

Supporting Information for

Original article

Discovery of an orally active VHL-recruiting PROTAC that achieves robust HMGCR degradation and potent hypolipidemic activity *in vivo*

Guoshun Luo^a, Zhenbang Li^a, Xin Lin^a, Xinyu Li^b, Yu Chen^c, Kun Xi^b, Maoxu Xiao^a, Hanlin Wei^a, Lizhe Zhu^{b,*}, Hua Xiang^{a,*}

^a*State Key Laboratory of Natural Medicines and Department of Medicinal Chemistry, China Pharmaceutical University, Nanjing 21198, China*

^b*School of Life and Health Sciences and Warshel Institute for Computational Biology, the Chinese University of Hong Kong, Shenzhen 518172, China*

^c*Key Laboratory of Smart Drug Delivery, Ministry of Education School of Pharmacy, Fudan University, Shanghai 201203, China*

*Corresponding authors. Tel./fax: +86 25 83271096 (Hua Xiang); +86 755 23519577 (Lizhe Zhu).

E-mail addresses: xianghua@cpu.edu.cn (Hua Xiang); zhulizhe@cuhk.edu.cn (Lizhe Zhu).

Received 29 August 2020; received in revised form 22 October 2020; accepted 30 October 2020

Table of contents

Figure S1 Dose-HMGCR inhibition curves of PROTACs and lovastatin.....	3
Figure S2 Cytotoxicity against HepG2 cells.	4
Figure S3 Chemical stability of 16b , 21b and lovastatin in saline (pH = 7.4).	5
Figure S4 Original immunoblots in HepG2 cells.	6
Figure S5 Effects of lovastatin, 21b and 21c on cellular HMGCR and cholesterol.....	7
Figure S6 Time-course study of 21c on HMGCR in Insig-silenced HepG2 cells.....	7
Procedures for <i>in silico</i> modeling of PROTAC (21c)-mediated ternary complex	8
References	9
Figure S7 Overview of the modeling approach.....	10
Table S1 Physicochemical and drug-likeness properties of 21b and 21c	11
Table S2 <i>In vitro</i> metabolic stability of 21b , 21c and lovastatin in mouse liver microsomes.	12
Figure S8 Curves for lovastatin, 21b and 21c in mouse microsomes.....	12
Figure S9 Time course of plasma concentrations for 21b and metabolite 21c	13
Table S3 Pharmacokinetic properties of 21b in mice (60 mg/kg, <i>p.o.</i>).....	13
Figure S10 Effects of 21b in MFD-induced hypercholesterolemia mice.....	14
Scheme S1 Synthesis of intermediate 9 and 15a–15c	15
Characterization. NMR, HRMS and HPLC spectra of target compounds	19

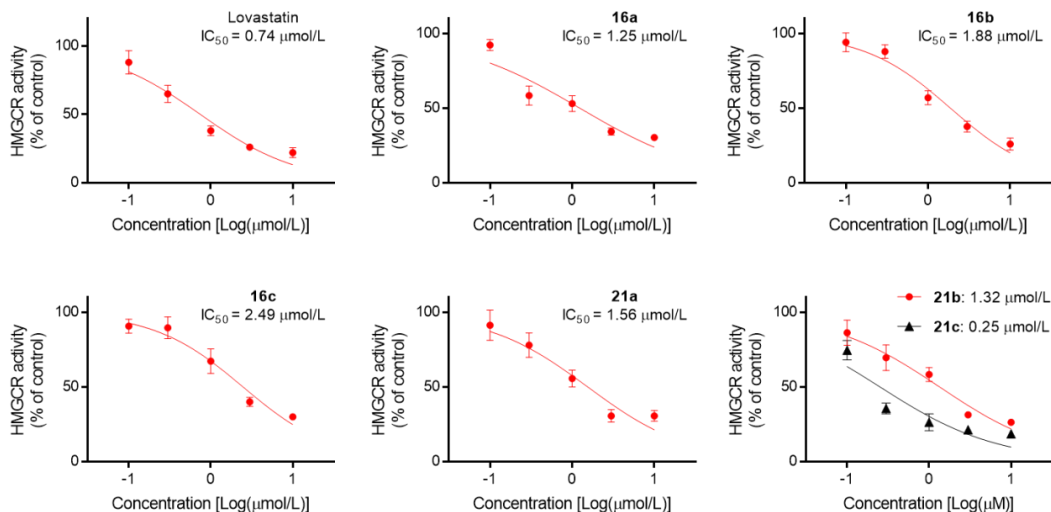


Figure S1 Dose-HMGCR inhibition curves of PROTACs and lovastatin.

HMGCR inhibition by lovastatin, **16a–16b**, **21a–21c** at range concentrations (0.1, 0.3, 1, 3 and 10 $\mu\text{mol/L}$) were determined by enzymatic assays (Biovision #K588-100). IC_{50} values, the concentration leading to 50% inhibition relative to DMSO (100%) were calculated by fitting the data to a dose-inhibition curve using GraphPad Prism. The data are represented as mean \pm SD from triplicate experiments.

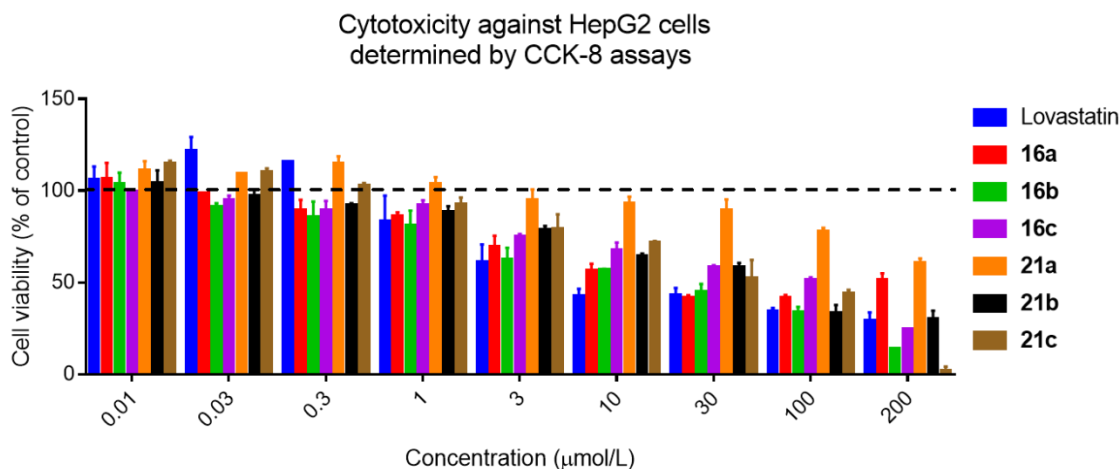


Figure S2 Cytotoxicity against HepG2 cells.

HepG2 cells were seeded in 96-well plates at a density of 5000/100 μL /well overnight, then were treated with lovastatin or test compounds (0.01, 0.03, 0.3, 1, 3, 10, 30, 100 and 200 $\mu\text{mol/L}$), respectively. After incubation for 48 h, the growth of the cells was evaluated by Cell Counting Kit-8 (CCK-8, KGA317, KeyGEN Biotech) and incubated for 2 h. Absorbance at 450 nm was evaluated in a microporous plate detection system. All tested compounds didn't show significant

cytotoxicity on HepG2 cells at concentrations up to 1 $\mu\text{mol/L}$. The cytotoxicity of compound lovastatin, **16a–16c**, and **21a–21c** on HepG2 cells over co-incubation for 48 h were determined by CCK-8 assay. Percentage cell viability relative to DMSO (100%) are represented as mean \pm SD from triplicate experiments.

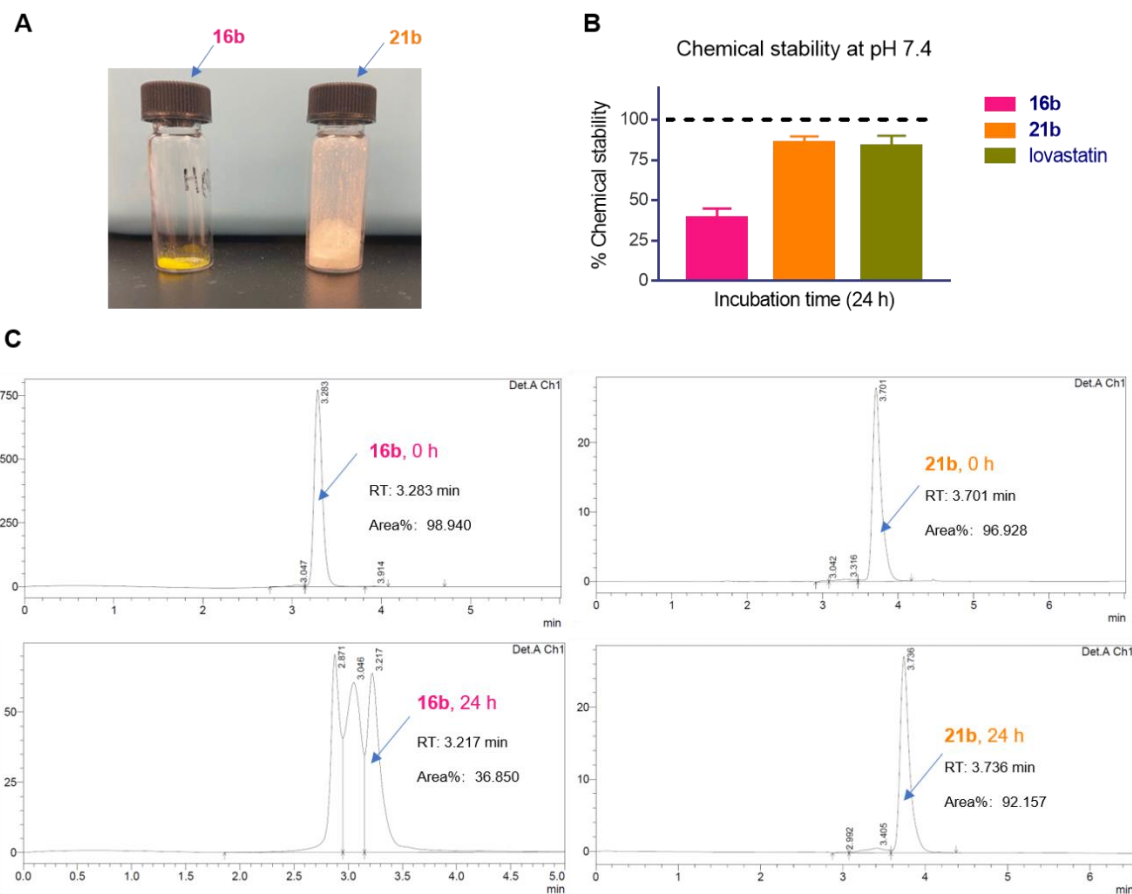


Figure S3 Chemical stability of **16b**, **21b** and lovastatin in saline (pH = 7.4).

(A) Photograph of compounds **16b** (yellow solid) and **21b** (white solid). (B) Drug stability data at pH 7.4. Acetonitrile solutions of **16b**, **21b** or lovastatin (10 mmol/L) were mixed with phosphate buffers (pH 7.4) and incubated for 24 h at 37 °C. Subsequently, aliquots were analyzed by HPLC and normalized to standard acetonitrile solutions. Percentage drug stability relative to standard solution (incubation time = 0 h) are represented as mean \pm SD from triplicate experiments. (C) Representative HPLC results. HPLC conditions are as follows: Discovery[®] 504971 column (C18, 250 mm \times 4.6 mm, 5 μm); temperature, 25 °C; injection volume, 5 μL ; isocratic flow, rate, 1 mL/min; solvent, 90% MeCN in H₂O.

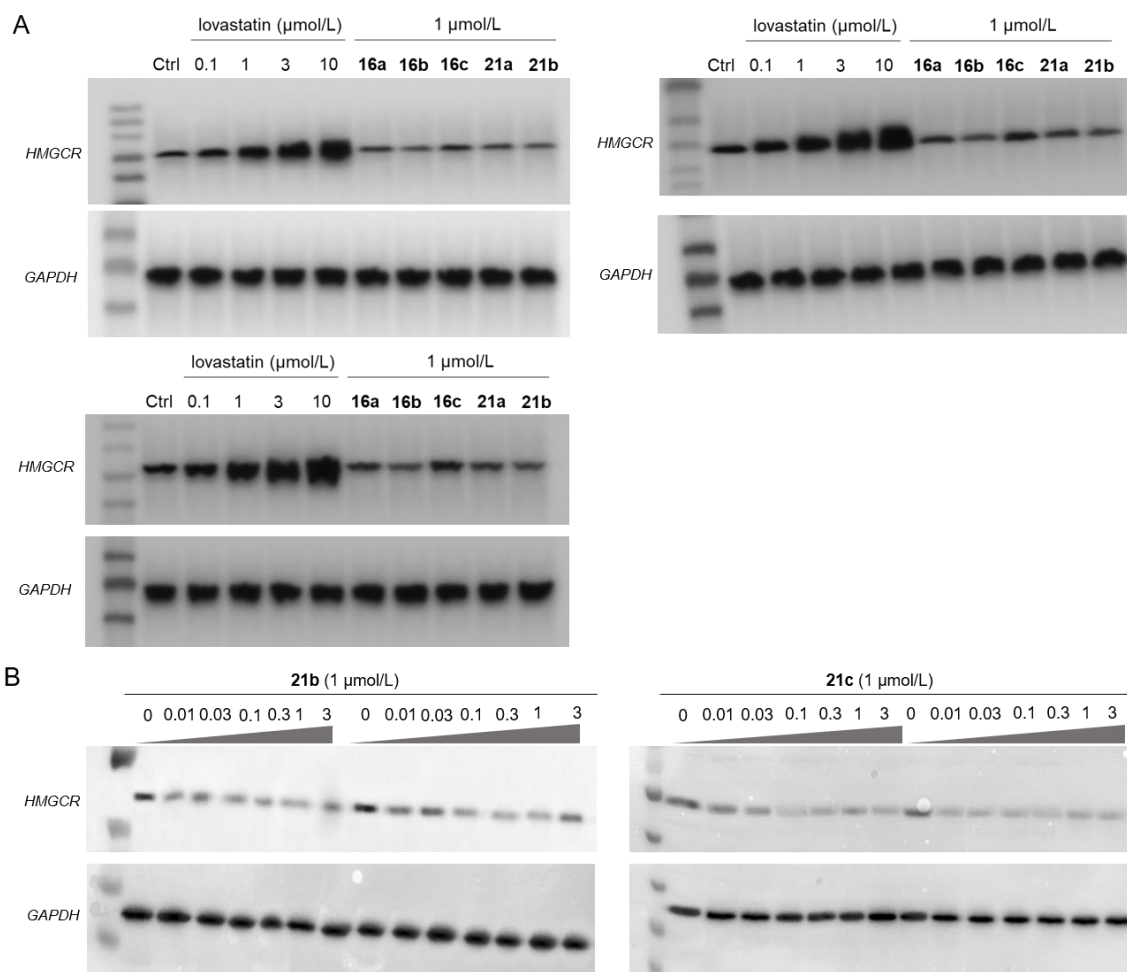


Figure S4 Original immunoblots in HepG2 cells.

Effect of lovastatin and PROTACs on HMGCR expression in HepG2 cells. (A) Cells were treated with DMSO, lovastatin (0.1, 1, 3 and 10 $\mu\text{mol/L}$) or **16a–16c**, **21a–21b** (1 $\mu\text{mol/L}$) for 16 h and harvested for immunoblotting with GAPDH as the internal reference. Results are from 3 independent experiments. (B) Cells were treated with DMSO, **21b** or **21c** (0.01, 0.03, 0.1, 0.3, 1 and 3 $\mu\text{mol/L}$) for 16 h and harvested for immunoblotting with GAPDH as the internal reference. Results are from two repeated experiments.

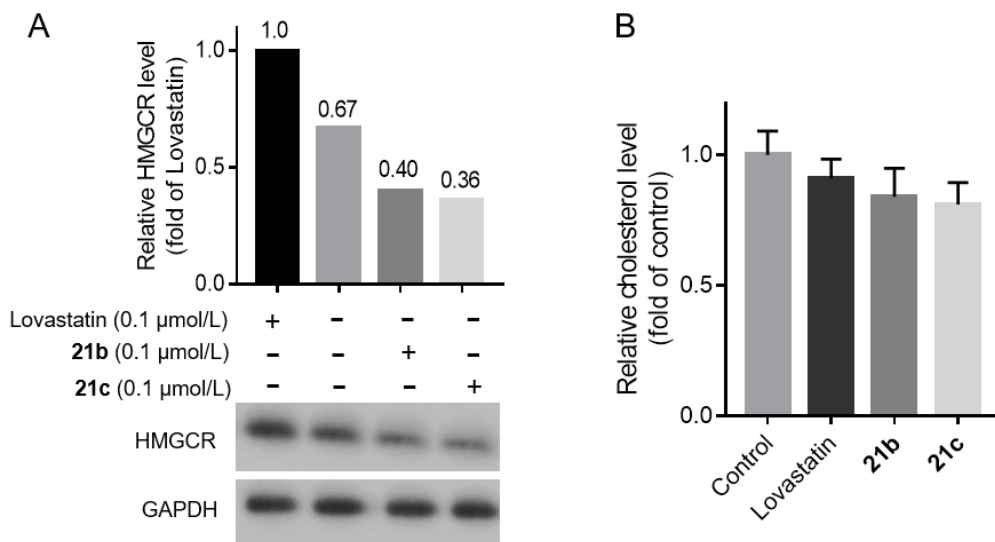


Figure S5 Effects of lovastatin, **21b** and **21c** on cellular HMGCR and cholesterol.

(A) Comparison of the effects of lovastatin, **21b** and **21c** on HMGCR expression in HepG2 cells after 16 h treatment. The data are presented as %HMGCR remaining relative to the lovastatin-control. (B) HepG2 cells were incubated with lovastatin (0.1 $\mu\text{mol/L}$), **21b** (0.1 $\mu\text{mol/L}$) and **21c** (0.1 $\mu\text{mol/L}$) for 24 h, and intracellular cholesterol levels relative to control were determined and presented as mean \pm SD from three independent experiments.

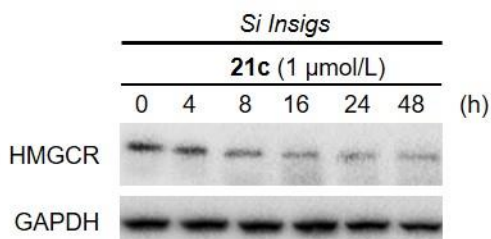


Figure S6 Time-course study of **21c** on HMGCR in Insig-silenced HepG2 cells.

Compound **21c** induces HMGCR degradation in a time-dependent manner. Insig-silenced HepG2 cells treated with **21c** at 1 $\mu\text{mol/L}$, were examined at indicated times.

Procedures for *in silico* modeling of PROTAC (21c)-mediated ternary complex

Protein–protein docking were carried out using the Rosetta software suite, which is free for academic use (www.rosettacommons.org). RDKit, an open-source cheminformatics software, version 2030.03.1, was employed to generate conformer. Computational tools for building PROTAC ternary models are implemented in python. Below we summarize our computational pipeline for building models of PROTAC ternary complexes.

Step 1: preparing the starting structure

To model the ternary structure for HMGCR, VHL and PROTAC **21c**. two complex structures (HMGCR–simvastatin acid: 1HW9 and VHL–VH032: 4W9H) were initially downloaded from the RCSB PDB (<http://www.rcsb.org/>), which were then manually combined by PyMol to make them oriented towards one another (starting structure, Fig. S7A). And then Rosetta were employed to generate the ligand params and to prepack the starting structure for docking.

```
$ROSETTA3/scripts/python/public/molfile_to_params.py -n NM1 -p NM1 NM1.mol2
$ROSETTA3/scripts/python/public/molfile_to_params.py -n NM2 -p NM2 NM2.mol2
$ROSETTA3/bin/docking_prepack_protocol.linuxgccrelease -s 1hw9vhlComplex.pdb -use_input_sc -extra_res_fa
NM1.params NM2.params
```

Step 2: Rosetta protein–protein docking

Here, we used the docking_protocol.mpi.linuxgccrelease program to output 10,000 diverse binding modes, the flag “-partners ABX_YC” flag specifies that the small-molecule ligands must move together with their paired proteins. The “-dock_pert” flag was employed to define the search space. Following that, InterfaceAnalyzer.mpi.linuxgccrelease was carried out to evaluate the results. The docking decoy will be selected (Fig. S7B) when the value of the packstate ≥ 0.5 and $dG_{\text{separated}}/dSASA \times 100 \leq -1.0$ (packstate means the Rosetta’s packing statistic score for the interface (0 = bad, 1 = perfect), $dG_{\text{separated}}/dSASA \times 100$ means that separated binding energy per unit interface area $\times 100$ to make units fit in score file)

```
$ROSETTA3/bin/docking_protocol.mpi.linuxgccrelease -s 1h9wvhlcomplexPrepack.pdb
-nstruct 100000 -use_input_sc -dock_pert 5 20 -partners ABX_YC -ex1 -ex2aro -extra_res_fa NM1.params NM2.params
-out:file:scorefile score.sc -score:docking_interface_score -ignore_unrecognized_res -load_PDB_components false -
out:path:all output_files
$ROSETTA3/bin/InterfaceAnalyzer.mpi.linuxgccrelease -in:file:l output_files/pdblist -fixedchains A B X -
load_PDB_components false -extra_res_fa NM1.params NM2.params -use_input_sc -compute_packstat true
-tracer_data_print false -out:file:score_only pack_input_score.sc -pack_input true -pack_separated true -
add_regular_scores_to_scorefile true -atomic_burial_cutoff 0.01 -sasa_calculator_probe_radius 1.4 -
pose_metrics::interface_cutoff 8.0
```

Step 3: generating linker conformations

Independently from docking step, RDKit was used to generate 10,000 conformations with threshold value larger than 1.5 and 4000 conformers were selected according to the energy ranging from low to high (Fig. S7C). The code used in RDKi is shown in attached file: Linker conformation generator.py.

Step 4: building ternary models

We used a custom python script¹ to evaluate the RMSD of the stubs in given linker conformer relative to their location in a docked model. If the RMSD of the stubs passes a specified cutoff value, the script then builds and outputs the complete model of the ternary complex (Fig. S7D). Then we chose a best result according to the score of docking result and the energy of ligand conformer for molecular dynamics simulation (Fig. S7E and F).

Step 5: molecular dynamics simulation

We employed molecular docking, molecular dynamics simulation to find the stable ternary structure. Before the MD simulation, we employed TIP3P water model to build the water box and added some chlorine or sodium ions to neutralize the system. AMBER18 package² was employed to run the MD simulations with leaprc.protein.ff14SB³ as the force field for the protein, and General Amber Force field (GAFF)⁴ as the force field for PROTAC **21c**. We first performed energy minimization to obtain a low energy starting conformation for the subsequent MD simulations. Four thousand steps of steepest descent method were initially employed followed by six thousand steps of conjugate gradient method. The whole system (protein, ligand, water, ions) was minimized, followed by minimization on the solutes (protein and ligand) only. With the Langevin thermostat applied, the system was heated under canonical ensemble from 0 to 303 K for 300 ps, with the force constant for the harmonic restraint set to be 10.0 kcal mol⁻¹Å⁻². The system was then equilibrated for 10 ns under NPT conditions (constant pressure = 1.0 bar). The relaxation time for barostat bath was 2.0 ps. Finally, the production simulation was run for 100 ns under NPT with periodic boundary conditions. The time step was set to be 2 fs and bonds connected with hydrogen atoms were constrained using SHAKE algorithm. The long-range electrostatics was handled by the particle-mesh Ewald (PME) method.⁵ The cut-off value for short range interactions was set to be 8.0 Å.

The “ptraj” tool in AMBER 18 was applied to calculate the pairwise RMSD (root mean square deviation) of the atoms to analyze the stability of the complexes during the simulation and for ensuring the basic analysis of trajectories of the sampling method. The k-means method installed “cpptraj” tool was used to cluster the trajectory. Here, we conducted a 500 ns simulation to get the stable ternary structure. Apparently, the conformation ensemble can be separated into three parts. The initial steady state lasted about 100 frames before the atoms stably oscillating around their new positions. And then the new conformation stabilized from 100 to 800 frames (Fig. S7G). Finally,

the whole structure had no significant changes during the rest of the simulation time (Fig. S7G). Based on this we clustered the trajectory into three groups (Fig. S7H), select the most stable mode (named HMGCR–21c–VHL_3) for detail analysis.

References

1. Bai N, Kirubakaran P, Karanicolas J. Rationalizing PROTAC-mediated ternary complex formation using Rosetta. *bioRxiv* 2020. Available from: <https://doi.org/10.1101/2020.05.27.119347>.
2. Salomon-Ferrer R, Case DA, Walker RC. An overview of the Amber biomolecular simulation package. *Wires Comput Mol Sci* 2013;**3**:198–210
3. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J Chem Theory Comput* 2015;**11**:3696–713.
4. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. *J Comput Chem* 2004;**25**:1157.
5. Darden T, York D, Pedersen L. Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *J Chem Phys* 1993;**98**:10089–92.

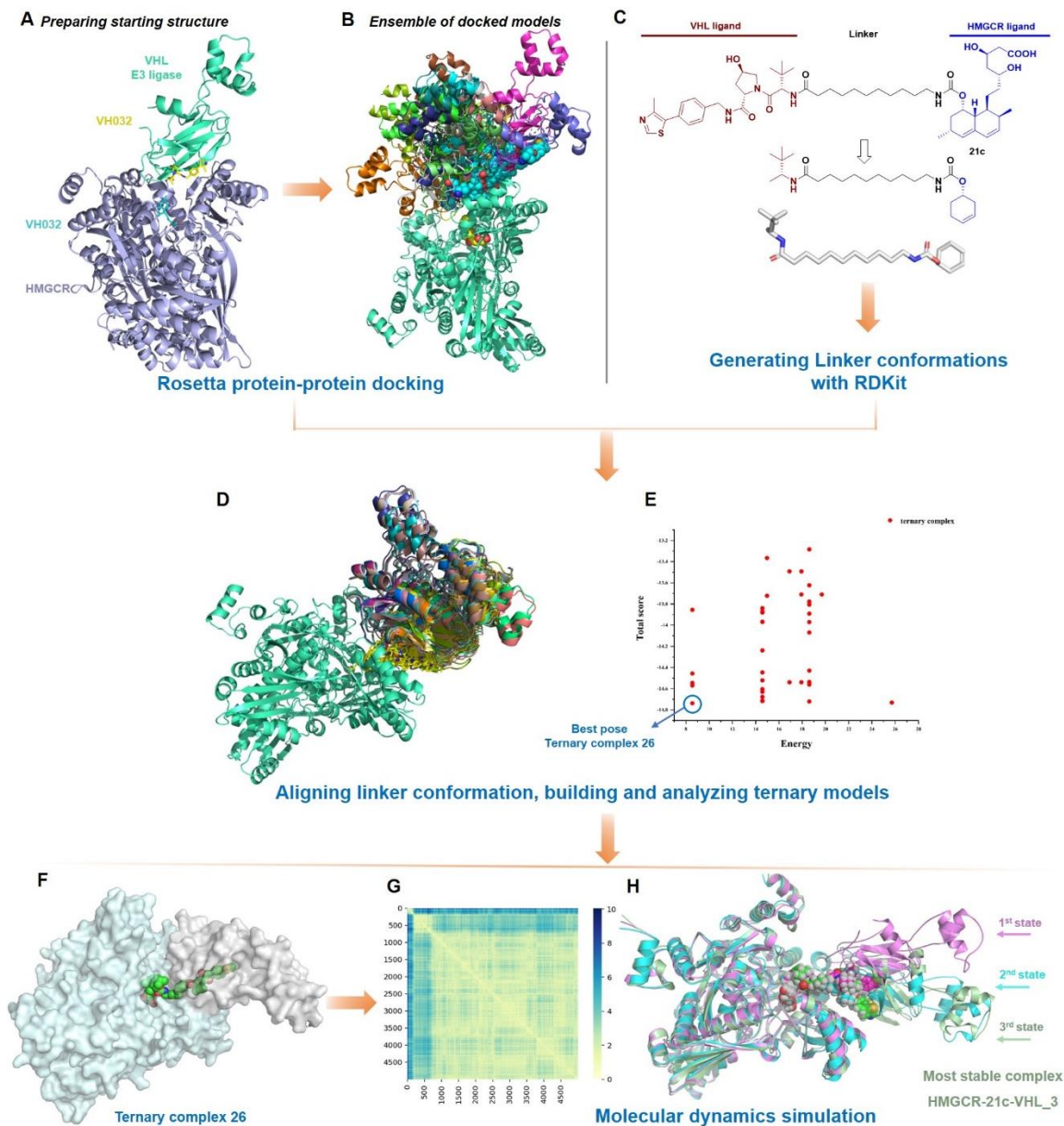


Figure S7 Overview of the modeling approach.

(A) The target protein HMGCRC bound to simvastatin acid and the E3 ligase VHL bound to VH032 are initially restricted to orient towards one another. (B) Parts of the docking results through Rosetta. (C) Based on **21c**, candidate linker conformations are generated using RDKit. (D) Complete ternary complex models. (E) Analysis of docking score and ligand conformer energy. (F) Structure of the pose with the lowest score and energy. (G) Molecular dynamics simulation. (H) Clustering structures of three dynamic states.

Table S1 Physicochemical and drug-likeness properties of **21b** and **21c**.

Property	21b	21c
MW ^a	959.54 (>500)	977.55 (>500)
HBA ^b	11 (>10)	12 (>10)
HBD ^c	5	7
tPSA (Å ²) ^d	159.12	184.47
MlogP ^e	6.88 (>5)	6.56 (>5)
MlogS (mg/L) ^f	1.75	1.92
Blood–brain barrier penetration (BBB) ^g	0.38	0.14
<i>In vivo</i> Caco-2 cell permeability (nm/s) ^h	27.98	18.22
Human intestinal absorption (HIA, %) ⁱ	90.96	85.98
Plasma protein binding (%) ^j	91.56	91.06

^aMW: molecular weight. ^bHBA: hydrogen-bond acceptors. ^cHBD: hydrogen-bond donors. ^dtPSA: topological polar surface area. ^eMlogP: calculated logarithm of the octanol-water partition coefficient. ^fMlogS: calculated water solubility. ^{a–f}Properties were predicted by Molsoft (<http://molsoft.com/mprop/>). ^{g–j}Properties were predicted by PreADMET (<https://preadmet.bmdrc.kr/>).

Table S2 *In vitro* metabolic stability of **21b**, **21c** and lovastatin in mouse liver microsomes.

Test article	Species		Percent remaining (%)					$t_{1/2}$ (min)	CL_{int} (mL/min/kg)
			0 min	5 min	15 min	30 min	45 min		
21b	Mouse	Mean	100.00	41.22	8.23	4.34	1.42	4.18	415.81
		RSD	0.05	0.09	0.14	0.80	1.41		
21c	Mouse	Mean	100.00	108.69	80.72	86.51	70.88	87.50	62.37
		RSD	0.02	0.01	0.07	0.01	0.11		
Lovastatin	Mouse	Mean	100.00	0.85	BQL	BQL	BQL	0.73	7516.09
		RSD	0.02	0.27	N/A	N/A	N/A		

All the data are mean of duplicate.

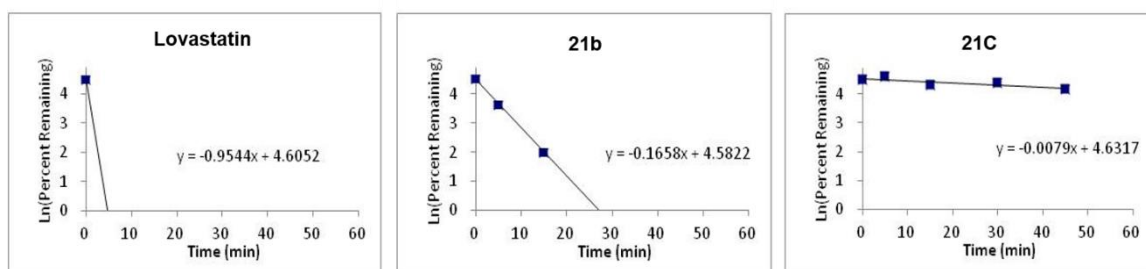


Figure S8 Curves for lovastatin, **21b** and **21c** in mouse microsomes.

Time course of remaining concentrations for lovastatin, **21b** and **21c** in mouse liver microsomes after incubation with NADPH at 37 °C for indicated times.

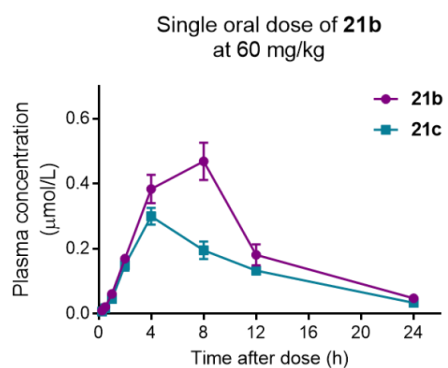


Figure S9 Time course of plasma concentrations for **21b** and metabolite **21c**.

Time course of plasma concentrations for **21b** and metabolite **21c** over 24 h after a single oral administration of **21b** at 60 mg/kg in C57BL/6 mice.

Table S3 Pharmacokinetic properties of **21b** in mice (60 mg/kg, *p.o.*)^a.

PK parameter	Unit	Mean ± SD	
		21b ^b	21c ^c
Kel	h ⁻¹	0.13 ± 0.01	0.11 ± 0.01
<i>t</i> _{1/2}	h	5.1 ± 0.4	6.2 ± 0.6
<i>T</i> _{max}	h	8.0	4.0
<i>C</i> _{max}	µmol/L	0.47 ± 0.04	0.29 ± 0.03
AUC _{0–24}	µmol·h/L	5.0 ± 0.41	3.3 ± 0.16
AUC _{0–inf}	µmol·h/L	5.4 ± 0.47	3.6 ± 0.18
MRT _{0–24}	h	8.8 ± 0.1	8.6 ± 0.2
MRT _{0–inf}	h	10.2 ± 0.4	10.7 ± 0.8

^aProdrug **21b** was dosed via single oral route at indicated concentrations. ^bPK parameters of parent **21b**. ^cPK parameters of metabolite **21c**. All data are expressed as the mean ± SD, *n*=5.

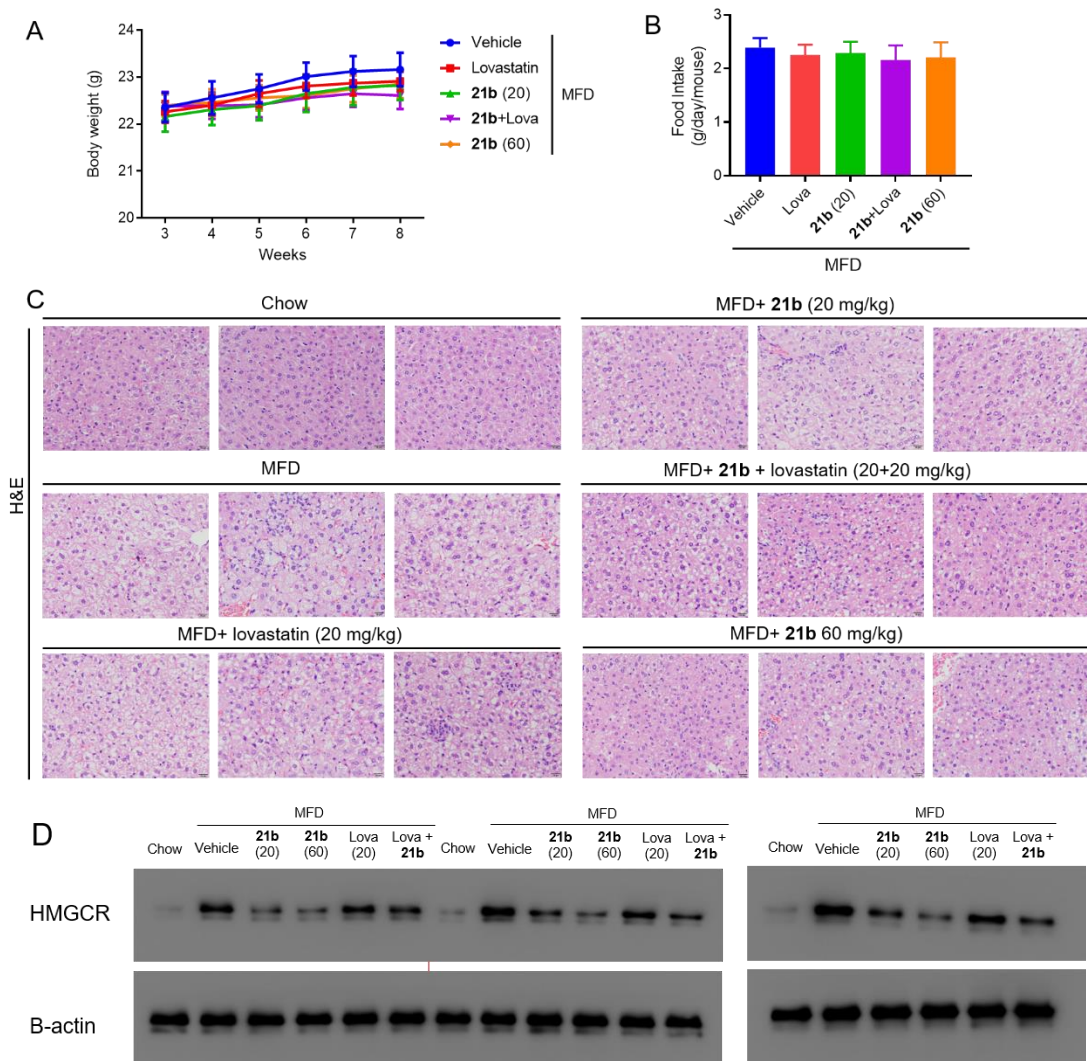
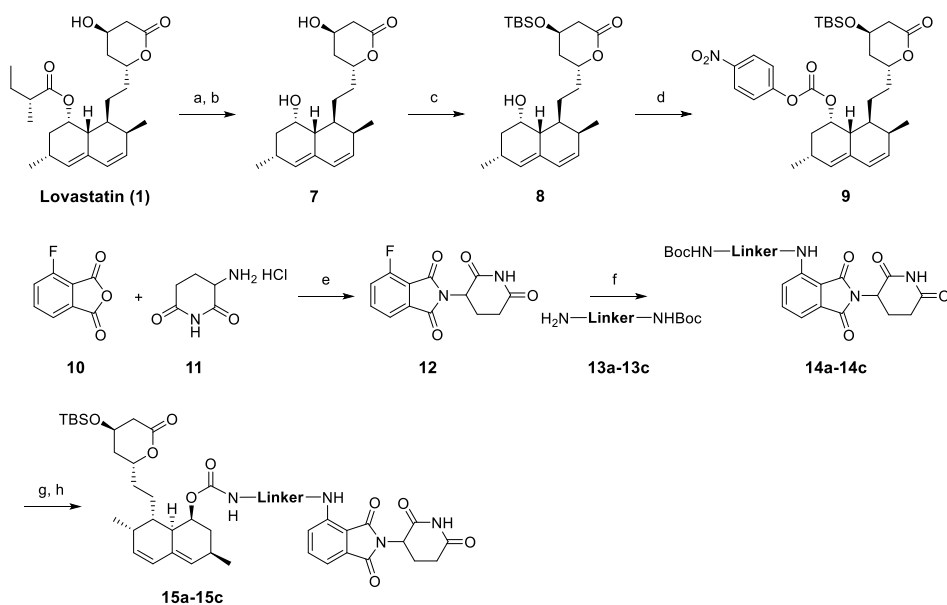


Figure S10 Effects of **21b** in MFD-induced hypercholesterolemia mice.

C57BL/6J mice ($n = 6$ per group) were randomly grouped and allowed *ad libitum* access to water and MFD diet for 8 weeks. Mice were orally administrated with vehicle, **21b** (20 mg/kg/day), lovastatin (20 mg/kg/day), or both (20+20 mg/kg/day) or **21b** (60 mg/kg) from 3th to 8th week. Body weight (A) and food intake (B) were examined during the 5-week treatment. Livers of randomly selected mice ($n=3$ per group) were examined by 3 separate H&E staining experiments (scale bar = 20 μ m) (C) and 3 separate Western blotting experiments (D), respectively.



Scheme S1 Synthesis of intermediate **9** and **15a–15c**.

Reagents and conditions: (a) KOH, H₂O/MeOH, reflux, 12 h; (b) 6 mol/L HCl, rt, 6 h, 45% for two steps; (c) TBSCl, imidazole, CH₂Cl₂, rt, 6 h, 84%; (d) *p*-nitrophenyl chloroformate, DMAP, pyridine, rt, 16 h, 64%; (e) NaOAc, AcOH, 12 h, reflux, 70%; (f) *N,N*-diisopropylethylamine, DMF, 90 °C, 12 h, 35%–45%; (g) TFA, DCM, rt, 0.5 h; (h) **9**, DMAP, pyridine, rt, 16 h, 60%–75%.

(4R,6R)-4-((tert-Butyldimethylsilyl)oxy)-6-(2-((1S,2S,6R,8S,8aR)-8-hydroxy-2,6-dimethyl-1,2,6,7,8,8a-hexahydronaphthalen-1-yl)ethyl)tetrahydro-2H-pyran-2-one (8)

To a solution of H₂O/MeOH (1:5, 132 mL) were added lovastatin **1** (18 g, 44.6 mmol) and KOH (25.2 g, 449 mmol). The reaction was stirred under reflux for 12 h. Then, the reaction mixture was evaporated under the reduced pressure to remove MeOH. To the residue were added H₂O (500 mL), CH₂Cl₂ (100 mL), and 6 mol/L HCl solution to adjust pH = 2. The mixture was stirred at room temperature for 6 h and then neutralized with saturated NaHCO₃ solution. After extraction with CH₂Cl₂, the solvent was removed to provide the crude product **7** orange oil (6.4 g, 45% yield). The crude product dissolved in anhydrous CH₂Cl₂ (85 mL) was treated with TBSCl (4.2 g, 30 mmol) and imidazole (3.4 g, 49.6 mmol). The mixture was stirred at rt for 6 h, and then extracted with CH₂Cl₂. The combined organic layer was washed with brine, concentrated under vacuum, and then purified by silica gel column chromatography using ethyl acetate/petroleum ether to afford compound **8** as a white solid (7.3 g, 84% yield). ¹H NMR (CDCl₃, 400 MHz) δ 5.99 (dd, *J* = 19.0, 9.6 Hz, 1H), 5.8–5.80 (m, 1H), 5.55 (d, *J* = 25 Hz, 1H), 4.76–4.55 (m, 1H), 4.35–4.25 (m, 1H), 3.9 (s, 1H), 2.50–2.33 (m, 2H), 2.30–2.25 (m, 2H), 2.15–2.14 (m, 1H), 2.02–1.55 (m, 14H), 1.16 (d, *J* = 7.6 Hz, 3H), 0.92–0.88 (m, 12H), 0.08 (t, *J* = 2.4 Hz, 6H).

(1*S*,3*R*,7*S*,8*S*,8*aR*)-8-(2-((2*R*,4*R*)-4-((*tert*-Butyldimethylsilyl)oxy)-6-oxotetrahydro-2*H*-pyran-2-yl)ethyl)-3,7-dimethyl-1,2,3,7,8,8*a*-hexahydronaphthalen-1-yl (4-nitrophenyl) carbonate (**9**)

To a three-necked bottle were added compound **8** (7.3 g, 16.8 mmol), *p*-nitrophenyl chloroformate (34 g, 168 mmol) and DMAP (17 g, 84 mmol) in anhydrous pyridine (100 mL), and the resulting solution was stirred at rt for 16 h. After completion, pyridine was removed under vacuum, and the residue was extracted with CH₂Cl₂ and washed with 1 mol/L HCl solution, saturated NaHCO₃ solution and brine, and then purified by silica gel column chromatography using ethyl acetate/petroleum ether to afford compound **9** as a white solid (6.4 g, 64% yield). ¹H NMR (400 MHz, CDCl₃) δ 8.33–8.20 (m, 2H), 7.46–7.36 (m, 2H), 6.00 (dd, *J* = 19.0, 9.7 Hz, 1H), 5.80 (ddd, *J* = 19.8, 9.5, 6.1 Hz, 1H), 5.55 (d, *J* = 24.7 Hz, 1H), 4.77–4.53 (m, 1H), 4.39–4.24 (m, 1H), 3.78 (s, 1H), 2.67–2.49 (m, 4H), 2.49–2.33 (m, 2H), 2.25 (ddd, *J* = 18.8, 13.6, 3.0 Hz, 2H), 2.10 (dt, *J* = 15.5, 7.8 Hz, 1H), 2.05–1.64 (m, 9H), 1.91–1.33 (m, 10H), 1.33–1.22 (m, 1H), 1.18 (d, *J* = 7.4 Hz, 2H), 1.08 (t, *J* = 9.7 Hz, 2H), 0.92 (dd, *J* = 10.2, 4.8 Hz, 4H), 0.90 (s, 6H), 0.87 (s, 9H), 0.09 (t, *J* = 2.4 Hz, 6H). ¹³C NMR (101 MHz, CDCl₃) δ 170.4, 170.3, 155.7, 152.3, 145.3, 133.2, 133.0, 131.2, 129.4, 129.3, 128.3, 128.2, 125.3, 122.0, 75.9, 75.4, 74.5, 72.4, 63.6, 54.6, 39.3, 37.5, 37.4, 36.8, 36.6, 36.2, 36.1, 32.7, 32.5, 32.3, 32.2, 30.8, 27.4, 27.5, 25.7, 25.7, 23.8, 23.5, 22.5, 22.5, 17.9, 13.9, -4.9, -4.9. MS (ESI) *m/z*: 600.1 [M+H]⁺.

2-(2,6-Dioxopiperidin-3-yl)-4-fluoroisoindoline-1,3-dione (**12**)

A solution of 4-fluoroisobenzofuran-1,3-dione **10** (1 g, 6 mmol), 2,6-dioxopiperidine-3-ammonium chloride **11** (1.1 g, 6.6 mmol) and potassium acetate (1.8 g, 18 mmol) in AcOH (20 mL) was reflux for 12 h. The mixture was concentrated under reduced pressure to remove AcOH, and the residue was purified by column chromatography using CH₂Cl₂/MeOH to give intermediate **12** as a white solid (1.2 g, 70% yield). ¹H-NMR (400 MHz, DMSO-*d*₆) δ: 11.15 (s, 1H), 7.97–7.90 (m, 1H), 7.80–7.67 (m, 2H), 5.15 (dd, *J* = 13.0, 5.4 Hz, 1H), 2.94–2.80 (m, 1H), 2.65–2.46 (m, 2H), 2.10–2.04 (m, 1H). MS (ESI) *m/z*: 227.1 [M+H]⁺.

tert-Butyl (3-(2-(2-(3-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisoindolin-4-yl)amino)propoxy)ethoxy)ethoxy)propyl)carbamate (**14a**)

To a round-bottom flask was added amine **13a** (0.32 g, 1.0 mmol), anhydrous DMA (3 mL) and DIPEA (0.39 g, 3.0 mmol) were added. The reaction mixture was stirred at 90 °C for 10 min under an inert atmosphere of nitrogen, then **12** (0.28 g, 1.0 mmol) in anhydrous DMA (0.5 mL) was added, and the mixture stirred at 90 °C for 12 h. After completion, the reaction mixture was diluted with

water and extracted with ethyl acetate. The combined organic layer was washed with brine, dried over Na₂SO₄, concentrated under vacuum, and then purified by silica gel column chromatography using ethyl CH₂Cl₂/MeOH to afford compound **14a** as a yellow oil (0.2 g, 35% yield). ¹H NMR (400 MHz, DMSO-*d*₆) δ 11.10 (s, 1H), 7.57 (t, *J* = 7.1 Hz, 1H), 7.09 (d, *J* = 8.6 Hz, 1H), 7.01 (d, *J* = 6.8 Hz, 1H), 6.72 (t, *J* = 5.1 Hz, 1H), 6.65 (t, *J* = 5.8 Hz, 1H), 5.04 (dd, *J* = 12.9, 5.1 Hz, 1H), 3.57–3.39 (m, 14H), 3.00–2.80 (m, 3H), 2.65–2.54 (m, 2H), 2.08–1.98 (m, 1H), 1.85–1.75 (m, 2H), 1.65–1.50 (m, 2H), 1.35 (s, 9H). MS (ESI) *m/z*: 577.2 [M+H]⁺. Compounds **14b** and **14c** were synthesized according to the procedure for **14a**.

tert-Butyl (2-(2-(2-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)amino)ethoxy)ethoxy)ethyl)carbamate (**14b**)

Yellow oil (0.22 g, 45% yield). ¹H NMR (400 MHz, DMSO-*d*₆) δ 11.05 (s, 1H), 7.57 (t, *J* = 7.1 Hz, 1H), 7.1 (d, *J* = 8.6 Hz, 1H), 7.02 (d, *J* = 6.9 Hz, 1H), 6.72–6.64 (m, 1H), 6.58 (t, *J* = 5.8 Hz, 1H), 5.05 (dd, *J* = 12.8, 5.3 Hz, 1H), 3.60–3.36 (m, 10H), 3.07–2.81 (m, 3H), 2.63–2.52 (m, 2H), 2.06–1.98 (m, 1H), 1.35 (s, 9H). MS (ESI) *m/z*: 505.1 [M+H]⁺.

tert-Butyl (3-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)amino) propyl) carbamate (**14c**)

Yellow solid (0.16 g, 40% yield). ¹H NMR (400 MHz, DMSO-*d*₆) δ 11.06 (s, 1H), 7.61 (t, *J* = 7.9 Hz, 1H), 7.23 (dd, *J* = 14.6, 7.8 Hz, 2H), 6.70 (s, 1H), 5.06 (d, *J* = 8.7 Hz, 1H), 3.58–3.36 (m, 2H), 3.14 (d, *J* = 7.1 Hz, 2H), 3.02 (s, 3H), 2.84–2.80 (m, 1H), 2.58–2.50 (m, 2H), 2.00–1.92 (m, 1H), 1.33 (s, 9H). MS (ESI) *m/z*: 431.1 [M+H]⁺.

(1*S*,3*R*,7*S*,8*S*,8*aR*)-8-(2-((2*R*,4*R*)-4-((*tert*-Butyldimethylsilyl)oxy)-6-oxotetrahydro-2*H*-pyran-2-yl)ethyl)-3,7-dimethyl-1,2,3,7,8,8*a*-hexahydronaphthalen-1-yl (3-(2-(2-(3-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)amino)propoxy)ethoxy)ethoxy)propyl)carbamate (**15a**)

To a solution of intermediate carbamate **14a** (0.35 g, 0.6 mmol) in DCM (4 mL) was added TFA (2 mL), and the mixture was stirred at rt for 30 min. Then, the reaction mixture was evaporated under reduced pressure. The crude product was dissolved in anhydrous pyridine (3 mL), to which was added carbonate **9** (0.18 g, 0.3 mmol), and DMAP (0.14 g, 1.2 mmol). The reaction mixture was stirred at rt for 16 h. After completion, pyridine was removed under vacuum, and the residue was extracted with EA and washed with 1 mol/L HCl solution, saturated NaHCO₃ solution and brine, and then purified by silica gel column chromatography to afford compound **15a** as a green solid (0.17 g, 60% yield). ¹H NMR (400 MHz, CDCl₃) δ 7.58–7.45 (m, 1H), 7.37 (d, *J* = 2.5 Hz,

1H), 7.14–7.04 (m, 2H), 6.95 (d, $J = 8.6$ Hz, 1H), 6.47 (s, 1H), 5.83–5.73 (m, 1H), 5.52 (s, 1H), 5.21 (s, 1H), 4.97–4.87 (m, 1H), 4.65 (s, 1H), 4.30 (d, $J = 3.6$ Hz, 1H), 3.72–3.48 (m, 9H), 3.43 (dd, $J = 12.5, 6.5$ Hz, 2H), 3.37–3.14 (m, 2H), 2.94–2.68 (m, 3H), 2.68–2.51 (m, 2H), 2.40 (d, $J = 26.7$ Hz, 2H), 2.25 (dd, $J = 13.5, 6.1$ Hz, 1H), 2.11 (ddd, $J = 20.0, 10.9, 6.1$ Hz, 2H), 1.95 (dt, $J = 12.5, 6.4$ Hz, 2H), 1.88 (s, 1H), 1.76 (s, 3H), 1.66 (d, $J = 19.1$ Hz, 8H), 1.45 (s, 3H), 1.39 (s, 6H), 1.31 (d, $J = 2.9$ Hz, 9H), 0.98 (s, 9H), 0.86 (d, $J = 6.9$ Hz, 9H), 0.12 (s, 6H). MS (ESI) m/z : 937.1 [M+H]⁺. Compounds **15b** and **15c** were synthesized according to the procedure for **15a**.

(1S,3R,7S,8S,8aR)-8-(2-((2R,4R)-4-((tert-Butyldimethylsilyl)oxy)-6-oxotetrahydro-2H-pyran-2-yl)ethyl)-3,7-dimethyl-1,2,3,7,8,8a-hexahydronaphthalen-1-yl (2-(2-(2-((2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)amino)ethoxy)ethoxy)ethyl)carbamate (15b)

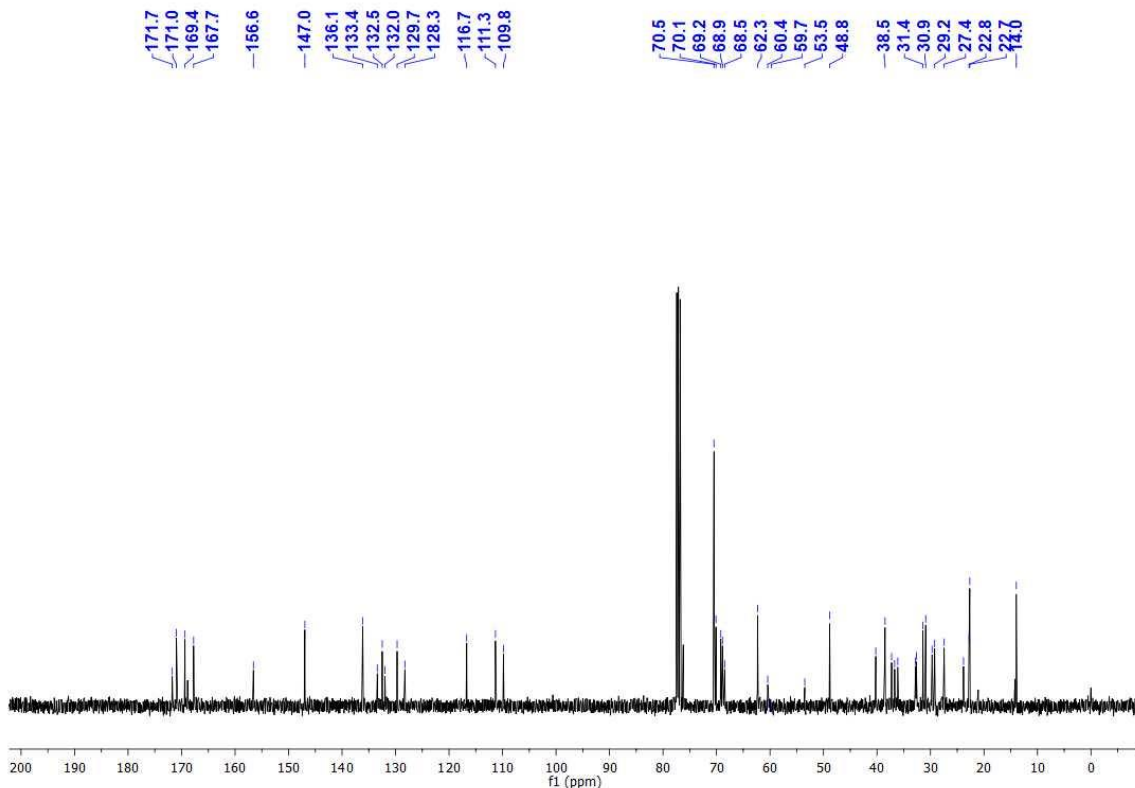
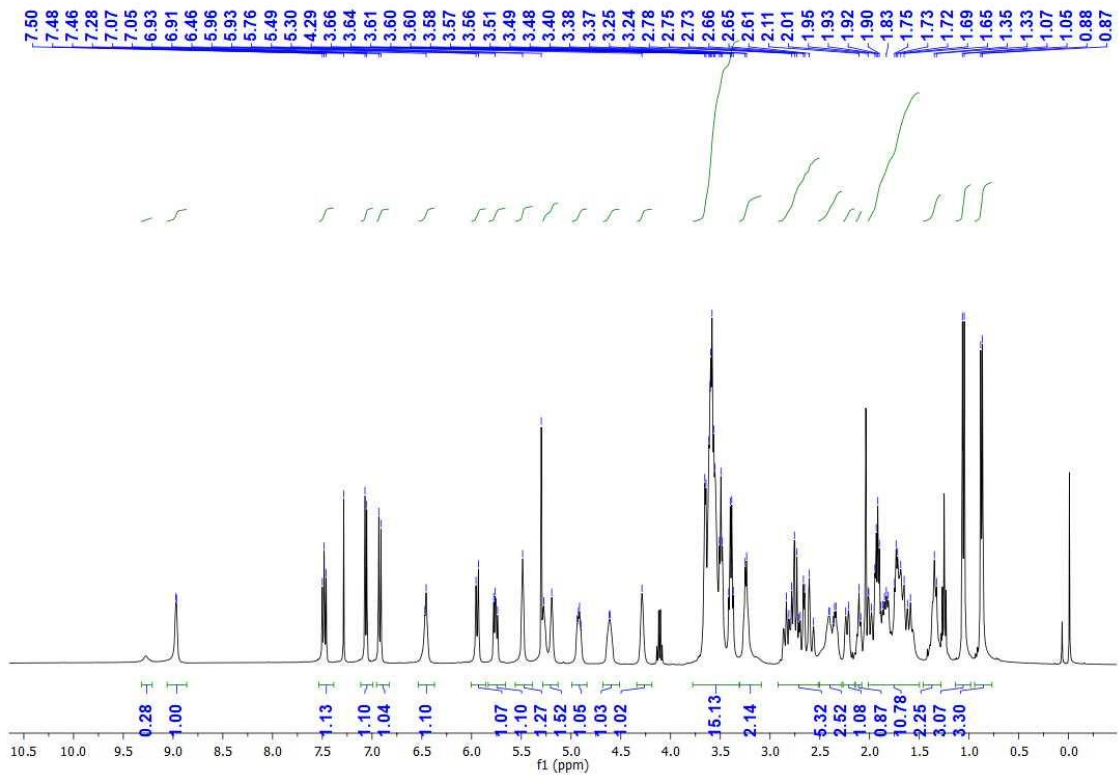
Yellow solid (0.16 g, 63% yield). ¹³C NMR (101 MHz, CDCl₃) δ 8.29 (s, 1H), 7.58–7.44 (m, 1H), 7.11 (t, $J = 5.6$ Hz, 1H), 6.94 (t, $J = 5.5$ Hz, 1H), 6.46 (s, 1H), 5.97 (d, $J = 9.7$ Hz, 1H), 5.78 (dd, $J = 9.5, 6.0$ Hz, 1H), 5.52 (s, 1H), 5.22 (d, $J = 7.0$ Hz, 2H), 4.93 (dd, $J = 12.1, 5.4$ Hz, 1H), 4.64 (s, 1H), 4.34–4.22 (m, 1H), 3.74–3.47 (m, 13H), 3.47–3.34 (m, 5H), 3.34–3.13 (m, 2H), 2.95–2.70 (m, 7H), 2.60 (dd, $J = 11.4, 7.0$ Hz, 2H), 2.40 (t, $J = 8.1$ Hz, 5H), 2.26 (d, $J = 11.6$ Hz, 1H), 2.19–2.10 (m, 2H), 2.05–1.82 (m, 8H), 1.81–1.57 (m, 15H), 1.44 (d, $J = 4.2$ Hz, 1H), 1.35 (s, 3H), 1.30 (s, 4H), 1.27 (t, $J = 3.5$ Hz, 13H), 1.09 (d, $J = 7.4$ Hz, 4H), 0.90 (s, 11H), 0.09 (d, $J = 1.8$ Hz, 6H). MS (ESI) m/z : 865.1 [M+H]⁺.

(1S,3R,7S,8S,8aR)-8-(2-((2R,4R)-4-((tert-Butyldimethylsilyl)oxy)-6-oxotetrahydro-2H-pyran-2-yl)ethyl)-3,7-dimethyl-1,2,3,7,8,8a-hexahydronaphthalen-1-yl (3-((2-(2,6-dioxopiperidin-3-yl)-1,3-dioxoisindolin-4-yl)amino)propyl)carbamate (15c)

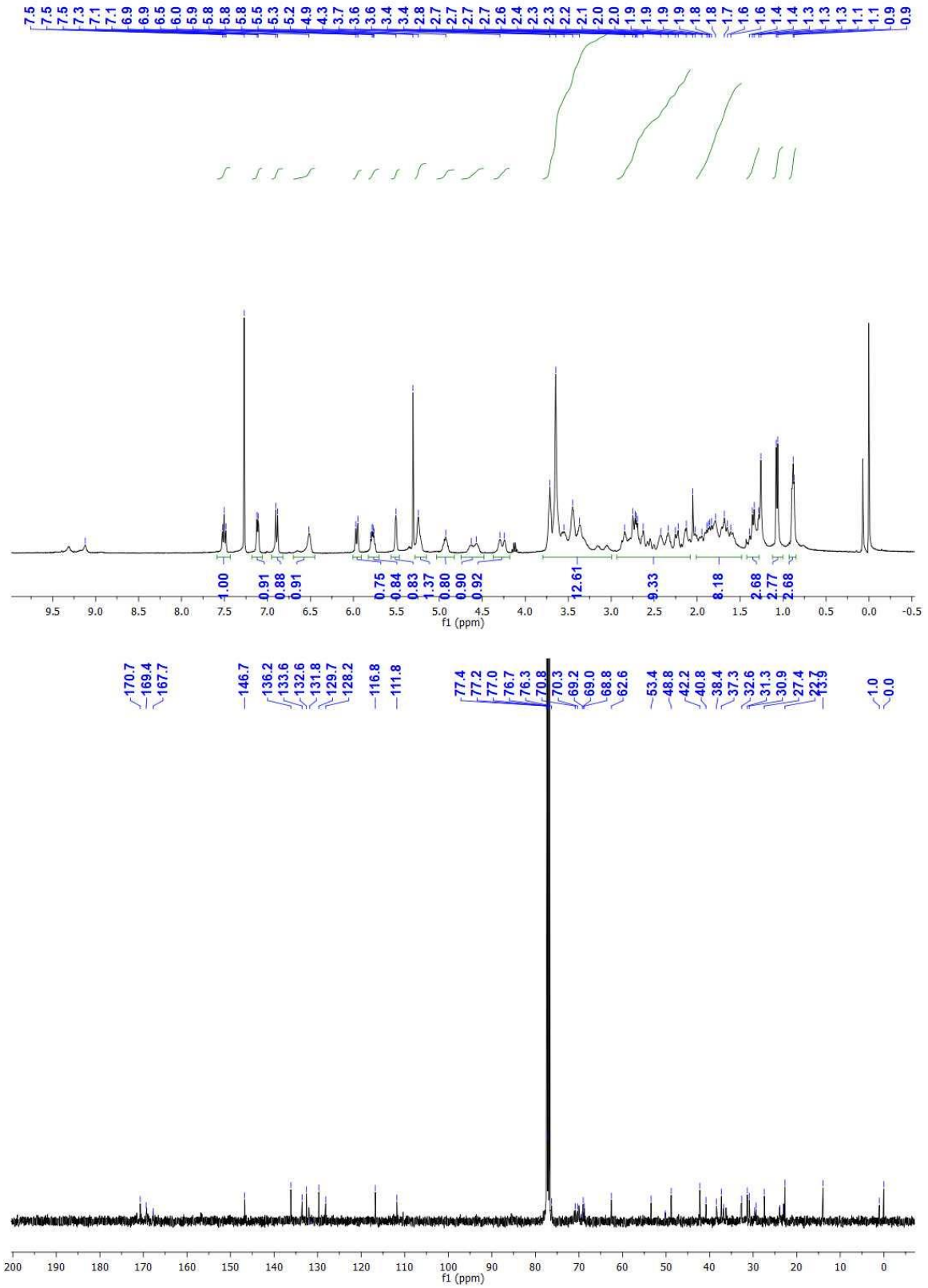
Yellow solid (0.18 g, 75% yield). ¹H NMR (400 MHz, CDCl₃) δ 9.18 (s, 1H), 7.45–7.40 (m, 1H), 7.06 (dd, $J = 7.0, 2.5$ Hz, 1H), 6.85 (d, $J = 8.6$ Hz, 1H), 6.44 (s, 1H), 5.97 (d, $J = 9.7$ Hz, 1H), 5.85–5.59 (m, 1H), 5.52 (s, 1H), 5.36–5.12 (m, 2H), 5.00–4.80 (m, 1H), 4.62 (s, 1H), 4.23 (s, 1H), 3.49 (s, 1H), 3.25 (s, 4H), 2.90–2.66 (m, 3H), 2.58 (dt, $J = 30.1, 11.4$ Hz, 2H), 2.46–2.15 (m, 4H), 2.08 (s, 1H), 1.81–1.72 (m, 8H), 1.44–1.25 (m, 11H), 1.06 (d, $J = 7.3$ Hz, 3H), 0.89 (d, $J = 7.0$ Hz, 3H), 0.09 (s, 6H). MS (ESI) m/z : 791.1 [M+H]⁺.

Characterization. NMR, HRMS and HPLC spectra of target compounds.

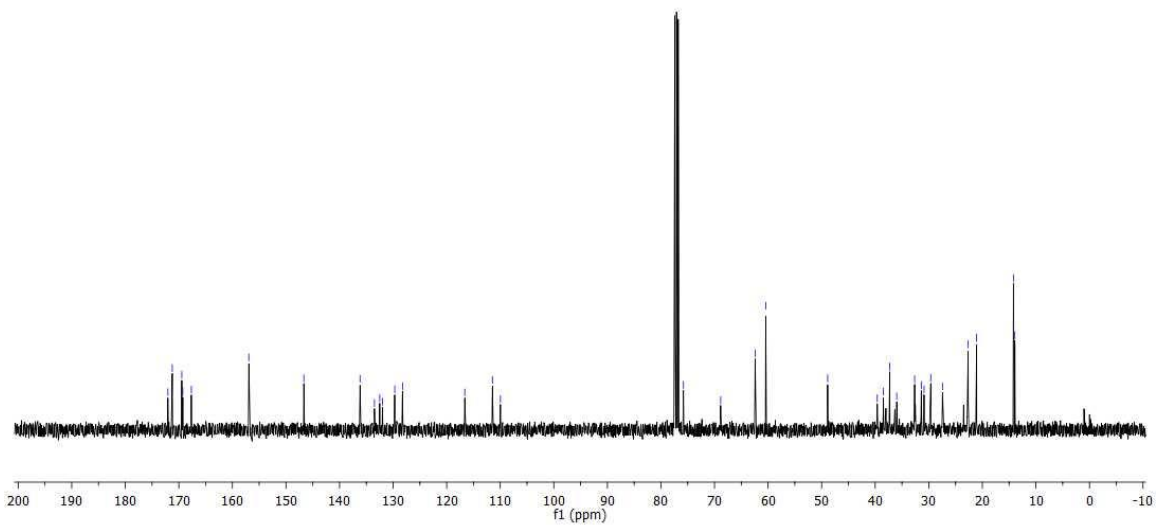
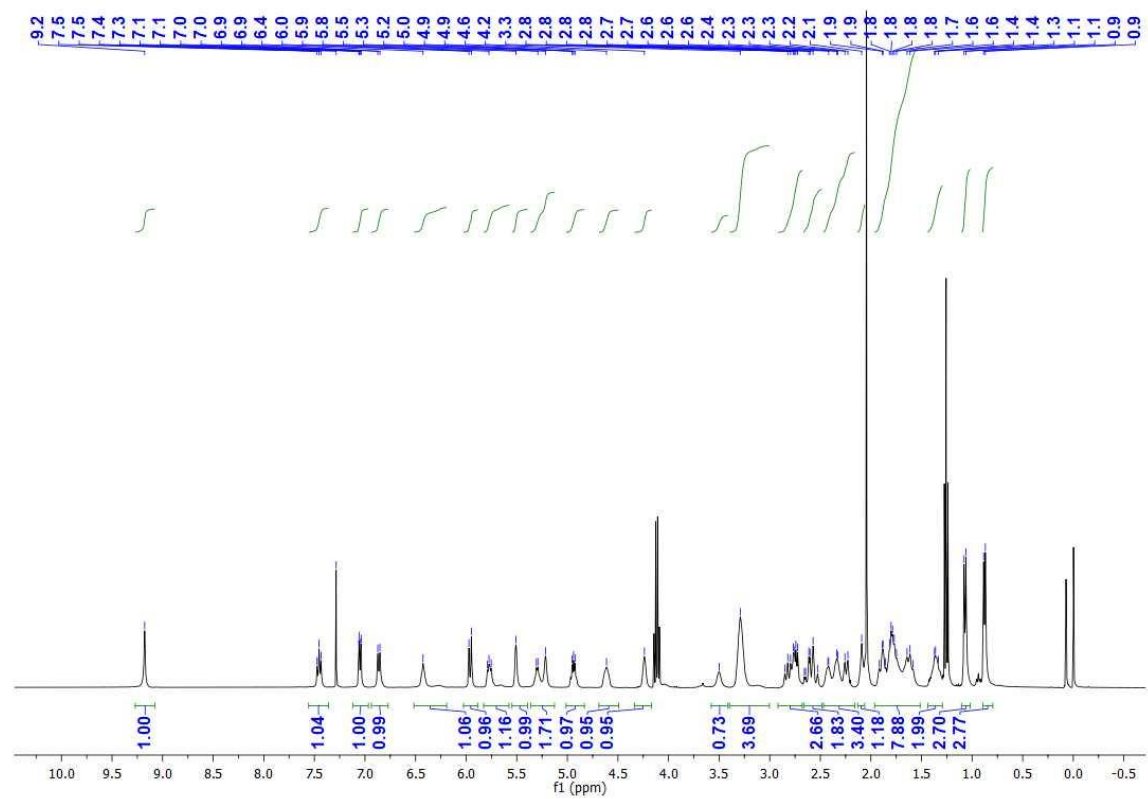
NMR spectrum of **16a**



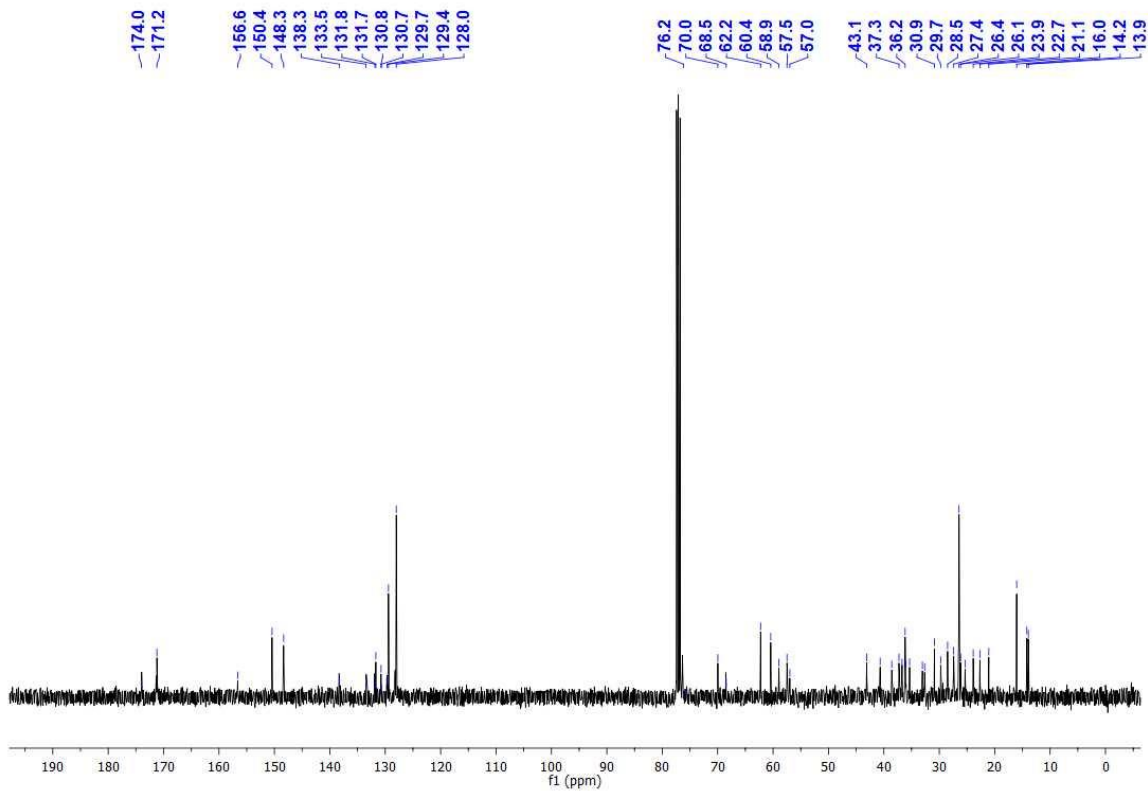
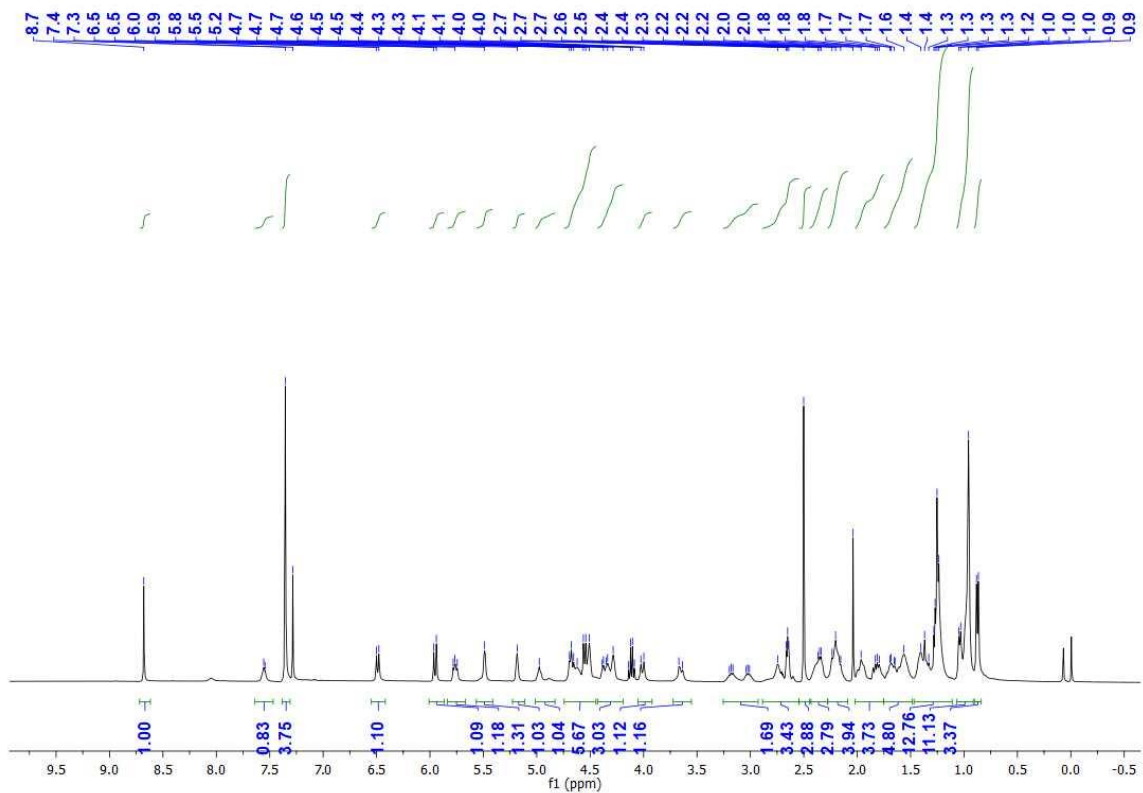
NMR spectrum of **16b**



NMR spectrum of **16c**

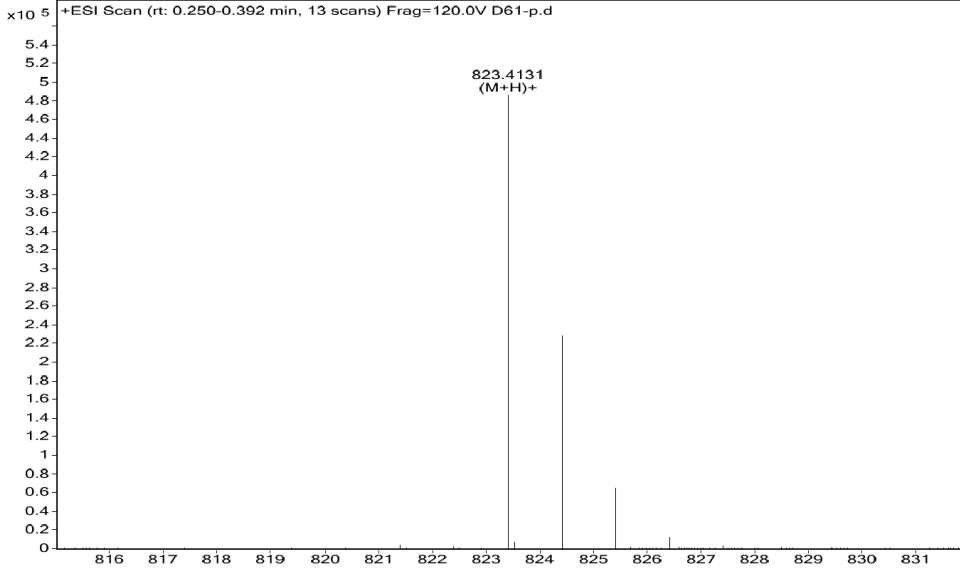


NMR spectrum of **21a**



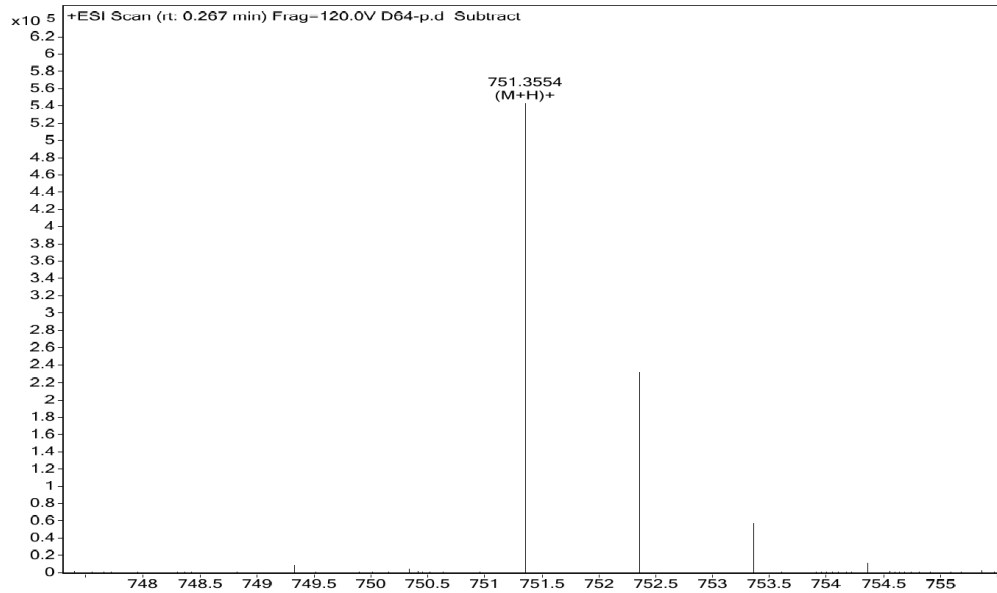
HRMS spectrum of 16a

Sample Name		Position	p2e8	Instrument Name	Instrument 1	User Name	G6520B-PC\Admin
Inj Vol	0.5	InjPosition		SampleType	Sample	IRM Calibration Status	Success
Data Filename	D61-p.d	ACQ Method	20110418-MSonly-p.m	Comment		Acquired Time	12/20/2019 12:11:12 PM



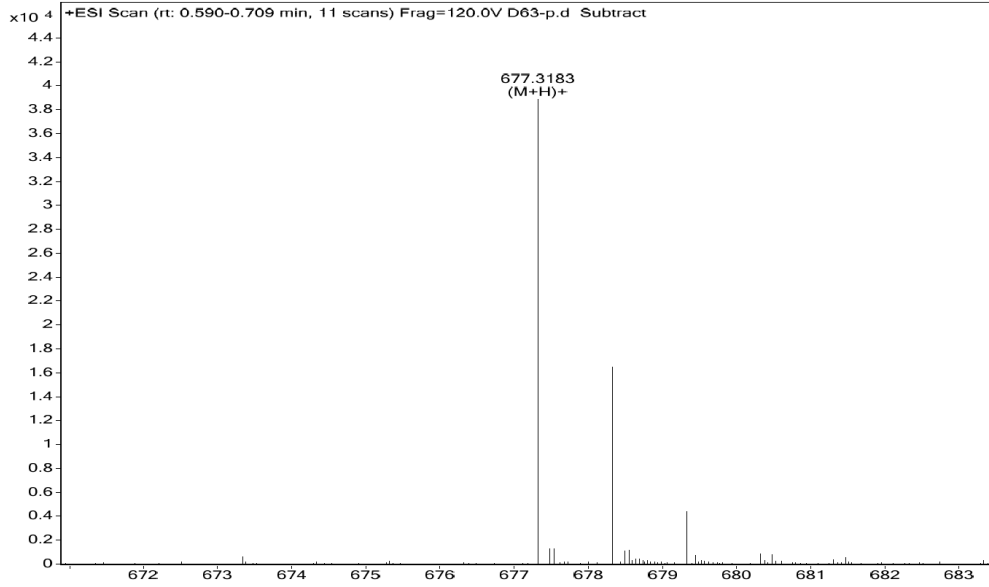
HRMS spectrum of 16b

Sample Name		Position	p2F2	Instrument Name	Instrument 1	User Name	G6520B-PC\Admin
Inj Vol	0.5	InjPosition		SampleType	Sample	IRM Calibration Status	Success
Data Filename	D64-p.d	ACQ Method	20110418-MSonly-p.m	Comment		Acquired Time	12/20/2019 12:25:11 PM



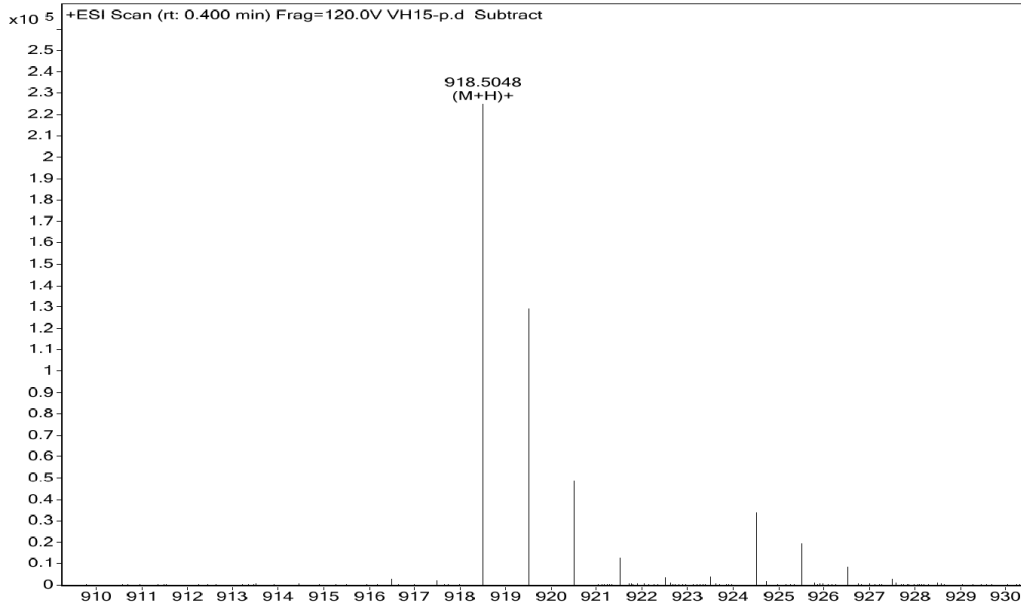
HRMS spectrum of 16c

Sample Name	Position	p2F1	Instrument Name	Instrument 1	User Name	G6520B-PC\Admin
Inj Vol	InjPosition	0.5	SampleType	Sample	IRM Calibration Status	Success
Data Filename	ACQ Method	D63-p.d	20110418-MSonly-p.m	Comment	Acquired Time	12/20/2019 12:22:10 PM



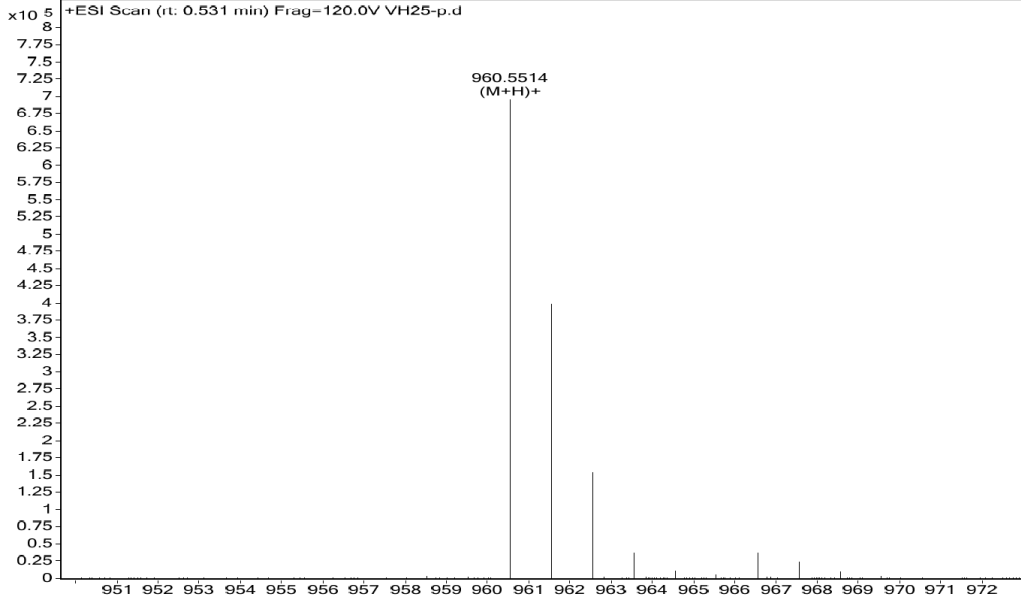
HRMS spectrum of 21a

Sample Name	Position	P2b3	Instrument Name	Instrument 1	User Name	G6520B-PC\Admin
Inj Vol	InjPosition	0.1	SampleType	Sample	IRM Calibration Status	Success
Data Filename	ACQ Method	VH15-p.d	20110418-MSonly-p.m	Comment	Acquired Time	4/29/2020 3:11:00 PM

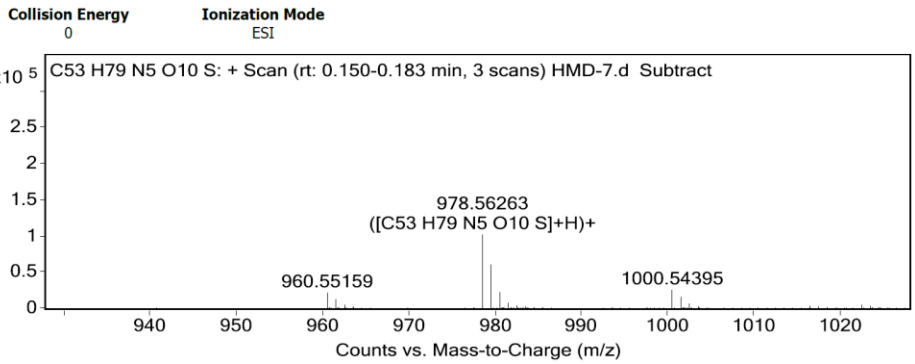


HRMS spectrum of 21b

Sample Name P2b4 **Instrument Name** Instrument 1 **User Name** G6520B-PC\Admin
Inj Vol 0.1 **InjPosition** Sample **IRM Calibration Status** Success
Data Filename VH25-p.d **ACQ Method** 20110418-HISonly-p.m **Comment** **Acquired Time** 4/29/2020 3:13:07 PM



HRMS spectrum of 21c

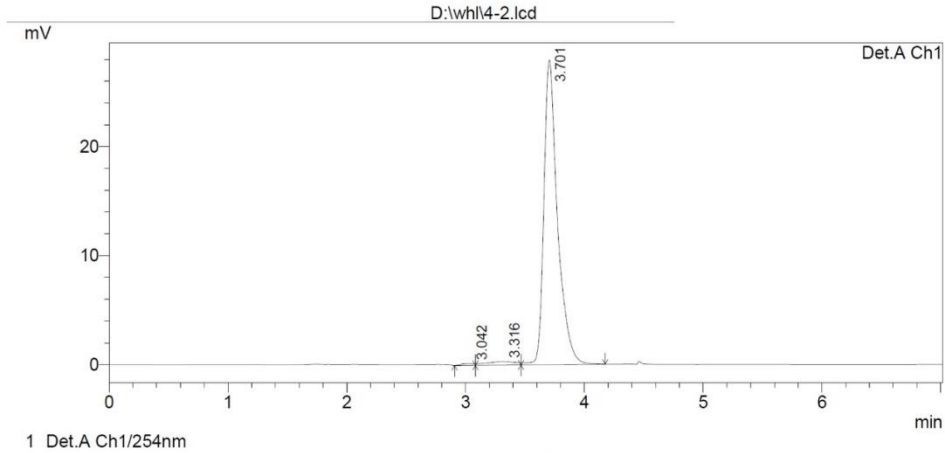


Formula Calculator Results

Formula	Best	Mass	Tgt Mass	Diff (ppm)	Ion Species	CalculatedMz
C53 H79 N5 O10 S	TRUE	977.55533	977.55476	-0.58	C53 H80 N5 O10 S	978.56204

HPLC spectrum of **21b**

==== Shimadzu LCsolution



Detector A Ch1 254nm

峰#	保留时间	面积	高度	Area %	拖尾因子
1	3.042	1332	161	0.579	0.000
2	3.316	5736	326	2.493	0.000
3	3.701	223035	27953	96.928	1.571
Total		230103	28440	100.000	