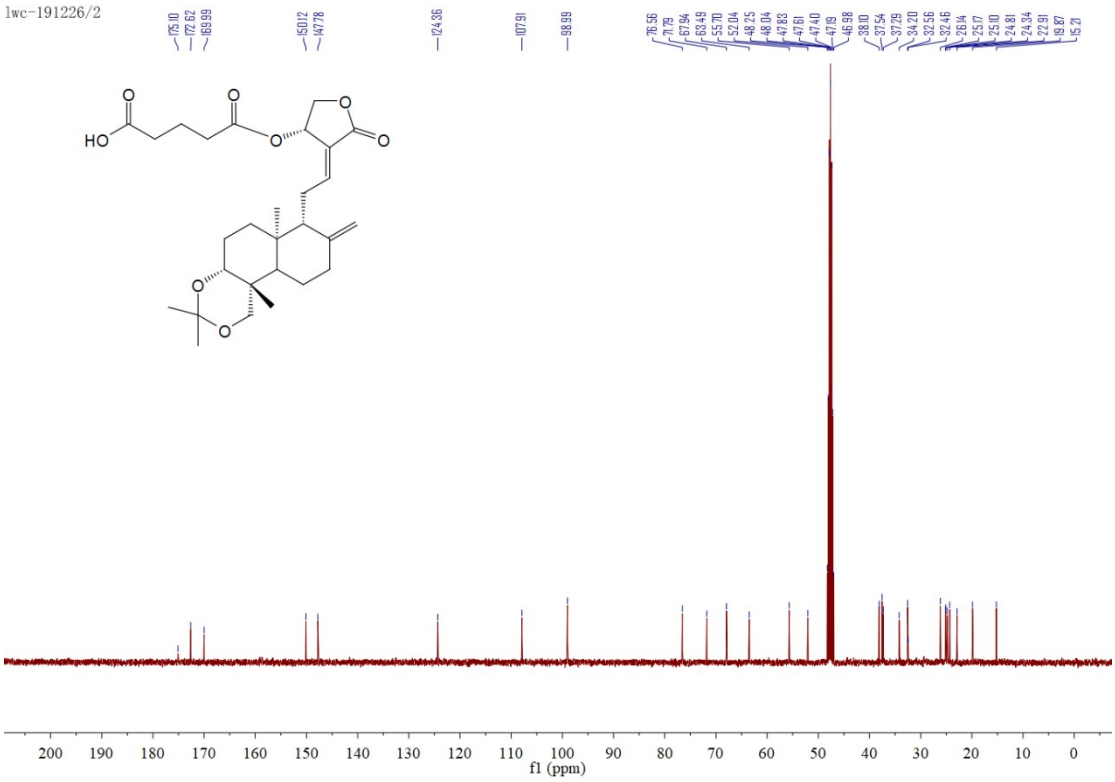


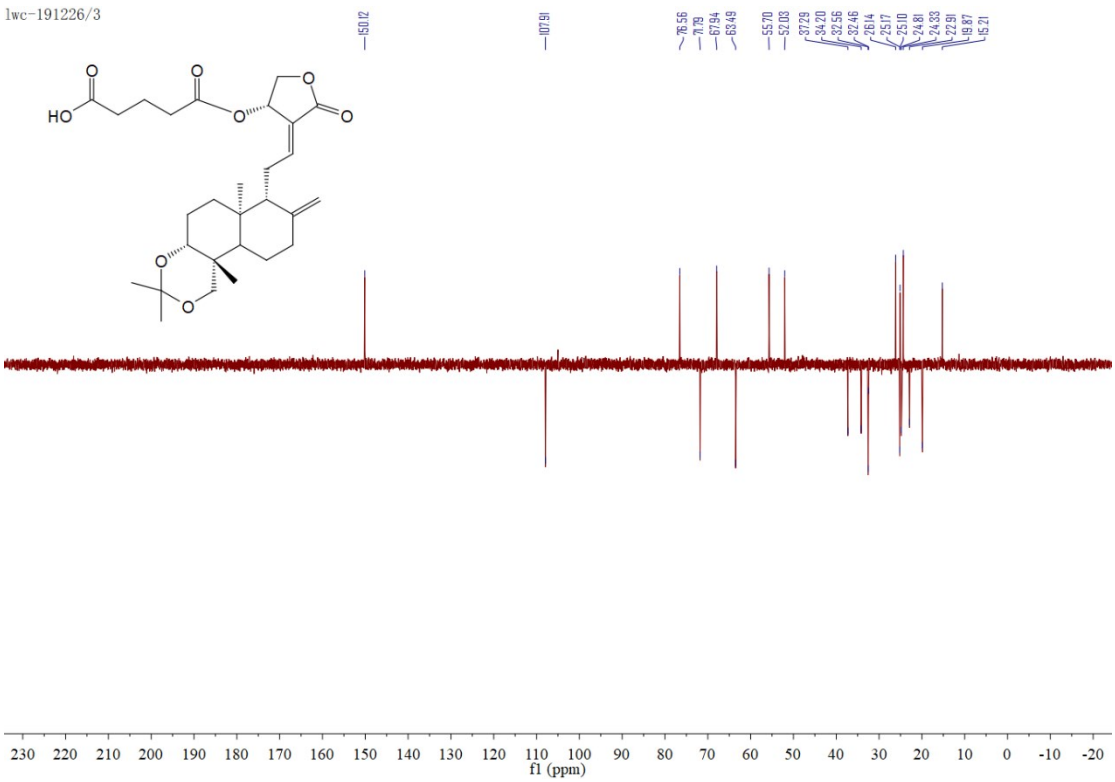


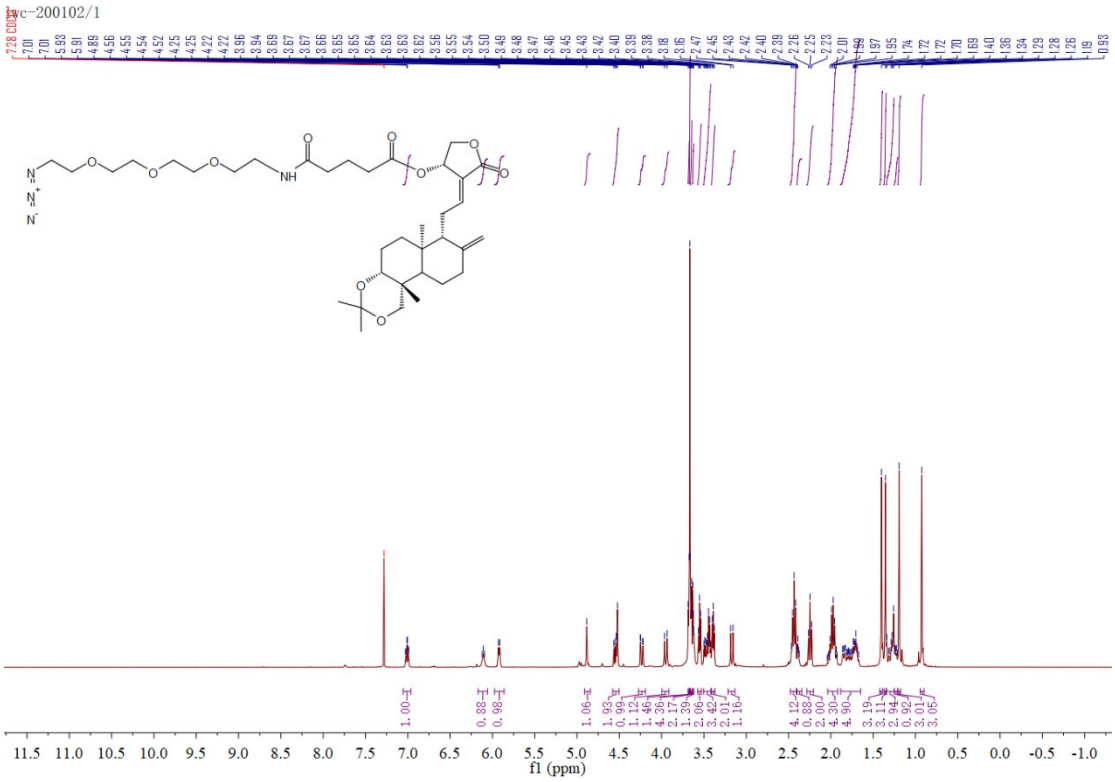

lwc-1912122/3
61.98
60.16
55.49
51.86
40.57
33.61
28.30
28.22
24.66
f1 (ppm)

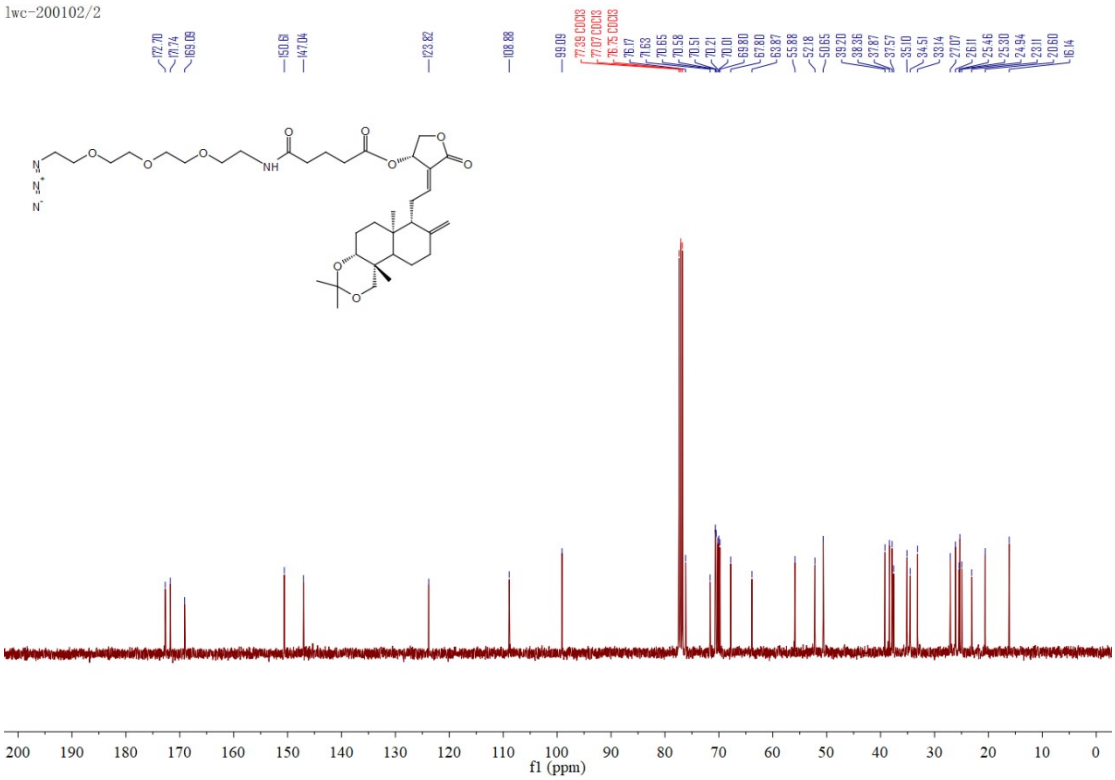

f1 (ppm)

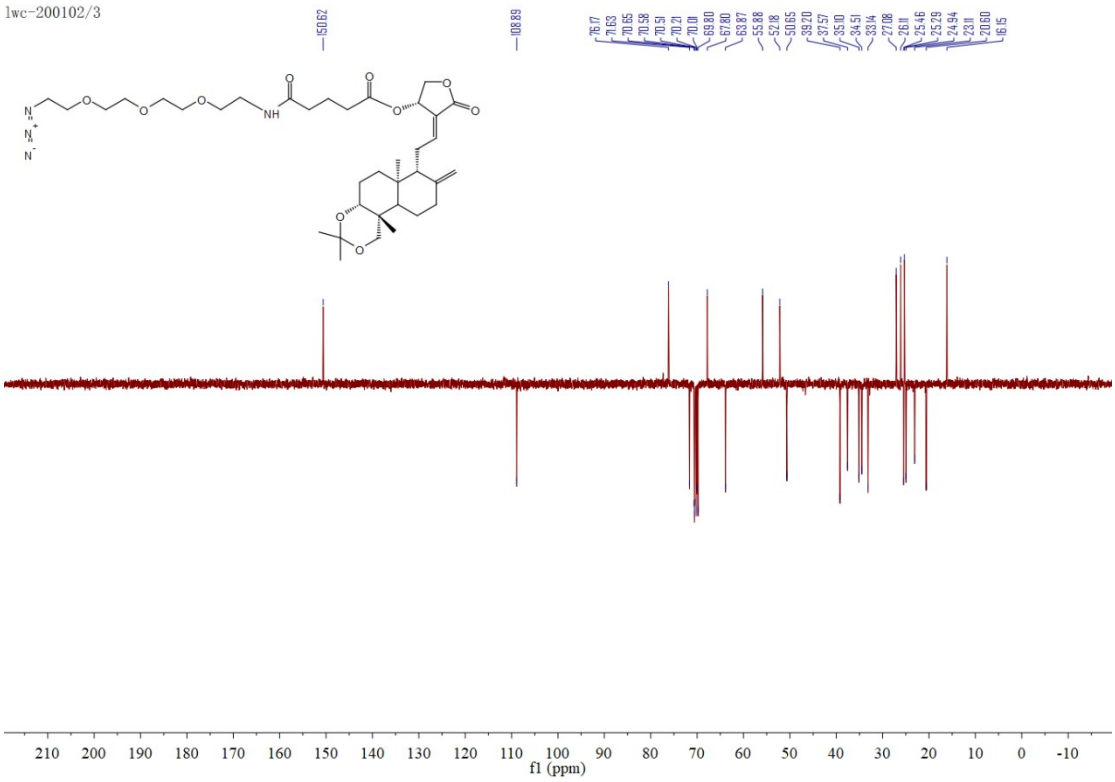

lwc-191226/1
f1 (ppm)

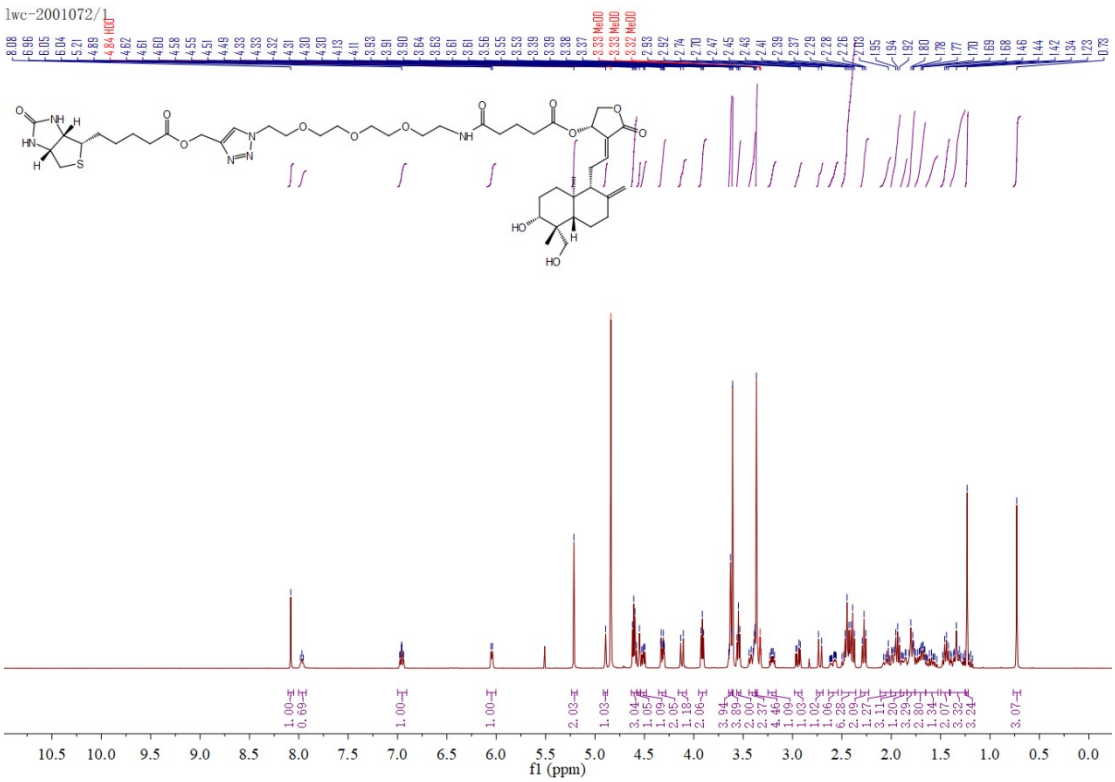

lwc-191226/2
f1 (ppm)

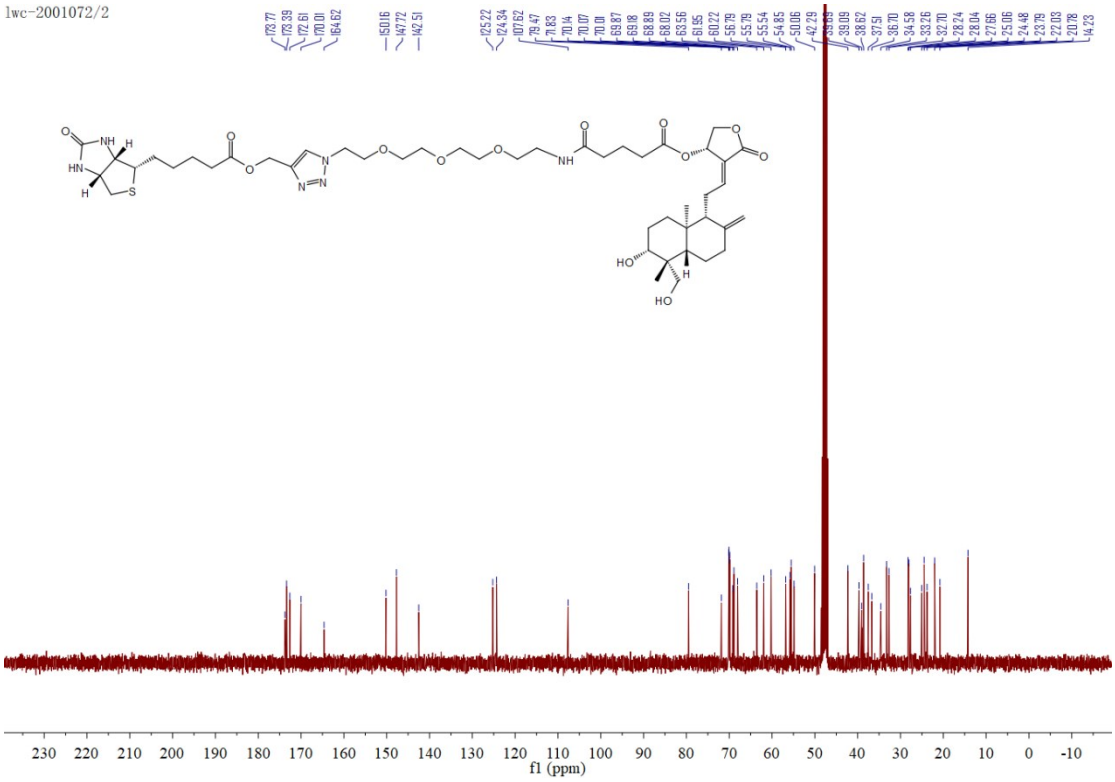

lwc-191226/3
150.12
107.91
76.56
76.79
67.94
63.49
55.70
52.03
37.29
34.20
32.56
32.46
26.14
25.17
25.10
24.81
24.33
22.91
19.87
15.21
HO
O
O
O
O
O
O
O
O
230 220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10 -20
f1 (ppm)


lwc-200102/1
7.28
7.01
7.01
7.01
5.93
5.91
4.89
4.56
4.55
4.54
4.52
4.25
4.22
4.22
3.96
3.94
3.89
3.87
3.87
3.86
3.85
3.84
3.83
3.82
3.56
3.55
3.54
3.50
3.49
3.48
3.47
3.46
3.45
3.43
3.42
3.40
3.39
3.38
3.18
3.16
2.47
2.45
2.43
2.42
2.40
2.39
2.26
2.25
2.23
2.01
1.99
1.97
1.95
1.74
1.72
1.72
1.70
1.69
1.40
1.36
1.34
1.29
1.28
1.26
1.18
0.93
NH
1.00
0.88
0.98
1.06
1.93
0.99
1.12
1.46
4.37
2.39
2.06
3.42
2.01
1.16
4.12
0.88
2.00
4.30
4.90
3.19
3.11
2.94
0.92
3.01
3.05
11.5 11.0 10.5 10.0 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 0.0 -0.5 -1.0
f1 (ppm)

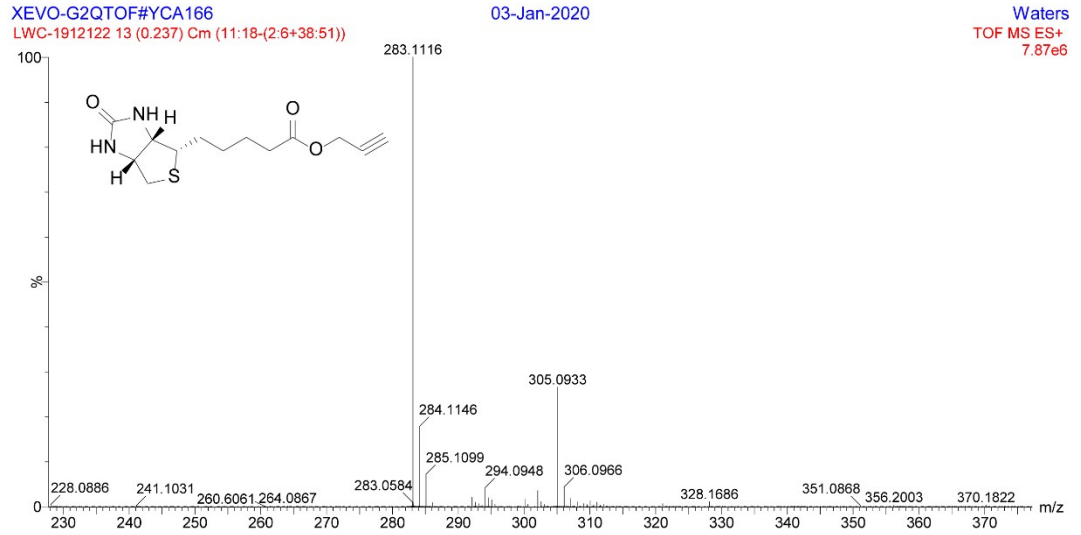

lwc-200102/2
172.70
170.74
169.08
150.61
147.04
123.82
108.88
99.09
77.39 CDCl3
77.07 CDCl3
76.75 CDCl3
76.17
71.63
70.65
70.58
70.51
70.21
70.01
69.80
67.80
63.87
55.88
52.18
50.65
39.20
38.36
37.87
37.57
35.10
34.51
33.14
27.07
26.11
25.46
25.30
24.94
23.11
20.60
16.14
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

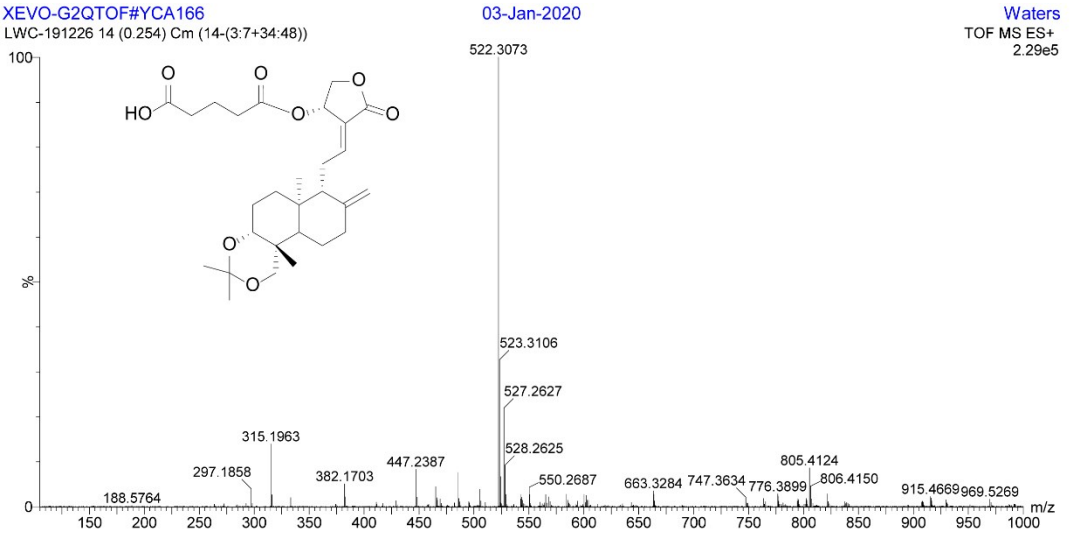

lwc-200102/3
150.62
108.89
76.17
71.63
70.65
70.58
70.51
70.21
70.01
69.80
67.80
63.87
55.88
52.18
50.65
39.20
37.57
35.10
34.51
33.14
27.08
26.11
25.46
25.29
24.94
23.11
20.60
16.15
210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

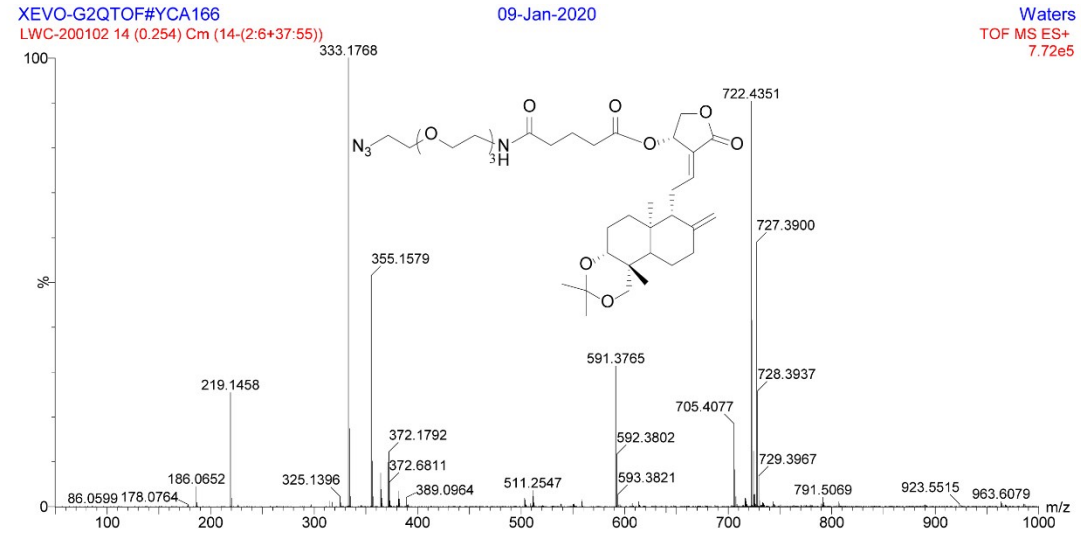

lwc-2001072/1
f1 (ppm)

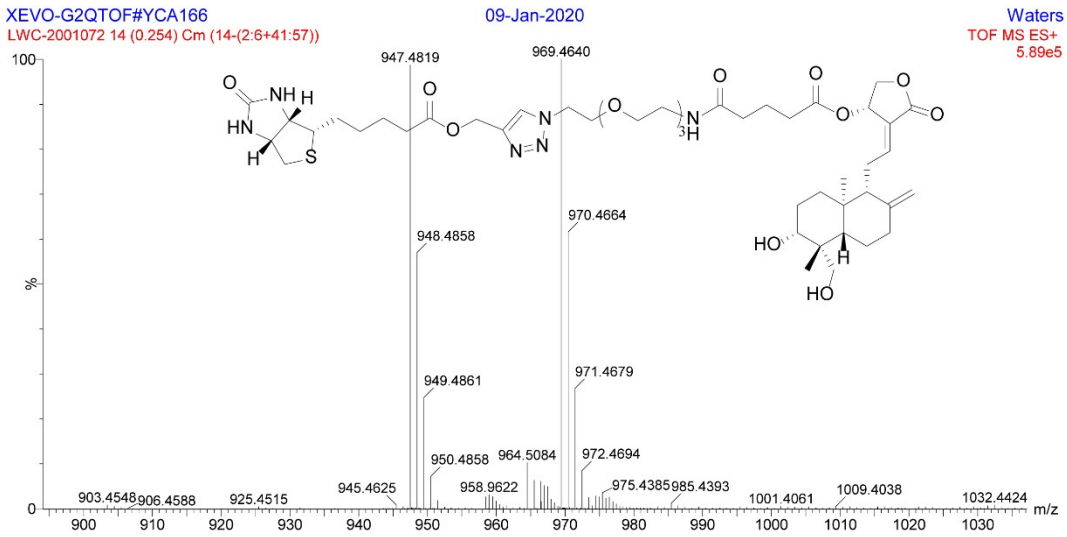

lwc-2001072/2
f1 (ppm)

3.5 HRMS spectra








XEVO-G2QTOF#YCA166
LWC-2001072 14 (0.254) Cm (14-(2:6+41:57))
09-Jan-2020
Waters
TOF MS ES+
5.89e5
100
%
0
947.4819
969.4640
970.4664
948.4858
971.4679
949.4861
950.4858
964.5084
972.4694
903.4548
906.4588
925.4515
945.4625
958.9622
975.4385
985.4393
1001.4061
1009.4038
1032.4424
900
910
920
930
940
950
960
970
980
990
1000
1010
1020
1030
m/z

3. References

1. D. Shivakumar, E. Harder, W. Damm, R. A. Friesner and W. Sherman, *J Chem Theory Comput*, 2012, **8**, 2553-2558.

2. M. Mendez, H. Matter, E. Defossa, M. Kurz, S. Lebreton, Z. Li, M. Lohmann, M. Lohn, H. Mors, M. Podeschwa, N. Rackelmann, J. Riedel, P. Safar, D. S. Thorpe, M. Schafer, D. Weitz and K. Breitschopf, *J Med Chem*, 2020, **63**, 2292-2307.

3. E. Ghidini, G. Marchini, A. M. Capelli, C. Carnini, V. Cenacchi, A. Fioni, F. Facchinetti and F. Rancati, *J Med Chem*, 2018, **61**, 4757-4773.