Pharmacophore-based virtual screening for identification of negative modulators of GLI1 as potential anticancer agents

Fabrizio Manetti,[1,2] Barbara Stecca,[3] Roberta Santini,[3] Luisa Maresca,[3] Giuseppe Giannini,[4] Maurizio Taddei,[1,2] and Elena Petricci[1]

[1]Dipartimento di Biotecnologie Chimica e Farmacia, Università di Siena, via Aldo Moro 2, I-53100 Siena, Italy
[2]Lead Discovery Siena, via Fiorentina 1, I-53100 Siena, Italy
[3]Istituto per lo Studio, la Prevenzione e la Rete Oncologica (ISPRO), Viale Pieraccini 6, I-50139 Firenze, Italy
[4]R&D, Alfasigma SpA, I-00071 Pomezia (Roma), Italy

COMPUTATIONAL DETAILS

A pharmacophoric-based virtual screening approach and molecular docking simulations were applied following computational protocols already described.[1,2] In brief, the structures of all the small molecules were sketched within Maestro (version 11.9, release 2019-1) and submitted to the LigPrep routine. The OPLS3e force field was applied, and Epik was used to generate the possible ionization and tautomeric states at pH 7 ± 2. Conformers were generated with the Systematic Pseudo Monte Carlo procedure (OPLS3e force field, water as the solvent, 5000 iteration Polak-Ribiere conjugate gradient minimizer, a cutoff root-mean-square deviation of 0.5 with 1000 steps, 25 kcal/mol cutoff relative to the global energy minimum to retain or discard conformations).

The three-dimensional pharmacophoric model for putative GLI-1 inhibitors was generated by means of the software Phase[3] by application of the common feature pharmacophoric hypothesis generation. This choice was based on the reduced number of GLI-1 inhibitors available (namely, glabrescione B and vismione E), that were both considered as active compounds. The resulting five-feature pharmacophore was constituted by two hydrogen bond acceptor (HBA) features, two hydrophobics (HY), and an aromatic ring RA) group. On the basis of the fitness core of Phase, the pharmacophoric-based virtual screening protocol led to prioritize 41 structurally different compounds that were purchased from Asinex[4] and AKos.[5]

The three-dimensional structure of the five-finger GLI/DNA complex was retrieved from the RCSB protein data bank (pdb code 2gli, 2.6 Å resolution) and used as the starting structure for molecular docking simulation. DNA and water molecules were removed and the Protein Preparation Wizard of the Maestro 2019-1 suite was applied to refine the original crystallographic structure by managing hydrogen atoms, bond orders, and water molecules, by using default values. SiteMap routine was applied for finding putative binding sites for GLI-1 inhibitors. Docking simulations were performed with Glide by using the standard precision (SP) scoring function with flexible ligand sampling and applying Epik state penalties. Receptor grids, whose size was adjusted manually to fit each binding site, were generated that accounted for conformational freedom of receptor hydroxyl groups within the gridbox.

EXPERIMENTAL

Chemistry

Glabrescione B was a gift from B. Botta and L. Di Marcotullio, University of Rome La Sapienza. GANT-61 (purity ≥ 98% by HPLC) was purchased from Sigma-Aldrich (Milan, Italy).

All the compounds purchased from either Asinex or Akos had a purity higher than 90%, as declared by vendors.

All reagents were used as purchased from commercial suppliers without further purification. The reactions were carried out in oven dried or flamed vessels. Solvents were dried and purified by conventional methods prior use. Flash column chromatography was performed with Aldrich silica gel 60, 0.040-0.063mm (230-400 mesh). Aldrich aluminum backed plates pre-coated with silica gel 60 (UV254) were used for analytical thin layer chromatography and were visualized by staining with a $KMnO_4$ or ninhydrin solution. NMR spectra were recorded at 25 °C and 400 MHz for ^{1}H and 100 MHz for [13]C. The solvent is specified for each spectrum. Splitting patterns are designated as s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet; br, broad. Chemical shifts (δ) are given in ppm relative to the resonance of their respective residual solvent peaks. Low resolution mass spectroscopy analyses were recorded by electrospray ionization on an Agilent 1100 Series instrument (flux: 9 L/min, T: 350 °C, P: 40 psi, frag: 70 eV).

Ethyl 5-cyclopropyl-7-(difluoromethyl)pyrazolo[1,5-a]pyrimidine-3-carboxylate 3. A mixture of **1** (1.79 mmol) and ethyl 3-amino-1*H*-pyrazole-4-carboxylate **2** (248 mg, 1.6 mmol) in 12N CH_3COOH (4 mL) in sealed tube was stirred at 110 °C over-night. The solvent was removed under vacuum and the solid obtained was diluted with AcOEt (20 mL), washed with H_2O (3 x 10 mL), and Brine (10 mL). Organic layers were dried over dry Na_2SO_4, filtered and the solvent removed under

reduced pressure obtaining the expected product in 56% yields. ^{1}H-NMR (400 MHz, CDCl$_3$, δ ppm): 8.61 (s, 1H), 7.18 (t, J = 53 Hz, 1H), 6.94 (s, 1H), 4.22 (q, J = 7.1 Hz, 1H); 2-08-2.14 (m, 2H); 1.26 (t, J = 7.12 Hz, 3H); 1.18-1.21 (m, 2H); 1.07-1.10 (m, 2H). ES-MS m/z: 304 [M+Na]$^+$. Anal. Calc. for C$_{13}$H$_{13}$F$_2$N$_3$O$_2$: C, 55.51; H, 4.66; F, 13.51; N, 14.94; O, 11.38; found: C, 55.46; H, 4.62; F, 13.58; N, 14.90.

General procedure for the synthesis of pyrazolo[1,5-a]pyrimidines SST0894-897. *The* ester **3** was dissolved a 1:1:1 mixture of EtOH, THF, and H$_2$O (9 mL). LiOH (58 mg, 2.4 mmol) was added and the reaction stirred at rt for 4 h. The mixture was treated with 4N HCl until pH 4-5 and diluted with CH$_2$Cl$_2$ (20 mL) and H$_2$O (20 mL). After extraction, the organic layer was dried over dry Na$_2$SO$_4$, filtered, and the solvent was removed under reduced pressure. The product obtained as a solid was dissolved (0.23 mmol) in dry DMF (1 mL) under nitrogen. The proper amine (0.23 mmol), HATU (88 mg, 0.23 mmol), and DIPEA (1.2 mL, 0.69 mmol) were added to the reaction mixture under nitrogen at 0 °C. The reaction mixture was stirred at rt under nitrogen overnight. The mixture was diluted with AcOEt (10 mL) and washed with H2O (3 x 5 mL), and Brine (8 mL). Organic layers were dried over dry Na$_2$SO$_4$ and filtered, and the solvent removed under reduced pressure. The crude obtained was purified by flash chromatography in a Buchi Sepacore system using PE/AcOEt gradients. The combustion analyses were within 0.5% of the theoretical values, confirming a purity of at least 95% for all tested compounds.

N-cyclohexyl-5-cyclopropyl-7-(difluoromethyl)pyrazolo[1,5-a]pyrimidine-3-carboxamide SST0894. Yield 63%. ^{1}H NMR (400 MHz, CDCl$_3$, δ ppm): 8.54 (s, 1H), 7.74 (d, J = 7.26 Hz, 1H); 7.27 (t, J = 53 Hz, 1H); 4.04-3.91 (m, 1H); 2.23-2.13 (m, 1H); 2.04-1.64 (m, 4H); 1.64-1.17 (m, 10H). ^{13}C NMR (101 MHz, CDCl3, δ ppm): δ 166.34, 160.46, 146.68, 146.47, 145.51, 138.40 (t, J = 27.9 Hz), 107.41 (d, J = 15.9 Hz), 105.05 (t, J = 35.0 Hz), 46.75, 32.72, 25.26, 24.02, 17.56, 12.87. ES-MS m/z: 357 [M$^+$Na]$^+$. Anal. Calc. for C$_{17}$H$_{20}$F$_2$N$_4$O: C, 61.07; H, 6.03; F, 11.36; N, 16.76; O, 4.78; found: C, 61.10; H, 6.04; F, 11.39; N, 16.74.

5-Cyclopropyl-7-(difluoromethyl)-N-phenethylpyrazolo[1,5-a]pyrimidine-3-carboxamide SST0895. Yield 25%. ^{1}H NMR (400 MHz, CDCl$_3$, δ ppm): 8.58 (s, 1H); 7.72 (s, 1H); 7.26 (t, J = 53 Hz, 1H); 7.41-7.12 (m, 6H); 3.78 (dd, J = 12.8, 6.5 Hz, 2 H); 2.92 (t, J = 6.7 Hz, 2H); 2.15-2.03 (m, 1H), 1.20-1.92 (m, 4H). ^{13}C NMR (101 MHz, CDCl$_3$, δ ppm): 166.56, 161.56, 146.68, 145.53, 138.73, 128.30, 128.18, 125.97, 109.81, 107.83, 104.74 (d, J = 56.9 Hz), 39.60, 35.55, 29.25, 24.02, 17.68, 12.62. ES-MS m/z: 359 [M$^+$Na]$^+$. Anal. Calc. for C$_{19}$H$_{18}$F$_2$N$_4$O: C, 64.04; H, 5.09; F, 10.66; N, 15.72; O, 4.49; found: C, 64.06; H, 5.09; F, 10.62; N, 15.78.

N,N-dibenzyl-5-cyclopropyl-7-(difluoromethyl)pyrazolo[1,5-a]pyrimidine-3-carboxamide SST0896. Yield 78%. ^{1}H NMR (400 MHz, CDCl$_3$, δ ppm): 8.40 (s, 1H); 7.40-7.08 (m, 12H); 4.72 (bs, 2H), 4.59 (bs, 2H), 2.13-2.09 (m, 1H); 1.18-1.52 (m, 4H). ^{13}C NMR (101 MHz, CDCl3, δ ppm): 166.16, 164.48, 146.35, 144.84, 137.80, 136.45, 128.21, 127.36, 126.97, 109.97, 107.57, 105.72 (d, J = 56.0 Hz), 51.37, 46.69, 17.74, 12.52. ES-MS m/z: 446 [M$^+$Na]$^+$. Anal. Calc. for C$_{25}$H$_{22}$F$_2$N$_4$O: C, 69.43; H, 5.13; F, 8.79; N, 12.96; O, 3.70; found C, 69.49; H, 5.18; F, 8.80; N, 12.91.

N-(benzo[d][1,3]dioxol-5-ylmethyl)-5-cyclopropyl-7-(difluoromethyl)pyrazolo[1,5-a]pyrimidine-3-carboxamide SST0897. Yield 89%. ^{1}H NMR (400 MHz, CDCl$_3$, δ ppm): 8.53 (s, 1H); 7.97 (bs, 1H); 7.24 (t, J = 52.9 Hz, 1H); 7.05 (s, 1H); 6.80-6.70 (m, 3H); 5.87 (bs, 2H); 4.50 (d, J = 5.3 Hz, 2H); 2.17- 2.06 (m, 1H); 1.28-1.08 (m, 4H). ^{13}C NMR (101 MHz, CDCl$_3$, δ ppm): 166.77, 161.18, 147.51, 146.44, 146.43, 145.63, 138.46, 132.05, 120.37, 107.91, 107.80, 104.85, 104.65, 100.59, 42.65, 17.59, 12.74. ES-MS m/z: 409 [M$^+$Na]$^+$. Anal. Calc. for C$_{19}$H$_{16}$F$_2$N$_4$O$_3$: C, 59.07; H, 4.17; F, 9.83; N, 14.50; O, 12.42; found: C, 59.00; H, 4.18; F, 9.85; N, 14.48.

Biology

Cell lines and treatments

Human melanoma cells A375 (CRL-1619), medulloblastoma cells DAOY (HTB-186), and murine NIH3T3 cells (CRL-1658) were obtained from ATCC (Manassas, VA). Cells were maintained in Dulbecco's modified Eagle's medium (DMEM) (Euroclone, Milan, Italy) (A375 and NIH3T3) or Eagle's Minimum Essential Medium (EMEM) (Euroclone) (DAOY) supplemented with 10% fetal bovine serum (FBS), 1% penicillin-streptomycin, 1% glutamine (Lonza, Basel, Switzerland). MCF10A cells were kindly provided by Dr. Silvo Conticello (CRL-ISPRO, Florence, Italy) and were maintained in DMEM/F12 (Euroclone) supplemented with 5% horse serum, 1% penicillin-streptomycin, 1% glutamine, 100 ng/mL cholera toxin, 10 μg/mL insulin, 20 ng/mL hEGF and 500 ng/mL hydrocortisone. Cells were periodically screened for mycoplasma contamination by PCR. A375, DAOY, and MCF10A cells were plated in 24-well or 96-well plates, treated with increasing concentrations of each compound for 72 h in complete medium supplemented with 1% FBS and viable cells counted by trypan blue-exclusion or crystal violet staining. All experimental points were set up in triplicate wells and all experiments were repeated at least three times. IC_{50} were calculated using GraphPad Prism version 6.0 (GraphPad software).

Luciferase reporter assays

NIH3T3 cells were transfected with the GLI-responsive luciferase reporter plasmid (8x3'Gli-BS)[6] in combination with renilla luciferase pRL-TK reporter vector (Promega, Madison, WI, USA) to normalize luciferase activities, and treated with SAG[7] for 48 h to activate the Hedgehog pathway. Cells were treated with DMSO, compounds (putative GLI inhibitors), glabrescione B or GANT61 for 48 h and luminescence was measured using the Dual-Glo Luciferase Assay System (Promega) and the GloMax® 20/20 Luminometer (Promega).

Lentiviral vector

Lentiviruses were produced in HEK-293T cells. Lentiviral vectors pLKO.1-puro (LV-c), pLKO.1-shGLI1 (LV-shGLI1), and pLKO.1-shGLI2 (LV-shGLI2) were already described.[8]

Quantitative real-time PCR

Total RNA was isolated with TRIzol reagent (Thermo Fisher Scientific, Waltham, MA) and subjected to DNase I treatment (Merck, Darmstadt, Germany). Reverse transcription was performed with high capacity cDNA reverse transcription kit (Thermo Fisher Scientific). Quantitative real-time PCR (qPCR) was carried out using Luna Universal qPCR Master Mix (New England Biolabs, Ipswich, MA) in a Rotorgene-Q (Qiagen, Hilden, Germany). Primer sequences are available upon request.
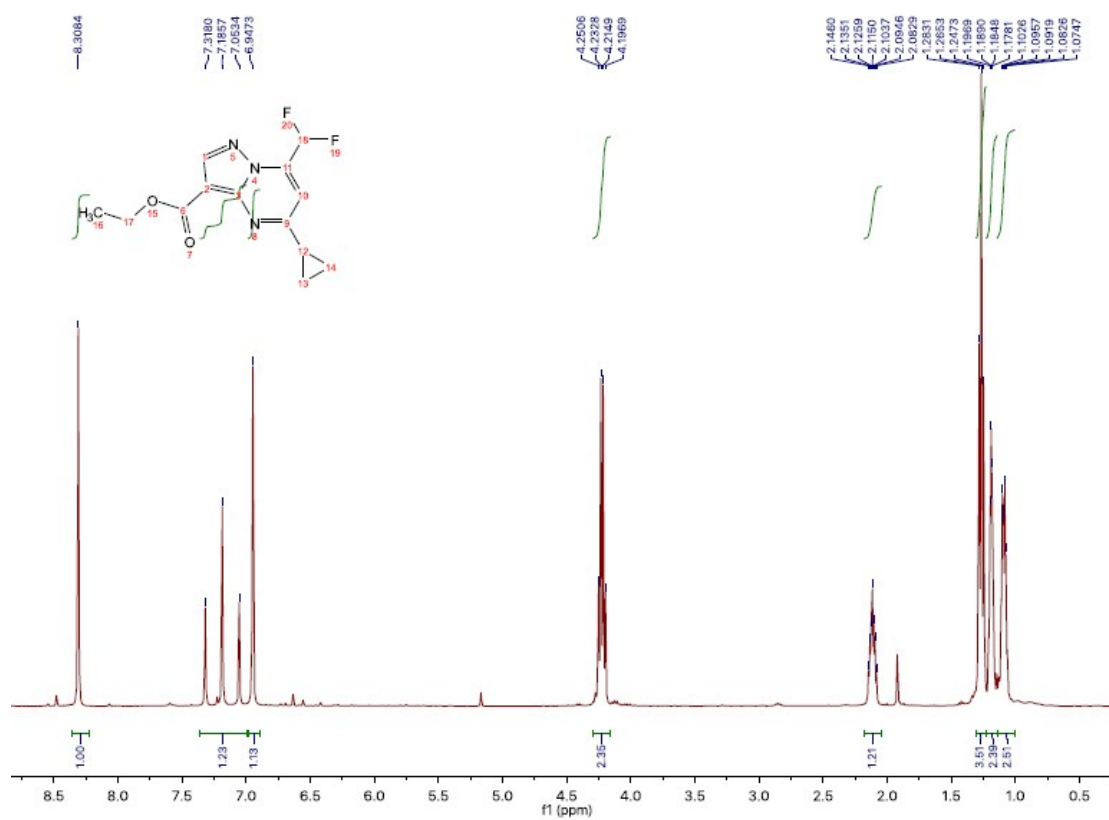
Western blot analysis

Cells were lysed in ice-cold RIPA buffer supplemented with protease and phosphatase inhibitors, centrifuged at 4 °C, and supernatant was collected, as previously described.[9,10] Equal amounts of proteins were resolved by SDS polyacrylamide gel electrophoresis, transferred into nitrocellulose membranes, and incubated in blocking buffer. The following antibodies were used: mouse anti-GLI1 (#2643) (Cell Signaling Technology, Danvers, MA), goat anti-GLI2 (AF-3635, R&D Systems, Minneapolis, MN), mouse anti-HSP90 (sc-13119) (Santa Cruz Biotechnology, Santa Cruz, CA), and the anti-mouse HRP-conjugated (Cell Signaling). Blotted membranes were developed using SuperSignal West Femto (ThermoFisher Scientific, Italy) and imaged with ChemiDoc™ Imaging Systems (Bio-Rad, Italy). Images were recorded as TIFF files for quantification with ImageJ software.
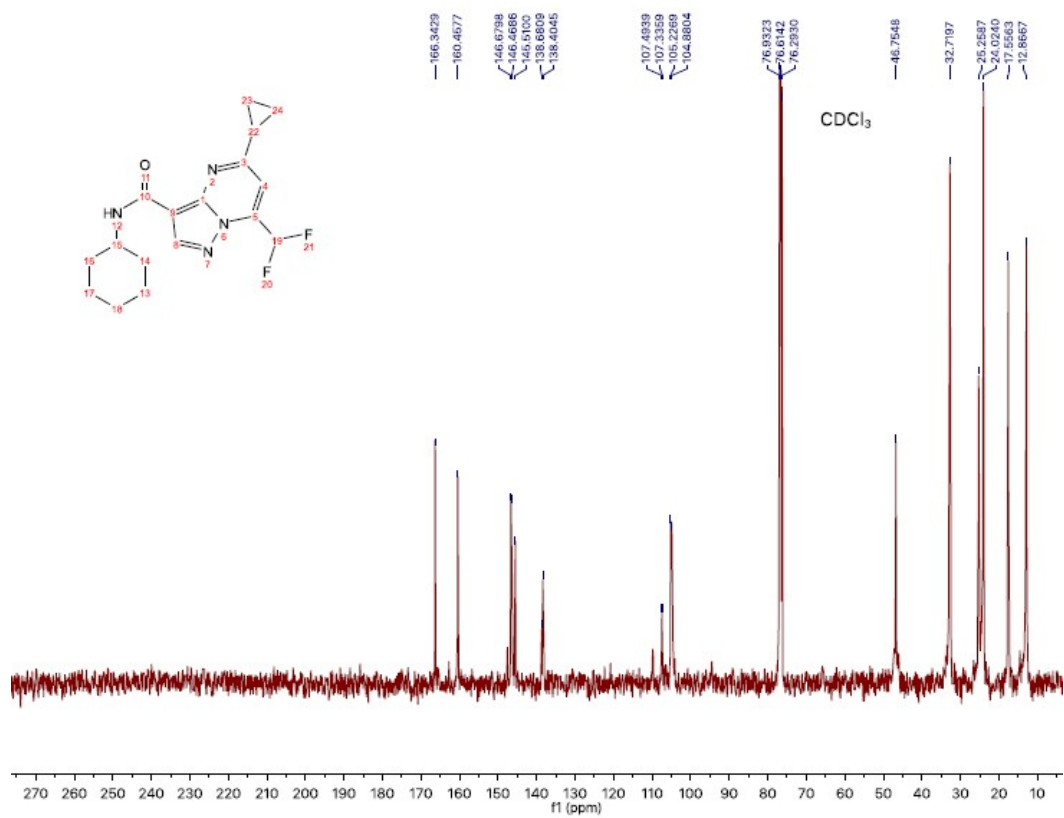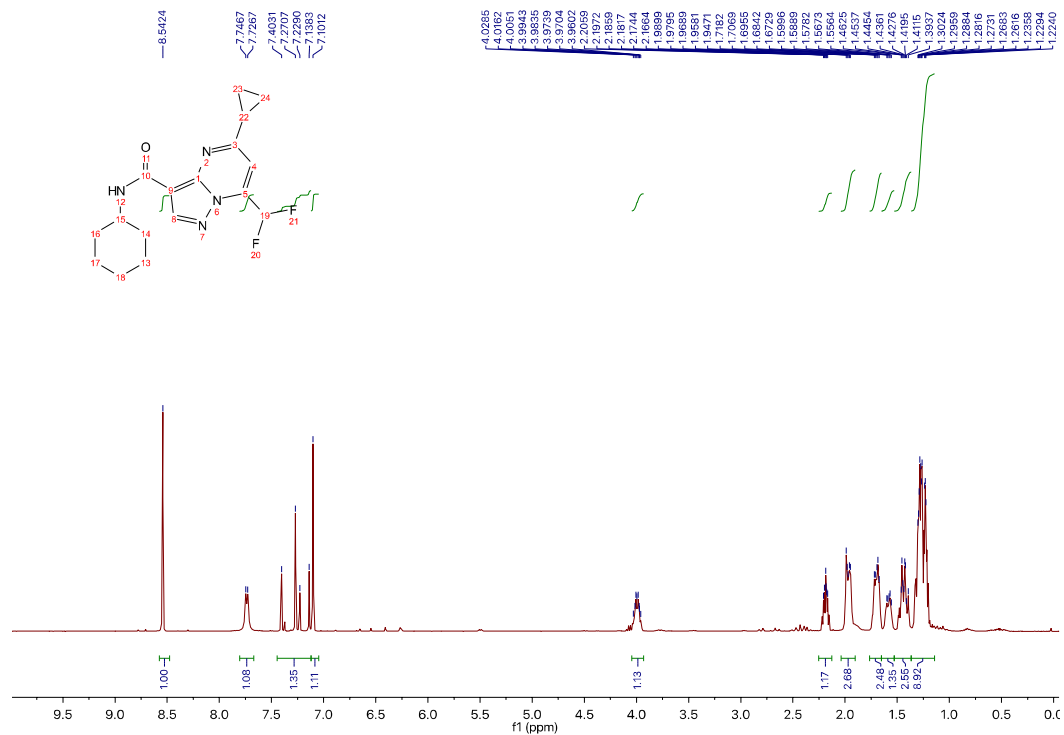
Statistical analysis

Data represent mean ± SD or mean ± SEM values calculated on at least 3 independent experiments. p Values were calculated using Student's t-test or one-way ANOVA. A two-tailed value of $p<0.05$ was considered statistically significant.
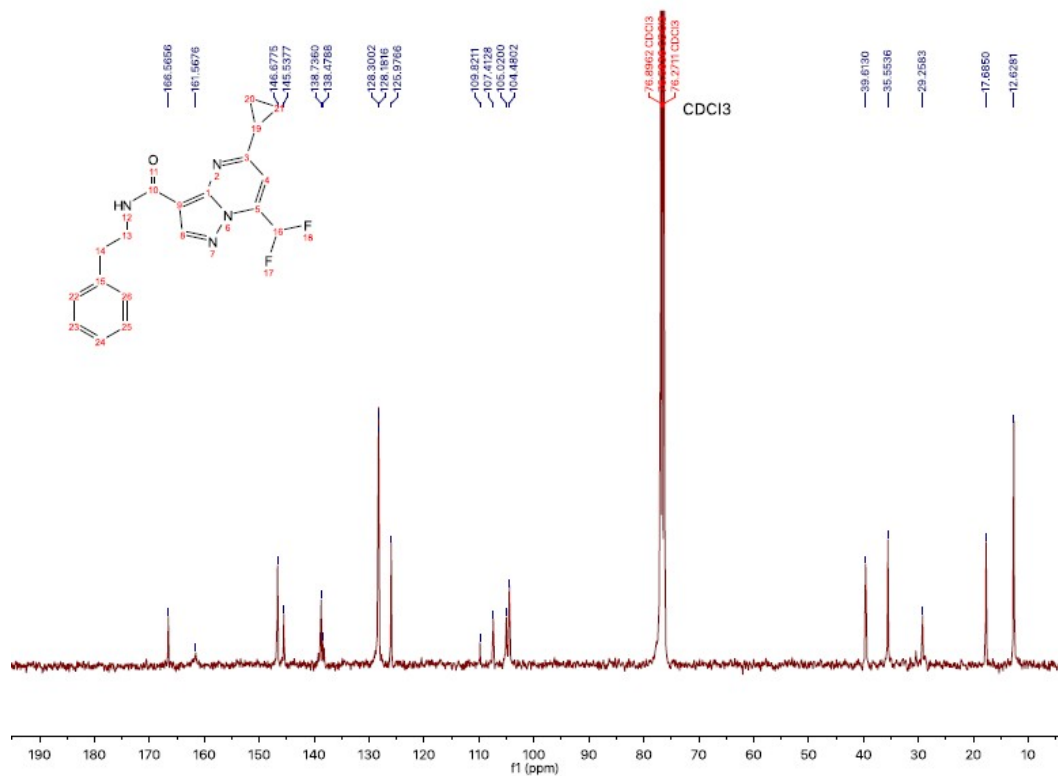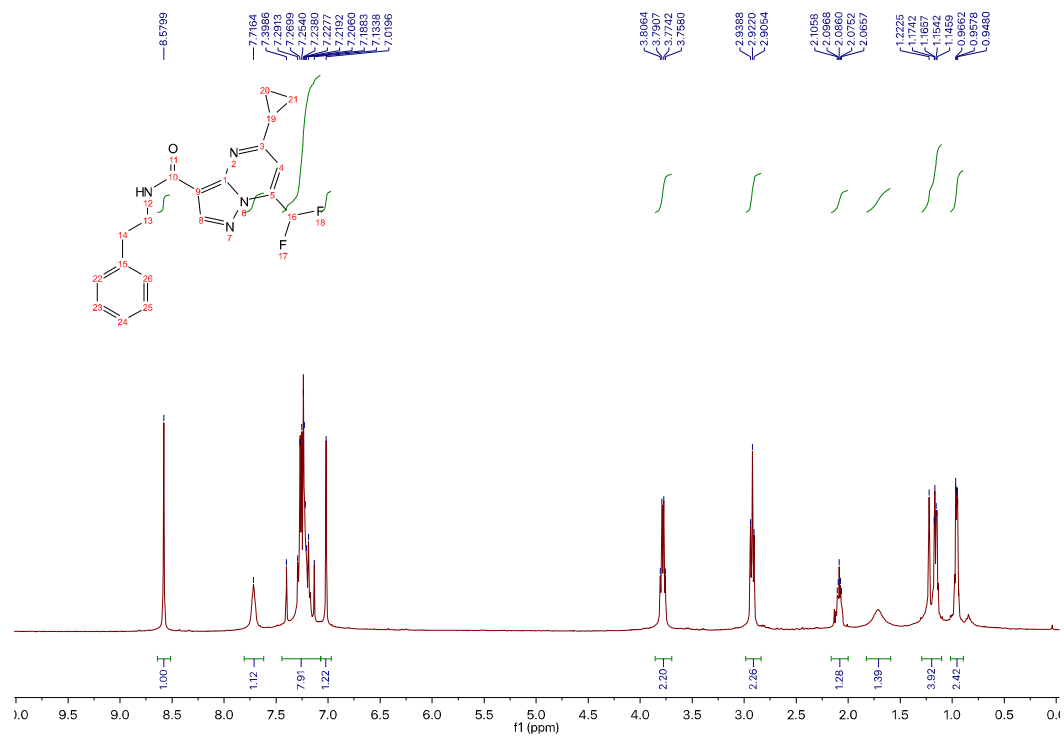
¹H and ¹³C NMR of synthesized compounds.

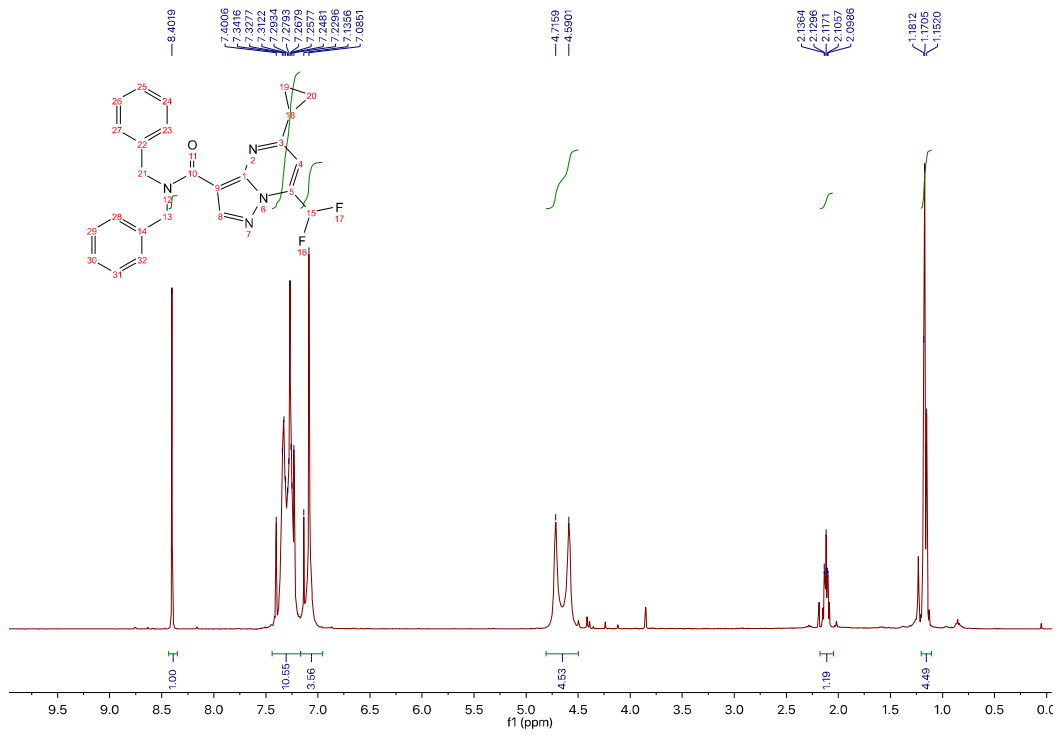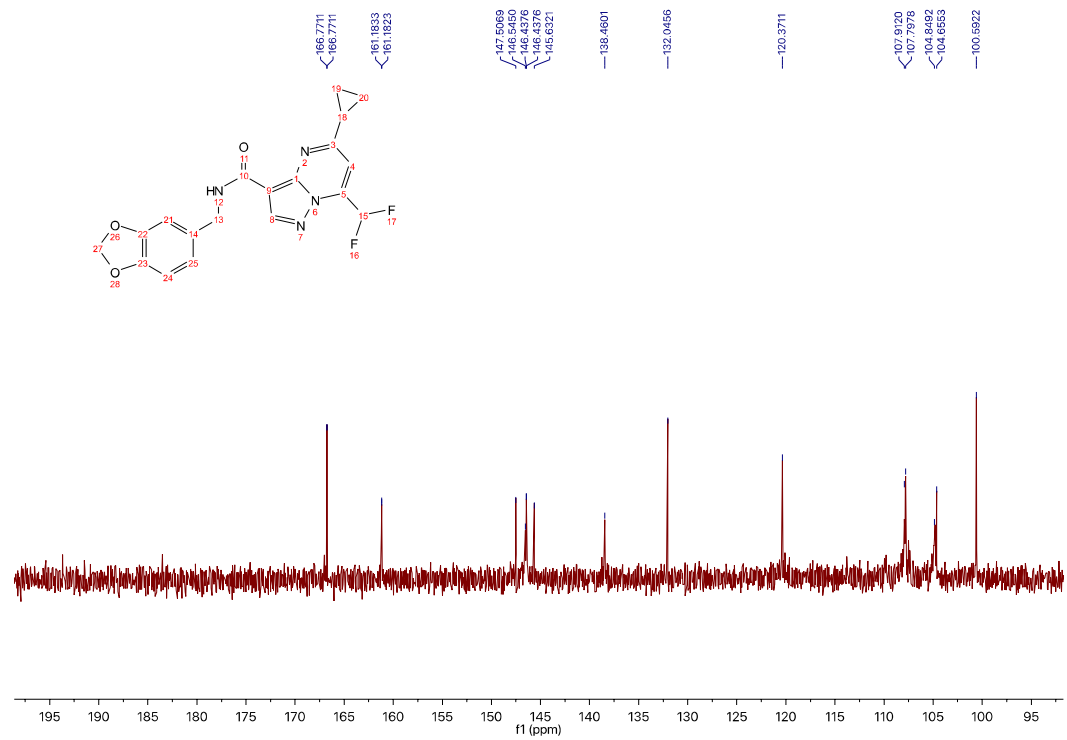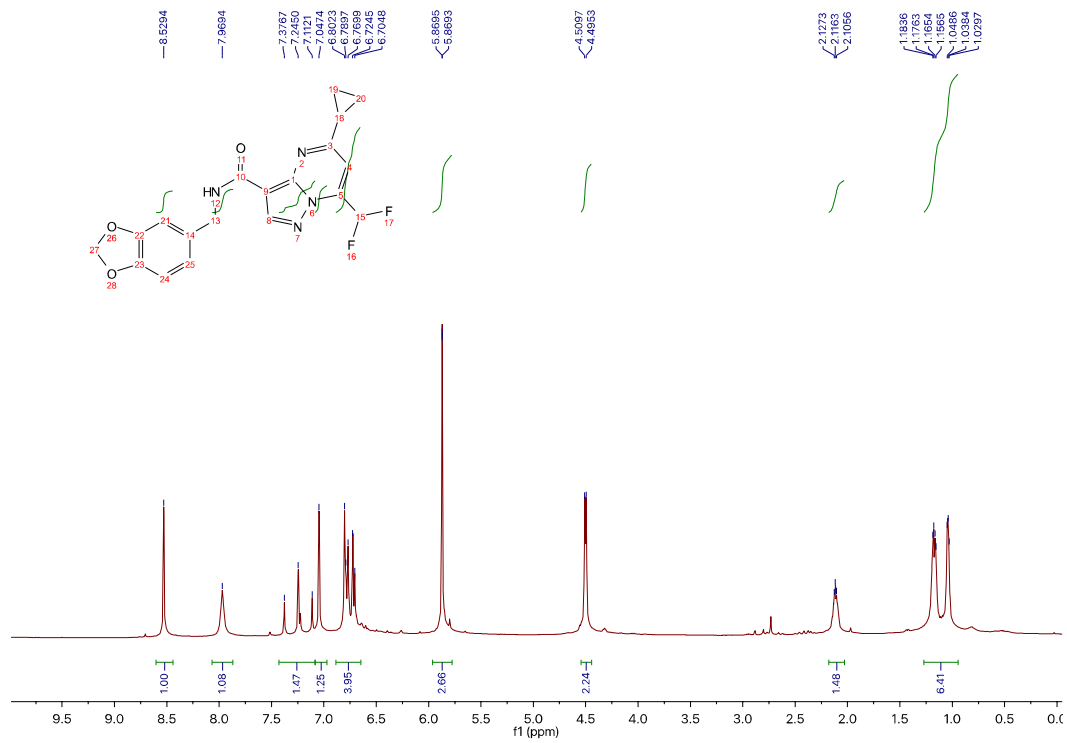
Compound **3**.

Compound **SST0894**.

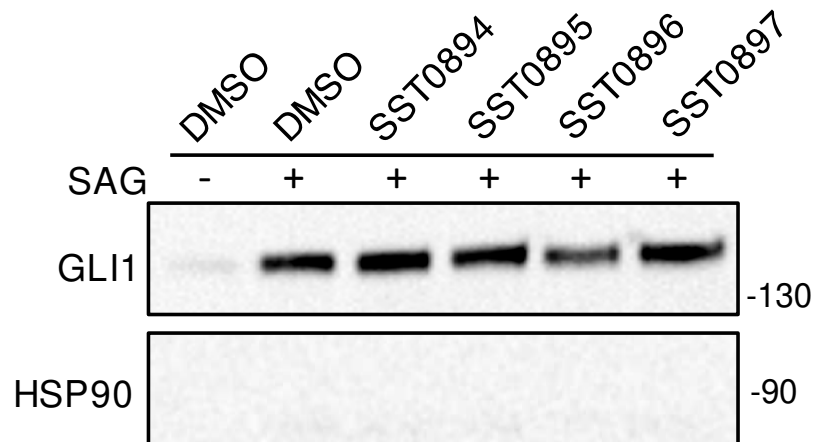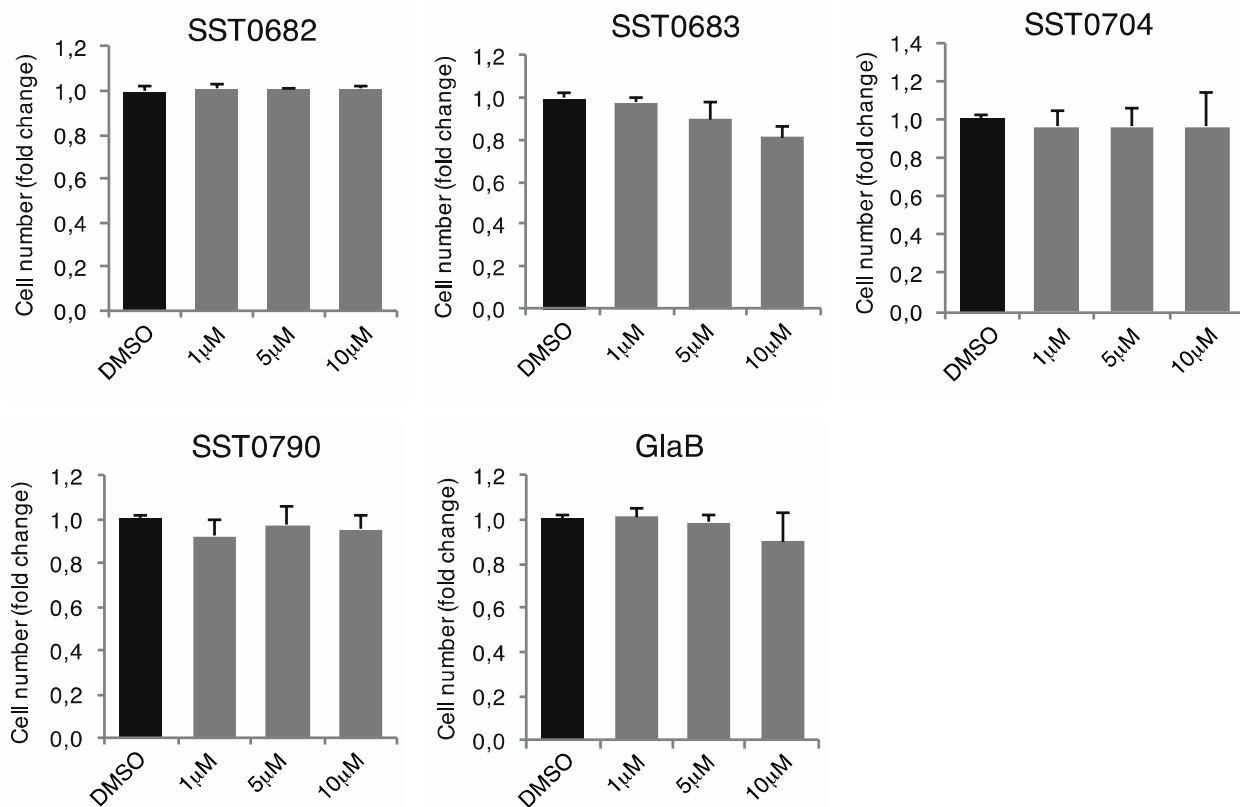Compound **SST0895**.
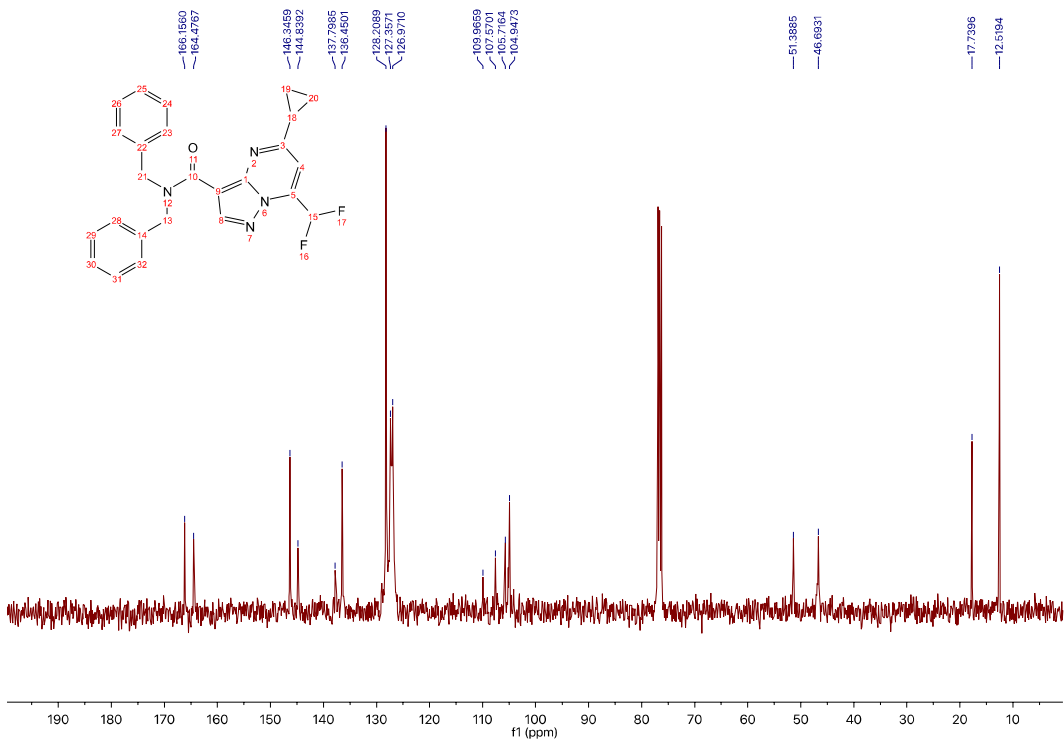
Compound **SST0896**.

Compound **SST0897**.

A

B

Supporting Information Figure 1. A) Western blot analysis of GLI1 protein in NIH3T3 cells treated with 100 nM SAG and DMSO or compounds SST0894, SST0895, SST0896, SST0897 (5 µM) for 48 h. HSP90 was used as loading control. B) Effects of compounds on MCF10A cell viability. Mammary epithelial MCF10A cells were treated with SST0682, SST0683, SST0704, SST0790 and glabrescione B (GlaB) at the indicated doses for 72 h. Data are shown as mean ± SD of at least three independent experiments.
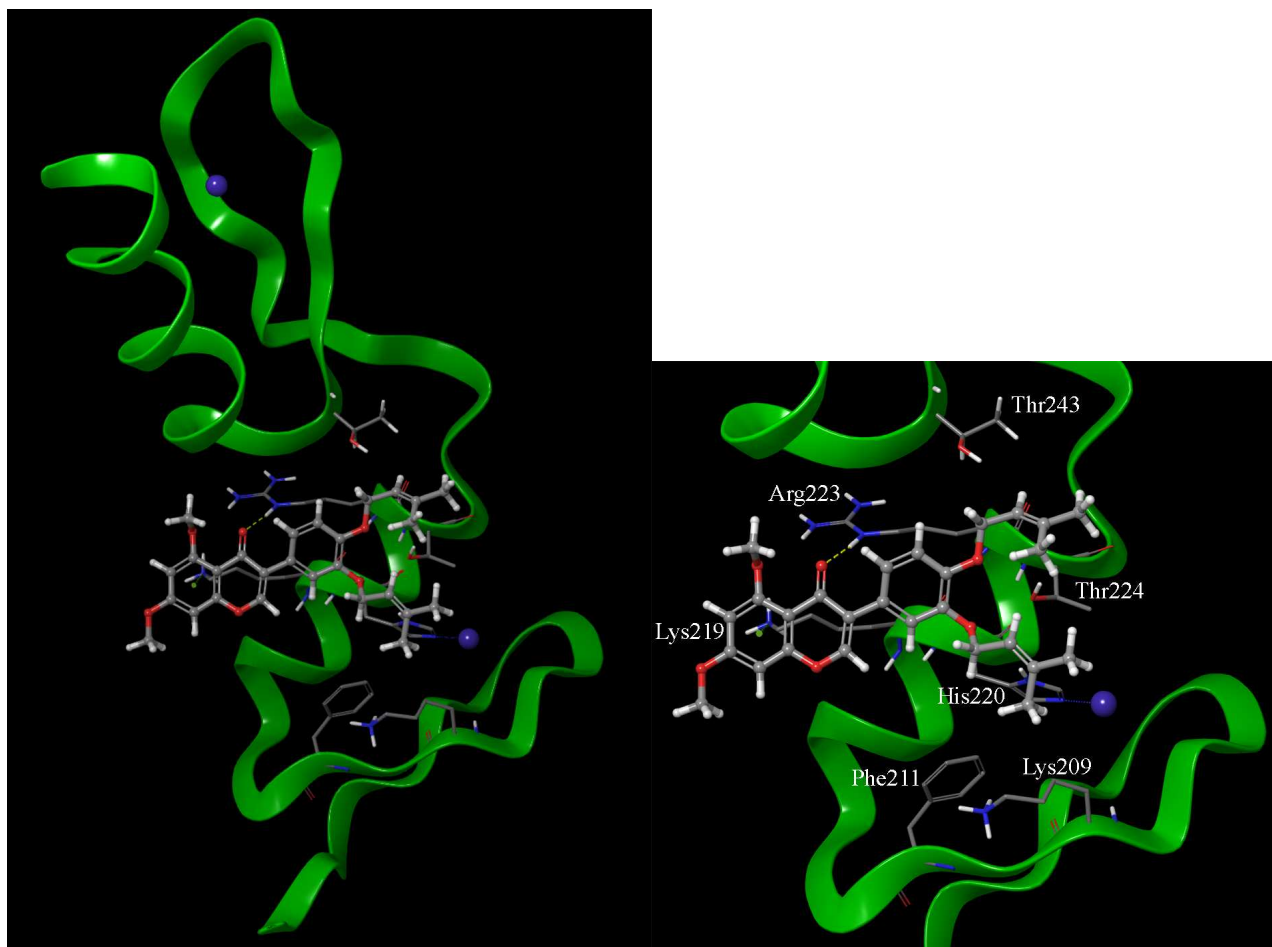
Supporting Information Figure 2. Best-ranked docking pose of glabrescione B within GLI1 protein. Left: the small molecule is accommodated in proximity of the zinc ion of the zinc finger 4, in a region between zinc finger 4 and 5. For the sake of clarity, zinc finger 1 and 2 are omitted. Right: details of the interaction pathway between glabrescione B and GLI1. The condensed phenyl ring of the chromene moiety is involved in a π-cation interaction with the terminal ammonium group of Lys219, the 4-carbonyl oxygen of the ligand and the guanidino group of Arg223 make a hydrogen bond These two amino acids are already described as crucial residues for binding of glabrescione B to GLI1.[11] Moreover, the prenyl chains are in the region occupied by the ethyl chain of the ester moiety of thiophene derivatives (Figure 5) and by the cyclopropyl moiety of **SST0704** (Figure 6). A prenyl side chain further stabilizes the complex by hydrophobic interactions with the side chains of Thr224 and Thr243, the polymethylenic chain of Lys209, and the aromatic moiety of Phe211. The structure of glabrescione B is represented by thick sticks, amino acid side chains by thin sticks, the zinc ion as a blu sphere, hydrogen bonds as dotted yellow lines, π-cation interaction as dotted green lines.

Supporting Information Table 1. Conversion of the protein data bank amino acid numbering into the full length GLI numbering.

Protein data bank structure[a]	Full length structure[b]
Lys207	Lys338
Lys209	Lys340
Phe211	Phe342
Lys219	Lys350
His220	His351
Arg223	Arg354
Thr224	Thr355
Thr243	Thr374

[a]Residue numbering found in the 155 amino acid structure 2gli stored within the protein data bank.
[a]Residue numbering found in the primary sequence of the 1106 amino acid zinc finger protein GLI1 deposited within the UNIPROT database (code P08151, https://www.uniprot.org/uniprot/P08151).

REFERENCES

1. Bhakta, S.; Scalacci, N.; Maitra, A.; Brown, A. K.; Dasugari, S.; Evangelopoulos, D.; McHugh, T. D.; Mortazavi, P. N.; Twist, A.; Petricci, E.; Manetti, F.; Castagnolo, D. Design and synthesis of 1-((1,5-bis(4-chlorophenyl)-2-methyl-*1H*-pyrrol-3-yl)methyl)-4-methylpiperazine (BM212) and *N*-adamantan-2-yl-*N'*-((*E*)-3,7-dimethylocta-2,6-dienyl)ethane-1,2-diamine (SQ109) pyrrole hybrid derivatives: discovery of potent antitubercular agents effective against multidrug-resistant Mycobacteria. *J. Med. Chem.* **2016**, *59*, 2780-2793.

2. Badolato, M.; Carullo, G.; Perri, M.; Cione, E.; Manetti, F.; Di Gioia, M. L.; Brizzi, A.; Caroleo, M. C.; Aiello, F. Quercetin/oleic acid-based G-protein-coupled receptor 40 ligands as new insulin secretion modulators. *Future Med. Chem.* **2017**, *9*, 1873-1885.

3. Dixon, S. L.; Smondyrev, A. M.; Rao, S. N. PHASE: a novel approach to pharmacophore modeling and 3D database searching. *Chem. Biol. Drug. Des.* **2006**, *67*, 370-372.

4. Asinex Corporation, www.asinex.com.

5. AKos Consulting & Solutions GmbH, http://www.akosgmbh.de.

6. Sasaki, H.; Hui, C.; Nakafuku, M.; Kondoh, H. A binding site for Gli proteins is essential for HNF-3β floor plate enhancer activity in transgenics and can respond to Shh in vitro. *Development* 1997, *124*, 1313-1322.

7. Chen, J.K.; Taipale, J.; Young, K.E.; Maiti, T.; Beachy, P.A. Small molecule modulation of Smoothened activity. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14071-14076.

8. Santini, R.; Vinci, M.C.; Pandolfi, S.; Penachioni, J.Y.; Montagnani, V.; Olivito, B.; Gattai, R.; Pimpinelli, N.; Gerlini, G.; Borgognoni, L.; Stecca, B. Hedgehog-GLI signaling drives self-renewal and tumorigenicity of human melanoma-initiating cells. *Stem Cells* **2012**, *30*, 1808-1818.

9. Pandolfi, S.; Montagnani, V.; Lapucci, A.; Stecca, B. HEDGEHOG/GLI-E2F1 axis modulates iASPP expression and function and regulates melanoma cell growth. *Cell Death Differ.* **2015**, *22*, 2006-2019.

10. Pietrobono, S.; Santini, R.; Gagliardi, S.; Dapporto, F.; Colecchia, D.; Chiariello, M.; Leone, C.; Valoti, M.; Manetti, F.; Petricci, E.; Taddei, M.; Stecca, B. Targeted inhibition of Hedgehog-GLI signaling by novel acylguanidine derivatives inhibits melanoma cell growth by inducing replication stress and mitotic catastrophe. *Cell Death Dis.* **2018**, *9*, 142.

11. Infante, P.; Mori, M.; Alfonsi, R.; Ghirga, F.; Aiello, F.; Toscano, S.; Ingallina, C.; Siler, M.; Cucchi, D.; Po, A.; Miele, E.; D'Amico, D.; Canettieri, G.; De Smaele, E.; Ferretti, E.; Screpanti, I.; Uccello Barretta, G.; Botta, M.; Botta, B.; Gulino, A.; Di Marcotullio, L. Gli1/DNA interaction is a druggable target for Hedgehog-dependent tumors. *EMBO J.* **2015**, *34*, 200.217.