Supporting Information

Chemical Space Mimicry for Drug Discovery

William Yuan,[§ †] Dadi Jiang,[†] Dhanya K. Nambiar,[†] Lydia P. Liew,[£] Michael P. Hay,[£] Joshua Bloomstein,[†] Peter Lu,[†] Brandon Turner,[†] Quynh-Thu Le,[†] Robert Tibshirani,[†] Purvesh Khatri,[†] Mark G. Moloney,[‡] and Albert C. Koong[†]*

§ Trinity College, University of Oxford, Oxford, OX1 3BH, United Kingdom

† Stanford University School of Medicine, Stanford, CA 94305, United States

£ Auckland Cancer Society Research Centre, Faculty of Medical and Health Sciences, The University of Auckland, Auckland, 1023, New Zealand

‡ Chemistry Research Laboratory, University of Oxford, Oxford, OX1 3TA, United Kingdom

*E-mail: akoong@stanford.edu

Contents:

Methods

- **Library Generation**
- **Library Characterization**
- **UPR Bioactivity Characterization**
- **VEGFR-2 Docking and Bioactivity Characterization**
- **Synthesis of Compounds 1-5**
- **NMR Spectra of Compounds 1-5**

Supplementary Figures

Table S1. SMILES, name, and structure of MIMICS generated, newly identified UPR inhibitors.

Table S2. SMILES, name, and structure of MIMICS generated, FDA approved compounds

Methods

Library Generation

Lists of SMILES corresponding to sets of interest were enumerated with each string separated by a newline character. The order of molecules within the list was randomized before use. The resultant list of SMILES was input directly to char-RNN, and the network trained until the validation loss between epochs converged and the raw epoch value exceeded 1. The Long Term Short Memory RNN architecture was used. Network size and layer number were scaled with input size. For the bioactives library, (input size = 880,000 molecules) a size of 350 and layer count of 4 were used. For the VEGFR2 library, (input size = 25,000 molecules) a size of 100 and a layer count of 2 were used. In both cases, dropout was set to 0.5 and sequence length was set to 40. Subsequent checkpoints were retained for parallel character generation. Blocks of text, at least one million characters long, were generated from sampling the created checkpoints, and formed the raw output. Sampling temperature was set to 1. Default values were used for all other parameters. Cleaning of output occurred as in Figure S1.

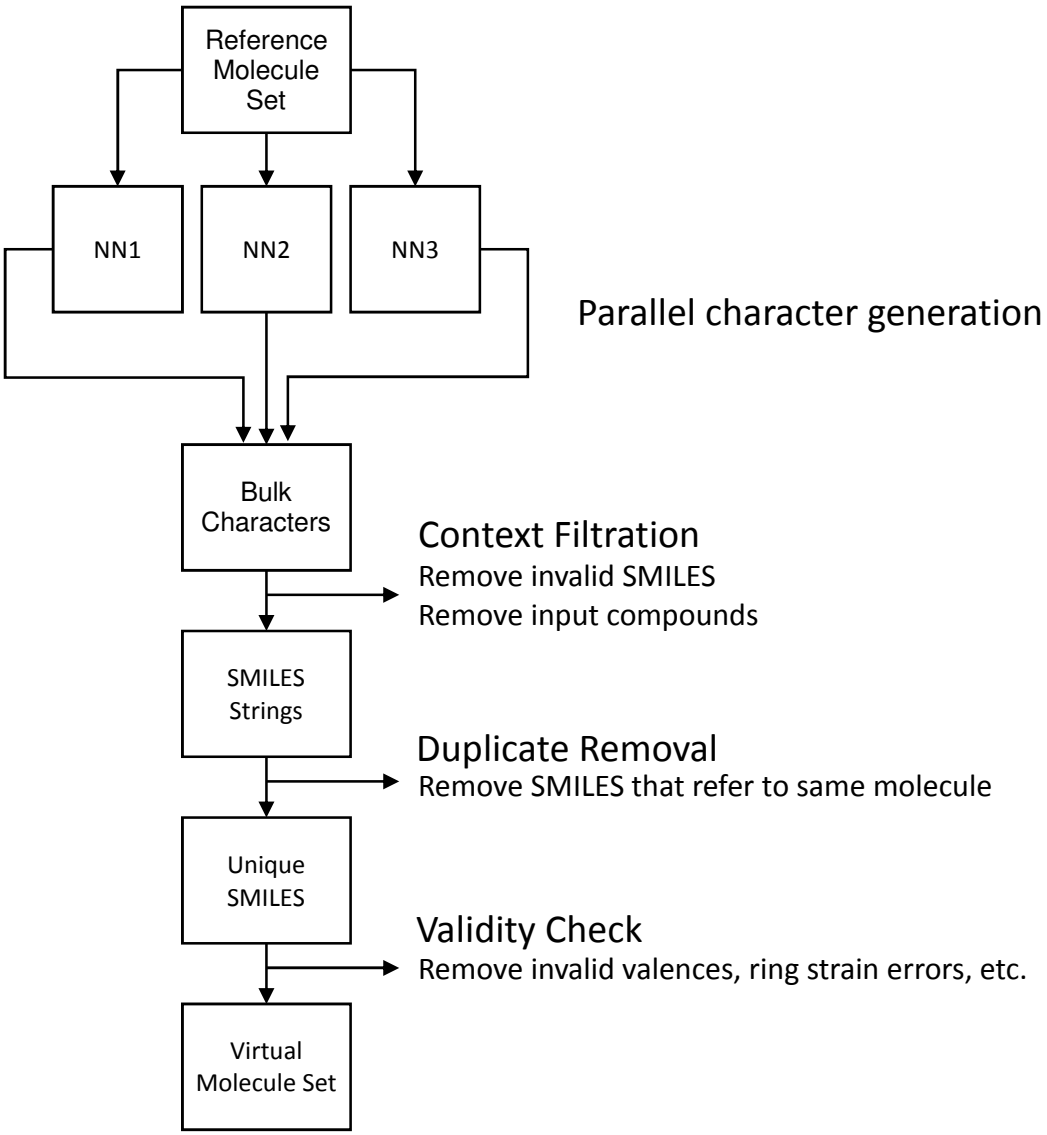


Figure S1. Enumeration of a set molecules from a reference. Compounds of interest, in the form of SMILES strings, are input into char-RNN, a freely available neural network. char-RNN is used to generate characters based on its analysis of the input set which are then processed back into meaningful SMILES strings. This collection of SMILES represents the generated set.

Library Characterization

Bemis-Murcko (BM) Clustering[1] was used to compute the structural diversity of MIMICS and input sets such that potential analogues of a molecule can be condensed into a

single framework. Summing the number of unique frameworks needed to enumerate a set was used to compare the MIMICS generated libraries relative to the input set.

Because of the methodology by which molecules are generated, it is not possible to enumerate the pathway from which any particular MIMICS molecule was created. To demonstrate that MIMICS generated truly novel structures, a subset of MIMICS molecules was selected, and for each molecule, the nearest neighbor in the input set (determined by Tanimoto similarity) was identified. The Tanimoto distance for each of these nearest neighbor pairs was computed and the distribution of scores examined. The same procedure was repeated for a sample of input molecules themselves, with the self-similarity term removed. This analysis provides a description of the novelty of the MIMICS molecules compared to the input set. High scoring population is used as a measure of novelty. Nearest neighbor analyses on both BM clusters and the molecules themselves were conducted using the OpenBabel[2] FP2 fingerprint.

Normalized principal moment of inertia (PMI) ratios were used as a structural similarity metric. 3D structures were calculated with generated SMILES strings and the three principal moments of inertia for each molecule (I1, I2, and I3, such that $I1 \leq I2 \leq I3$) were expressed as the ratios I1/I3 and I2/I3. Plotting these ratios against each other results in a triangular projection defined by the points (0,1), (0.5, 0.5), and (1,1), corresponding to rod-like, disk-like, and sphere-like molecules respectively[3]. The construction of a dedicated “ring-like” MIMICS implementation, with input molecules sourced from the ZINC[4] library, demonstrated the ability for MIMICS to preserve different shape distributions (Figure S2 and S3).

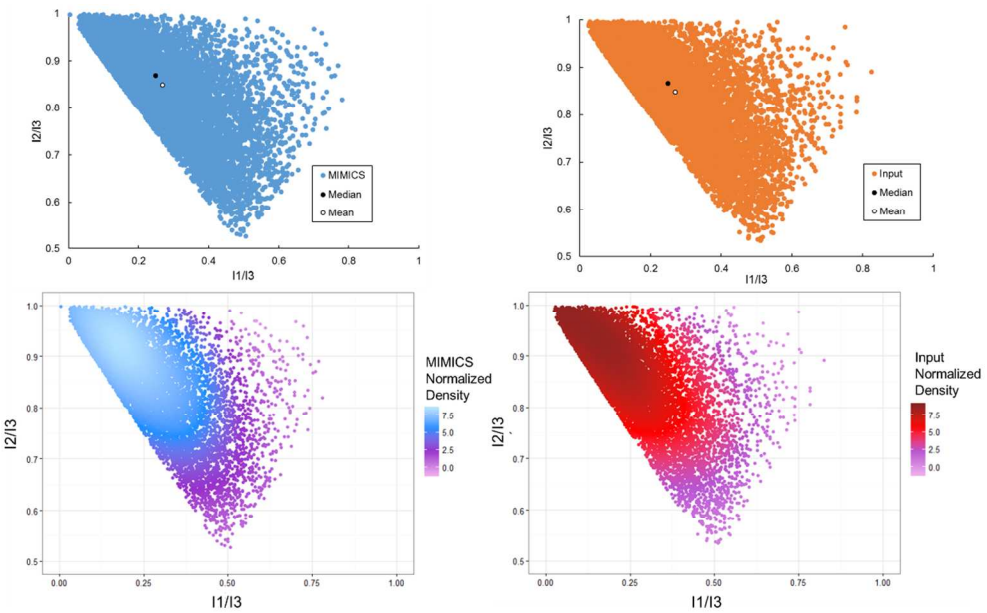


Figure S2. Normalized Principal Moment of Inertia Comparison for Bioactives Library: mean, median, and distribution (Top) are preserved in MIMICS generated sets, along with compound distribution density (Bottom).

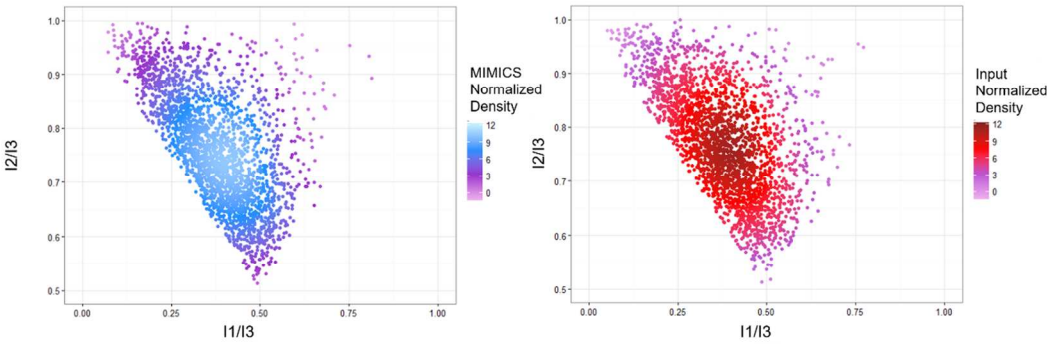


Figure S3. Normalized Principal Moment of Inertia Comparison for an Intermediate-Disk Conformer Library: MIMICS is capable of preserving different kinds of distributions.

UPR Bioactivity Confirmation

Main and control screen cell lines, in this case, HT1080 human fibrosarcoma cells expressing XBP1 or ATF4-luciferase and a CMV-luciferase reporter construct,

respectively were the subject of a primary compound screen. Cells were treated with 300 nM thapsigargin to activate XBP1. Varying concentrations of each molecule were tested in the presence of thapsigargin to determine bioactivity. Inhibitory effect was measured in terms of reduction of observed XBP1 spliced product relative to reduction of CMV control product. Compounds with demonstrated effect were then tested over a series of eight concentrations[5,6]. The results of this assay took the form of dose response curves for each identified compound at each tested concentration.

VEGFR2 Docking and Bioactivity Confirmation

MIMICS molecules were docked against the 1YWN[7] VEGFR-2 crystal structure using Vina Autodock[8]. Grid center was set at (-2.181, 32.597, 19.385) and grid size was set at (20, 29,16). Compounds were ranked by computed binding energy and high ranking compounds were selected for synthesis based on manual synthetic accessibility analysis.

HUVEC cells were obtained and cultured in EGM-2 medium supplemented with 5% fetal bovine serum (FBS) and growth supplements (EGM-2 MV bullet kit) (Lonza, Walkersville, MD) under standard culture conditions (37℃, 95% humidified air and 5% CO2).

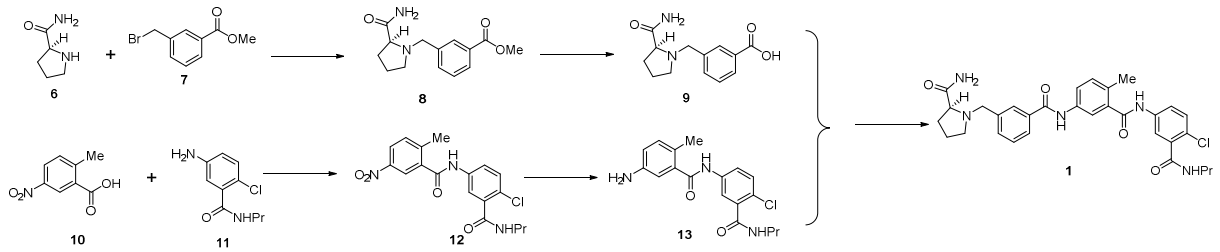
HUVEC[9] Tube formation Assay: HUVECs (20,000 cells per well) were seeded on matrigel-coated 96-well culture plate. To examine the effect of drugs on in vitro tube formation, HUVECs were simultaneously seeded and treated with different drug concentrations and tube formation was observed periodically over time under a phase contrast microscope. The tubular structure formation was imaged and quantified by

counting the number of closed rings formed in each treatment group 6 hours post treatment under 4X objective.

Synthesis of Compounds 1–5

General procedures. All final products were analysed by reverse-phase HPLC, (ZORBAX Eclipse XDB C8 5 μm column, 4.6 × 150 mm; Agilent Technologies) using an Agilent Technologies 1260 Infinity equipped with a diode-array detector. Mobile phases were gradients of 80% acetonitrile/20% H_2O (v/v) in 45 mM ammonium formate at pH 3.5 and 0.8 mL/min. Final compound purity was determined by monitoring at 330 ± 50 nM and was >95%. Melting points were determined on an Electrothermal 2300 Melting Point Apparatus. NMR spectra were obtained on a Bruker Avance 400 spectrometer at 400 MHz for ^{1}H and 100 MHz for ^{13}C spectra. Spectra were obtained in $(CD_3)_2SO$ or $CDCl_3$, and were referenced to Me_4Si. Chemical shifts and coupling constants were recorded in units of ppm and Hz, respectively. Low resolution mass spectra were gathered by direct injection of methanolic solutions into an Agilent 6120 mass spectrometer using an atmospheric pressure chemical ionization (APCI) mode with a fragmentor voltage of 50 V and a drying gas temperature of 250 °C. High resolution mass spectra (HRMS) were measured on an Agilent Technologies 6530 Accurate-Mass Quadrupole Time of Flight (Q-TOF) LC / MS interfaced with an Agilent Jet Stream Electrospray Ionization (ESI) source allowing positive or negative ions detection. Organic solutions were dried over $MgSO_4$ and solvents were evaporated under reduced pressure on a rotary evaporator. Thin-layer chromatography was carried out on aluminium-backed silica gel plates (Merck 60 F_{254}) with visualization of components by UV light (254 nm) or exposure to I_2. Column chromatography was carried out on silica gel (Merck 230–400 mesh). DCM refers to dichloromethane, DIPEA refers to diisopropylethylamine, DMAP refers to 4-dimethylaminopyridine, DMF refers to dimethylformamide, EDCI refers to 1-ethyl-3-(dimethylaminopropyl)carbodiimide, EtOAc refers to ethyl acetate, EtOH refers to ethanol, HATU refers to 2-(7-aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate, MeCN refers to acetonitrile, MeOH refers to methanol, pet. ether refers to petroleum ether boiling fraction 40–60 °C.

(*S*)-1-(3-((3-((4-Chloro-3-(propylcarbamoyl)phenyl)carbamoyl)-4-methylphenyl)carbamoyl)benzyl)pyrrolidine-2-carboxamide (1). (Compound 1081)



Methyl (S)-3-((2-carbamoylpyrrolidin-1-yl)methyl)benzoate (8). Methyl 3-(bromomethyl)benzoate (**7**) (602 mg, 2.63 mmol) was added to a stirred solution of L-prolinamide (**6**) (300 mg, 2.63 mmol) and K_2CO_3 (545 mg, 3.94 mmol) in DMF (5 mL) and the mixture was stirred at 20 °C for 24 h. The resulting mixture was diluted with EtOAc (50 mL) and washed with water (50 mL), then washed with brine (50 mL), dried and the solvent evaporated. The crude residue was triturated with EtOAc/pet. ether to give methyl ester **8** (544 mg, 79%) as a white solid: mp 147–149 °C; ^{1}H NMR ($CDCl_3$) δ 7.94–7.97 (m, 2H), 7.48 (d, *J* = 7.6 Hz, 1H), 7.41 (t, *J* = 7.6 Hz, 1H), 7.16 (br s, 1H), 5.22 (br s, 1H), 3.99 (d, *J* = 13.1 Hz, 1H), 3.93 (s, 3H), 3.54 (d, *J* = 13.1 Hz, 1H), 3.20 (dd, *J* = 10.1 Hz, 1H), 3.00–3.04 (m, 1H), 2.31–2.38 (m, 1H), 2.21–2.29 (m, 1H), 1.91–1.99 (m, 1H), 1.73–1.86 (m, 2H); *m/z* 263.2 (MH^+, 100%).

(S)-3-((2-carbamoylpyrrolidin-1-yl)methyl)benzoic acid (9). 2N NaOH (5 mL) was added to a stirred solution of methyl ester **8** (644 mg, 2.45 mmol) in MeOH (5 mL) and stirred at 20 °C for 17 h. MeOH was removed under reduced pressure, the aqueous layer was saturated with NaCl and the product precipitated from the salt solution. The white solids were isolated by filtration, washed with cold water and dried to give acid **9** (255 mg, 42%) as a white solid: mp 239–241 °C; ^{1}H NMR (DMSO-d_6) δ 13.17 (br s, 1H), 9.75 (br s, 1H), 8.13 (br s, 1H), 7.98–7.99 (m, 2H), 7.75 (d, *J* = 7.2 Hz, 1H), 7.61 (br s, 1H), 7.55 (t, *J* = 7.7 Hz, 1H), 4.40–4.50 (m, 2H), 4.10–4.16 (m, 1H), 3.48 (br s, 1H), 3.29 (obscured br s, 1H), 2.45–2.47 (obscured m, 1H), 2.06–2.08 (m, 1H), 1.79–1.90 (m, 2H); *m/z* 249.1 (MH^+, 100%).

2-Chloro-5-(2-methyl-5-nitrobenzamido)-*N*-propylbenzamide (12). 5-Amino-2-chloro-*N*-propylbenzamide (**11**) (400 mg, 1.88 mmol) was added to a solution of 2-

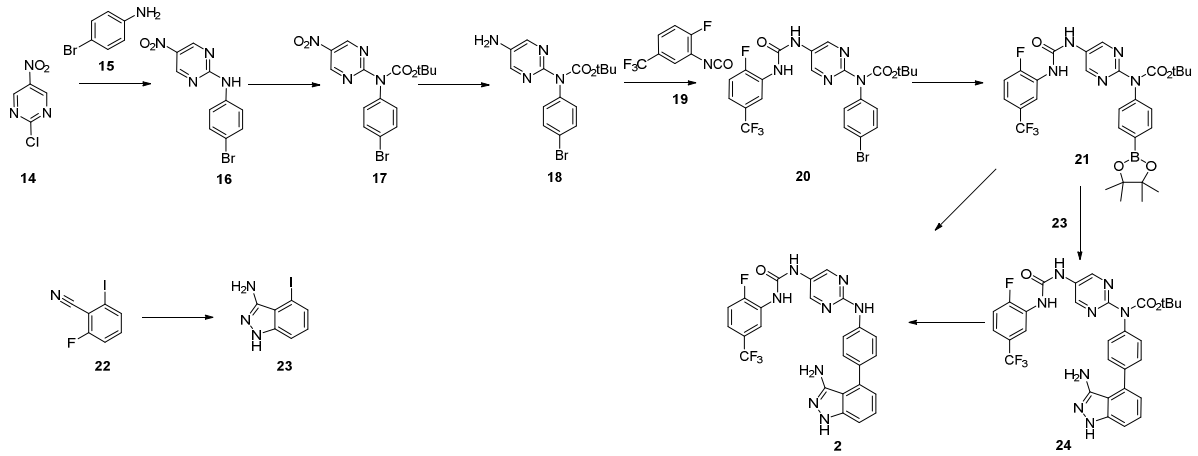
methyl-5-nitrobenzoic acid (**10**) (358 mg, 1.97 mmol), HATU (751 mg, 1.97 mmol) and DIPEA (0.82 mL, 4.70 mmol) in DMF (5 mL). The reaction mixture was stirred at 20 °C for 16 h. The resulting mixture was diluted with EtOAc (50 mL), washed with water (50 mL) and brine (50 mL), dried and the solvent evaporated. The crude residue was purified by column chromatography, eluting with a gradient (30–40%) of EtOAc/pet. ether to give amide **10** (755 mg, quant.) as a white solid: mp 117–120 °C; ^{1}H NMR ($CDCl_3$) δ 8.36 (d, *J* = 2.3 Hz, 1H), 8.23 (dd, *J* = 8.4, 2.3 Hz, 1H), 8.00 (br s, 1H), 7.92 (dd, *J* = 8.7, 2.3 Hz, 1H), 7.81 (d, *J* = 2.3 Hz, 1H), 7.47 (d, *J* = 8.4 Hz, 1H), 7.43 (d, *J* = 8.7 Hz, 1H), 6.32 (br t, *J* = 5.7 Hz, *H), 3.38 (m, 2H), 2.60 (s, 3H), 1.63 (qt, *J* = 7.4, 7.1 Hz, 2H), 0.99 (t, *J* = 7.4 Hz, 3H); *m/z* 376.1 (MH^+, 100%).

2-Chloro-5-(2-methyl-5-aminobenzamido)-*N*-propylbenzamide (13). $SnCl_2.2H_2O$ (2.27 g, 10.0 mmol) was added to a solution of amide **12** (755 mg, 2.01 mmol) in EtOAc (40 mL) and the reaction mixture was heated at reflux temperature for 17 h. The resulting mixture was cooled to 20 °C and was basified with sat. $NaHCO_3$ (100 mL), and was extracted with EtOAc (2 × 100 mL). The combined organic layers were washed with brine (50 mL), dried and the solvent evaporated. The resulting off white residue was triturated with EtOAc/pet. ether, filtered, and dried to give the amide **13** (475 mg, 68%) as a white solid mp 188–190 °C; ^{1}H NMR ($CDCl_3$) δ 7.86 (d, *J* = 8.7 Hz, 1H), 7.76 (d, *J* = 2.4 Hz, 1H), 7.58 (br s, 1H), 7.39 (d, *J* = 8.7 Hz, 1H), 7.05 (d, *J* = 8.1 Hz, 1H), 6.80 (d, *J* = 2.5 Hz, 1H), 6.71 (dd, *J* = 8.1, 2.5 Hz, 1H), 6.32 (br s, 1H), 3.68 (br s, 2H), 3.43 (m, 2H), 2.36 (s, 3H), 1.66 (qt, *J* = 7.6, 7.1 Hz, 2H), 1.00 (t, *J* = 7.6 Hz, 3H); *m/z* 346.2 (MH^+, 100%).

(*S*)-1-(3-((3-((4-Chloro-3-(propylcarbamoyl)phenyl)carbamoyl)-4-methylphenyl)carbamoyl)benzyl)pyrrolidine-2-carboxamide (1). Acid **9** (211 mg, 0.85 mmol), HATU (323 mg, 0.85 mmol) and DIPEA (0.47 mL, 2.70 mmol) were stirred in DMF (5 mL) as a suspension at 20 °C for 30 min. Amine **13** (267 mg, 0.77 mmol) was added to the reaction mixture and the mixture was stirred at 20 °C for 17 h. The resulting mixture was diluted with EtOAc (50 mL), washed with water (50 mL) and brine (50 mL), dried and concentrated *in vacuo*. The crude residue was purified by column chromatography, eluting with 10% MeOH/DCM, to give amide **1** (208 mg, 47%) as an off-white solid: mp 146–149 °C; ^{1}H NMR (DMSO-d_6) δ 10.57 (br s, 1H), 10.32 (br s, 1H),

8.44 (t, *J* = 5.8 Hz, 1H), 7.90–7.91 (m, 2H), 7.76–7.86 (m, 4H), 7.61 (d, *J* = 7.6 Hz, 1H), 7.45–7.50 (m, 2H), 7.31 (d, *J* = 8.5 Hz, 1H), 7.27 (d, *J* = 2.9 Hz, 1H), 7.09 (d, *J* = 2.9 Hz, 1H), 3.92 (d, *J* = 13.3 Hz, 1H), 3.49 (d, *J* = 13.3 Hz, 1H), 3.19 (td, *J* = 7.0, 5.8 Hz, 2H), 2.97–3.00 (m, 1H), 2.86–2.91 (m, 1H), 2.35 (s, 3H), 2.23–2.29 (m, 1H), 2.04–2.12 (m, 1H), 1.66–1.77 (m, 3H), 1.52 (qt, *J* = 7.4, 7.0 Hz, 2H), 0.92 (t, *J* = 7.4 Hz, 3H); ^{13}C NMR (DMSO-d_6) δ 175.8, 167.9, 166.1, 165.5, 139.3, 138.0, 137.6, 136.9 (2), 134.7, 132.0, 130.8, 130.2, 129.8, 128.3, 127.7, 126.3, 123.7, 121.6, 121.1, 119.3, 119.0, 67.1, 58.4, 53.0, 40.7, 29.8, 23.2, 22.2, 18.7, 11.4; *m/z* 576.3 (MH^+, 100%); HRMS calcd for $C_{31}H_{35}ClN_5O_4$ $(MH)^+$ *m/z* 576.2372); found 576.2377. HPLC purity 97.3 %.

1-(2-((4-(3-amino-1*H*-indazol-4-yl)phenyl)amino)pyrimidin-5-yl)-3-(2-fluoro-5-(trifluoromethyl)phenyl)urea (2) (Compound 8683).



***N*-(4-bromophenyl)-5-nitropyrimidin-2-amine (16).** A mixture of 2-chloro-5-nitropyrimidine (**14**) (500 mg, 3.13 mmol) and *p*-bromoaniline (**15**) (533 mg, 3.13 mmol) were stirred with K_2CO_3 (433 mg, 3.13 mmol) in MeCN (20 mL) for 6 h. An additional portion of 2-chloro-5-nitropyrimidine (**14**) (100 mg, 0.63 mmol) was added, and stirred for another 17 h. The resulting mixture was diluted with EtOAc (50 mL), washed with water (50 mL) and brine (50 mL), dried and the solvent evaporated. The crude residue was purified by trituration with EtOH to give amine **16** (640 mg, 69%) as a yellow solid: mp 193–194 °C; ^{1}H NMR (DMSO-d_6) δ 10.95 (s, 1H), 9.25 (s, 2H), 7.74 (ddd, *J* = 8.9, 3.0, 2.0 Hz, 2H), 7.56 (ddd, *J* = 8.9, 3.0, 2.0 Hz, 2H); *m/z* 295.0 (MH^+, 100%).

***tert*-Butyl (4-Bromophenyl)(5-nitropyrimidin-2-yl)carbamate (17).** Di-*tert*-butyl Dicarbonate (340 mg, 1.56 mmol) was added to a solution of amine **16** (460 mg, 1.56

mmol) and DMAP (19 mg, 0.16 mmol) in DCM (10 mL). The resulting mixture was stirred at 20 °C for 4 h before being diluted with DCM (50 mL). The organic layer was washed with water (50 mL) and brine (50 mL), dried and concentrated *in vacuo* to give carbamate **17** (614 mg, quant.) as a yellow solid which was used without further purification: mp 94–96 °C; ^{1}H NMR ($CDCl_3$ MHz) δ 9.29 (s, 2H), 7.58 (ddd, *J* = 8.6, 2.9, 2.0 Hz, 2H), 7.08 (ddd, *J* = 8.6, 2.9, 2.0 Hz, 2H), 1.47 (s, 9H); *m/z* 394.1 ($M\text{-}H^-$, 100%).

***tert*-Butyl (5-Aminopyrimidin-2-yl)(4-bromophenyl)carbamate (18).** Iron powder (74 mg, 1.32 mmol) was added to a stirred solution of carbamate **17** (130 mg, 0.44 mmol) in a mixture of ethanol (4.2 mL) and acetic acid (0.3 mL). The reaction mixture was heated to 80 °C for 21 h. The reaction mixture was cooled to 20 °C, was diluted with EtOAc (20 mL) and was basified with sat. $NaHCO_3$ (20 mL). The resulting mixture was filtered, and the organic layer was dried and concentrated *in vacuo*. The crude residue was purified by column chromatography, eluting with EtOAc, to give aminopyrimidine **18** (105 mg, 88%) as a yellow solid: mp 193–194 °C; ^{1}H NMR ($CDCl_3$) δ 8.17 (s, 2H), 7.43 (ddd, *J* = 8.8, 2.9, 2.1 Hz, 2H), 7.13 (ddd, *J* = 8.8, 2.9, 2.1 Hz, 2H), 3.74 (s, 2H), 1.44 (s, 9H); *m/z* 363.1 ($M\text{-}H^-$, 100%).

***tert*-Butyl (4-Bromophenyl)(5-(3-(2-fluoro-5-(trifluoromethyl)phenyl)ureido)pyrimidin-2-yl)carbamate (20).** 2-Fluoro-5-(trifluoromethyl)phenyl isocyanate (74 µL, 0.51 mmol) was added to a stirred solution of aminopyrimidine **19** (186 mg, 0.51 mmol) in DCM (5 mL). The reaction mixture was stirred at 20 °C for 18 h. The resulting white solid was removed by filtration, washed with DCM and dried under reduced pressure to give urea **20** (250 mg, 86%) as a white solid: mp >305 °C; ^{1}H NMR (DMSO-d_6) δ 9.46 (br s, 2H), 8.89 (s, 2H), 8.53 (dd, *J* = 7.2, 2.1 Hz, 1H), 7.48–7.57 (m, 3H), 7.41–7.48 (m, 1H), 7.10 (ddd, *J* = 8.8, 3.1, 2.1 Hz, 2H), 1.39 (s, 9H); *m/z* 570.1 (MH^+, 100%).

***tert*-Butyl (5-(3-(2-Fluoro-5-(trifluoromethyl)phenyl)ureido)pyrimidin-2-yl)(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)carbamate (21)** $PdCl_2$(dppf) (8.4 mg, 0.011 mmol) was added to a purged suspension of bromide **20** (131 mg, 0.23 mmol), bis(pinacolato)diboron (70 mg, 0.27 mmol) and KOAc (56 mg, 0.57 mmol) in dioxane (3 mL). The reaction mixture was further purged with N_2 for 5 min, sealed and heated at 100 °C for 5 h. The mixture was cooled to 20 °C and was diluted with EtOAc

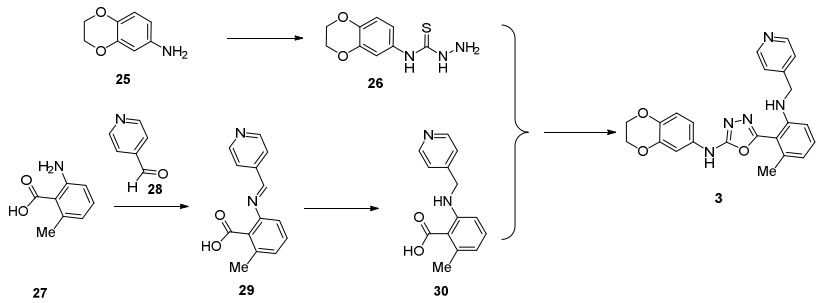
(50 mL), washed with water (50 mL) and brine (50 mL), dried and the solvent evaporated. The crude residue was purified by column chromatography, eluting with a gradient (40–60%) of EtOAc/pet. ether, to give boronate ester **21** (111 mg, 79%) as a pale yellow solid: mp 119–122 °C; ^{1}H NMR (DMSO-d_6) δ 9.54 (br s, 1H), 9.18 (br s, 1H), 8.90 (s, 2H), 8.54 (dd, J = 7.2, 2.1 Hz, 1H), 7.63 (dd, J = 8.5, 2.0, 1.8 Hz, 2H), 7.50–7.55 (m, 1H), 7.43–7.47 (m, 1H), 7.11 (dd, J = 8.5, 2.0, 1.8 Hz, 2H), 1.39 (s, 9H), 1.29 (s, 12H); *m/z* 616.2 ($M\text{-}H^-$, 100%).

4-Iodo-1*H*-indazol-3-amine (23). Hydrazine hydrate (1.01 g, 20.2 mmol) was added to a suspension of 2-fluoro-6-iodobenzonitrile (**22**) (500 mg, 2.02 mmol) and $NaHCO_3$ (255 mg, 3.04 mmol) in EtOH (10 mL) and the resulting mixture was heated at 80 °C for 17 h. The mixture was cooled to 20 °C and diluted with EtOAc (50 mL), washed with water (50 mL) and brine (50 mL), dried and concentrated *in vacuo*. The crude residue was purified by column chromatography, eluting with 75% EtOAc/pet. ether, to give amine **23** (319 mg, 61%) as a pale brown solid: mp 152–153 °C (lit. mp 152-154 °C); ^{1}H NMR (DMSO-d_6, 400 MHz) δ 11.78 (br s, 1H), 7.34 (dd, J = 7.2, 0.5 Hz, 1H), 7.28 (dd, J = 8.3, 0.5 Hz, 1H), 6.93 (dd, J = 8.3, 7.2 Hz, 1H), 5.04 (s, 2H); *m/z* 260.1 (MH^+, 100%).

***tert*-Butyl (4-(3-Amino-1*H*-indazol-4-yl)phenyl)(5-(3-(2-fluoro-5-(trifluoromethyl) phenyl)ureido)pyrimidin-2-yl)carbamate (24).** Boronate ester **21** (449 mg, 0.73 mmol) and iodide **23** (188 mg, 0.73 mmol) were dissolved in a mixture of toluene/EtOH (1:1, 4 mL) and the solution was purged with N_2. Sodium carbonate (43 mg, 2.75 mmol) was dissolved in water (1 mL) and was added to the reaction mixture. $PdCl_2$(dppf) (5.4 mg, 0.07 mmol) was added to the mixture and the mixture was purged with N_2 for further 5 min. The reaction vessel was sealed and heated to 95 °C for 18 h. The reaction mixture was cooled to 20 °C and was diluted with EtOAc (50 mL), washed with water (50 mL) and brine (50 mL), dried and concentrated *in vacuo*. The crude residue was purified by column chromatography, eluting with a gradient (2–10%) of MeOH/DCM, to give carbamate **24** (214 mg, 47%) as a white solid: mp 318–321 °C; ^{1}H NMR (DMSO-d_6) δ 11.78 (br s, 1H), 9.71 (br s, 1H), 9.34 (br s, 1H), 8.94 (s, 2H), 8.54 (dd, J = 7.3, 2.1 Hz, 1H), 7.50–7.54 (m, 1H), 7.43–7.46 (m, 3H), 7.26–7.29 (m, 4H), 6.81 (m, 1H), 4.31 (br s, 2H), 1.42 (s, 9H); *m/z* 621.2 ($M\text{-}H^-$, 100%).

1-(2-((4-(3-Amino-1*H*-indazol-4-yl)phenyl)amino)pyrimidin-5-yl)-3-(2-fluoro-5-(trifluoromethyl)phenyl)urea (2). 2M HCl (6 mL) was added to a solution of carbamate **24** (92 mg, 0.15 mmol) in EtOH (6 mL) and the mixture was heated to 80 °C for 23 h. The mixture was cooled to 20 °C and the EtOH was removed under reduced pressure. The aqueous layer was basified with sat. $NaHCO_3$ solution and the product was extracted with EtOAc (50 mL). The organic layer was washed with brine (50 mL), dried and the solvent evaporated. The crude product was triturated with DCM to give urea **2** (61 mg, 79%) as an off-white solid: mp 212–214 °C; ^{1}H NMR (DMSO-d_6) δ 11.69 (br s, 1H), 9.79 (br s, 1H), 9.13 (br s, 2H), 8.63 (s, 2H), 8.58 (dd, *J* = 7.3, 2.2 Hz, 1H), 7.90 (ddd, *J* = 8.7, 2.6, 1.9 Hz, 2H), 7.48–7.53 (m, 1H), 7.38–7.43 (m, 3H), 7.22–7.29 (m, 2H), 6.79 (dd, *J* = 6.30, 1.5 Hz, 1H), 4.34 (s, 2H); ^{13}C NMR (DMSO-d_6) δ 156.2, 153.6 (d, *J* = 247.8 Hz), 152.6, 149.8 (2), 148.2, 142.1, 140.3, 135.7, 131.7, 129.0 (2), 128.6 (d, *J* = 1.1 Hz), 126.2, 126.0, 125.6, 123.9 (q, *J* = 270.1), 119.5, 119.0, 117.9 (2), 116.9, 116.2 (d, *J* = 20.6 Hz), 110.6, 108.4; *m/z* 523.2 (MH^+, 100%); HRMS calcd for $C_{25}H_{19}F_4N_8O$ $(MH)^+$ *m/z* 523.1613; found 523.1619. HPLC purity 95.2 %.

***N*-(2,3-Dihydrobenzo[*b*][1,4]dioxin-6-yl)-5-(2-methyl-6-((pyridin-4-ylmethyl)amino)phenyl)-1,3,4-oxadiazol-2-amine (3) (Compound 284).**



***N*-(2,3-Dihydrobenzo[*b*][1,4]dioxin-6-yl)hydrazinecarbothioamide (26).** CS_2 (0.69 mL, 11.45 mmol) was added to a stirred mixture of 2,3-dihydrobenzo[*b*][1,4]dioxin-6-amine (**25**) (1.73 g, 11.45 mmol) and NaOH (550 mg, 13.7 mmol) in dry DMF (15 mL) and the mixture was stirred at 20 °C for 1 h. Hydrazine hydrate (1.1 mL, 34.4 mmol) was added and the mixture was stirred at 70 °C for 1 h. The reaction mixture was cooled to 20 °C, diluted with ice/water (150 mL) and stirred vigorously for 1 h. The resulting precipitate was filtered and air-dried to give the carbothioamide **26** (1.15 g, 45%) as a

grey solid: mp 127–129 °C; ^{1}H NMR (DMSO-d_6) δ 9.43 (br s, 1H), 8.99 (br s, 1H), 7.24 (br s, 1H), 6.93 (br d, J = 8.6 Hz, 1H), 6.76 (d, J = 8.6 Hz, 1H), 4.71 (br s, 2H), 4.18–4.25 (m, 4H); *m/z* 226.2 (MH^+, 100%).

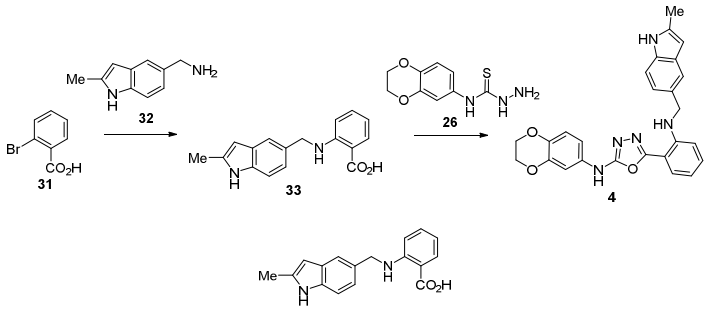
2-Methyl-6-((pyridin-4-ylmethylene)amino)benzoic Acid (29). 2-Amino-6-methylbenzoic acid (**27**) (350 mg, 2.31 mmol) and 4-pyridinecarboxaldehyde (**28**) (165 mg, 1.54 mmol) were stirred in MeOH (5 mL) at 20 °C for 23 h. The resulting white precipitate was removed by filtration, washed with MeOH (2 mL) to give imine **29** (257 mg, 46%) as an off white solid which was used without further purification: mp 159–162 °C; ^{1}H NMR (DMSO-d_6) δ 8.70–8.71 (m, 2H), 7.84 (d, J = 1.4 Hz, 1H), 7.61 (dd, J = 4.6, 1.4 Hz, 2H), 7.33 (t, J = 7.7 Hz, 1H), 6.82 (d, J = 7.7 Hz, 1H), 6.74 (d, J = 7.7 Hz, 1H), 6.24 (d, J = 1.6 Hz, 1H), 2.54 (s, 3H); *m/z* 241.2 (MH^+, 100%).

2-Methyl-6-((pyridin-4-ylmethyl)amino)benzoic Acid (30). $NaBH(OAc)_3$ (266 mg, 1.25 mmol) was added to a solution of imine **29** (201 mg, 0.84 mmol) in acetic acid (5 mL) and the reaction mixture was stirred at 20 °C for 18 h. The mixture was diluted with brine (20 mL) and extracted with EtOAc (3 × 20 mL). The combined organic fractions were dried and concentrated *in vacuo*. The crude residue was triturated with EtOAc to give acid **30** (180 mg, 89%) as an off-white solid: mp 161–163 °C; ^{1}H NMR (DMSO-d_6) δ 13.05 (br s, 1H), 8.48 (d, J = 5.6 Hz, 2H), 7.30 (d, J = 5.6 Hz, 2H), 7.03 (t, J = 7.8 Hz, 1H), 6.44 (d, J = 7.8 Hz, 1H), 6.31 (d, J = 7.8 Hz, 1H), 4.44 (s, 2H), 2.35 (s, 3H), NH not observed; *m/z* 243.2 (MH^+, 100%).

***N*-(2,3-Dihydrobenzo[*b*][1,4]dioxin-6-yl)-5-(2-methyl-6-((pyridin-4-ylmethyl)amino)phenyl)-1,3,4-oxadiazol-2-amine (3).** A mixture of carbothioamide **26** (154 mg, 0.69 mmol), acid **30** (166 mg, 0.69 mmol) and EDCI (394 mg, 2.06 mmol) were stirred in DCM (5 mL) at 20 °C for 4 h. The resulting mixture was diluted with DCM (50 mL), washed with sat. $NaHCO_3$ (50 mL) and then brine (50 mL). The organic layer was dried and concentrated *in vacuo*. The crude residue was purified by column chromatography, eluting with EtOAc, to give amine **3** (24 mg, 8%) as an off-white solid: mp 203–206 °C; ^{1}H NMR (DMSO-d_6) δ 10.33 (br s, 1H), 8.49 (d, J = 5.9 Hz, 2H), 7.30–7.35 (m, 3H), 7.25 (d, J = 2.6 Hz, 1H), 7.09 (t, J = 7.9 Hz, 1H), 7.02 (dd, J = 8.8, 2.6 Hz, 1H), 6.84 (d, J = 8.8 Hz, 1H), 6.55 (d, J = 7.9 Hz, 1H), 6.40 (d, J = 7.9 Hz, 1H), 4.48 (d, J = 6.1 Hz, 2H), 4.20–4.25 (m, 4H), 2.34 (s, 3H); ^{13}C NMR (DMSO-d_6) δ 159.6, 156.5,

149.6 (2), 149.1, 147.1, 143.4, 138.4, 138.3, 132.6, 131.2, 122.0 (2), 118.5, 117.2, 110.4, 108.9, 107.3, 106.2, 64.3, 63.9, 45.2, 21.5; *m/z* 416.1 (MH^+, 100%); HRMS calcd for $C_{23}H_{22}N_5O_3$ $(MH)^+$ *m/z* 416.1717; found 416.1721; HPLC purity 97.7%.

_N_-(2,3-Dihydrobenzo[_b_][1,4]dioxin-6-yl)-5-(2-(((2-methyl-1_H_-indol-5-yl)methyl)amino)phenyl)-1,3,4-oxadiazol-2-amine (4) (Compound 217).

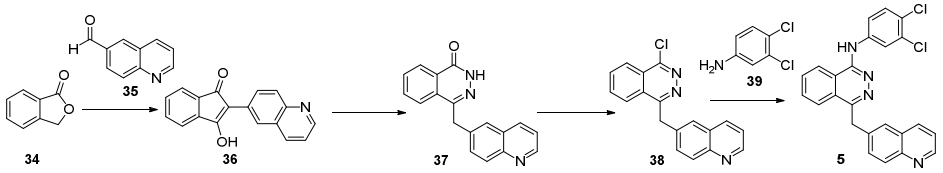


2-(((2-Methyl-1_H_-indol-5-yl)methyl)amino)benzoic Acid (33). A mixture of 2-bromobenzoic acid (**31**) (228 mg, 1.13 mmol), 1-(2-methyl-1*H*-indol-5-yl)methylamine (**32**) (200 mg, 1.25 mmol), $K_3PO_4 \cdot H_2O$ (784 mg, 3.40 mmol) and CuI (22 mg, 0.11 mmol) in ethylene glycol (127 µL, 2.27 mmol) and *n*-butanol (4 mL) was heated at 100 °C in a sealed tube for 4 h. The mixture was cooled to 20 °C and was diluted with 1N HCl (50 mL) and extracted with EtOAc (3 × 50 mL). The organic layer was dried and concentrated *in vacuo*. The crude residue was purified by column chromatography, eluting with a gradient (20–50%) of EtOAc/pet. ether, to give acid **33** (133 mg, 42%) as an off-white solid: mp 137–140 °C; ^{1}H NMR (DMSO-d_6) δ 10.86 (br s, 1H), 7.78 (dd, *J* = 7.9, 1.4 Hz, 1H), 7.36 (s, 1H), 7.28 (td, *J* = 7.9, 1.4 Hz, 1H), 7.22 (d, *J* = 8.4 Hz, 1H), 6.99 (dd, *J* = 7.9, 1.4 Hz, 1H), 6.74 (d, *J* = 8.4 Hz, 1H), 6.52 (td, *J* = 7.9, 1.4 Hz, 1H), 6.07 (m, 1H), 4.41 (s, 2H), 2.36 (s, 3H), 2 NH resonances not observed; *m/z* 279.2 ($M-H^-$, 100%).

_N_-(2,3-Dihydrobenzo[_b_][1,4]dioxin-6-yl)-5-(2-(((2-methyl-1_H_-indol-5-yl)methyl)amino)phenyl)-1,3,4-oxadiazol-2-amine (4). Carbothioamide **26** (133 mg, 0.59 mmol) was added to a solution of acid **33** (166 mg, 0.59 mmol) and EDCI (125 mg, 0.65 mmol) in DCM (15 mL) and the resulting mixture was stirred for 30 min. An additional portion of EDCI (249 mg, 1.30 mmol) was added to the reaction mixture and this was allowed to stir at 20 °C for 21 h. The resulting mixture was diluted with DCM

(50 mL), washed with water (50 mL) and brine (50 mL), dried and the solvent evaporated. The crude residue was purified by column chromatography, eluting with a gradient (30–100%) of EtOAc/pet. ether, to give amine **4** (28 mg, 10%) as an off-white solid: mp (EtOAc/pet. ether) 174–177 °C; ^{1}H NMR (DMSO-d_6) δ 10.89 (br s, 1H), 10.44 (br s, 1H), 7.75 (t, J = 5.2 Hz, 1H), 7.58 (dd, J = 7.9, 1.5 Hz, 1H), 7.41 (s, 1H), 7.29 (td, J = 7.9, 1.5 Hz, 1H), 7.24 (d, J = 8.3 Hz, 1H), 7.19 (d, J = 2.6 Hz, 1H), 7.03 (dd, J = 7.9, 1.5 Hz, 1H), 7.00 (dd, J = 8.7, 2.6 Hz, 1H), 6.89 (d, J = 8.3 Hz, 1H), 6.83 (d, J = 8.7 Hz, 1H), 6.71 (t, J = 7.9 Hz, 1H), 6.07 (s, 1H), 4.52 (d, J = 5.2 Hz, 2H), 4.19–4.24 (m, 4H), 2.36 (s, 3H); ^{13}C NMR (DMSO-d_6) δ 158.4, 158.1, 146.3, 143.4, 138.5, 136.0, 135.4, 132.4, 131.7, 128.8(2), 126.7, 119.8, 117.8, 117.2, 115.2, 111.4, 110.6, 110.4, 106.2, 105.3, 99.0, 64.3, 63.9, 47.1, 13.4; *m/z* 454.2 (MH^+, 100%); HRMS calcd for $C_{26}H_{24}N_5O_3$ $(MH)^+$ *m/z* 454.1874; found 454.1879; HPLC purity 99.4%.

***N*-(3,4-Dichlorophenyl)-4-(quinolin-6-ylmethyl)phthalazin-1-amine (5) (Compound 2113).**



3-Hydroxy-2-(quinolin-6-yl)-1*H*-inden-1-one (36). A solution of NaOMe in dry MeOH (prepared by dissolving sodium metal (1.03 g, 44.7 mmol) in MeOH (50 mL)) was added to a stirred solution of phthalide (**34**) (2.0 g, 14.9 mmol) and quinoline-6-carbaldehyde (**35**) (2.34 g, 14.9 mmol) in dry MeOH (100 mL) at 0 °C. The mixture was stirred at 20 °C for 15 min and then heated at reflux temperature for 3 h. The mixture was cooled and the solvent removed by evaporation. The residue was diluted with water (100 mL) and then extracted with Et_2O (3 × 50 mL). The aqueous fraction was acidified with HOAc and the mixture stirred at 5 °C for 1 h. The resulting suspension was filtered, washed with water and air-dried to give indenone **36** (2.93 g, 72%) as a red powder: ^{1}H NMR (DMSO-d_6, 400 MHz) δ 13.00 (br s, 1H), 8.35 (br d, J = 6.7 Hz, 1H), 7.87–7.97 (m, 2H), 7.65–7.80 (m, 3H), 7.48–7.55 (m, 1H), 7.42 (m, 1H), 7.27–7.36 (m, 2H); *m/z* 274.2 (MH^+, 100%).

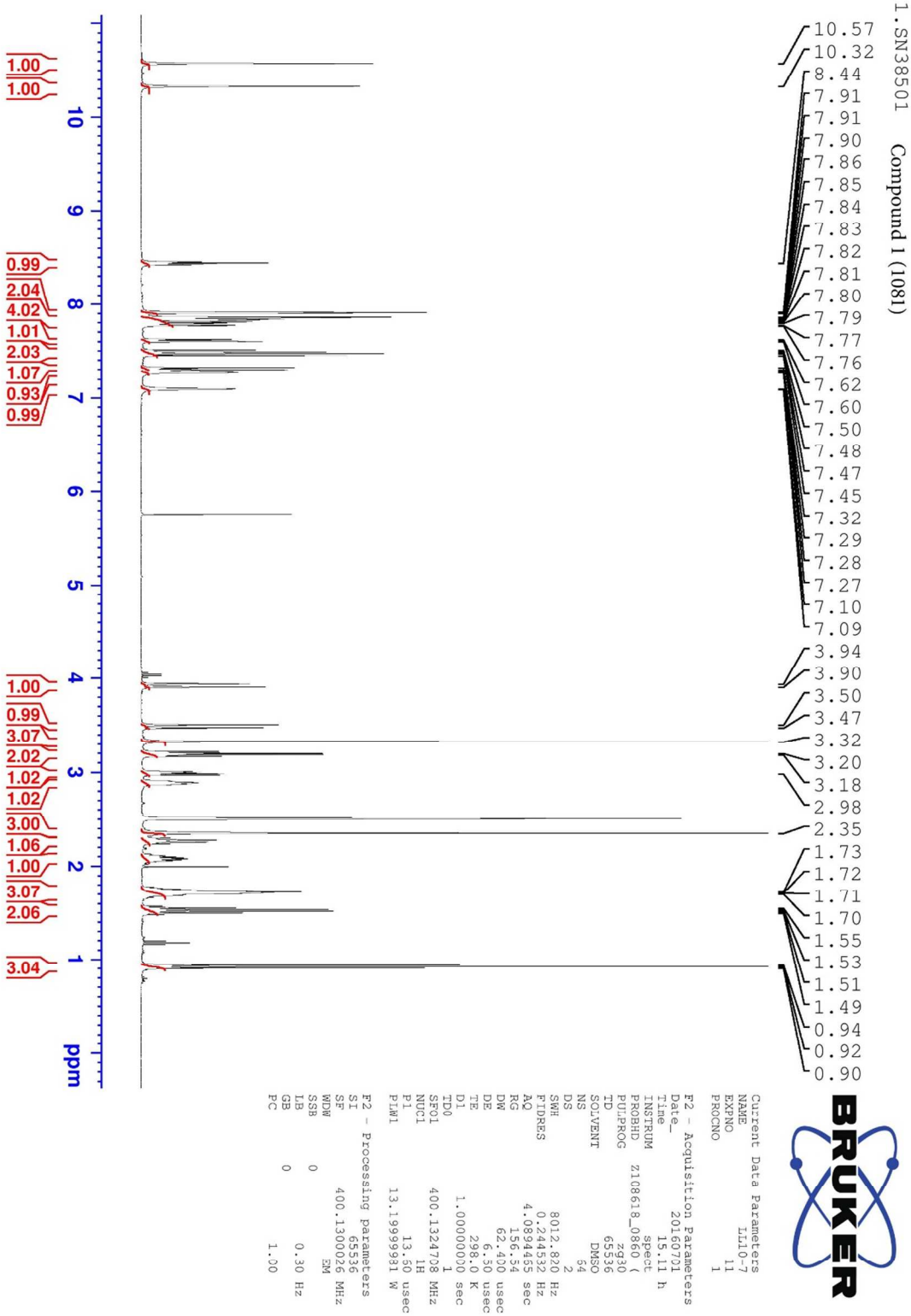
4-(Quinolin-6-ylmethyl)phthalazin-1(2*H*)-one (37). A solution of indenone **36** (2.93 g, in hydrazine hydrate (10 mL) was stirred at 110 °C for 16 h. The mixture was stood at 5 °C for 4 h and the precipitate filtered and washed with cold EtOH (3 mL) and air-dried. The material was purified by column chromatography, eluting with 3% MeOH/DCM, to give the phthalazinone **37** (1.27 g, 41%) as a tan powder: mp (MeOH) 236–238 °C; ^{1}H NMR (DMSO-d_6) δ 12.62 (br s, 1H), 8.84 (dd, J = 4.2, 1.7 Hz, 1H), 8.23–8.27 (m, 2H), 7.93–8.00 (m, 2H), 7.78–7.88 (m, 3H), 7.73 (dd, J = 8.7, 2.0 Hz, 1H), 7.48 (dd, J = 8.3, 4.2 Hz, 1H), 4.52 (s, 2H); m/z 292.1 (MH^+, 100%); Anal. calcd for $C_{18}H_{13}N_3O$·¼CH_3OH: C, 74.22; H, 4.78; N, 14.23%. Found: C, 74.18; H, 4.68; N, 14.36%.

1-Chloro-4-(quinolin-6-ylmethyl)phthalazine (38). $POCl_3$ (1 mL) was added drop wise to a stirred solution of phthalazinone **37** (326 mg, 1.13 mmol) in MeCN (5 mL). The resulting mixture was stirred at 20 °C for 16 h, and a second portion of $POCl_3$ (2 mL) was added to the reaction mixture. After 48 hours, the reaction mixture was poured on ice, basified with sat. $NaHCO_3$, and extracted with EtOAc (3 × 50 mL). The crude residue was purified by column chromatography, eluting with EtOAc, to give chlorophthalazine **38** (267 mg, 77%) as a white solid: mp 193–196 °C; ^{1}H NMR (DMSO-d_6) δ 8.84 (dd, J = 4.2, 1.8 Hz, 1H), 8.42–8.44 (m, 1H), 8.31–8.33 (m, 1H), 8.26–8.28 (m, 1H), 8.08–8.15 (m, 2H), 7.95 (d, J = 8.7 Hz, 1H), 7.87 (d, J = 1.8 Hz, 1H), 7.77 (dd, J = 8.5, 1.8 Hz, 1H), 7.48 (dd, J = 8.5, 4.2 Hz, 1H), 4.94 (s, 2H); *m/z* 306.1 (MH^+, 100%).

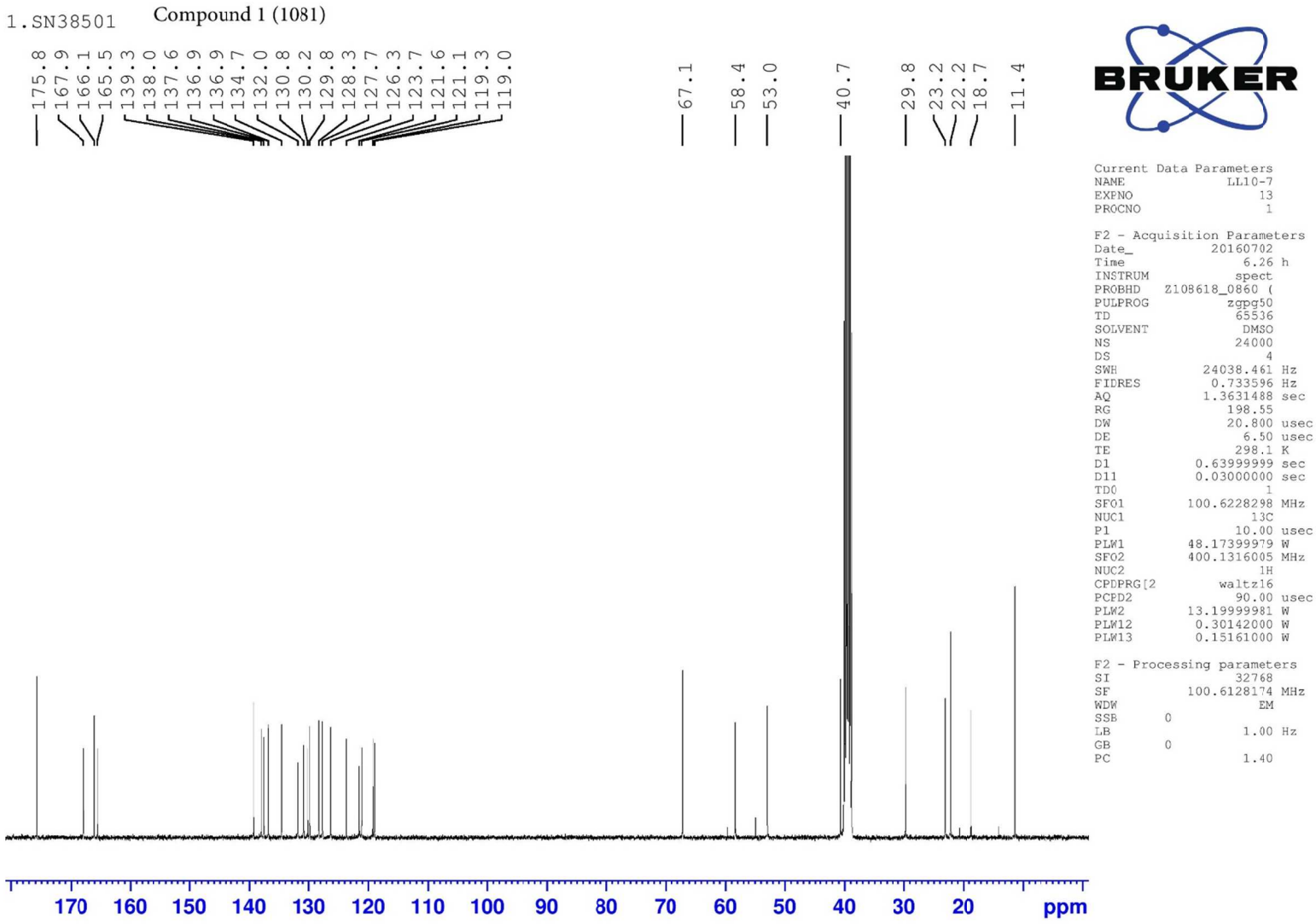
***N*-(3,4-Dichlorophenyl)-4-(quinolin-6-ylmethyl)phthalazin-1-amine (5).** A mixture of chlorophthalazine **38** (90 mg, 0.29 mmol) and 3,4-dichloroaniline (**39**) in *i*PrOH was heated to 90 °C in a sealed tube for 22 h. The mixture was cooled to 20 °C and the pale orange solid was removed by filtration and washed with EtOH. The orange solid partitioned between (EtAOc (10 mL) and 0.1 M HCl (80 mL). The aqueous layer was removed, concentrated under reduced pressure and basified with ammonium hydroxide solution (1 mL). This solution was then diluted with water (20 mL) and extracted with EtOAc (2 × 20 mL). The combined organic fraction was dried and concentrated *in vacuo*. The residue was triturated with EtOAc/pet. ether to give amine **5** (52 mg, 41%) as an off-white solid: mp 199–202 °C; ^{1}H NMR (DMSO-d_6) δ 9.42 (s, 1H), 8.82 (dd, J = 4.2, 1.0 Hz, 1H), 8.59 (d, J = 7.9 Hz, 1H), 8.50 (d, J = 2.5 Hz, 1H), 8.27 (d, J = 8.4, 1.0

Hz, 1H), 8.22 (d, J = 8.1, 0.8 Hz, 1H), 7.90–8.00 (m, 4H), 7.88 (d, J = 1.0 Hz, 1H), 7.77 (dd, J = 8.7, 2.0 Hz, 1H), 7.60 (d, J = 8.8 Hz, 1H), 7.47 (dd, J = 8.4, 4.2 Hz, 1H), 4.79 (s, 2H); ^{13}C NMR (DMSO-d_6) δ 153.0, 151.7, 150.1, 146.6, 141.0, 137.5, 135.6, 132.4, 131.8, 130.8, 130.6, 130.2, 129.0, 127.9, 126.9, 126.1, 125.1, 122.8, 122.7, 121.5, 121.2, 120.2, 118.8, 38.4; *m/z* 431.1 (MH^+, 100%); HRMS calcd for $C_{24}H_{17}Cl_2N_4$ $(MH)^+$ *m/z* 431.0825; found 431.0832; HPLC purity 94.9%.

NMR Spectra of Compounds 1-5

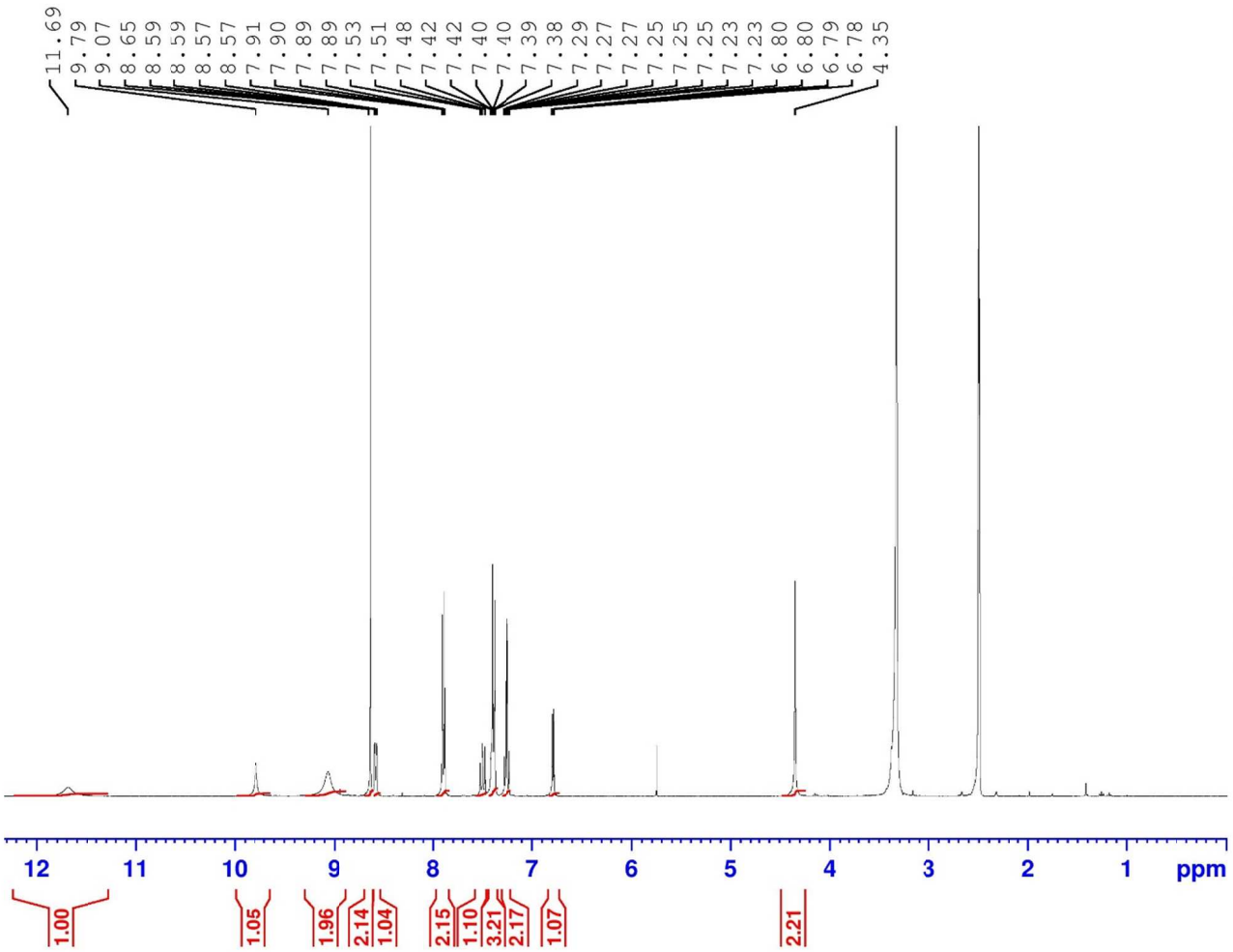


Spectra 1: 1H NMR Spectra of Compound 1



Spectra 2: 13C NMR Spectra of Compound 1

2.SN38502 Compound 2 (8683)



BRUKER

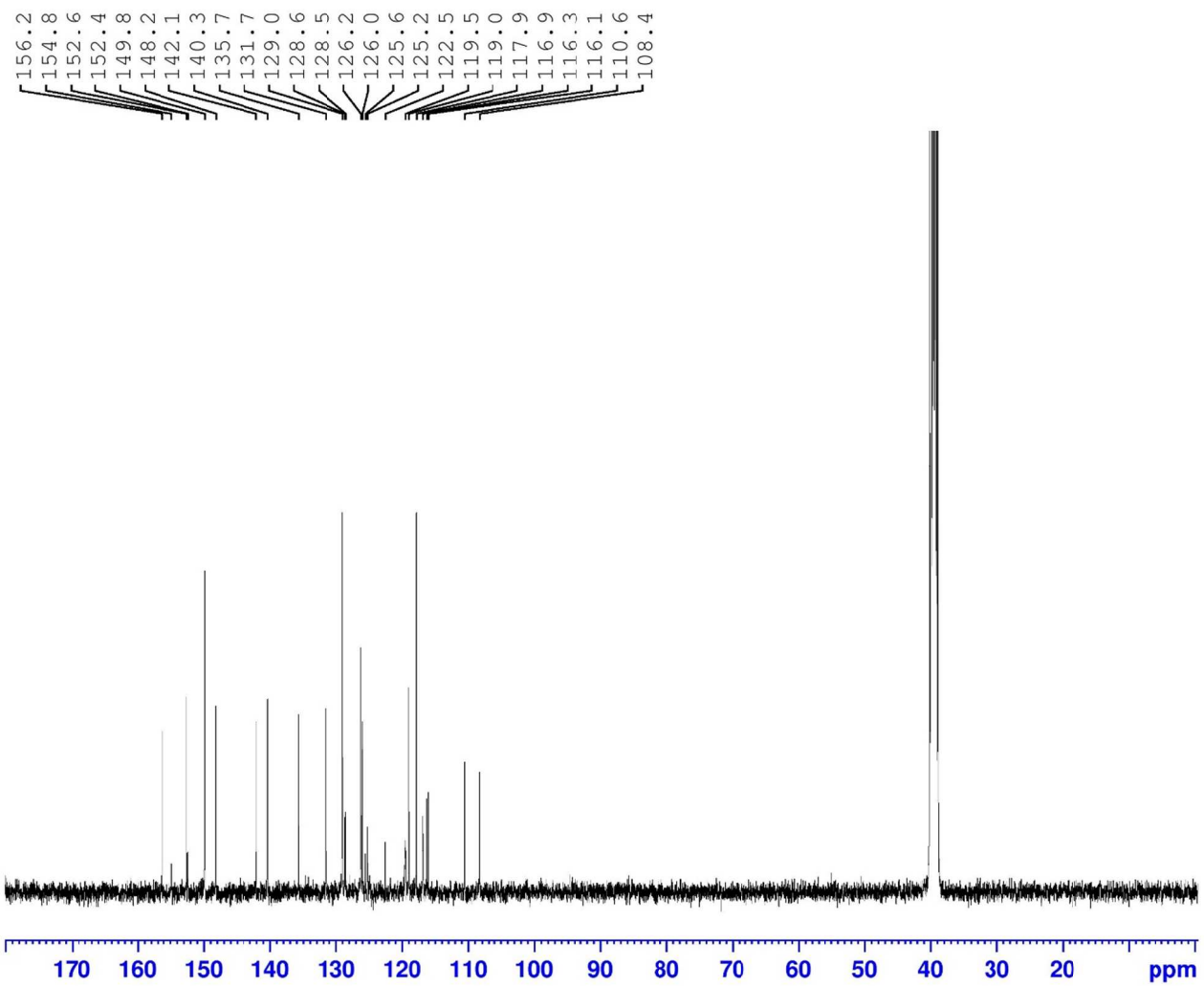
```
Current Data Parameters
NAME                  LL10-25
EXPNO                      21
PROCNO                      1

F2 - Acquisition Parameters
Date_                20160812
Time                    18.01 h
INSTRUM                  spect
PROBHD    Z108618_0860 (
PULPROG                  zg30
TD                      65536
SOLVENT                  DMSO
NS                         64
DS                          2
SWH                   8012.820 Hz
FIDRES                0.244532 Hz
AQ                    4.0894465 sec
RG                       198.55
DW                       62.400 usec
DE                        6.50 usec
TE                       298.0 K
D1                 1.00000000 sec
TD0                         1
SF01               400.1324708 MHz
NUC1                       1H
P1                       13.60 usec
PLW1               13.19999981 W

F2 - Processing parameters
SI                      65536
SF                 400.1300025 MHz
WDW                        EM
SSB          0
LB                       0.30 Hz
GB           0
PC                       1.00
```

Spectra 3: 1H NMR Spectra of Compound 2

2.SN38502 Compound 2 (8683)

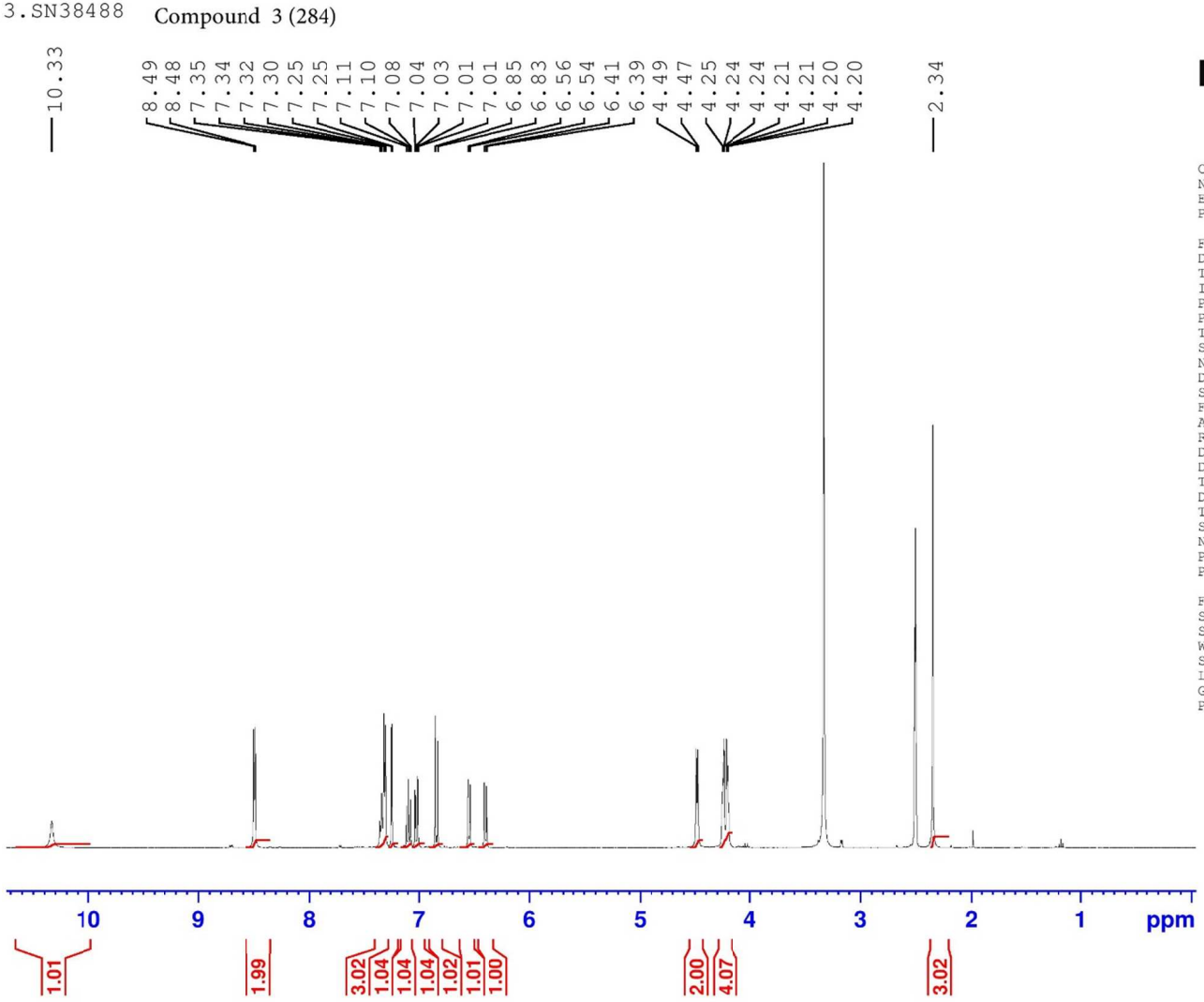



```
Current Data Parameters
NAME                  LL10-25
EXPNO                      22
PROCNO                      1

F2 - Acquisition Parameters
Date_                20160813
Time                     5.31 h
INSTRUM                  spect
PROBHD    Z108618_0860 (
PULPROG                 zgpg50
TD                       65536
SOLVENT                   DMSO
NS                       20000
DS                           4
SWH                24038.461 Hz
FIDRES              0.733596 Hz
AQ                  1.3631488 sec
RG                     198.55
DW                     20.800 usec
DE                      6.50 usec
TE                     298.0 K
D1                0.63999999 sec
D11               0.03000000 sec
TD0                          1
SF01             100.6228298 MHz
NUC1                      13C
P1                     10.00 usec
PLW1             48.17399979 W
SF02             400.1316005 MHz
NUC2                      1H
CPDPRG[2              waltz16
PCPD2                  90.00 usec
PLW2             13.19999981 W
PLW12             0.30142000 W
PLW13             0.15161000 W

F2 - Processing parameters
SI                       32768
SF               100.6128164 MHz
WDW                         EM
SSB          0
LB                      1.00 Hz
GB           0
PC                      1.40
```

Spectra 4: 13C NMR Spectra of Compound 2



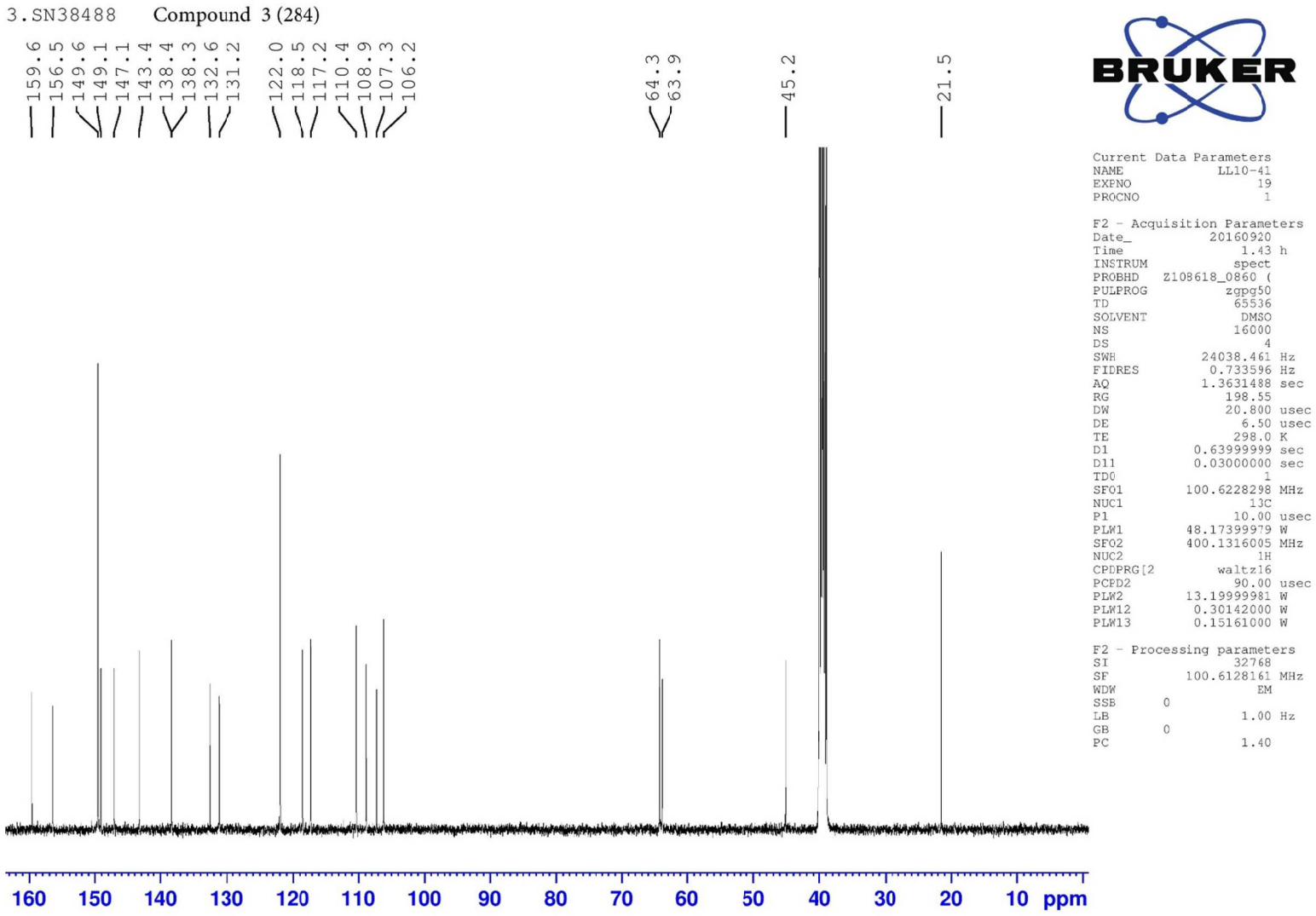
BRUKER

```
Current Data Parameters
NAME                  LL10-41
EXPNO                      18
PROCNO                      1

F2 - Acquisition Parameters
Date_               20160919
Time                   16.31 h
INSTRUM                 spect
PROBHD     Z108618_0860 (
PULPROG                 zg30
TD                     65536
SOLVENT                 DMSO
NS                         64
DS                          2
SWH                 8012.820 Hz
FIDRES              0.244532 Hz
AQ                  4.0894465 sec
RG                     198.55
DW                     62.400 usec
DE                      6.50 usec
TE                     298.0 K
D1                  1.00000000 sec
TD0                         1
SF01             400.1324708 MHz
NUC1                       1H
P1                     13.60 usec
PLW1            13.19999981 W

F2 - Processing parameters
SI                     65536
SF               400.1300026 MHz
WDW                        EM
SSB          0
LB                      0.30 Hz
GB           0
PC                      1.00
```

Spectra 5: 1H NMR Spectra of Compound 3



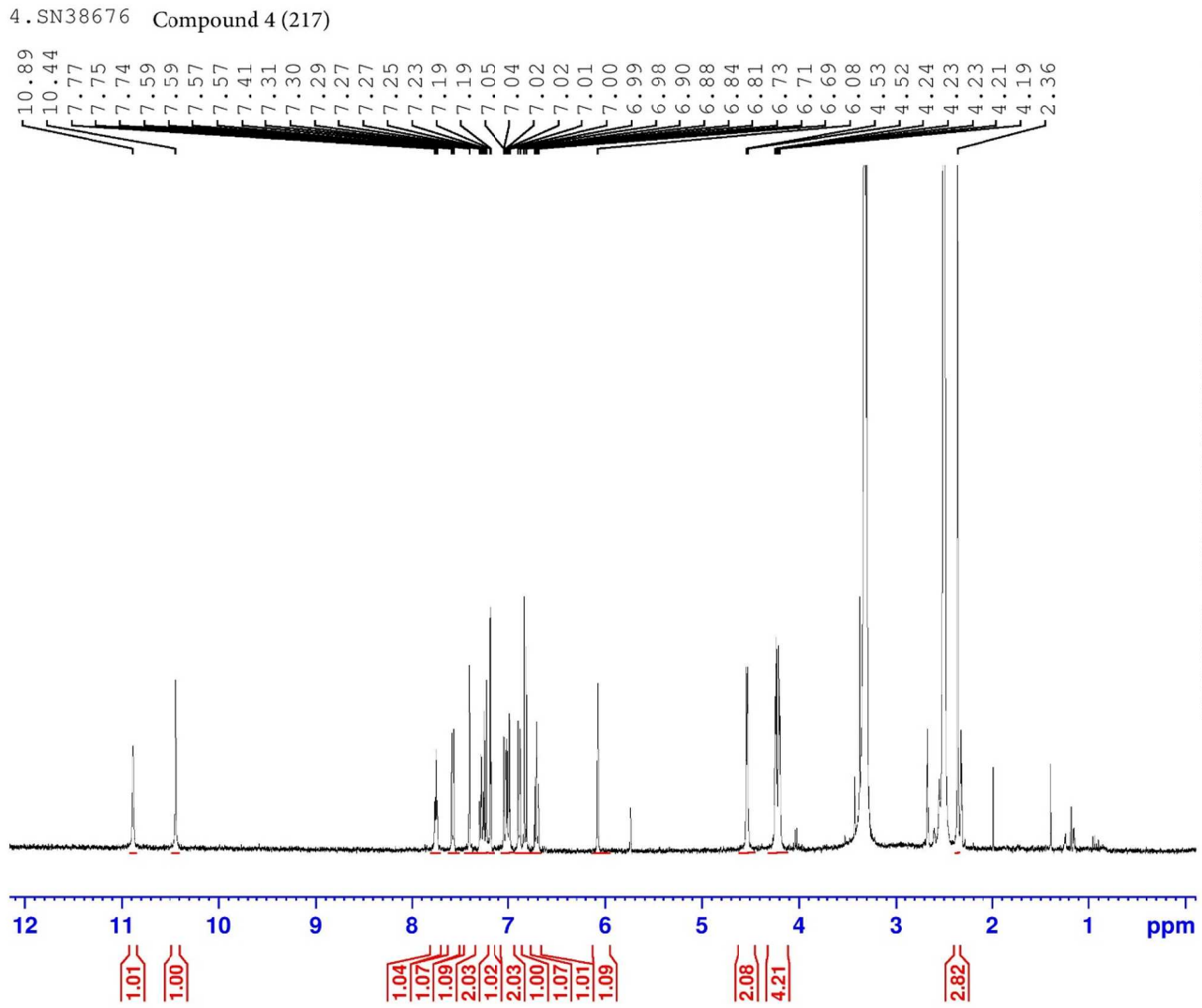
BRUKER

Current Data Parameters
NAME LL10-41
EXPNO 19
PROCNO 1

F2 - Acquisition Parameters
Date_ 20160920
Time 1.43 h
INSTRUM spect
PROBHD Z108618_0860 (
PULPROG zgpg50
TD 65536
SOLVENT DMSO
NS 16000
DS 4
SWH 24038.461 Hz
FIDRES 0.733596 Hz
AQ 1.3631488 sec
RG 198.55
DW 20.800 usec
DE 6.50 usec
TE 298.0 K
D1 0.63999999 sec
D11 0.03000000 sec
TD0 1
SF01 100.6228298 MHz
NUC1 13C
P1 10.00 usec
PLW1 48.17399979 W
SF02 400.1316005 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 90.00 usec
PLW2 13.19999981 W
PLW12 0.30142000 W
PLW13 0.15161000 W

F2 - Processing parameters
SI 32768
SF 100.6128161 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40

Spectra 6: 13C NMR Spectra of Compound 3



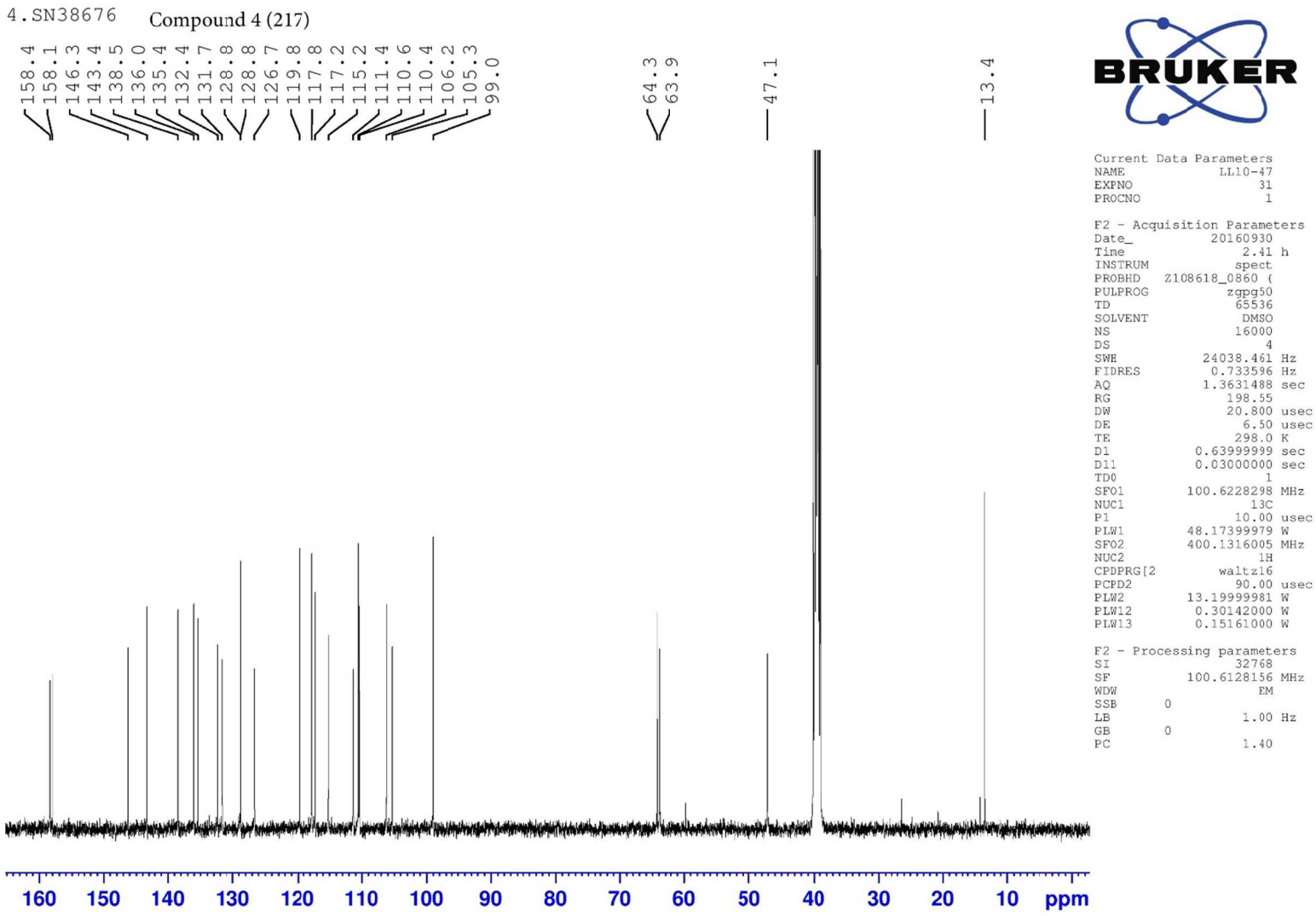
BRUKER

Current Data Parameters
NAME LL10-47
EXPNO 5
PROCNO 1

F2 - Acquisition Parameters
Date_ 20160929
Time 12.37 h
INSTRUM spect
PROBHD Z108618_0860 (
PULPROG zg30
TD 65536
SOLVENT DMSO
NS 256
DS 2
SWH 8012.820 Hz
FIDRES 0.244532 Hz
AQ 4.0894465 sec
RG 198.55
DW 62.400 usec
DE 6.50 usec
TE 298.0 K
D1 1.00000000 sec
TD0 1
SFO1 400.1324708 MHz
NUC1 1H
P1 13.60 usec
PLW1 13.19999981 W

F2 - Processing parameters
SI 65536
SF 400.1300027 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

Spectra 7: 1H NMR Spectra of Compound 4



BRUKER

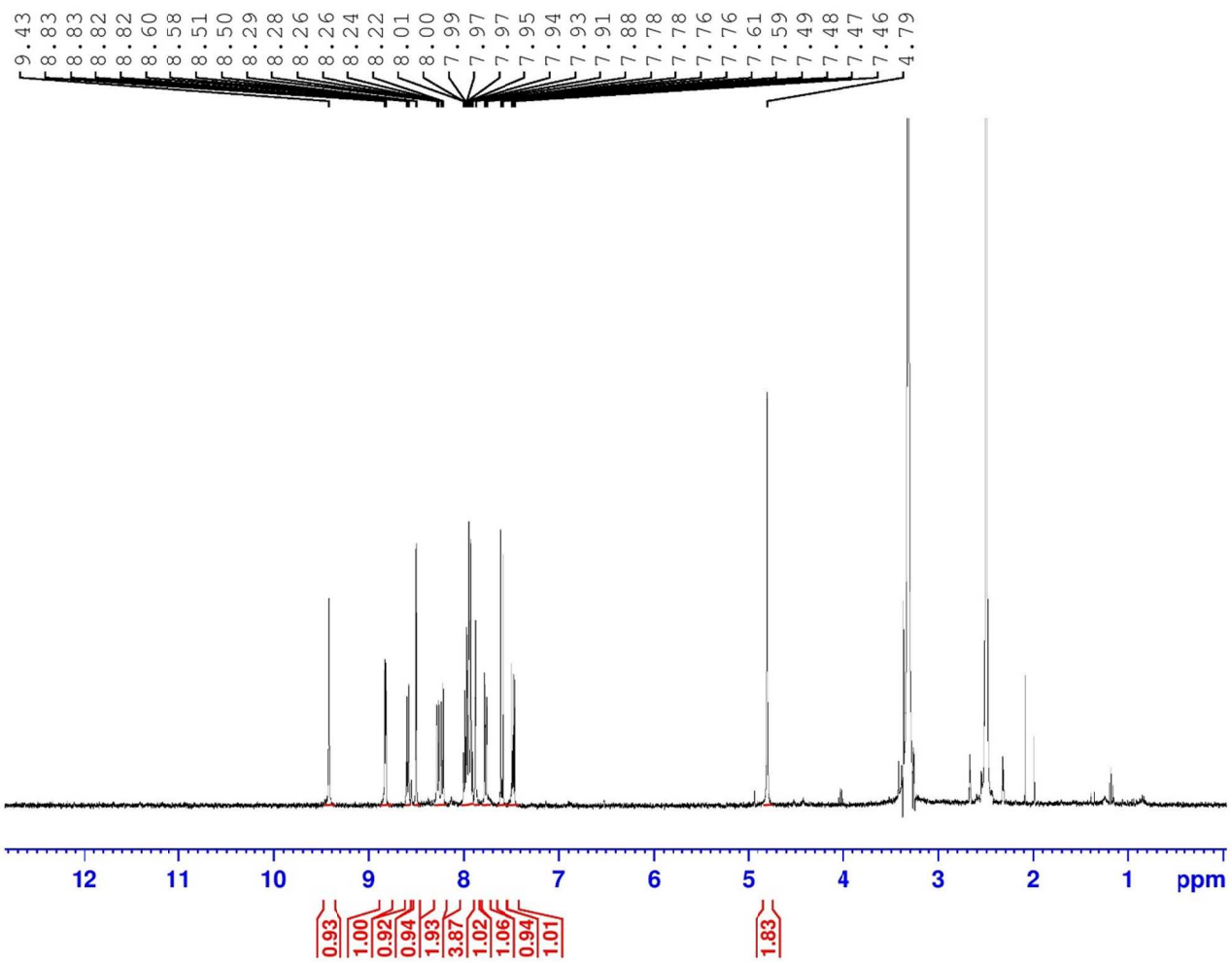
```
Current Data Parameters
NAME                  LL10-47
EXPNO                      31
PROCNO                      1

F2 - Acquisition Parameters
Date_               20160930
Time                    2.41 h
INSTRUM                spect
PROBHD    Z108618_0860 (
PULPROG                zgpg50
TD                      65536
SOLVENT                  DMSO
NS                      16000
DS                          4
SWH                24038.461 Hz
FIDRES              0.733596 Hz
AQ                  1.3631488 sec
RG                     198.55
DW                     20.800 usec
DE                      6.50 usec
TE                     298.0 K
D1                0.63999999 sec
D11               0.03000000 sec
TD0                         1
SF01             100.6228298 MHz
NUC1                      13C
P1                     10.00 usec
PLW1             48.17399979 W
SF02             400.1316005 MHz
NUC2                      1H
CPDPRG[2              waltz16
PCPD2                  90.00 usec
PLW2             13.19999981 W
PLW12             0.30142000 W
PLW13             0.15161000 W

F2 - Processing parameters
SI                      32768
SF               100.6128156 MHz
WDW                        EM
SSB          0
LB                      1.00 Hz
GB           0
PC                      1.40
```

Spectra 8: 13C NMR Spectra of Compound 4

5.SN38677 Compound 5 (2113)



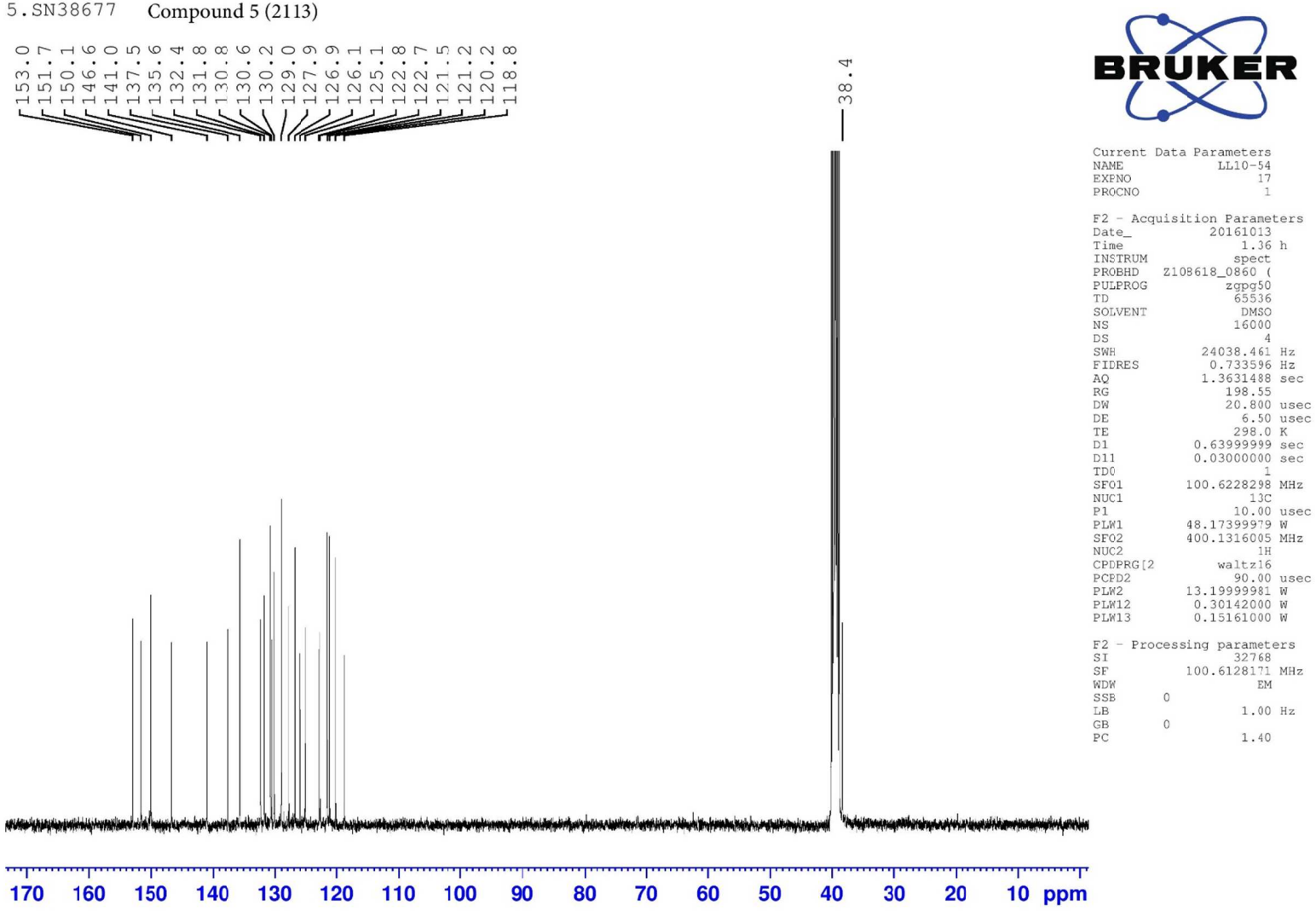
BRUKER

```
Current Data Parameters
NAME                  LL10-54
EXPNO                      11
PROCNO                      1

F2 - Acquisition Parameters
Date_               20161011
Time                    14.25 h
INSTRUM                 spect
PROBHD    Z108618_0860 (
PULPROG                 zg30
TD                     65536
SOLVENT                 DMSO
NS                        64
DS                         2
SWH                  8012.820 Hz
FIDRES               0.244532 Hz
AQ                   4.0894465 sec
RG                     198.55
DW                     62.400 usec
DE                      6.50 usec
TE                     298.0 K
D1                1.00000000 sec
TD0                        1
SF01           400.1324708 MHz
NUC1                      1H
P1                     13.60 usec
PLW1           13.19999981 W

F2 - Processing parameters
SI                     65536
SF             400.1300027 MHz
WDW                       EM
SSB       0
LB                     0.30 Hz
GB        0
PC                     1.00
```

Spectra 9: 1H NMR Spectra of Compound 5



BRUKER

Current Data Parameters
NAME LL10-54
EXPNO 17
PROCNO 1

F2 - Acquisition Parameters
Date_ 20161013
Time 1.36 h
INSTRUM spect
PROBHD Z108618_0860 (
PULPROG zgpg50
TD 65536
SOLVENT DMSO
NS 16000
DS 4
SWH 24038.461 Hz
FIDRES 0.733596 Hz
AQ 1.3631488 sec
RG 198.55
DW 20.800 usec
DE 6.50 usec
TE 298.0 K
D1 0.63999999 sec
D11 0.03000000 sec
TD0 1
SF01 100.6228298 MHz
NUC1 13C
P1 10.00 usec
PLW1 48.17399979 W
SF02 400.1316005 MHz
NUC2 1H
CPDPRG[2 waltz16
PCPD2 90.00 usec
PLW2 13.19999981 W
PLW12 0.30142000 W
PLW13 0.15161000 W

F2 - Processing parameters
SI 32768
SF 100.6128171 MHz
WDW EM
SSB 0
LB 1.00 Hz
GB 0
PC 1.40

Spectra 10: 13C NMR Spectra of Compound 5

Supplementary Figures

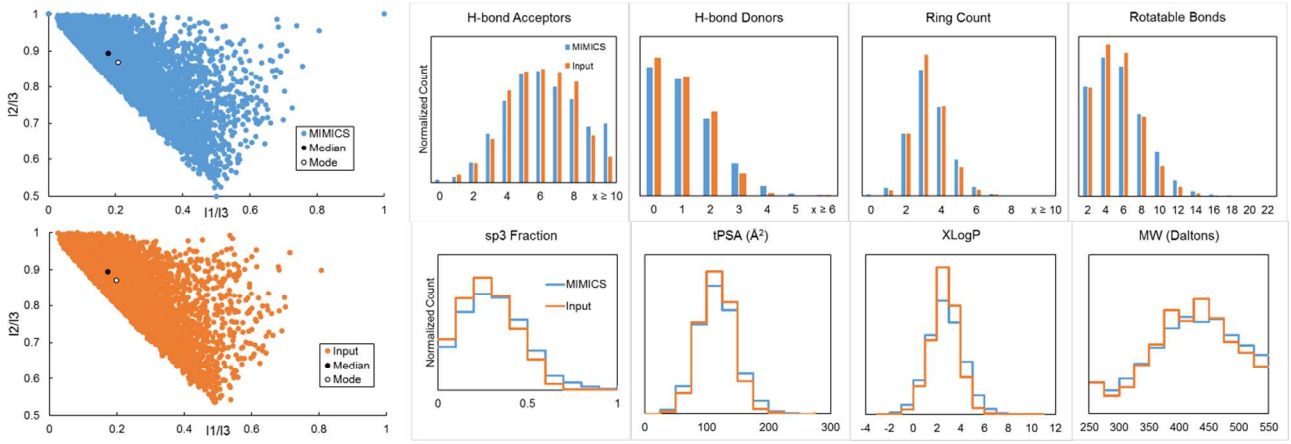


Figure S4. Comparison of MIMICS VEGFR2 inhibitors with input VEGFR2 inhibitors. MIMICS (blue) and input molecules (red) are compared structurally and descriptively.

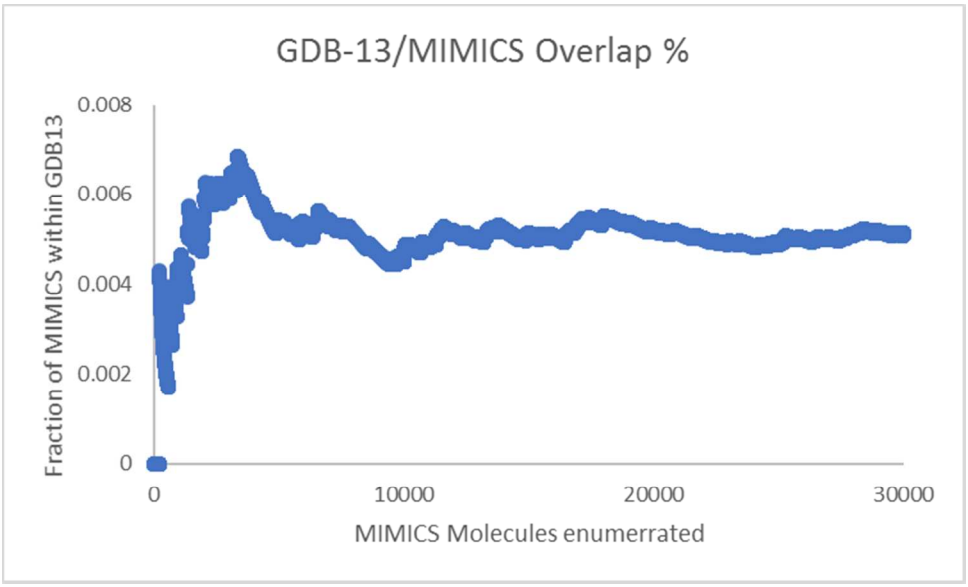
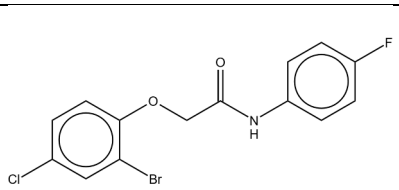
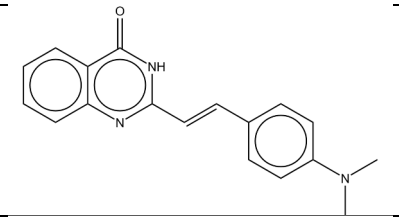
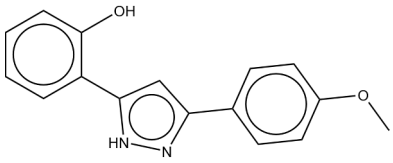
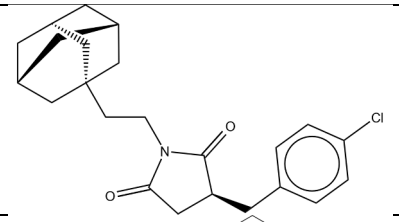
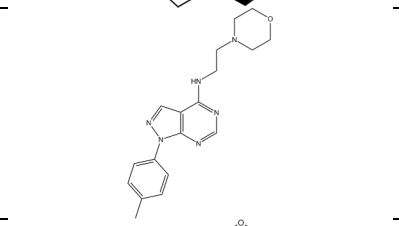
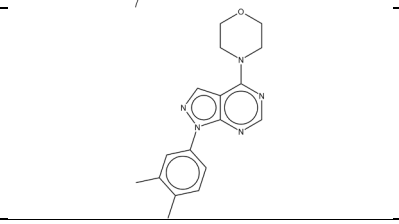


Figure S5. Percent of Bioactive MIMICS molecules (Figures 1, 2) contained within GDB13[11] as a function of MIMICS molecules enumerated.

Table S1. SMILES, name, and structure of MIMICS generated, newly identified UPR inhibitors. MIMICS only interacted with SMILES strings, which were converted into chemical structure after generation. None of the identified compounds were present in MIMICS input.

SMILES	Name	Structure
Fc1ccc(NC(=O)COc2ccc(Cl)cc2Br)cc1	STF-19945	
CN(C)c1ccc(\C=C\c2nc3ccccc3c(=O)[nH]2)cc1	STF-035627	
COc1ccc(cc1)-c1cc([nH]n1)-c1ccccc1O	STF-060280	
Clc1ccc(C[C@H]2CC(=O)N(CC[C@]34C[C@H]5C[C@H](C[C@H](C5)C3)C4)C2=O)cc1	STF-92416	
CC1=CC=C(C=C1)N1N=CC2=C1N=CN=C2NCCN1CCOCC1	STF-123094	
Cc1ccc(cc1C)-n1ncc2c(ncnc12)N1CCOCC1	STF-123066	

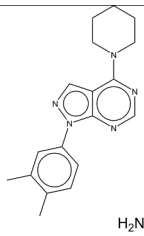
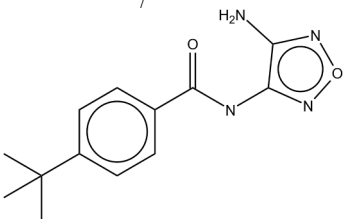
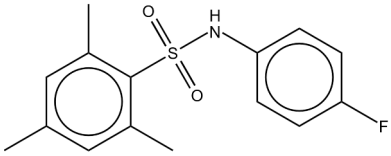
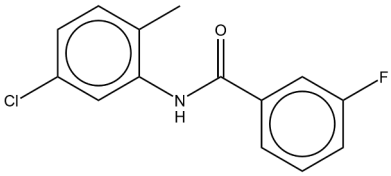
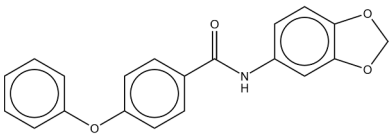
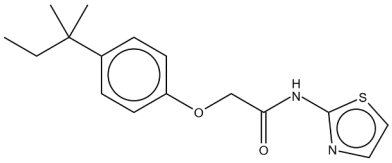
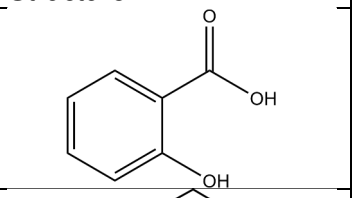
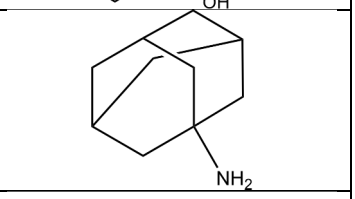
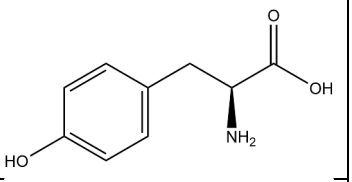
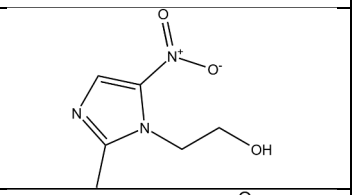
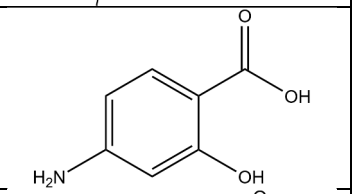
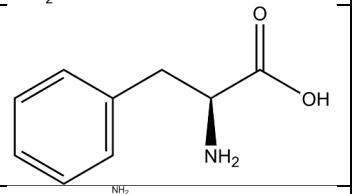
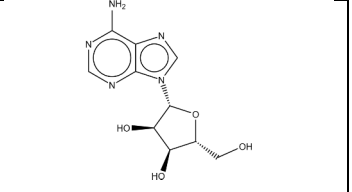
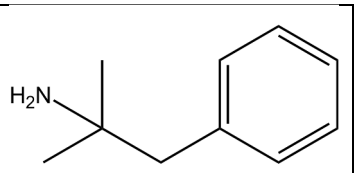
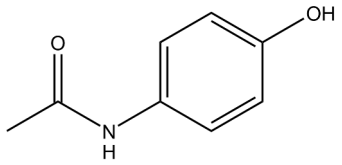
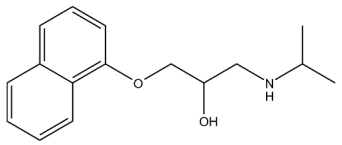
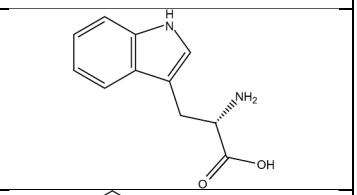
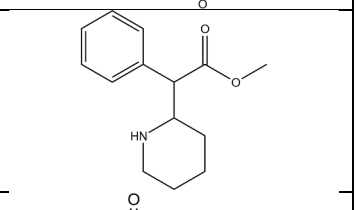
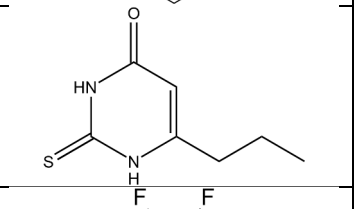
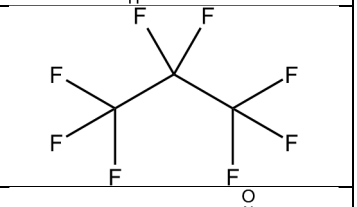
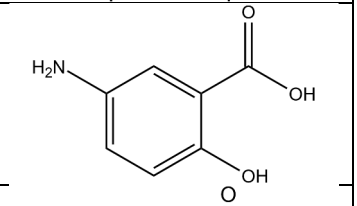
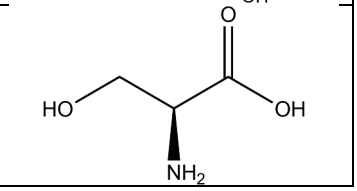
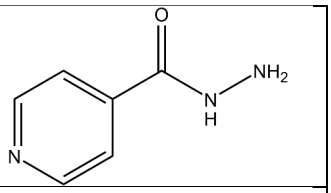
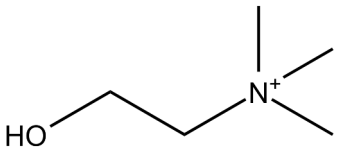
Cc1ccc(cc1C)-n1ncc2c(ncnc12)N1CCCCC1	STF-123067	
CC(C)(C)c1ccc(cc1)C(=O)Nc1nonc1N	STF-092146	
Cc1cc(C)c(c(C)c1)S(=O)(=O)Nc1ccc(F)cc1	STF-047078	
Cc1ccc(Cl)cc1NC(=O)c1cccc(F)c1	STF-046304	
O=C(Nc1ccc2OCOc2c1)c1ccc(Oc2ccccc2)cc1	STF-021898	
O=C(Nc1ccc2OCOc2c1)c1ccc(Oc2ccccc2)cc1	STF-018832	

Table S2. SMILES, name, and structure of MIMICS generated, FDA approved compounds. These compounds were present in the MIMICS generated bioactive library.

SMILES	Name	Structure
OC(=O)C1=CC=CC=C1O	Salicylic acid	
NC12CC3CC(CC(C3)C1)C2	Amantadine	
N[C@@H](CC1=CC=C(O)C=C1)C(O)=O	L-Tyrosine	
CCCCCCCCCCCCOCCOCCOCCOCCOCCOCCOCC OCCOCCOCCO	Polidocanol	
CC1=NC=C(N1CCO)[N+]([O-])=O	Metronidazole	
NC1=CC=C(C(O)=O)C(O)=C1	Aminosalicylic Acid	
N[C@@H](CC1=CC=CC=C1)C(O)=O	L-Phenylalanine	
Nc1ncnc2n(cnc12)[C@@H]1O[C@H](CO)[C@@H](O)[C@H]1O	Adenosine	

CC(C)(N)CC1=CC=CC=C1	Phentermine	
CC(=O)NC1=CC=C(O)C=C1	Acetaminophen	
CC(C)NCC(O)COC1=C2C=CC=CC2=CC=C1	Propranolol	
N[C@@H](CC1=CNC2=C1C=CC=C2)C(O)=O	L-Tryptophan	
COC(=O)C(C1CCCCN1)C1=CC=CC=C1	Methylphenidate	
CCCC1=CC(=O)NC(=S)N1	Propylthiouracil	
FC(F)(F)C(F)(F)C(F)(F)F	Perflutren	
NC1=CC=C(O)C(=C1)C(O)=O	Mesalazine	
N[C@@H](CO)C(O)=O	L-Serine	

NNC(=O)C1=CC=NC=C1	Isoniazid	
C[N+](C)(C)CCO	Choline	

References

(1) Bemis Murcko clustering computed using JKlustor version 15.10.5.0 (http://www.chemaxon.com)

(2) O'boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. **2011**, *J. Cheminf.*, 3–33.

(3) Sauer W. H. B.; Schwarz M. K. Molecular shape diversity of combinatorial libraries: a prerequisite for broad bioactivity. **2003***, J. Chem. Inf. Comput. Sci.*, 43, 987–1003

(4) Irwin, John J.; Brian K. Shoichet. "ZINC – A Free Database of Commercially Available Compounds for Virtual Screening. **2005,** *J. Chem. Inf. Model*, 45., 177-82.

(5) Spiotto, M. T.; Banh, A.; Papandreou, I.; Cao, H.; Galvez, M. G.; Gurtner, G. C.; Denko, N. C.; Le, Q. T.; Koong, A. C. Imaging the unfolded protein response in primary tumors reveals microenvironments with metabolic variations that predict tumor growth., **2009**, *Cancer Res.*, 70(1), 78–88.

(6) Jiang, D.; Niwa, M.; Koong, A. C. Targeting the IRE1–XBP1 branch of the unfolded protein response in human diseases, **2015**. *Semin. Cancer Biol.*, 48–56.

(7) Miyazaki, Y.; Matsunaga, S.; Tang, J.; Maeda, Y.; Nakano, M.; Philippe, R. J.; Shibahara, M.; Liu, W.; Sato, H.; Wang, L.; Nolte, R. T.Novel 4-amino-furo[2,3-d]pyrimidines as Tie-2 and VEGFR2 Dual Inhibitors **2005**, *Bioorg. Med. Chem. Lett.* 15, 2203– 2207

(8) Trott, O.; Olson, A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading., **2010,** *J. Comput. Chem.*, 31 (2010) 455-461

(9) Ko, J. M.; Lung, M. L. In vitro Human Umbilical Vein Endothelial Cells (HUVEC) Tube-formation Assay., **2012,** *Bio-protoc.* 2(18): e260.

(10) Kruger, A.W.; Rozema, M.J.; Chu-Kung, A.; Gandarilla, J.; Haight, A.R.; Kotecki, B.J.; Richter, S.M.; Schwartz, A.M.; Wang, Z. The Discovery and Development of a Safe, Practical Synthesis of ABT-869. *Org. Process Res. Develop.* **2009**, *13*, 1419-1425.

(11) Blum L. C.; Reymond J.-L. 970 Million Druglike Small Molecules for Virtual Screening in the Chemical Universe Database GDB-13. *J. Am. Chem. Soc.*, **2009**, 131, 8732-8733.