Electronic Supporting Information

Structure-based discovery of a specific TLR1-TLR2 small molecule agonist from ZINC drug library database

Zhipeng Chen, Xiaohong Cen, Junjie Yang, Xiaoshan Tang, Kai Cui and Kui Cheng*

Guangdong Provincial Key Laboratory of New Drug Screening and Guangzhou Key Laboratory of Drug Research for Emerging Virus Prevention and Treatment, School of Pharmaceutical Sciences, Southern Medical University, Guangzhou, 510515, China.

Phone: +86 20 6164 7192, Fax: +86 20 6164 8533

Email address: Chengk@smu.edu.cn

General Chemistry Methods.

All of the solvents and reagents used were obtained commercially and used as received. The solvents were removed under reduced pressure using standard rotary evaporators. Compounds were purified using flash chromatography (FC) (Silica gel 60 or alumina-B, 200-400 mesh, Sorbent Tech.) or recrystallization. Purity for all final compounds was confirmed to be greater than 97% purity by ^{1}H NMR, ^{1}H NMR spectra were recorded at 400 MHz in $CDCl_3$ using residual solvent peaks ($CDCl_3$, δ 7.28) as the internal standard. ^{13}C NMR spectra were recorded at 101 MHz in $CDCl_3$ using residual $CHCl_3$ (77.16 ppm) as the internal reference.

Virtual Screening.

The ZNIC drug database (10,519,614 small molecules) was docked into the TLR1 and TLR2 binding domain (PDB: 2Z7X) [1] using Glide 7.4. The molecules are created, as appropriate, with multiple protonation and tautomeric states. The TLR1/2 conformations were prepared using standard Glide protocols. This includes addition of hydrogens, restrained energy-minimizations of the protein structure with the Optimized Potentials for Liquid Simulations-All Atom (OPLS-AA) force field, and finally setting up the Glide grids using the Protein and Ligand Preparation Module. Base on the interaction between Pam3CSK4 with TLR1 and TLR2[1], (Fig. S1) the potential binding pocket is defined in the critical residues included pocket (Fig. S2).

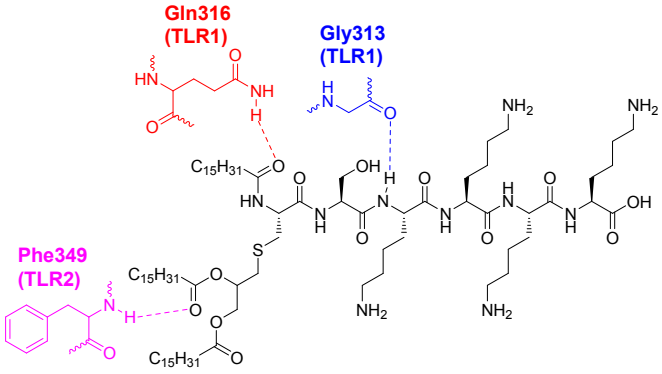


Fig. S1. The Chemical structure of Pam3CSK4, and the critical hydrogen binding interacting between Pam3CSK4 and TLR1/2.

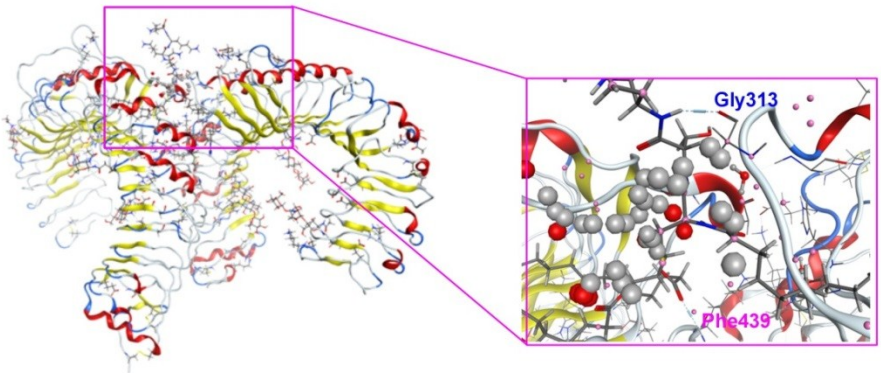


Fig. S2. Defined the potential binding pocket of small molecule to TLR1/2. The pocket including PHE312, GLY313, PHE314, PRO315, GLN316 in TLR1 and PHE322, TYR323, LEU324, PHE325, TYR326, ASP327, PHE349, LEU350, PRO352 in TLR2.

All 10.5 million compounds were first docked and ranked using High Throughput Virtual Screening (HTVS) Glide in above defined pocket, continued with standard precision (SP) Glide for the top 10000 compounds. The resultant top 5000 compounds were then docked using the more accurate and computationally intensive extra-precision (XP) mode. Initial top-ranked 100 compounds were selected out.

The selection of the candidate molecules was based on the following criteria: (1) Complementarity exists between the ligand and the active site of TLR1/2. (2) Reasonable chemical structure and pose are in the active site of TLR1/2. Some unusually highly scored molecules were found to have many rotatable bonds (such as long aliphatic structures), which were excluded for further evaluation. (3) There is formation of at least one hydrogen bond between the ligand and the important residues of TLR1/2 (such as GLY313, GLN316 from TLR1, or PHE349 from TLR2) etc). (4) Protonation state and the tautomeric form of the ligand have to be acceptable. As a result, 100 candidate compounds can meet the above criteria. In order to achieve good chemical diversity, the resulting 100 candidates were subsequently filtered by chemical diversity and binding energy. Consequently, 14 potential TLR1/2 agonists were designated and purchased for vitro assay.

QUANTI-Blue SEAP Assay.

Cells were cultured in 96-well plates (4×10^4 cells per well) of 200 μL DMEM (supplemented with 10% FBS, 10 × penicillin/streptomycin, and 10 × L-glutamine) at 37 °C for 24 hours before drug treatment on the first day. In the next 24 hours of treatment, medium was removed from the 96-well plate and substituted with 200 μL of DMEM containing indicated concentrations of compounds in the half maximal effective concentration (EC50) determinations; or 0.5 μM of SMU127, 0.77 nM (1 ng/ml) Pam_3CSK_4 according to the manufacturer's instruction (InvivoGen), as well as different antibodies (0 to 10 μg/mL), including anti-hTLR1–IgG, anti-hTLR2–IgA, or anti-hTLR6–IgA (InvivoGen) in the antibody experiments. A sample buffer (50 μL) was collected and transferred from each well of the cell culture supernatants to a transparent 96-well plate (Thermo Scientific). Each well was treated with 50 μL of QUANTI-Blue (InvivoGen) buffer and incubated at 37 °C for 0.5-1 hour. Then measure the purple color by using a plate reader at an absorbance of 620 nm (A620).

U937 Cell Transfection and NF-kB–GFP Reporter Assay.

Human macrophage U937 cells [American Type Culture Collection (ATCC) CRL-1593.2] were cultured in RPMI 1640 medium (10% FBS, penicillin (100 U/ml), and streptomycin (100 mg/ml). By using the commercially available pGreenFire plasmid (System Biosciences) was stably inserted an NF-κB–GFP reporter. Briefly, HEK 293T cells (ATCC CRL-3216) were transfected by using a polyethylenimine/DNA ratio of 6:1 accompanied with the pGreenFire vector (4.33 μg) and the pREV (4.33 μg), pMDL (4.33 μg) and pVSVg (2 μg) viral packaging plasmids. Then harvest viral particles from the medium after 48 to 72 hours transfection and concentrated by using an 8.5% PEG-8000 (polyethylene glycol, molecular weight 8000) and 10 mM NaCl solution. The concentrated virus and polybrene (8 μg/mL) were added to U937 cells for 48 hours. Then, select the stably transfected cells by using U937 growth medium supplemented with puromycin (1 mg/ml). After selection process finished, using a MoFlo Cytomation (Beckman Coulter) fluorescence-activated cell sorter to sort the cells for GFP expression. After sorting for insertion, cells were treated with a TLR1/2 agonist Pam_3CSK_4 [66 nM (100 ng/mL), InvivoGen] and sorted for activation. Then between the untreated and the treated cells the top 10% of activated cells were collected for each sort until no further peak separation was achieved. The sorted cells were seeded in six-well plates at density of 1×10^6 cells per well with 3 ml growth medium of RPMI 1640 medium [supplemented with 10% FBS, penicillin (100 U/ml), and streptomycin (100

mg/mL)] and the indicated concentrations of compound and Pam_3CSK_4 for 24 hours at 37 °C in a 5% CO_2 humidified incubator. 24 hours later before the flow cytometry analysis, the cells of each well were mixed and then use propidium iodide to stain 200 μL of cells containing medium for 10 minutes.

ELISA Assay

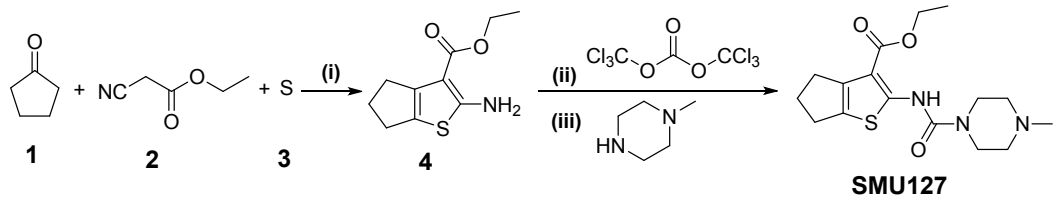
Peripheral blood mononuclear cells (PBMC) were seeded in 12-well plates at a density of 2.5 × 10^5 cells per well with 0.5 mL of medium [RPMI 1640 medium supplemented with 10% FBS, penicillin (100 U/mL) and streptomycin (100 mg/mL)]. The cells were treated with indicated concentration of SMU127 and incubated for 24 hours at 37 °C in a 5% CO_2 humidified incubator. The cell culture supernatants were collected and froze at -80 °C until measurement. The level of cytokine TNF-α were determined using recombinant cytokine standards, cytokine-specific capture antibodies and detection antibodies according to the commercially available ELISA kit (BD Biosciences) with each sample for duplicate.

Western blotting analysis.

In the first day, 3 mL of HEK-Blue hTLR2 cells were seeded in 6-well plate (Thermo Scientific) at density of 1.5 × 10^6 per well in 3 mL DMEM (supplemented with 10% FBS, 10 × penicillin/streptomycin, and 10 × L-glutamine) and incubated for 24 hours. In the second day, removed and replaced the medium with DMEM medium only to 3mL totally and treated the cells with different concentration of SMU127 and incubated for 24 hours. In the third day, the cells were harvested and lysed with 150 μL cell lysis buffers (PIPA mixed with PMSF at ratio of 50:1 before use, Boster). Cell lysates of equal amount were denatured, separated by sodium dodecyl sulfate polycrylamide gel electrophoresis (SDS-PAGE), and transferred to nitrocellulose membrane. Then use PBS (containing 5% nonfat milk) to block the membrane for 2 h , and then incubated with the primary antibody of TLR2 overnight at 4 °C, followed with incubating by a horseradish peroxidase conjugated secondary antibody. The immunoblots were visualized by enhanced exposure machine. During the operation, the following primary and secondary antibodies were employed: mouse-GAPDH (1:2000, Solarbio, M1000110), rabbit-TLR2 (1:2000, Cell Signaling, 12276), goat-anti-mouse-GAPDH-HRP (1:2500, Boster, BA1050) and goat-anti-rabbit-TLR2-HRP (1:1000, Solarbio, SE134).

***In Vivo* antitumor effects of SMU127.**

Female BABL/c mice at 6 weeks of age (17–19 g) were obtained from Southern Medical University. The animal experiments conformed to the Southern Medical University's Committee on the Care and Use of Laboratory Animals Guidelines, according to protocol No. SYXK (Yue) 2016-0167, approved by the Laboratory Animal Center of Guangdong Province. Before the beginning of the study, the mice were allowed to adapt to the environment for 7 days. All of mice were raised under standard conditions with a 12 h light-dark cycle at 22 ± 1 °C and 55 ± 5% humidity with food and water provided ad libitum. Female BABL/c mice were inoculated subcutaneously with 5 × 10^5 4T1 breaster cancer cells. After the formation of a solid tumor with a volume of about 50 mm^3, the tumor bearing mice were randomly divided into 2 groups with seven mice in each group. Then, treatment with groups of mice by PBS or SMU127 (0.1 mg, i.p.) once every 2 days. At the end of the experiment, the mice were sacrificed. The tumor diameters were measured with calipers, and the tumor volume was calculated by the formula V (mm^3) =length×width×hight.



Scheme S1. Synthesis Routes of Compound SMU127. Reagents and conditions: (i) diethylamine, ethanol, ultrasonic, rt, 1h; (ii) dimethyl carbonate, 90°C, rf, 8h; (iii) 1-methylpiperazine, dimethyl carbonate, rt.

Synthesis of intermediate **4**

The mixture of cyclopentanone **1** (1 mmol), ethyl cyanoacetate **2** (1.1 mmol), sulfur **3** (1.1 mmol), and diethylamine (1.1 mmol) in ethanol (5 mL) was kept at room temperature in an ultrasonic cleaner for 1 hour. After completion of the reaction, the solvent was evaporated to dryness under reduced pressure. The crude product was purified by flash column chromatography on silica gel (petroleum ether (PE) : ethyl acetate (EA) = 10 : 1) to provide the product ethyl 2-amino-5,6-dihydro-4*H*-cyclopenta[*b*]thiophene-3-carboxylate (**4**) as yellow solid with 95 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) *δ* 5.79 (s, 2H), 4.27 (q, 2H), 2.85 (t, 2H), 2.73 (t, 2H), 2.31-2.35 (m, 2H), 1.34 (t, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.6, 165.7, 142.5, 121.1, 102.6, 59.3, 30.7, 28.8, 27.2, 14.4. ESI-MS: m/z 212.1 ($[M+H]^+$).

Synthesis of **SMU127**

The mixture of intermediate **4** (1 mmol) with bis(trichloromethyl)carbonate (0.35 mmol), and dimethyl carbonate (5 mL) was stirring at 90 °C for 8 hours (monitored by TLC). The mixture was cooled to room temperature. Without further purification, 1-methylpiperazine (1.1 mmol) was added and stirring at room temperature. After completion of the reaction, the solvent was evaporated to dryness under reduced pressure. The crude product was purified by flash column chromatography on silica gel (petroleum ether (PE) : ethyl acetate (EA) = 1 : 4) to provide the product ethyl 2-(4-methylpiperazine-1-carbothioamido)-5,6-dihydro-4H-cyclopenta[*b*] thiophene-3-carboxylate **SMU127** as white solid with 90 % yield. ^{1}H NMR (400 MHz, $CDCl_3$) *δ* 10.82 (s, 1H), 4.30 (q, 2H), 3.59 (t, 4H), 2.81-2.90 (m, 4H), 2.48 (t, 4H), 2.35-2.38 (m, 5H), 1.37 (t, 3H). ^{13}C NMR (101 MHz, $CDCl_3$) *δ* 166.7, 155.4, 152.8, 140.8, 130.2, 106.1, 60.1, 54.4, 45.9, 43.6, 30.3, 28.8, 27.7, 14.2. ESI-MS: m/z 338.2 ($[M+H]^+$).

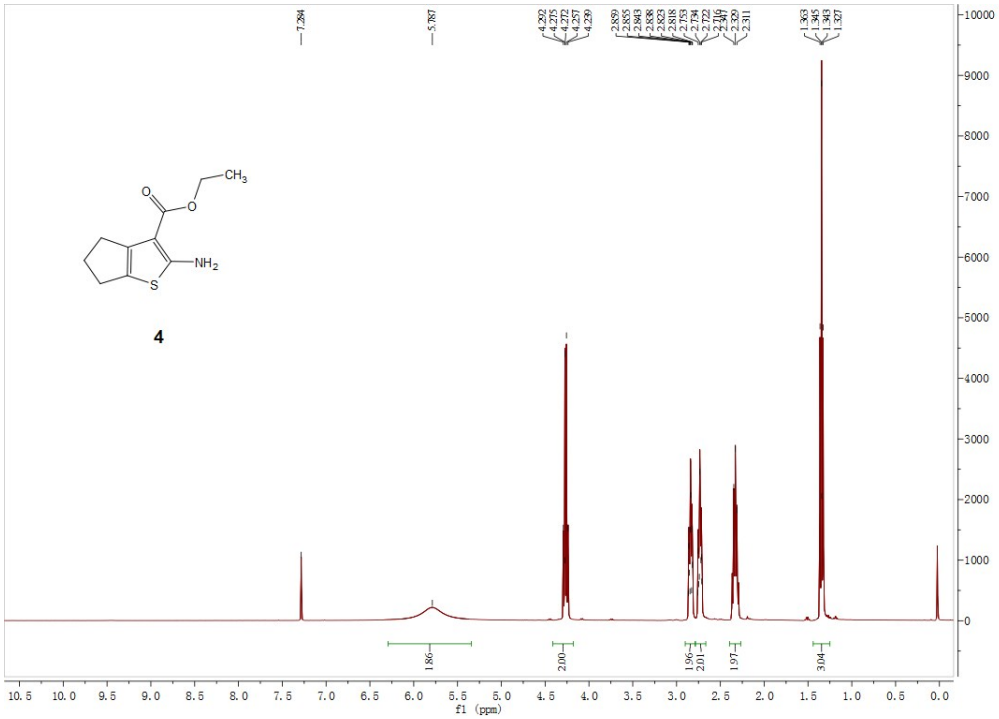


Fig. S3 ^{1}H NMR spectrum (400 MHz, $CDCl_3$) of **4.**

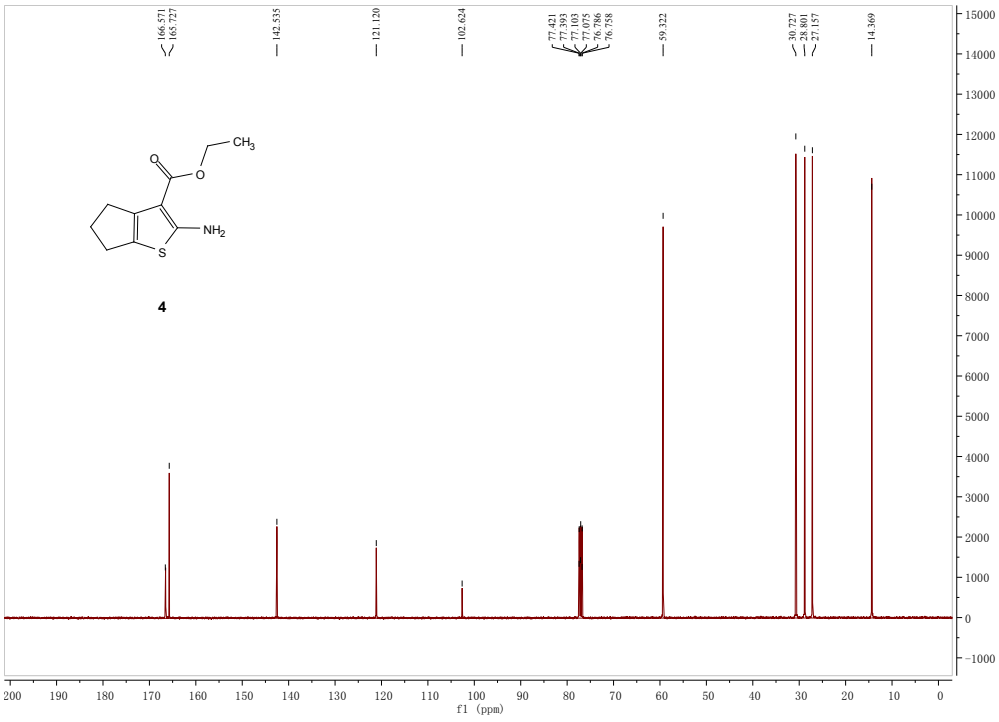


Fig. S4 ^{13}C NMR spectrum (101 MHz, $CDCl_3$) of **4.**

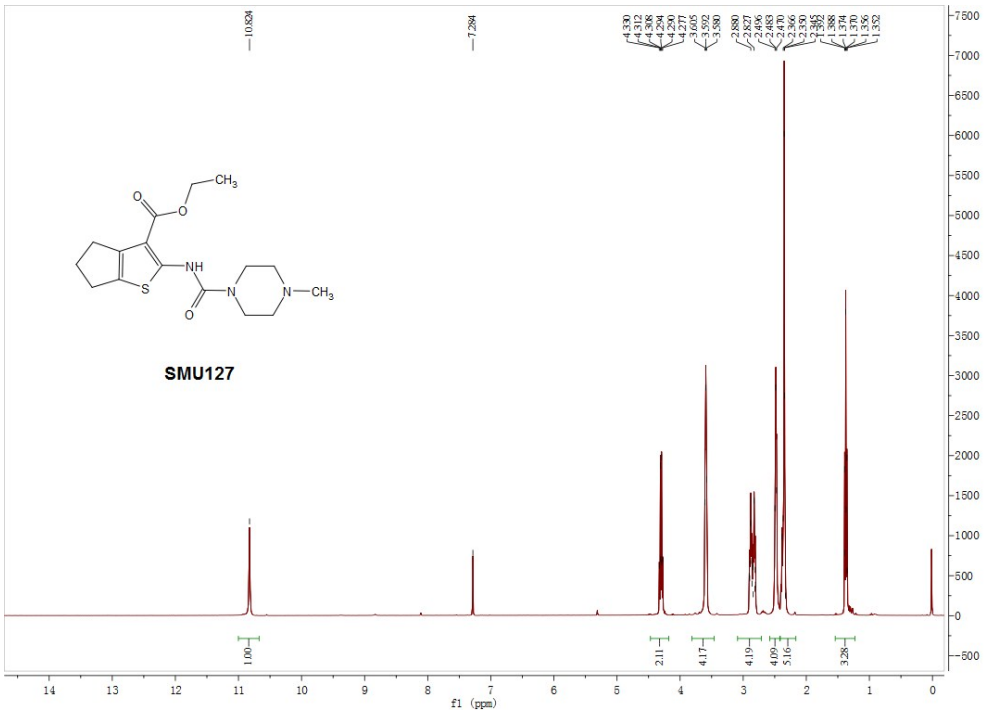


Fig. S5 ^{1}H NMR spectrum (400 MHz, $CDCl_3$) of **SMU127.**

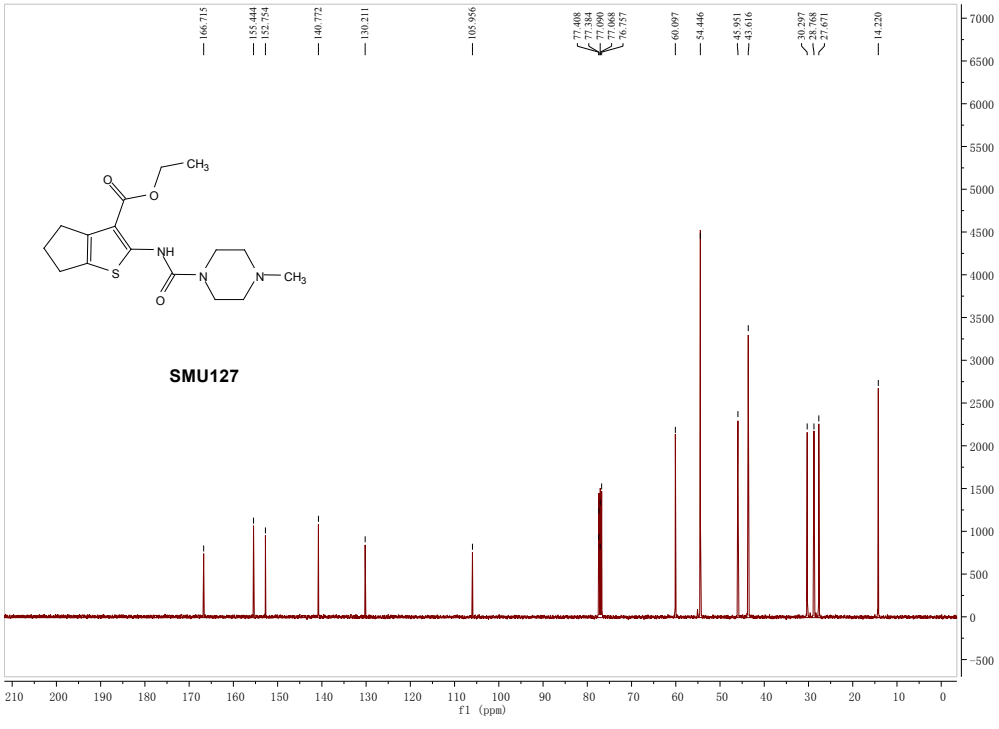


Fig. S6 ^{13}C NMR spectrum (101 MHz, $CDCl_3$) of **SMU127.**

References:

[1] S.J. Mi, S.E. Kim, Y.H. Jin, E.L. Mi, H.M. Kim, S.G. Paik, H. Lee, J.O. Lee, *Cell* **2007**, *130*, 1071-1082.