Electronic Supplementary Material (ESI) for RSC Medicinal Chemistry.

Synthetic inhibitor leads of human tropomyosin receptor kinase A (*h*TrkA)

Govindan Subramanian,[1,*] Rajendran Vairagoundar,[1] Scott J. Bowen,[1] Nicole Roush,[1] Theresa Zachary,[1] Christopher Javens,[1] Tracey Williams,[1] Ann Janssen,[1] and Andrea Gonzales[1]

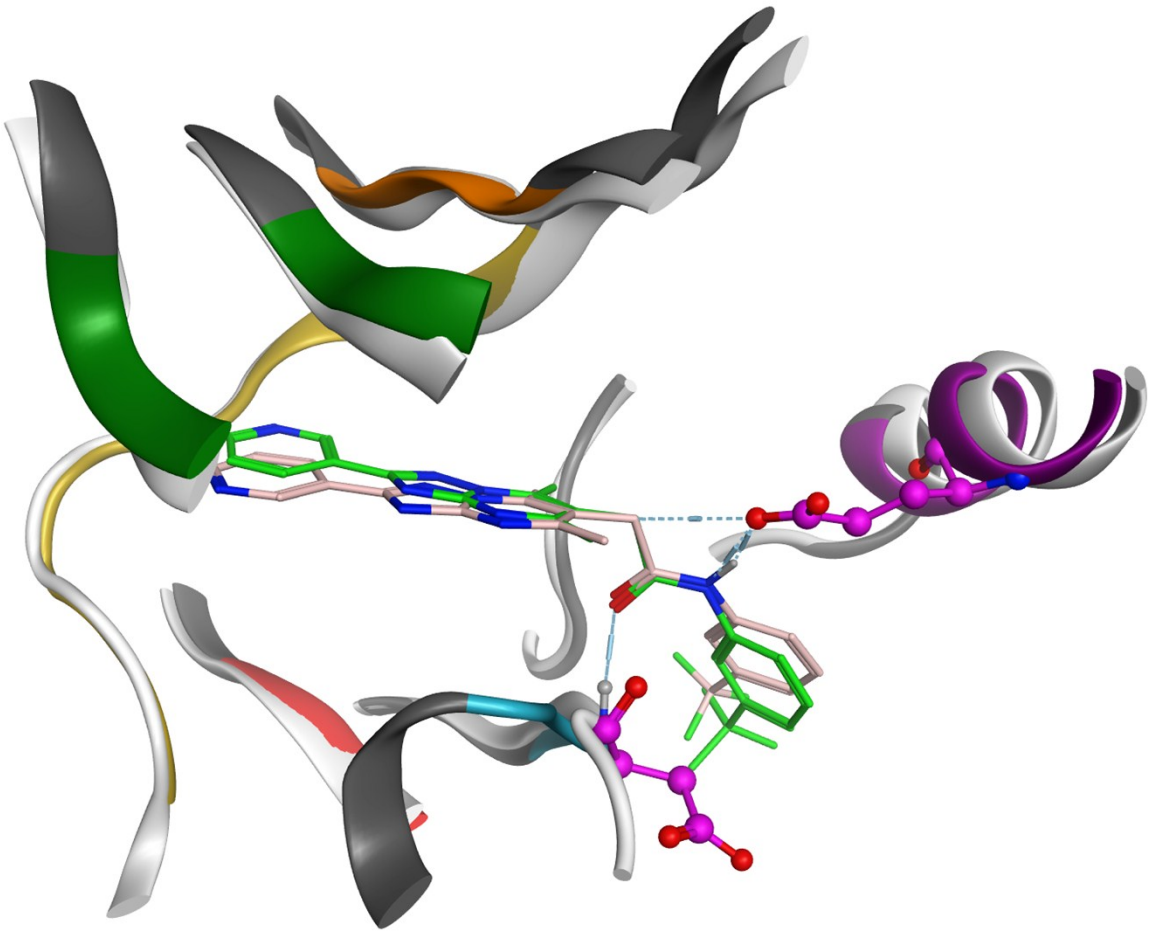
Supporting Information

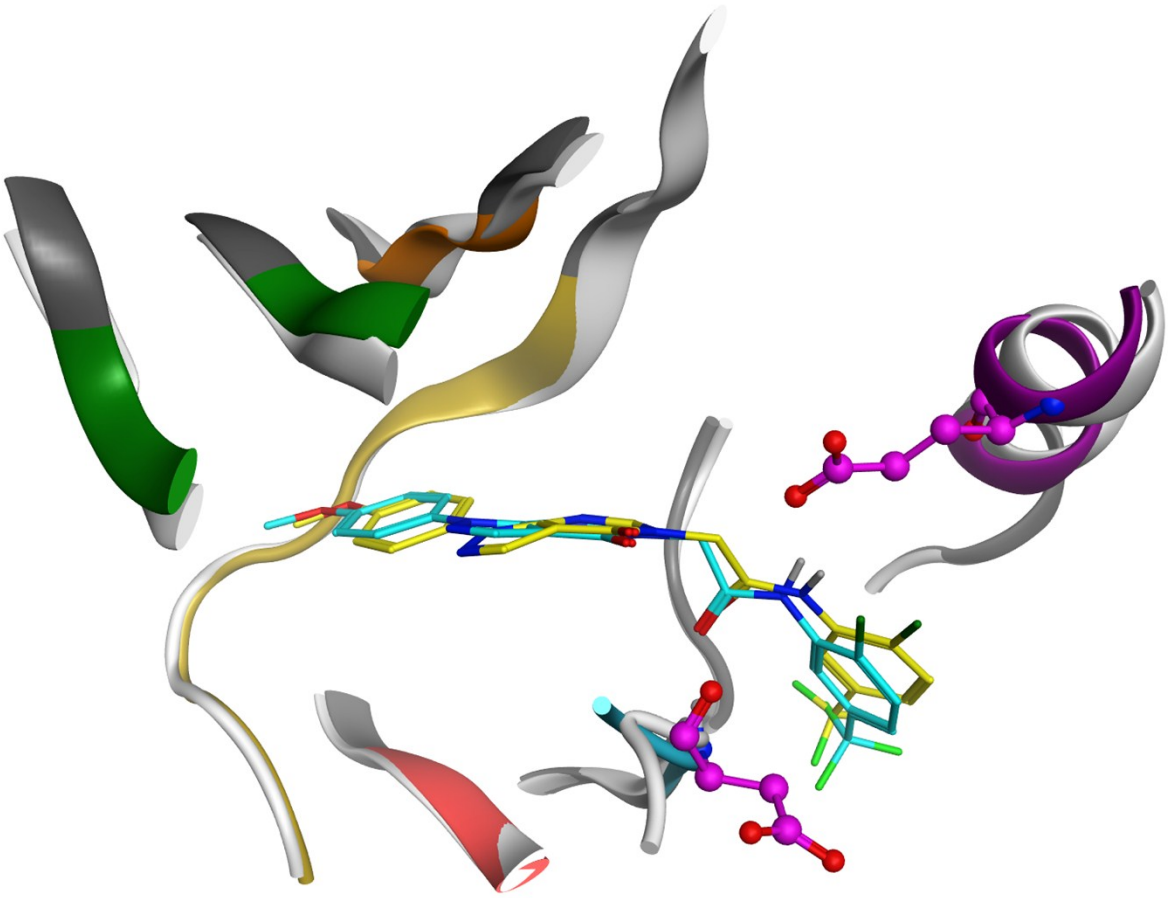
[1] Veterinary Medicine Research & Development (VMRD), Zoetis, 333 Portage Street, Kalamazoo, MI49007, USA.

Method 1S (Met. 1S):

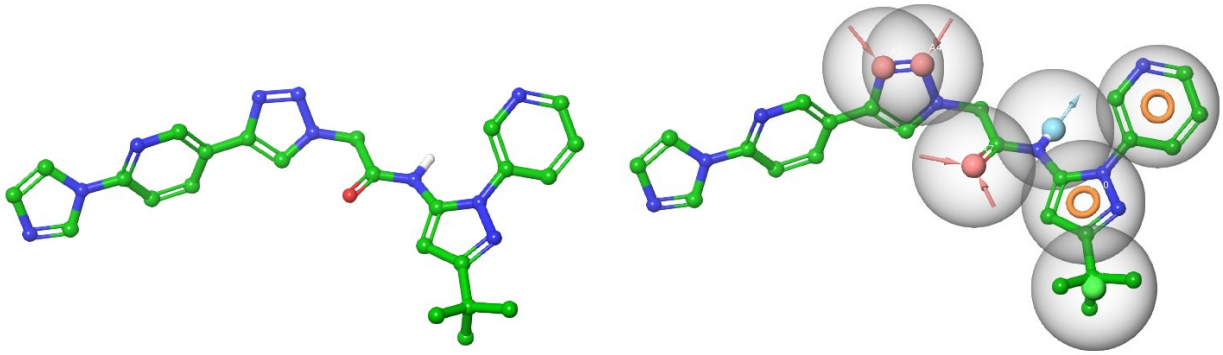
Small molecule collection: ~9 million lead-like small molecules from the ZINC database were cleaned for the inclusion of appropriate polar and apolar hydrogens, formal molecular charges, and energy minimized to a local minimum using MOE (Chemical Computing Group: Montreal, Canada).

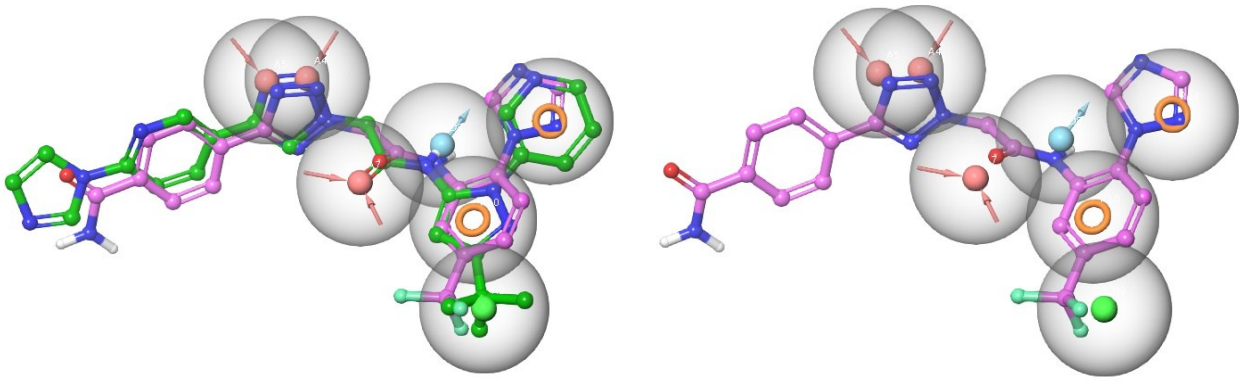
GLIDE docking: The PDB entry 3V5Q was downloaded from the protein data bank (http://www.rcsb.org) and protein prepared using the default parameters in MAESTRO (Schrodinger: New York, USA). The orthosteric ligand in the bound complex was used to define the binding site. Two interaction constraints for the ligand to H-bond with the kinase were imposed as part of the GLIDE virtual screening workflow. This would include the backbone amide NH of D668 and the sidechain carboxylate of E590 to interact with the appropriate ligand functional groups. The superposition of the docked pose for **1** (pink; GScore: -9.13) to the x-ray structure of **1** (green) and the docked pose of **3** (yellow; GScore: -9.17) to the x-ray structure of **3** (cyan) bound *h*TrkA kinase domain (kinase colored ribbon) are shown respectively on the top and bottom below for comparison.





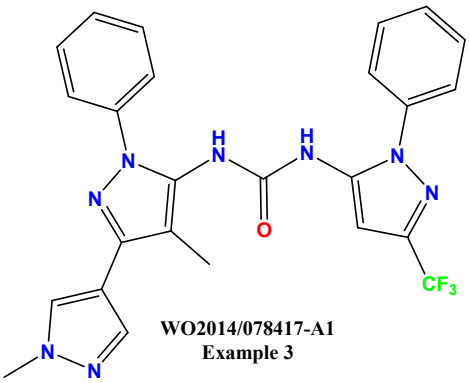
PHASE pharmacophore search: The energy minimized ZINC database structures were used to build a PHASE multi-conformational database and the same used for subsequent pharmacophore screening. The type 2 ligand (green ball-stick) from PDB entry 4PMM was extracted and 7 pharmacophore features comprising 3 H-bond acceptors, 1 H-bond donor, 2 ring aromatic, 1 hydrophobe as depicted below and ligand volume was used as the search query against the above prepared database. The fit to all the 7 pharmacophore features (Fitness score: 1.9969) of the disclosed ligand **2** (magenta stick) is shown for comparison.

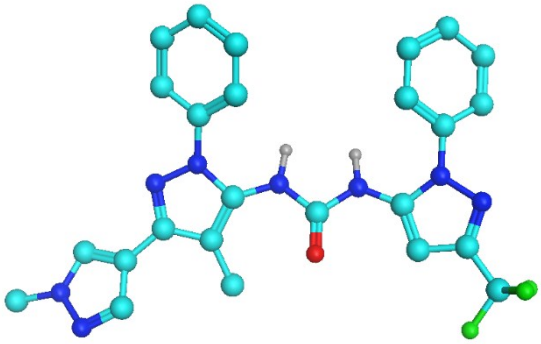


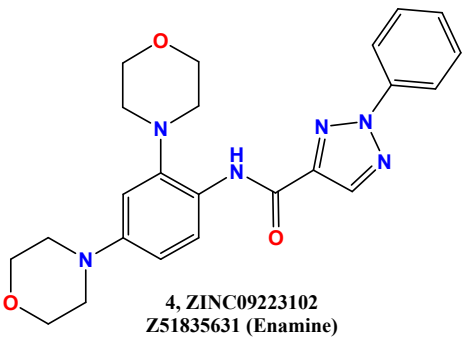


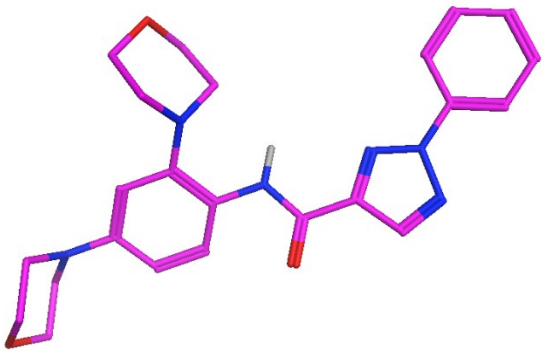
Method 2S (Met. 2S):

ROCS shape search: The 3D optimized structures of the ZINC database structures from MOE were then subjected to a multi-conformation search employing the OMEGA module (OpenEye Scientific Software: Sante Fe, NM) and a maximum of 200 conformations saved for each ligand. The resulting multi-conformational ligand database was used as input source to ROCS (OpenEye Scientific Software: Sante Fe, NM), a shape-based ligand fitting algorithm. The query molecule (example **3** from WO2014/078417A1 patent application) used for the shape and feature-based alignment, the singleton *in silico* hit (**4**) with a ComboScore of 1.238, and the ROCS overlay is shown below.









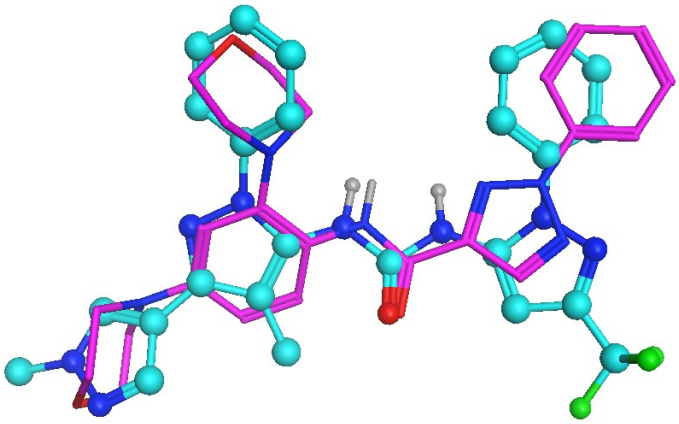
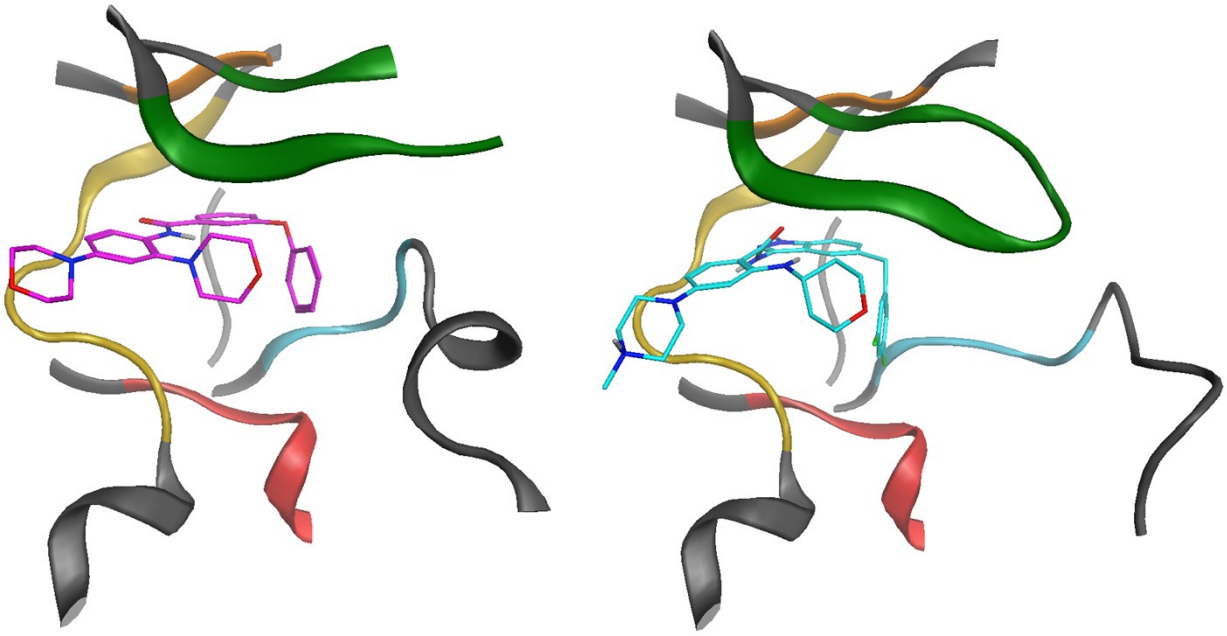


Figure 1S (Fig. 1S): Binding mode of 5 (left) and Entrectinib (right) in *h*TrkA kinase domain and the overlay of the ligands (bottom left) in the binding pocket and the kinase domain (bottom right) to compare the similarities.



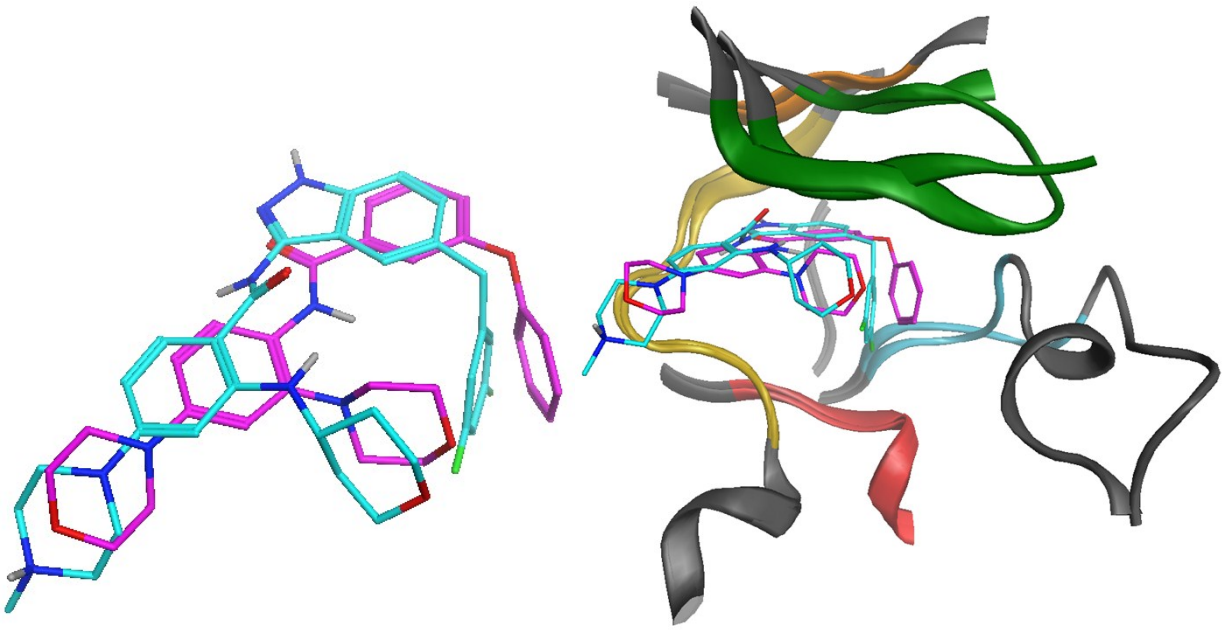


Table 1S. Data collection, processing, and refinement statistics for the kinase domain of *h*TrkA-ligand x-ray complexes along with the 3D rendering of the ligand in the kinase binding site.

	1[8]	**3**[9]	**5**[10]	**6**[11]
Data collection				
X-ray source[1]	PXI/X06SA (SLS)	PXI/X06SA (SLS)	PXI/X06SA (SLS)	I02 (DLS)[12]
Wavelength [Å]	1.0	1.0	0.9998	0.9282
Detector	EIGER X 16M	EIGER X 16M	Pilatus 6M	Pilatus 6M
Temperature [K]	100	100	100	100
Space group	P 3_1 1 2	P 3_1 1 2	I 2 2 2	P 3_1 1 2
Cell: a; b; c [Å] α; β; γ; [°]	52.11; 52.11; 226.75 90; 90; 120	52.09; 52.09; 227.16 90; 90; 120	131.4; 152.1; 156.4 90; 90; 90	52.66; 52.66; 233.54 90; 90; 120
Resolution [Å]	2.81 (3.06 – 2.81)	2.51 (2.76 – 2.51)	3.00 (3.25 – 3.00)	2.19 (2.44 – 2.19)
Unique reflections	8869 (1919)	12146 (2755)	30047 (6424)	18645 (5039)
Multiplicity	4.1 (4.0)	4.2 (4.3)	3.5 (3.5)	2.2 (2.2)
Completeness [%]	99.1 (96.4)	97.2 (90.2)	94.7 (96.2)	95.1 (94.4)
R_{sym} [%]	8.1 (42.7)	5.5 (41.9)	9.3 (54.8)	3.8 (44.0)
R_{meas} [%]	9.3 (49.1)	6.3 (47.5)	10.9 (63.6)	5.0 (57.6)
Mean(I)/sd	19.81 (4.74)	18.2 (4.34)	13.77 (2.41)	15.29 (2.39)
Refinement[3]				
Resolution [Å]	75.58 – 2.81	75.72 – 2.53	109.03 – 3.00	77.85 – 2.19
# of refls. (working /test)	8196 / 672	11226 / 906	29475 / 572	17176 / 1467
R_{cryst} [%]	19.9	20.7	25.0	22.0
R_{free} [%][4]	24.2	26.1	28.0	27.2
# atoms: Protein	2310	2323	6586	2318
# atoms: Water	19	24	1	35
# atoms: Ligand	31	33	102	31

Deviations from ideal geometry[5]				
Bond length [Å]	0.008	0.007	0.012	0.010
Bond angles [°]	1.24	1.26	1.68	1.43
Bonded B's [Å^2][6]	3.7	8.0	1.2	5.0
Ramachandran Plot[7]				
Most favored [%]	92.7	93.1	91.9	94.8
Additional allowed [%]	6.9	6.5	7.5	4.8
Generously allowed [%]	0.4	0.4	0.6	0.4
Disallowed [%]	0.0	0.0	0.0	0.0
PDB code	6PMB	6PMA	6PME	6PMC

[1] Swiss Light Source (SLS, Villigen, Switzerland)

[2] Values in parenthesis refer to the highest resolution bin

[3] Values as defined in REFMAC5, without sigma cut-off

[4] Test-set contains 7.6%, 7.5%, 1.9%, and 7.6% respectively of measured reflections

[5] Root mean square deviations from geometric target values

[6] Calculated with MOLEMAN

[7] Calculated with PROCHECK

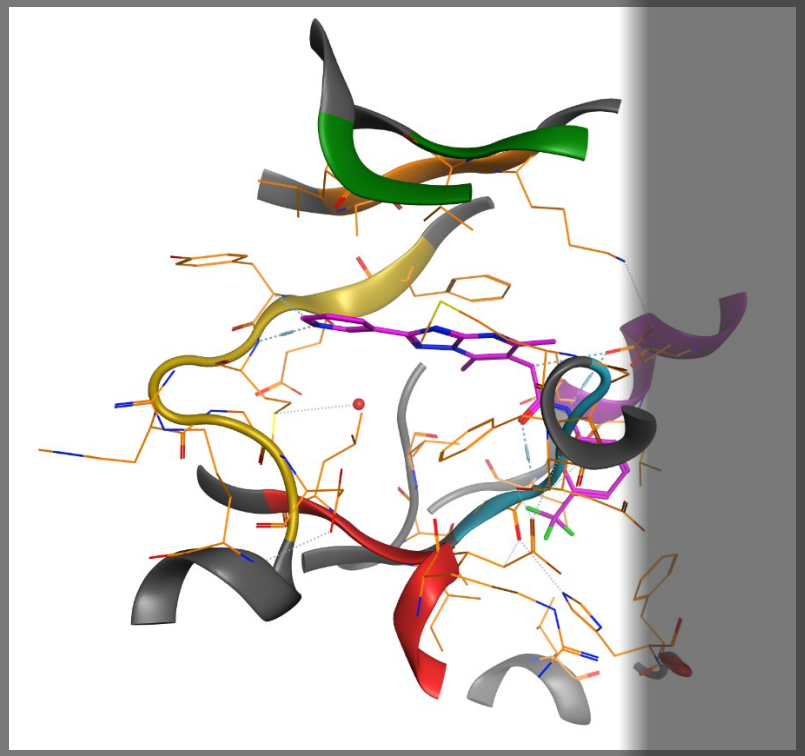
[8] amino acids defined by electron density include residues 500-795. Amino acids not defined by the electron density include residues 549-551, 609-613 and the side chains of residues 508, 521, 531, 537-538, 547-548, 555, 558-559, 566, 568, 607, 671, 675, 725, 759, 761, 773, 775, 780, and 793.

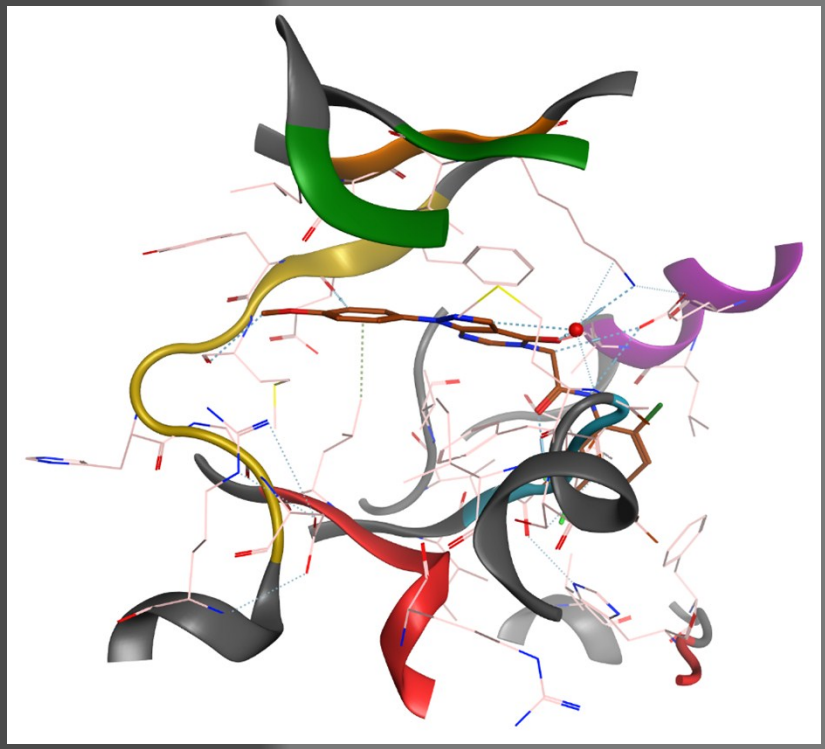
[9] amino acids defined by electron density include residues 500-795. Amino acids not defined by the electron density include residues 549-551, 609-611 and the side chains of residues 501, 508, 521, 523, 535, 537, 547-548, 554-555, 558-559, 566, 570, 583, 585, 607, 686, 691, 747, 761, 767, 775, and 780.

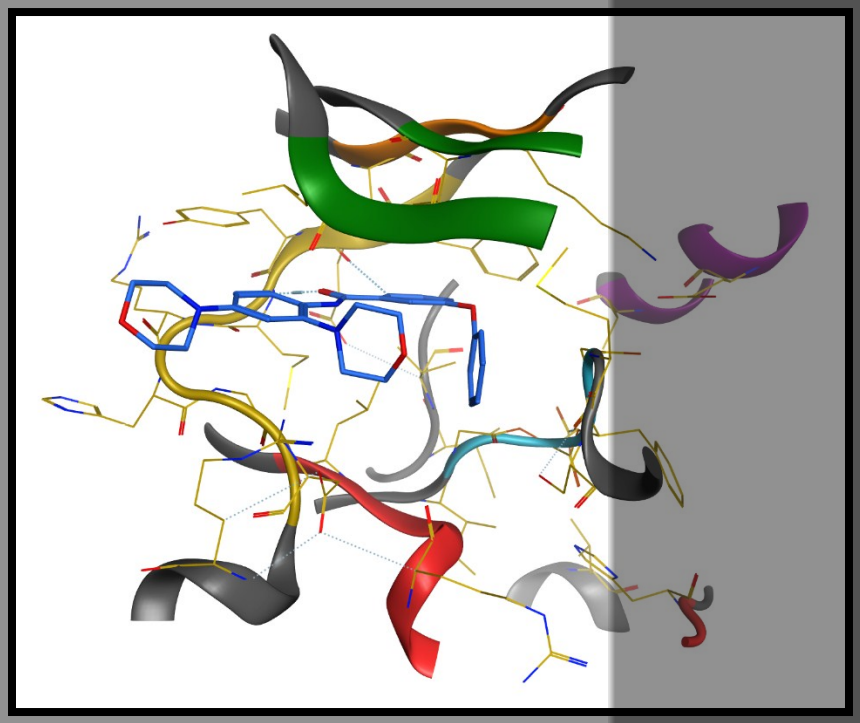
[10] amino acids defined by electron density include residues 500-788. Amino acids not defined by the electron density include residues 532-537, 608-619, 684-686 and the side chains of residues 544, 547, 570, 660, 669, 682, 703, 740, and 769.

[11] amino acids defined by electron density include residues 500-795. Amino acids not defined by the electron density include residues 606-612 and the side chains of residues 501, 508, 510, 521, 538, 546-547, 551-552, 554-555, 558-559, 566, 599, 638, 644, 671, 686, 688, 725, 750, and 761.

[12] Diamond Light Source (DLS, Oxford, England)







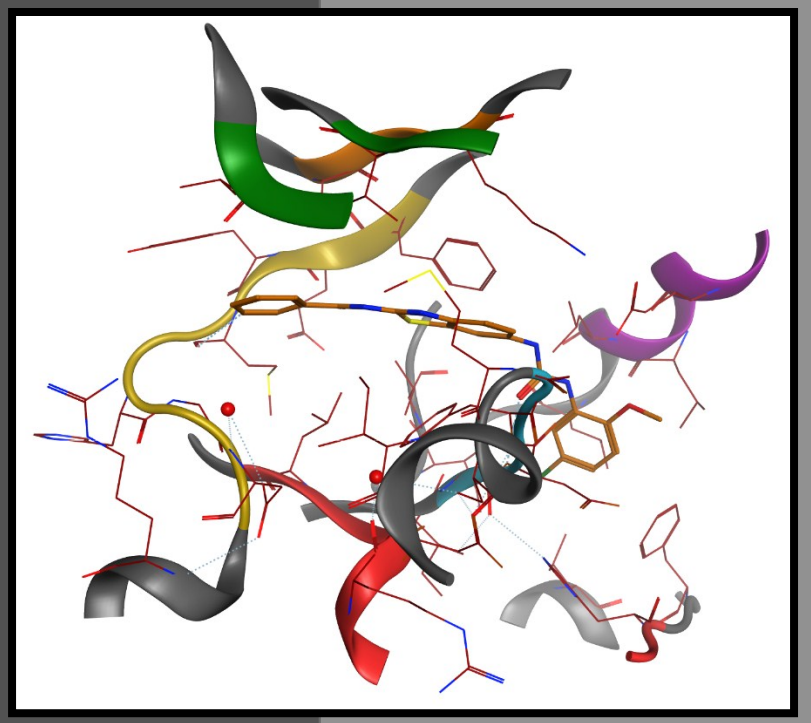


Table 2S. *In silico* modeled resistance mutational data for *h*TrkA-**inh** kinase domain complex. Clinically observed mutations highlighted in red, along with the other amino acid possibilities in that position based on CODON permutations. All solvent waters were removed prior to calculation performed using MOE2019.01 resistance scan.

Mutation	Δaffinity	Δstability	Δaffinity	Δstability	Δaffinity	Δstability	Δaffinity	Δstability
F589C	0.683	3.046	0.546	2.586	0.184	3.126	0.093	2.933
F589I	0.499	1.126	0.080	1.414	0.064	1.556	0.301	1.315
F589L	0.438	1.223	0.287	1.113	0.086	1.423	0.176	1.349
F589S	0.760	3.335	0.678	2.982	0.167	3.234	-0.020	3.471
F589Y	-0.078	0.730	0.116	0.616	-0.026	0.878	-0.177	1.117
F589V	0.620	1.819	0.402	1.530	0.155	2.200	0.342	1.845
G595A	-0.025	-0.409	-0.018	0.329	-0.074	0.322	0.062	0.224
G595R	-0.185	-1.296	-0.357	-1.389	-0.118	0.457	-0.076	0.112
G595D	-0.036	0.616	-0.158	0.535	0.070	1.856	-0.189	0.946
G595C	-0.054	-0.099	-0.107	-0.271	0.243	0.597	-0.060	0.509
G595E	-0.169	-0.513	-0.194	-0.097	0.182	1.495	-0.221	0.955
G595S	-0.007	0.388	0.018	0.512	-0.191	0.776	0.022	0.798
G595W	-0.163	-0.319	-0.325	-1.403	0.517	1.477	-0.390	-1.031
G595V	-0.176	-1.537	-0.148	-0.588	0.203	0.443	-0.182	-0.636
G667A	0.236	0.722	0.471	0.624	-0.009	0.247	0.376	0.451
G667R	0.320	1.116	0.589	0.025	0.365	1.499	0.243	-0.836
G667D	0.170	0.505	-0.032	0.511	0.014	1.072	0.153	0.850
G667C	-0.032	0.309	0.466	0.597	-0.067	0.596	0.271	0.561
G667E	-0.161	0.853	0.183	0.265	0.146	1.551	0.077	0.454
G667S	0.024	1.463	0.266	0.988	-0.045	0.600	0.166	0.763
G667W	0.122	1.572	1.366	1.859	0.524	2.737	-0.184	-0.543
G667V	0.485	0.906	0.446	0.203	0.008	0.408	0.193	0.357
Inhibitor	**1**		**3**		**5**		**6**	