Supporting Information

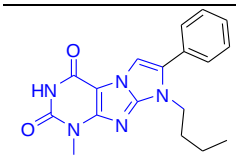
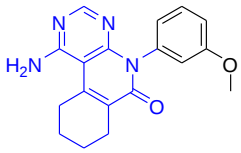
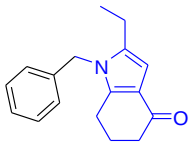
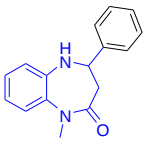
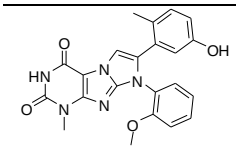
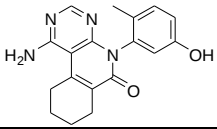
Current kinase inhibitors cover a tiny fraction of fragment space

Hongtao Zhao*, and Amedeo Caflisch*

Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

*To whom correspondence should be addressed. E-mail: h.zhao@bioc.uzh.ch (H.Z.); caflisch@bioc.uzh.ch (A.C.).

Table S1. Representative hits from fragment-based high-throughput virtual screening campaigns and leads derived from the initial hits.

	Protein	IC_{50} (μM)	LE[a]	Fragmentation	PDB code	Ref
	EphB4	1.5	0.32	DAIM	N.A.	[1]
	EphB4	5.2	0.30	In ref. 2	N.A.	[2]
	BRD4	7.0	0.37	In ref. 2	4PCI	[3]
	BRD4	7.5	0.37	In ref. 2	4PCE	[3]
Undisclosed	CREBBP	5.0	0.37	DAIM	4TQN	b
Undisclosed	CREBBP	4.0	0.39	In ref. 2	4TS8	c
Hit optimization						
	EphB4	0.002	-	-	4GK2	[4-5]
	EphB4	0.16	-	-	4G2F	[2]

[a] The ligand efficiency (LE) is the measured affinity divided by the number of non-hydrogen atoms and has units of kcal/mol per heavy atom. [b]Min Xu et al. unpublished results. [c]H. Zhao et al. unpublished results. Anchor fragments obtained by automatic fragmentation are colored in blue.

Table S2. SMILE strings of top 50 potential hinge-binding fragments for each of monocyclic, bicyclic and multicyclic systems in kinase inhibitors.

monocyclic	bicyclic	multicyclic
CNc1ncc(SC)s1	CNc1c(C#N)cnc2cc(OC)c(OC)cc21	CNc1nc2ccccc2n2cncc12
CNc1ncc(Cl)c(NC)n1	CNc1ncnc2cc(OC)c(OC)cc21	CCn1c2ccccc2c2c3c(c4c(c21)CCc1nn(C)cc1-4)C(=O)NC3
CNc1ccccn1	COc1cc2cncnc2cc1OC	CNc1ncc2c(n1)-c1c(c(C(N)=O)nn1C)CC2
Nc1ncccn1	COc1ccnc2cc(OC)c(OC)cc21	COc1cc2ncc3c(N)nccc3c2cc1OC
CNc1ccnc(NC)n1	C=C1c2ccccc2NC1=O	CNC(=O)Nc1cccc2c1C(=O)c1c[nH]nc1-2
CNc1nccc(N(C)C)n1	Cn1ccc2c1ncnc2N	CNc1[nH]nc2c1Cc1cc(OC)c(OC)cc1-2
c1ccncc1	CNc1ncc2ccc(=O)n(C)c2n1	Cc1cc2n[nH]c(=O)n2c2ccccc12
CNc1nccs1	CNc1ncnc2ccccc21	c1[nH]nc2c1Cc1ccccc1-2
CNCc1ccncc1	CNc1nc2c(ncn2C(C)C)c(NC)n1	CNc1nc2cc[nH]c(=O)c2c2cc(F)ccc21
CNc1ncccn1	Cn1ncc2c1ncnc2N	CNc1c(C#N)cnc2cc3cc(OC)c(OC)cc3cc21
O=C1C=CC(=O)N1	CNc1nccn2ccnc12	CN=Cc1c(O)[nH]c2ccc3ncsc3c12
CNc1cc[nH]n1	c1cc2ncccn2n1	CC(=O)Nc1cccc2c1C(=O)c1cn[nH]c1-2
CNc1ccncc1	Cn1cnc2ccccc21	CCn1c2ccc(COC)cc2c2c3c(c4c(c21)Cc1ccccc1-4)C(=O)NC3
CNc1ccncn1	c1nc2ccccc2s1	CNc1nc2scnc2c2c1ncn2C
c1cn[nH]c1	Cc1n[nH]c2ccccc12	CNc1ncc2c(n1)N(C)c1ccccc1C(=O)N2C
Cn1ccnc1	c1n[nH]c2ccccc12	Cn1cc2c(n1)CCc1c-2sc(NC(N)=O)c1C(N)=O
CNN=C1C=NNC1=O	Cc1cccc2[nH]cnc21	Cn1nc(C(N)=O)c2c1-c1c[nH]nc1CC2
CNc1cc[nH]c(=O)c1	CNc1ncnc2ccsc21	Cn1nc(C(N)=O)c2ccc3[nH]ncc3c21
Nc1cnon1	c1ccc2cnccc2c1	O=CNNC(=O)Nc1cccc2c1C(=O)c1c[nH]nc1-2
CNC(=O)Nc1cc(C(C)(C)C)on1	c1cc2cccnc2[nH]1	CNc1nc2cccnc2n2cncc12
C=Cc1cncc(C#N)c1NC	c1ccc2ncccc2c1	CNc1ncnc2c1NCc1cc(OC)c(OC)cc1O2
NC(=O)c1ccccc1	Nc1ccnc2ccnn21	Cc1ccc2c(c1)Cc1c[nH]nc1-2
c1c[nH]cn1	COc1ccnc2ccsc21	O=C1NC(=O)c2c1ccc1[nH]c3ccc(O)cc3c12
CCc1ccncc1	c1ccc2nccnc2c1	Cc1ccc2c(c1)Cc1cn[nH]c1-2
CNc1nnco1	Nc1nccn2cncc12	CNc1ncc2c(n1)N1CCC(=O)N1C=C2
CCn1cccn1	CC=CC(=O)Nc1cc2c(cc1OC)ncc(C#N)c2NC	Cc1ccc2c(c1)-c1[nH]ncc1C2
CNC(=O)c1ccncc1	CNc1ccnc2ccnn21	CCn1c2ccc(O)cc2c2c1ccc1c2C(=O)NC1=O
Nc1ccccn1	Cc1c[nH]c2ncncc12	CCn1c2ccc(NC(=O)NC)cc2c2c3c(c4c(c21)CCc1nn(C)cc1-4)C(=O)NC3
COc1cccnc1	c1cn2ccccc2n1	COc1cc2c(cc1Cl)NC(=O)Nc1cnc(C#N)c(n1)OCCCCCO2
CNc1ncc(C(F)(F)F)c(NC)n1	c1cn2cccnc2n1	CNc1cc2c(cc1Cl)NC(=O)Nc1cnc(C#N)c(n1)OCCCCCO2
CNC1=CC(=O)NC1=O	c1cc2n(n1)CCC2	O=C1Nc2ccccc2Nc2ccccc21
c1cncnc1	c1nc2ccccc2[nH]1	COc1ccc2c(c1)-c1[nH]ncc1C2
CC(=O)Nc1ccccn1	CNc1ncnc2cccc(OC)c21	CNCc1ccc2c(c1)Cc1cn[nH]c1-2
CC(=O)Nc1cc[nH]n1	O=C1NCCc2[nH]ccc21	CNC(=O)c1ccc2[nH]c3c(c2c1)C(C)CNC3=O
COc1ncccn1	CNc1ccc2nccn2n1	COc1cc2c(cc1OC)-c1n[nH]cc1C2
CNC(=O)c1cc(OC)ccn1	CNC(=O)c1cn2ncnc(NC)c2c1C	CCc1cc2cc[nH]c(=O)c2c2ccccc12
c1cscn1	CC(C)n1ncc2c1ncnc2N	CNc1nc2ccsc2n2c(C)cnc12
c1cocn1	c1cc2cncnc2[nH]1	CNc1nc2nc3c(cnn31)CCCCC(=O)Nc1cccc(c1)N2
COc1ccsc1C(N)=O	Nc1n[nH]c2ccccc12	CNc1ncc2c(n1)n1c3ccccc3nc1n(C)c2=O
CNc1cc(C)[nH]n1	CNc1ccc2c(cnn2C)c1	CCn1c2cc(OC)ccc2c2c3c(c4c5c(ccn5C)ccc4c21)C(=O)NC3=O

CNc1ncc([N+](=O)[O-])c(NC)n1	CNc1ncc2ccn(C)c2n1	CCc1cccc2c1[nH]c1c2c2c(c3c4c(ccn4C)ccc13)C(=O)NC2=O
CNc1cnccn1	CNc1ncnn2ccc(C)c12	CON=C1c2ccccc2NC1=C1C(=O)Nc2cc(Br)ccc21
CNC(=O)c1ccsc1NC(N)=O	CNc1ccc2[nH]ncc2c1	CNc1nc2[nH]ccc2c2c1ncn2C
CON=Cc1c(N)ncnc1OC	NC(=O)c1ccc2[nH]cnc2c1	CNC(=O)c1nn(C)c2c1C(C)(C)Cc1cnc(NC)nc1-2
N#Cc1cccnc1N	CNc1ncc2ccccc2n1	CC(=O)Nc1ccc2c(c1)CC1(C2)C(=O)NC(=O)N1C
C=C1C=NNC1=O	CCn1ncc2c1nc(SC)nc2NC	CN(C)Cc1nc2c3ccccc3sc2c(=O)[nH]1
CSc1cnc(NC=O)s1	CNc1ncnn2ccc(C(C)C)c12	COc1ccc2c(c1)CCn1ncc(C(N)=O)c1N2

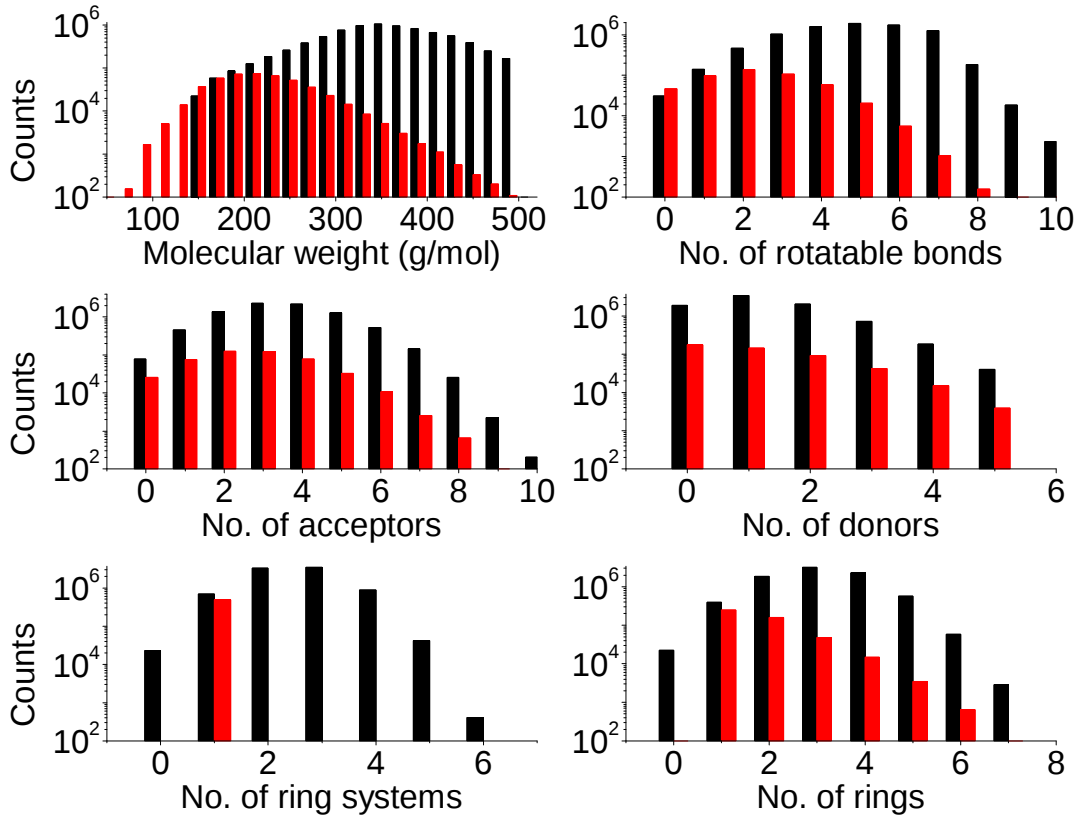


Figure S1. (A) Distribution of molecular properties of the 7.5 million compounds in the ZINC Drugs-Now library (black) and their 477,617 unique ring systems (red) obtained by automatic decomposition using the algorithm reported in Zhao et al., *ACS Med. Chem. Lett.* **2012**, 3, 834-838.

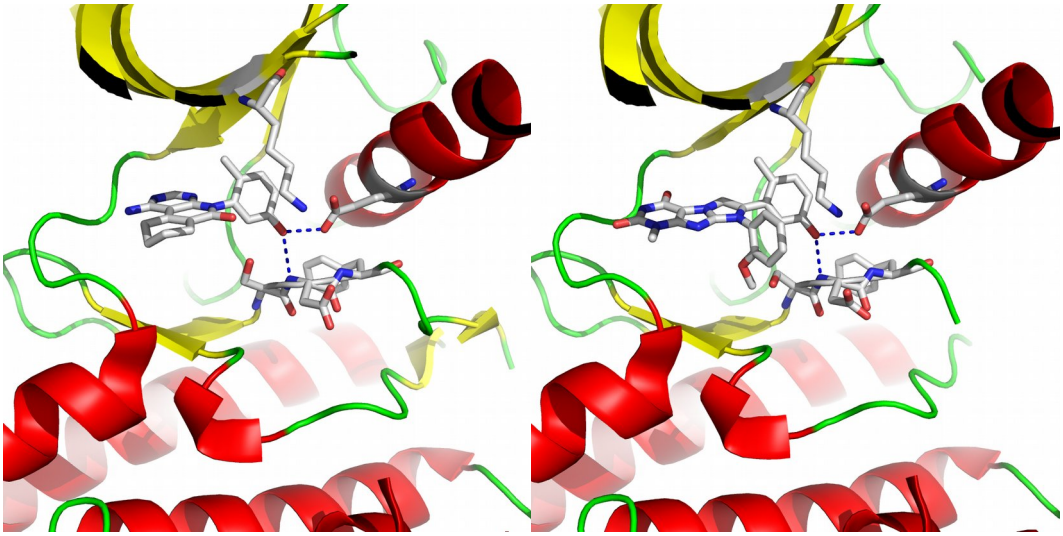


Figure S2. Native poses of two type $I_{1/2}$ kinase inhibitors in complex with EphA3. (Left) PDB code 4G2F and (Right) PDB code 4GK2. The characteristic hydrogen-bonding pattern between the phenol moiety of type $I_{1/2}$ inhibitors and the protein was highlighted by dotted blue lines.

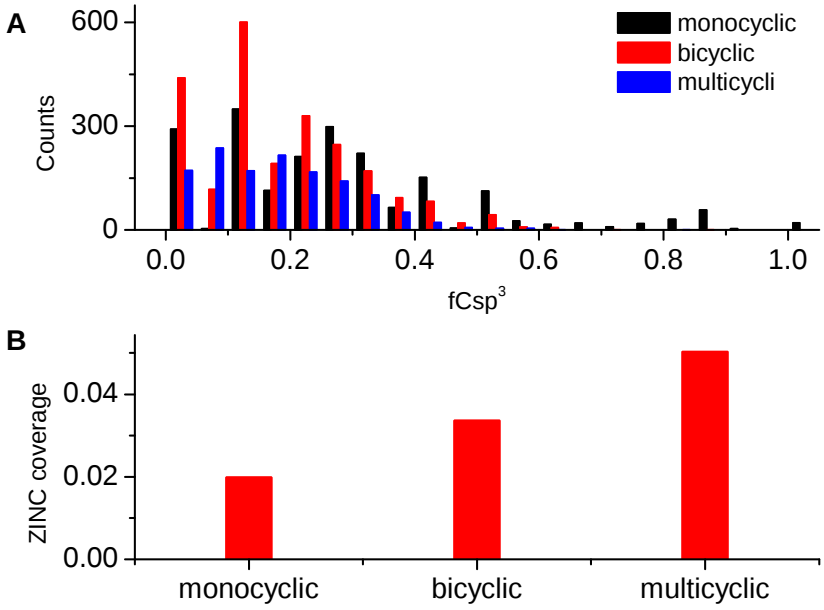


Figure S3. (**A**) Distribution of the fraction of sp3 carbon atoms (fCsp3) in monocyclic (black), bicyclic (red), and multicyclic (blue) putative hinge binding fragments obtained by automatic decomposition of kinase inhibitors. (**B**) ZINC coverage of putative hinge binding fragments calculated by comparing the number of heavy atoms (in the putative hinge-binding fragments from known kinase inhibitors and from the ZINC library) instead of the chemical structure.

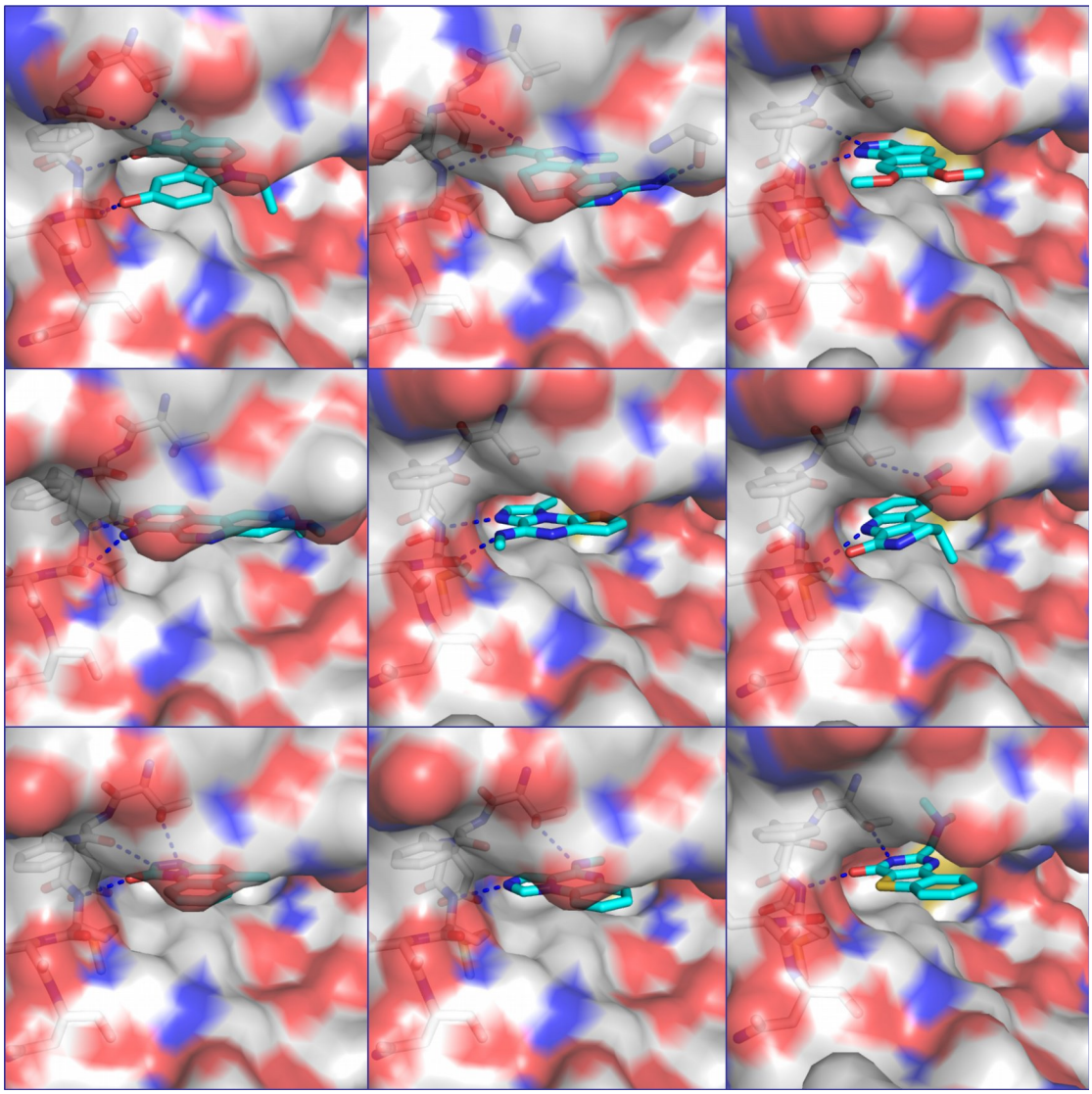


Figure S4. Predicted binding modes of top 9 multicyclic putative hinge-binding fragments in the ATP site of a receptor tyrosine kinase of EphB4 (PDB code 2VWX). Hydrogen bonds with the hinge region (shown in sticks) are shown in dotted blue lines. Acidic C-H···O hydrogen bonds are not shown.

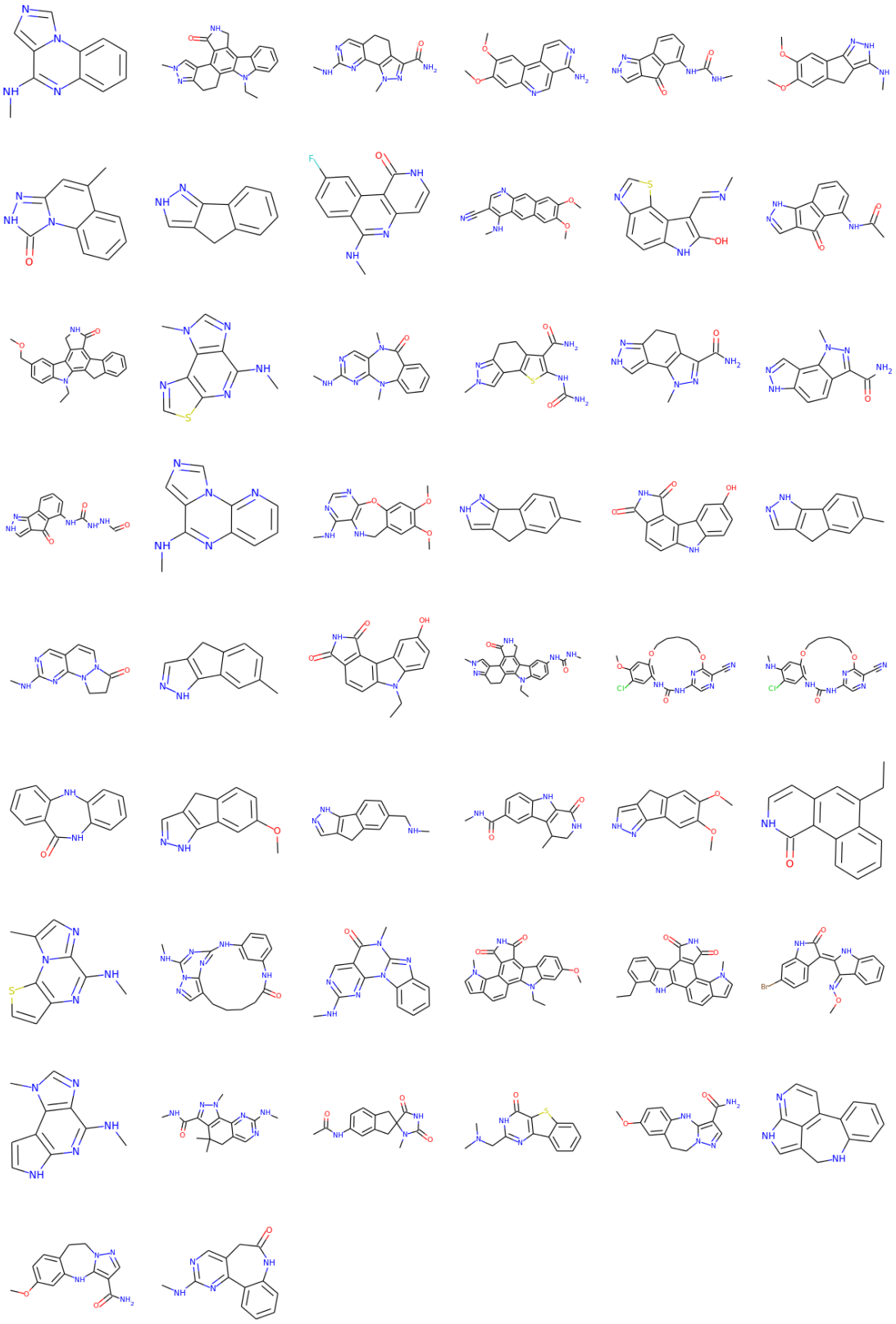


Figure S5. 2D representation of top 50 multicyclic putative hinge-binding fragments in kinase inhibitors.

References

[1] P. Kolb, C. B. Kipouros, D. Huang, A. Caflisch, *Proteins* **2008**, *73*(1), 11-18.

[2] H. T. Zhao, J. Dong, K. Lafleur, C. Nevado, A. Caflisch, *ACS Med. Chem. Lett.* **2012**, *3*(10), 834-838.

[3] H. Zhao, L. Gartenmann, J. Dong, D. Spiliotopoulos, A. Caflisch, *Bioorg. Med. Chem. Lett.* **2014**, *24*(11), 2493-2496.

[4] K. Lafleur, D. Huang, T. Zhou, A. Caflisch, C. Nevado, *J. Med. Chem.* **2009**, *52*(20), 6433-6446.

[5] K. Lafleur, J. Dong, D. Huang, A. Caflisch, C. Nevado, *J. Med. Chem.* **2013**, *56*(1), 84-96.