SpOTTing and designing promiscuous ligands for drug discovery

P. Schneider, M. Röthlisberger, D. Reker and G. Schneider

Electronic Supplementary Information

Chemical Communications

Experimental

Data collection. We extracted all compounds and activity data for machine learning from ChEMBL19 (www.ebi.ac.uk/chembl). Target IDs were manually assigned to target classes. For flagging of potential false-positives we used a list of 106 substructures (inSili.com LLC, Zurich, Switzerland) and counted cumulative flags for each compound.

Neural network model. We trained a feedforward network using own software, as described previously.\textsuperscript{4} The nonlinear network function contained weight vectors $w$ and $v$, and the neurons' bias values $\theta$ and $\odot$:

\[
\text{Prediction score } f(x) = \text{sigm} \left( \sum_{i=1}^{HID} w_i \text{sigm} \left( \sum_{j=1}^{IN} v_{ij} x_j + v_i \right) + \theta \right).
\]

where $\text{sigm}$ is the neuron activation function, and $x$ the input values ($IN = 210$; CATS2 descriptor)\textsuperscript{11}. The number of hidden neurons $HID$ was varied (Table 1). The model was optimized with a (1,500) evolution strategy and adaptive stepsize adjustment.\textsuperscript{1}

Self-organizing map. We used the MOLMAP software (inSili.com LLC, Zürich, Switzerland) for data projection onto a toroidal self-organizing map containing 20×18 clusters, with $2\times10^6$ update cycles, and the Gaussian neighbourhood kernel with linearly decaying update radius ($t_{\text{initial}} = 10$).\textsuperscript{2}

Synthesis and analytics. Building blocks and solvents were purchased from Sigma-Aldrich (www.sigmaaldrich.com) and used without further purification. Proton and carbon nuclear magnetic resonance ($^1$H and $^{13}$C NMR) spectra were recorded on a Bruker Avance 400 (400 and 100 MHz, respectively). Analytical LC-MS was carried out in a Hitachi LaChrom Ultra–Advion CMS system, equipped with a Nucleodur C$_{18}$ H Tec column, under a 5-$50\%$ gradient of acetonitrile: $\text{H}_2\text{O}$ (pH 3) in each solvent), and a total flow rate of 0.5 mL/min. Preparative HPLC was carried out on a Shimadzu LC-8A system, coupled to a Nucleodur 100-5 C$_{18}$ H Tec column and a SPD-20A UV/Vis detector. High-resolution mass spectrometry (HRMS) analysis was performed in positive ion mode on a Bruker Daltonics maXis ESI-QTOF device. Melting point (mp) analysis was done on a Büchi M-560 system.

We synthesized compound 1 by reductive amination.\textsuperscript{3} 1-methyl-1H-imidazole-2-carbaldehyde (0.5 mmol, 55.89 mg) and 1-bis(4-fluorophenyl)methylpiperazine (0.5 mmol, 148.07 mg) were dissolved in 5 mL 1,2-dichloroethane and stirred under nitrogen for 19 hours at room temperature. Sodium triacetoxysorobohydrde (0.7 mmol, 152.52 mg) was added, and the pH was adjusted to 4 with acetic acid. The reaction was stirred for another 29 hours and monitored by HPLC-MS, then quenched with 5 mL of saturated NaHCO$_3$. The crude product was extracted with three times 15 mL diethyl ether, washed with 30 mL brine, dried over MgSO$_4$ and filtered. The solvent was removed under a stream of nitrogen and the product was purified by preparative HPLC. White-brown amorphous solid (purity: 95%, 7.8 mg, 4%; re-synthesis of 46 mg, 6%), $mp = 49{\degree}C$. $^1$H-NMR (400 MHz, chloroform-d): $\delta$ 7.36 (dd, $J = 8.5, 5.3$ Hz, 4H), 7.11 (s, 1H), 7.03-6.91 (m, 5H), 4.36 (s, 1H), 4.00 (s, 2H), 3.85 (s, 3H), 2.83 (s, 4H), 2.60 (s, 1H) ppm. $^{13}$C NMR (101 MHz, chloroform-d): $\delta$ 163.23, 160.78, 136.88, 129.30 ($d, J = 7.9$ Hz), 124.37, 122.79, 115.69 ($d, J = 21.4$ Hz), 74.09, 52.35, 51.39, 50.30, 34.22 ppm. HRMS ($C_2H_2F_2N_4$) [M+H]$^+$ calc. 383.2042 Da, found 383.2042 Da.

Dynamic light scattering. Dynamic light scattering (90Plus Particle Size Analyzer, Brookhaven Instruments Corp., USA) was used to determine the colloidal aggregation potential of compound 1 in aqueous concentrations of 0.3-1.0 mM. For each concentration, the correlation function was recorded after 0, 15, 30, 45 and 60 minutes. Measurements were performed at 25 $^\circ$C, with default settings for water, and the dust filter parameter was set to 50.

Activity determination. All ligand binding assays were performed on Cerep (Celle l’Evescault, France) on a fee-for-service basis. The assay protocols can be found at URL: www.cerep.fr.

References


This journal is © The Royal Society of Chemistry 2015