Supporting Information

Discovery of a novel chemotype of tyrosine kinase inhibitors by fragment-based docking and molecular dynamics

Hongtao Zhao[†], Jing Dong[†], Karine Lafleur[‡], Cristina Nevado[*, ‡], and Amedeo Caflisch[*, †]

[†]*Department of Biochemistry and* [‡]*Department of Organic Chemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland*

Keywords: in silico screening; EphB4 kinase; angiogenesis; cancer; explicit solvent MD

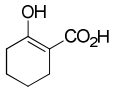
[*]To whom correspondence should be addressed. For C.N.: phone, (41) 446353945; fax, (41) 446353948; E-mail, nevado@oci.uzh.ch. For A.C.: phone, (41) 446355521; fax, (41) 446356862; E-mail, caflisch@bioc.uzh.ch.

Accession Code: The atomic coordinates and structure factors of EphA3 in complex with Compound **7** have been deposited with the Protein Data Bank as entry 4G2F.

Synthesis of Compound 7

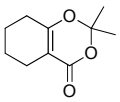
All reagents were used as received unless otherwise noted. Solvents were purchased in the best quality available, degassed by purging thoroughly with nitrogen and dried over activated molecular sieves of appropriate size. Alternatively, they were purged with argon and passed through alumina columns in a solvent purification system (Innovative Technology). Reactions were monitored by thin layer chromatography (TLC) using Merck TLC silica gel 60 F254. Flash column chromatography was performed over silica gel (230-400 mesh). NMR spectra were recorded on AV2 400 or AV2 500 MHz Bruker spectrometers. Chemical shifts are given in ppm. The spectra are calibrated to the residual ^{1}H and ^{13}C signals of the solvents. Multiplicities are abbreviated as follows: singlet (s), doublet (d), triplet (t), quartet (q), doublet-doublet (dd), quintet (quint), septet (sept), multiplet (m), and broad (br). Melting points were determined on a Büchi Melting Point B-540 instrument. The purity of all tested compounds was determined by HPLC on a Waters Acquity UPLC (Waters, Milford, USA) Top spectrometer using an Acquity BEH C18 HPLC column (1.7 µm, 1x50 mm, Waters) with a mixture of H_2O + 0.1% HCOOH (A) and CH_3CN + 0.1% HCOOH (B) solvent (0.1 ml flow rate, linear gradient from 5 to 98% B within 4 min followed by flushing with 98% B for 1 min). UV detection was set to 200-260 nm.

2-Hydroxycyclohex-1-enecarboxylic acid (9)



Ethyl 2-oxocyclohexanecarboxylate (1 mL, 6.3 mmol) was added to a solution of NaOH (275 mg, 6.9 mmol) in water (12 mL) at 0 °C. The resulting reaction mixture was stirred under ice-cooling for 3 h, and at room temperature for 12 h. The aqueous solution was washed with ether and acidified with concentrated HCl under ice cooling. The mixture was stirred for 45 min, and the formed precipitate was filtered off and dried to afford the expected product as a white solid (549 mg, 61%). ^{1}H NMR (300 MHz, $CDCl_3$): δ = 11.98 (s, 1H), 2.32-2.25 (m, 4H), 1.75-1.59 (m, 4H), OH not observed; ^{13}C NMR (100 MHz, $CDCl_3$): δ = 176.3, 174.7, 97.1, 29.3, 22.4, 22.3, 21.8; IR (film): $\tilde{\upsilon}$ = 3005, 2940, 2860, 1639, 1584, 1434, 1210, 795 cm^{-1}; MS (ESI): *m/z*: calcd for $C_7H_9O_3^-$: 141.2, found: 140.9.

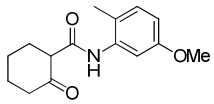
2,2-Dimethyl-5,6,7,8-tetrahydro-4H-benzo[d][1,3]dioxin-4-one (10)



To a solution of 2-hydroxycyclohex-1-enecarboxylic acid (406 mg, 2.8 mmol), acetone (417 µL, 5.6 mmol) and acetic anhydride (582 µL, 6.1 mmol) at -5 °C was added concentrated sulfuric acid (36 µL, 0.7 mmol). The reaction was stirred at 0 °C for 4 h, poured into a 10% solution of Na_2CO_3, and extracted with ether. The organic

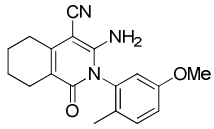
layer was dried over $MgSO_4$, filtered, and concentrated under reduced pressure to afford the desired product as a colorless solid (484 mg, 93%). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 2.29 (tt, J = 6.2 Hz, J = 1.8 Hz, 2H), 2.18 (tt, J = 6.3 Hz, J = 1.8 Hz, 2H), 1.76-1.71 (m, 2H), 1.69-1.66 (m, 8H); ^{13}C NMR (100 MHz, $CDCl_3$): δ = 164.8, 162.0, 105.1, 102.2, 27.4, 25.1, 21.9, 21.6, 21.1; IR (film): $\tilde{\upsilon}$ = 2996, 2939, 2859, 1719, 1649, 1401, 1296, 1202, 1024, 760 cm^{-1}; MS (ESI): *m/z*: calcd for $C_{10}H_{14}O_3Na^+$: 205.1, found: 205.0.

_N_-(5-Methoxy-2-methylphenyl)-2-oxocyclohexanecarboxamide (11)



A mixture of 2,2-dimethyl-5,6,7,8-tetrahydro-4H-benzo[d][1,3]dioxin-4-one (457 mg, 2.5 mmol) and 5-methoxy-2-methylaniline (343 mg, 2.5 mmol) in o-xylene (8 mL) was subjected to microwave irradiation for 5 min at 150 °C. The reaction mixture was cooled down, and the resulting precipitate removed by filtration over celite. The mixture was extracted with EtOAc and washed with HCl 10%. The organic layer was dried over $MgSO_4$, filtered, and concentrated under reduced pressure. Purification by column chromatography (gradient toluene/EtOAc 10:0 to 10:1) afforded the desired product as a white solid (308 mg, 46%). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 9.41 (s, 1H), 7.73 (d, J = 2.6 Hz, 1H), 7.06 (d, J = 8.4 Hz, 1H), 6.61 (dd, J = 8.4 Hz, J = 2.6 Hz, 1H), 3.79 (s, 3H), 3.35 (ddd, J = 11.1 Hz, J = 5.5 Hz, J = 0.9 Hz, 1H), 2.66-2.63 (m, 1H), 2.54-2.43 (m, 2H), 2.27 (s, 3H), 2.14-2.10 (m, 1H), 2.05-1.93 (m, 2H), 1.87-1.77 (m, 2H); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 212.0, 167.0, 158.3, 136.8, 130.8, 119.9, 110.7, 107.3, 55.6, 55.4, 42.5, 32.4, 27.6, 24.7, 17.0; IR (film): $\tilde{\upsilon}$ = 3227, 3122, 2939, 1709, 1496, 1281, 1122, 1037, 803 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{15}H_{19}NO_3Na^+$: 284.1257, found: 284.1257.

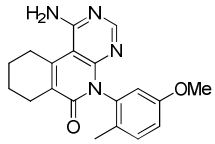
3-Amino-2-(5-methoxy-2-methylphenyl)-1-oxo-1,2,5,6,7,8-hexahydroisoquinoline -4-carbonitrile (12)



A mixture of N-(5-methoxy-2-methylphenyl)-2-oxocyclohexanecarboxamide (280 mg, 1.1 mmol), malononitrile (141 mg, 2.1 mmol), piperidine (63 µL, 0.6 mmol) in EtOH (3 mL) was heated to 100 °C for 2 h. The reaction mixture was cooled down, and the resulting precipitate was filtered off, affording the expected product as a white solid (273 mg, 82%). ^{1}H NMR (500 MHz, $CDCl_3$): δ = 7.32 (d, J = 8.5 Hz, 1H), 6.98 (dd, J = 8.5 Hz, J = 2.3 Hz, 1H), 6.68 (d, J = 2.3 Hz, 1H), 4.67 (s, 2H), 3.80 (s, 3H), 2.67-2.65 (m, 2H), 2.45-2.43 (m, 2H), 2.04 (s, 3H), 1.81-1.73 (m, 4H); ^{13}C NMR (125 MHz, $CDCl_3$): δ = 160.7, 159.3, 152.3, 147.1, 133.8, 132.7, 127.8, 117.2, 116.5, 116.2, 113.1, 72.1, 55.5, 28.2, 22.9, 22.0, 21.8, 16.3; IR (film): $\tilde{\upsilon}$ = 3419, 3308, 3208, 2869,

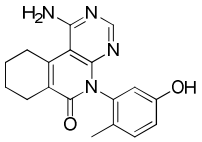
2202, 1662, 1606, 1531, 1448, 1242, 842 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{18}H_{19}N_3O_2Na^+$: 332.1370, found: 332.1370.

1-Amino-5-(5-methoxy-2-methylphenyl)-7,8,9,10-tetrahydropyrimido[4,5-c]isoquinolin-6(5H)-one (13)



3-Amino-2-(5-methoxy-2-methylphenyl)-1-oxo-1,2,5,6,7,8-hexahydroisoquinoline-4-carbonitrile (260 mg, 0.8 mmol) was dissolved in formamide (2.6 mL), and the resulting solution was heated to 210 °C for 7 h. The reaction was poured into ice water, and the resulting solid was filtered off. Purification by column chromatography (gradient EtOAc/CH_2Cl_2 1:1 to 2:1) afforded the expected product as a white solid (180 mg, 63%). ^{1}H NMR (500 MHz, DMSO-d_6): δ = 8.02 (s, 1H), 7.23 (d, J = 8.4 Hz, 1H), 7.11 (s, 2H), 6.89 (dd, J = 8.4 Hz, J = 2.6 Hz, 1H), 6.65 (d, J = 2.6 Hz, 1H), 3.71 (s, 3H), 3.04-3.03 (m, 2H), 2.47-2.46 (m, 2H), 1.76 (s, 3H), 1.73-1.70 (m, 4H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 161.0, 160.5, 158.0, 155.7, 153.7, 143.1, 137.9, 130.7, 126.9, 125.9, 114.1, 113.5, 97.6, 55.2, 28.6, 24.4, 21.9, 20.7, 16.3; IR (film): $\tilde{\upsilon}$ = 3421, 3144, 2941, 1626, 1532, 1216, 1032, 805 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{19}H_{20}N_4O_2Na^+$: 359.1479, found: 359.1479.

1-Amino-5-(5-hydroxy-2-methylphenyl)-7,8,9,10-tetrahydropyrimido[4,5-c]isoquinolin-6(5H)-one (7)



To a solution of 1-amino-5-(5-methoxy-2-methylphenyl)-7,8,9,10-tetrahydropyrimido[4,5-c]isoquinolin-6(5H)-one (127 mg, 0.4 mmol) in CH_2Cl_2 (4 mL) was added BBr_3 (1 M, 944 µL, 0.9 mmol). The reaction was stirred at 25 °C for 16 h, and poured into MeOH. The reaction mixture was evaporated several times EtOH, triturated with EtOAc, and the precipitate was filtered off. Recrystallization in EtOH and CH_2Cl_2 afforded the expected product as a white solid (34 mg, 28%). mp 262-265 °C; ^{1}H NMR (400 MHz, DMSO-d_6): δ = 9.33 (s, 1H), 8.04 (s, 1H), 7.15-7.09 (m, 3H), 6.72 (dd, J = 8.2 Hz, J = 2.5 Hz, 1H), 6.41 (d, J = 2.5 Hz, 1H), 3.07-3.01 (m, 2H), 2.46-2.43 (m, 2H), 1.74-1.71 (m, 7H); ^{13}C NMR (125 MHz, DMSO-d_6): δ = 160.9, 159.7, 155.9, 154.7, 153.5, 142.9, 137.4, 130.7, 126.3, 125.0, 115.4, 115.0, 97.6, 28.6, 24.4, 21.9, 20.7, 16.3; IR (film): $\tilde{\upsilon}$ = 3415, 3138, 2936, 2860, 1622, 1525, 1450, 1300, 1208, 804 cm^{-1}; HRMS (ESI): *m/z*: calcd for $C_{18}H_{18}N_4O_2Na^+$: 345.1322, found: 345.1323. HPLC purity 88.4%.

Protein Expression and Purification

A clone of the EphA3 kinase domain (residues: 606-947) was obtained from Prof. Sirano Dhe-Paganon's group[1] and expressed in Escherichia coli strain BL21 (DE3). Cells expressing EphA3 were induced with a 1 mM solution of isopropyl-beta-D-thiogalactopyranoside (IPTG) for 12 h at 15 °C. Cell pellets were resuspended in buffer A (50 mM Tris, pH 8.0, and 100 mM NaCl, supplemented with protease inhibitors) and lysed by sonication. After centrifugation at 15,000 rpm for 1 h, the soluble fraction of EphA3 was purified using HisTrap FF crude and HiTrap Q HP columns (GE Healthcare), followed by gel filtration chromatography (Superdex75; GE Healthcare). The appropriate fractions were combined and concentrated to ~10 mg/mL using Amicon filter devices (10 kDa as cutoff) in a storage solution (100 mM sodium chloride and 10 mM Tris-HCl pH 8.0, 5% glycerol). The resulting solution was aliquoted and stored at -80 °C for further usage.

Crystallization, Data Collection, and Structure Determination

Crystals of the EphA3 kinase domain were grown at 20 °C using the hanging drop vapor diffusion method. A 5 mM compound **7** (in 100 % DMSO) was added into the EphA3 protein to reach a final DMSO concentration of 10% (v/v) and the mixture was incubated on ice for 1 hour before crystallization. Then equal volumes of protein (with compound **7**) and reservoir solutions (0.1 M sodium cacodylate pH 6.5, 0.15 M ammonium sulfate, 22.5% PEG 3350) were mixed and crystals appeared after 1 to 2 days. The crystals were flash-frozen in liquid nitrogen without extra cryoprotectant for measurements.

Crystallography

Data sets were collected on a PILATUS 2M detector at the Swiss Light Source beamline X06DA of the Paul Scherrer Institute (Villigen, Switzerland) and indexed, integrated and scaled with the XDS[2] and CCP4[3] programs. The structures were solved by molecular replacement with PHASER[4] using the apo EphA3 kinase domain structure (PDB entry 2GSF) as a search model and refined with PHENIX[5].

Space group	**P1 21 1**	<I/σ(I)>	24.5(4.3)
Unit cell		R merge	0.044(0.416)
a (Å)	53.07	Completeness (%)	99.6(97.8)
b (Å)	38.24	Multiplicity	6.6(6.1)
c (Å)	75.84	**Refinement**	
alpha	90.00	Resolution range (Å)	39.07-1.70
beta	101.59	R factor/R free	0.179/0.208
gamma	90.00	Mean B factors (Å^2)	26.40
Resolution range (Å)	47.29 -1.70	RMS bonds (Å)	0.0106
Unique reflections	33115(4694)	RMS angles (°)	1.510

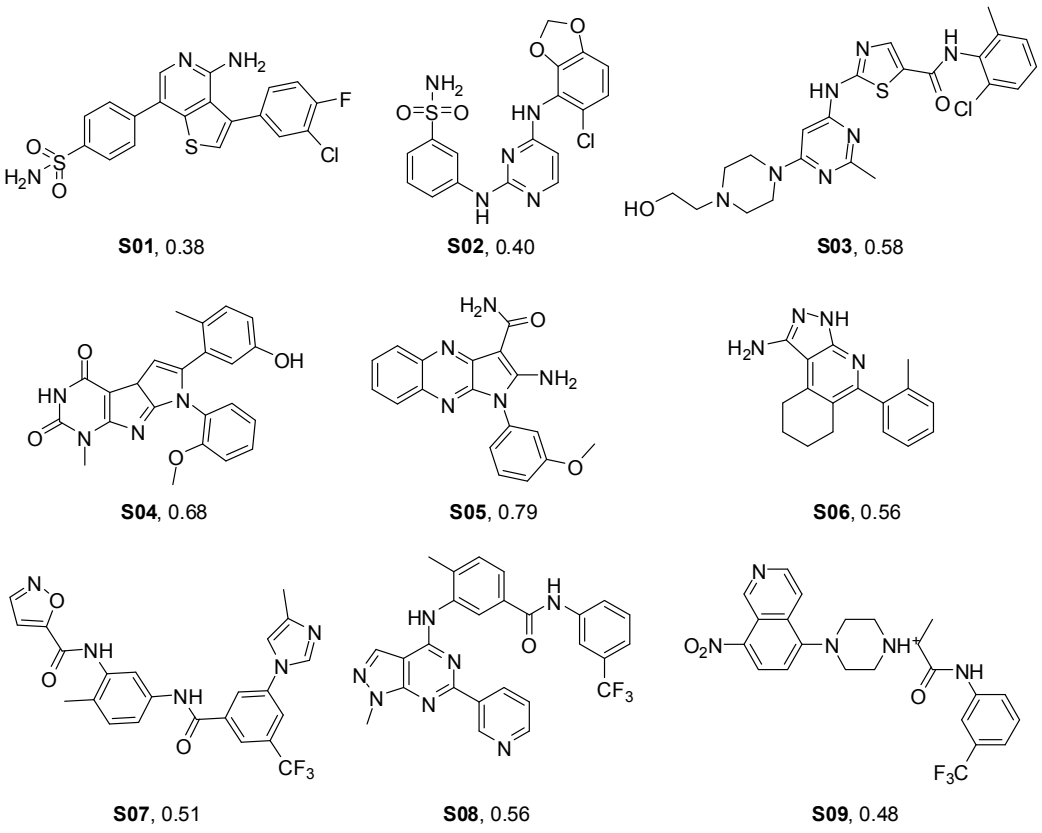


Figure S1. Previously reported Eph inhibitors.[1, 6-13] The values next to the compound number are the Tanimoto similarity index calculated by MACCS keys with compound **7**. Compound **S04**, **S05** and **S09** were discovered by our in silico approach.[9-10, 13]

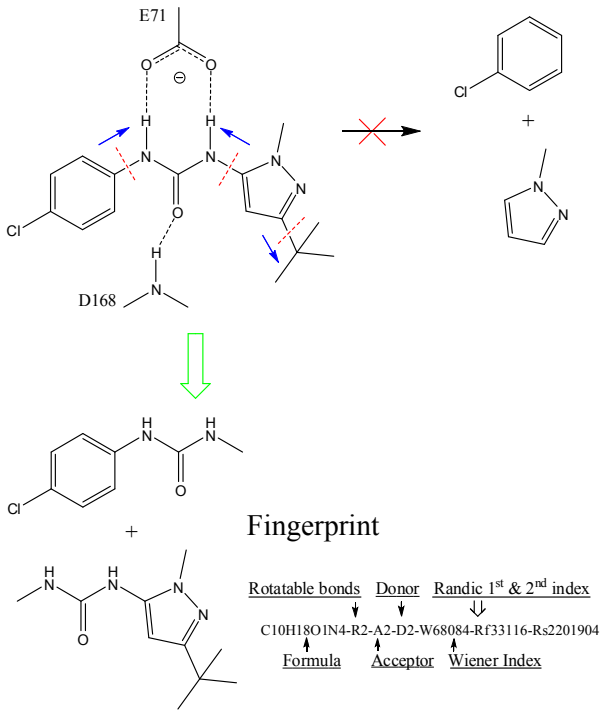


Figure S2. Decomposition of a known type II kinase inhibitor (PDB entry 1KV1) into anchor fragments. Over-decomposition could lead to very small fragments (top, right). Here, the following protocol is used to generate anchor fragments by decomposition of a compound library. Firstly, all rotatable bonds of a molecule are broken to obtain initial ring systems. Secondly, each ring system is extended at the cut until it reaches another ring or the length of the extension (number of heavy atoms) reaches three. In case the atom on the third level is not an sp^3 carbon, the elongation goes further until it reaches an sp^3 carbon or a ring. Thirdly, to keep the valence unchanged, hydrogen atoms are attached to carbon atoms while heteroatoms are capped by CH_3 groups. Lastly, to remove identical anchor fragments, a fingerprint consisting of molecular formula, number of rotatable bonds, acceptors and donors, and two 2D topological indices, i.e. modified Wiener[14] and Randic[15] index, is computed. The Wiener index is modified by taking valence and element difference into account as $W = \sum(v_i e_i \sum(d_{ij} e_j))$, wherein v_i is the valence of the i^{th} atom minus the number of hydrogen atoms attached to it, e_i is the atomic number, and d_{ij} is the distance between the i^{th} and j^{th} atom in the molecular graph. The modified Randic first and second index are ${}^1\chi = \sum p_i p_j$ and ${}^2\chi = \sum p_i p_j p_k$, respectively, with $p_i = \sum d_{ij}$.

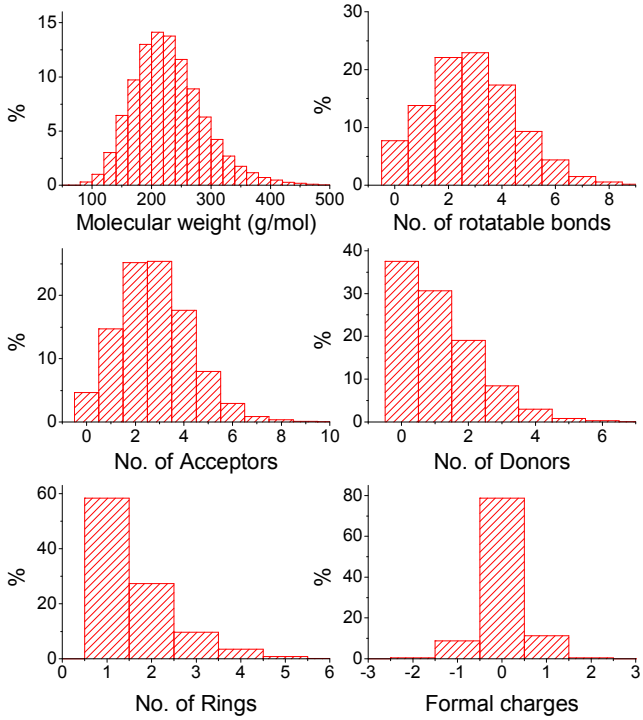


Figure S3. Distribution of physicochemical properties of the library of anchor fragments (563,774 fragments) obtained by decomposing the ZINC all-now compound library (9 million molecules) within 3 hours on a single commodity CPU by the in house developed software LIBO. The program LIBO is written in C++ language and is available upon request on the homepage of the last author.

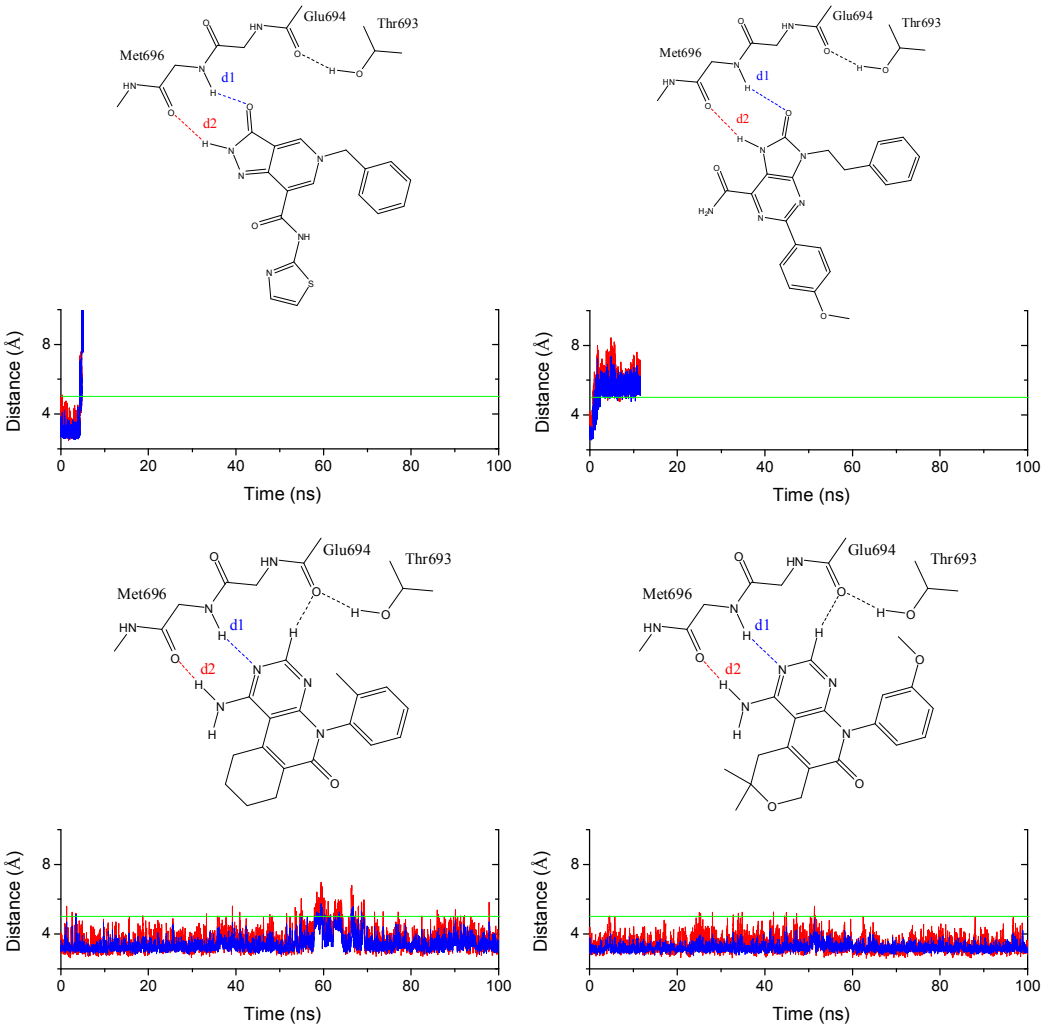


Figure S4. Screening by explicit solvent molecular dynamics. 2D representation of binding mode as well as MD time series of Hbonds of four scaffolds identified by the ALTA virtual screening approach. The two molecules in the top panels are discarded because their hydrogen bonds with the hinge region break apart within the first 10 ns.

References

1. Choi, Y.; Syeda, F.; Walker, J. R.; Finerty, P. J., Jr.; Cuerrier, D.; Wojciechowski, A.; Liu, Q.; Dhe-Paganon, S.; Gray, N. S. Discovery and structural analysis of Eph receptor tyrosine kinase inhibitors. *Bioorg Med Chem Lett* **2009,** 19, 4467-70.

2. Kabsch, W. Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants. *J Appl Crystallogr* **1993,** 26, 795-800.

3. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **1994,** 50, 760-3.

4. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser crystallographic software. *J Appl Crystallogr* **2007,** 40, 658-674.

5. Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **2002,** 58, 1948-54.

6. Miyazaki, Y.; Nakano, M.; Sato, H.; Truesdale, A. T.; Stuart, J. D.; Nartey, E. N.; Hightower, K. E.; Kane-Carson, L. Design and effective synthesis of novel templates, 3,7-diphenyl-4-amino-thieno and furo-[3,2-c]pyridines as protein kinase inhibitors and in vitro evaluation targeting angiogenetic kinases. *Bioorg Med Chem Lett* **2007,** 17, 250-4.

7. Bardelle, C.; Cross, D.; Davenport, S.; Kettle, J. G.; Ko, E. J.; Leach, A. G.; Mortlock, A.; Read, J.; Roberts, N. J.; Robins, P.; Williams, E. J. Inhibitors of the tyrosine kinase EphB4. Part 1: Structure-based design and optimization of a series of 2,4-bis-anilinopyrimidines. *Bioorg Med Chem Lett* **2008,** 18, 2776-80.

8. Sun, L.; Liang, C.; Shirazian, S.; Zhou, Y.; Miller, T.; Cui, J.; Fukuda, J. Y.; Chu, J. Y.; Nematalla, A.; Wang, X.; Chen, H.; Sistla, A.; Luu, T. C.; Tang, F.; Wei, J.; Tang, C. Discovery of 5-[5-fluoro-2-oxo-1,2- dihydroindol-(3Z)-ylidenemethyl]-2,4- dimethyl-1H-pyrrole-3-carboxylic acid (2-diethylaminoethyl)amide, a novel tyrosine kinase inhibitor targeting vascular endothelial and platelet-derived growth factor receptor tyrosine kinase. *J Med Chem* **2003,** 46, 1116-9.

9. Lafleur, K.; Huang, D.; Zhou, T.; Caflisch, A.; Nevado, C. Structure-based optimization of potent and selective inhibitors of the tyrosine kinase erythropoietin producing human hepatocellular carcinoma receptor B4 (EphB4). *J Med Chem* **2009,** 52, 6433-46.

10. Zhao, H.; Huang, D. Hydrogen bonding penalty upon ligand binding. *PLoS One* **2011,** 6, e19923.

11. van Linden, O. P.; Farenc, C.; Zoutman, W. H.; Hameetman, L.; Wijtmans, M.; Leurs, R.; Tensen, C. P.; Siegal, G.; de Esch, I. J. Fragment based lead discovery of small molecule inhibitors for the EPHA4 receptor tyrosine kinase. *Eur J Med Chem* **2012,** 47, 493-500.

12. Martiny-Baron, G.; Holzer, P.; Billy, E.; Schnell, C.; Brueggen, J.; Ferretti, M.; Schmiedeberg, N.; Wood, J. M.; Furet, P.; Imbach, P. The small molecule specific EphB4 kinase inhibitor NVP-BHG712 inhibits VEGF driven angiogenesis. *Angiogenesis* **2010,** 13, 259-67.

13. Zhao, H.; Huang, D.; Caflisch, A. Discovery of tyrosine kinase inhibitors by docking into an inactive kinase conformation generated by molecular dynamics. *ChemMedChem* **2012**.

14. Wiener, H. Structural determination of paraffin boiling points. *J Am Chem Soc* **1947,** 69, 17-20.

15. Randic, M. Characterization of molecular branching. *J Am Chem Soc* **1975,** 97, 6609-6615.