CHEMMEDCHEM
Chemistry & Drug Discovery

Supporting Information

Supplementary material

Table S1: Definition of unwanted groups

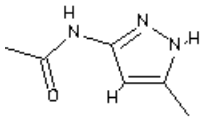
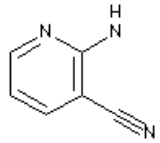
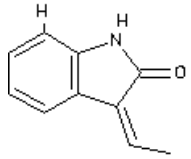
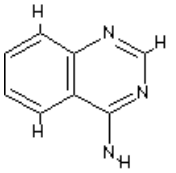
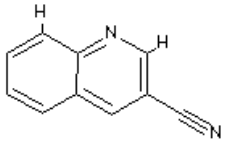
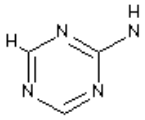
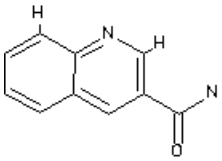
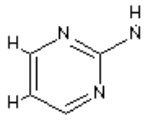
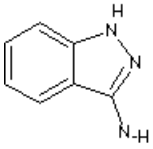
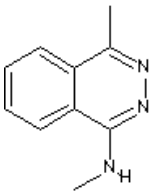
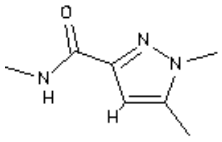
Name	Smarts pattern
> 2 ester groups	C(=O)O[C,H1].C(=O)O[C,H1].C(=O)O[C,H1]
2-halo pyridine	n1c([F,Cl,Br,I])cccc1
acid halide	C(=O)[Cl,Br,I,F]
acyclic C=C-O	C=[C!r]O
acyl cyanide	N#CC(=O)
acyl hydrazine	C(=O)N[NH2]
aldehyde	[CH1](=O)
Aliphatic long chain	[R0;D2][R0;D2][R0;D2][R0;D2]
alkyl halide	[CX4][Cl,Br,I]
amidotetrazole	c1nnnn1C=O
aniline	c1cc([NH2])ccc1
azepane	[CH2R2]1N[CH2R2][CH2R2][CH2R2][CH2R2][CH2R2]1
Azido group	N=[N+]=[N-]
Azo group	N#N
azocane	[CH2R2]1N[CH2R2][CH2R2][CH2R2][CH2R2][CH2R2][CH2R2]1
benzidine	[cR2]1[cR2][cR2]([Nv3X3,Nv4X4])[cR2][cR2][cR2]1[cR2]2[cR2][cR2][cR2]([Nv3X3,Nv4X4])[cR2][cR2]2
beta-keto/anhydride	[C,c](=O)[CX4,CR0X3,O][C,c](=O)
biotin analogue	C12C(NC(N1)=O)CSC2
Carbo cation/ anion	[C+,c+,C-,c-]
catechol	c1c([OH])c([OH,NH2,NH])ccc1
charged oxygen or sulfur atoms	[O+,o+,S+,s+]
chinone	C1(=[O,N])C=CC(=[O,N])C=C1
chinone	C1(=[O,N])C(=[O,N])C=CC=C1
conjugated nitrile group	C=[C!r]C#N
crown ether	[OR2,NR2]@[CR2]@[CR2]@[OR2,NR2]@[CR2]@[CR2]@[OR2,NR2]

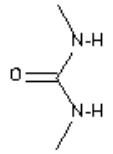
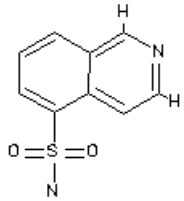
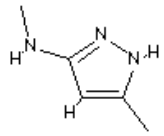
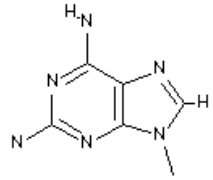
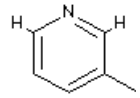
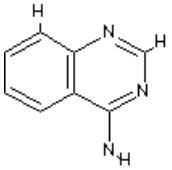
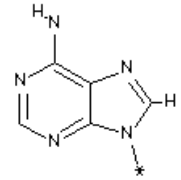
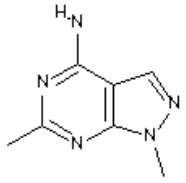
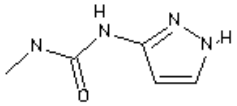
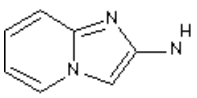
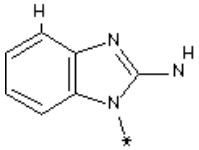
Name	Smarts pattern
cumarine	c1ccc2c(c1)ccc(=O)o2
cyanamide	N[CH2]C#N
cyanate / aminonitrile / thiocyanate	[N,O,S]C#N
cyanohydrins	N#CC[OH]
cycloheptane	[CR2]1[CR2][CR2][CR2][CR2][CR2][CR2]1
cycloheptane	[CR2]1[CR2][CR2]cc[CR2][CR2]1
cyclooctane	[CR2]1[CR2][CR2][CR2][CR2][CR2][CR2][CR2]1
cyclooctane	[CR2]1[CR2][CR2]cc[CR2][CR2][CR2]1
diaminobenzene	[cR2]1[cR2]c([N+0X3R0,nX3R0])c([N+0X3R0,nX3R0])[cR2][cR2]1
diaminobenzene	[cR2]1[cR2]c([N+0X3R0,nX3R0])[cR2]c([N+0X3R0,nX3R0])[cR2]1
diaminobenzene	[cR2]1[cR2]c([N+0X3R0,nX3R0])[cR2][cR2]c1([N+0X3R0,nX3R0])
diazo group	[N!R]=[N!R]
diketo group	[C,c](=O)[C,c](=O)
disulphide	SS
enamine	[CX2R0][NX3R0]
ester of HOBT	C(=O)Onnn
four member lactones	C1(=O)OCC1
halogenated ring	c1cc([Cl,Br,I,F])cc([Cl,Br,I,F])c1[Cl,Br,I,F]
halogenated ring	c1ccc([Cl,Br,I,F])c([Cl,Br,I,F])c1[Cl,Br,I,F]
heavy metal	[Hg,Fe,As,Sb,Zn,Se,se,Te,B,Si]
het-C-het not in ring	[NX3R0,NX4R0,OR0,SX2R0][CX4][NX3R0,NX4R0,OR0,SX2R0]
hydantoin	C1NC(=O)NC(=O)1
hydrazine	N[NH2]
hydroquinone	[OH]c1ccc([OH,NH2,NH])cc1
hydroxamic acid	C(=O)N[OH]
imine	C=[N!R]
imine	N=[CR0][N,n,O,S]
iodine	I

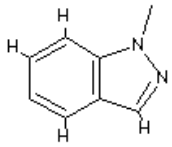
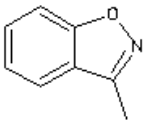
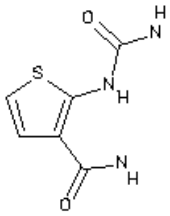
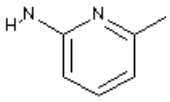
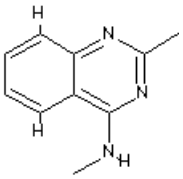
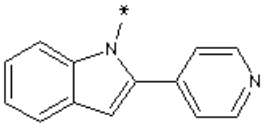
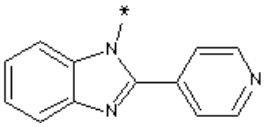
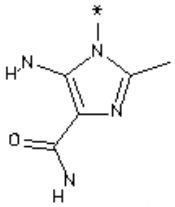
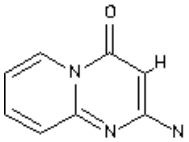
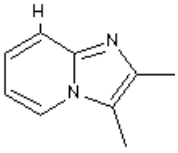
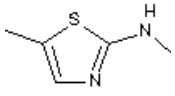
Name	Smarts pattern
isocyanate	N=C=O
isolated alkene	[$([CH2]),$([CH][CX4]),$(C([CX4])[CX4])]=[$([CH2]),$([CH][CX4]),$(C([CX4])[CX4])]
ketene	C=C=O
methylidene-1,3-dithiole	S1C=CSC1=S
Michael acceptor	C=!@CC=[O,S]
Michael acceptor	[$([CH]),$(CC)]#CC(=O)[C,c]
Michael acceptor	[$([CH]),$(CC)]#CS(=O)(=O)[C,c]
Michael acceptor	C=C(C=O)C=O
Michael acceptor	[$([CH]),$(CC)]#CC(=O)O[C,c]
N oxide	[NX2,nX3][OX1]
N-acyl-2-amino-5-mercapto-1,3,4-thiadiazole	s1c(S)nnc1NC=O
N-C-halo	NC[F,Cl,Br,I]
N-halo	[NX3,NX4][F,Cl,Br,I]
N-hydroxyl pyridine	n[OH]
nitro group	[N+](=O)[O-]
N-nitroso	[#7]-N=O
oxime	[C,c]=N[OH]
oxime	[C,c]=NOC=O
Oxygen-nitrogen single bond	[OR0,NR0][OR0,NR0]
perfluorinated chain	[CX4](F)(F)[CX4](F)F
peroxide	OO
phenol ester	c1ccccc1OC(=O)[#6]
phenyl carbonate	c1ccccc1OC(=O)O
phosphor	P
phthalimide	[cR,CR]~C(=O)NC(=O)~[cR,CR]
Polycyclic aromatic hydrocarbon	a1aa2a3a(a1)A=AA=A3=AA=A2
Polycyclic aromatic hydrocarbon	a21aa3a(aa1aaaa2)aaaa3

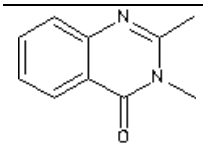
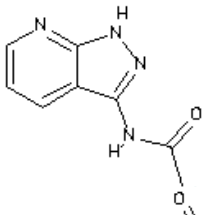
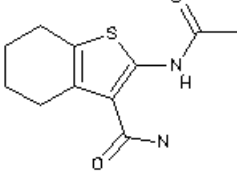
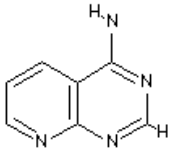
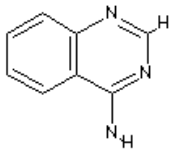
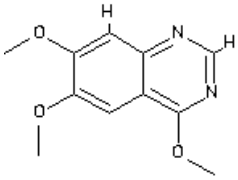
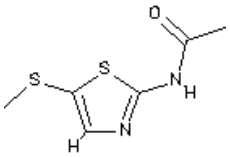
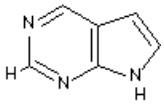
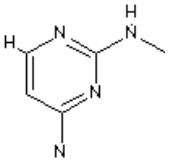
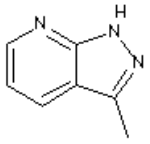
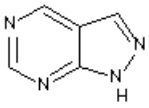
Name	Smarts pattern
Polycyclic aromatic hydrocarbon	a31a(a2a(aa1)aaaa2)aaaa3
polyene	[CR0]=[CR0][CR0]=[CR0]
quaternary nitrogen	[s,S,c,C,n,N,o,O]~[nX3+,NX3+](~[s,S,c,C,n,N])~[s,S,c,C,n,N]
quaternary nitrogen	[s,S,c,C,n,N,o,O]~[n+,N+](~[s,S,c,C,n,N,o,O])(~[s,S,c,C,n,N,o,O])~[s,S,c,C,n,N,o,O]
quaternary nitrogen	[*]=[N+]=[*]
saponine derivative	O1CCCCC1OC2CCC3CCCCC3C2
silicon halogen	[Si][F,Cl,Br,I]
stilbene	c1ccccc1C=Cc2ccccc2
sulfinic acid	[SX3](=O)[O-,OH]
Sulfonic acid	[C,c]S(=O)(=O)O[C,c]
Sulfonic acid	S(=O)(=O)[O-,OH]
sulfonyl cyanide	S(=O)(=O)C#N
sulfur oxygen single bond	[SX2]O
sulphate	OS(=O)(=O)[O-]
sulphur nitrogen single bond	[SX2H0][N]
Thiobenzothiazole	c12ccccc1(SC(S)=N2)
thiobenzothiazole	c12ccccc1(SC(=S)N2)
Thiocarbonyl group	[C,c]=S
thioester	SC=O
thiol	[S-]
thiol	[SH]
Three-membered heterocycle	*1[O,S,N]*1
triflate	OS(=O)(=O)C(F)(F)F
triphenyl methyl-silyl	[SiR0,CR0](c1ccccc1)(c2ccccc2)(c3ccccc3)
triple bond	C#C

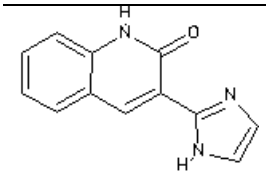
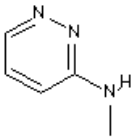
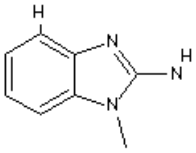
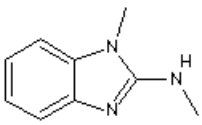
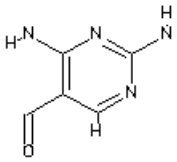
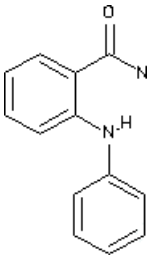
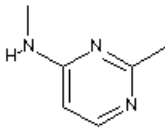
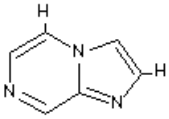
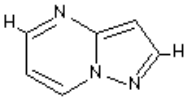
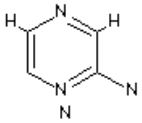
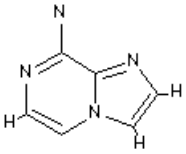
Table S2: Kinase inhibitor core fragments and their occurrence in the filtered library.

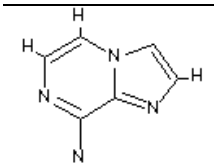
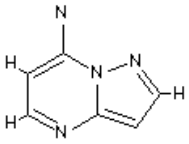
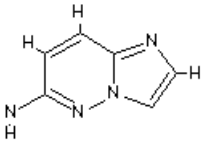
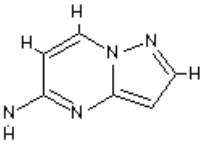
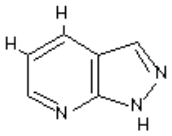
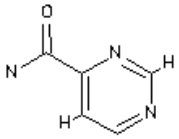
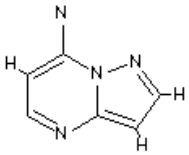
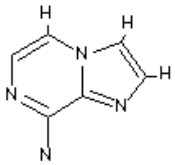
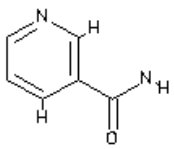
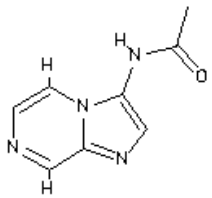
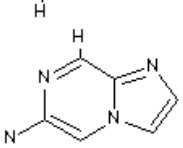
Core fragment	Hits
	51
	254
	633
	103
	2
	67
	50
	447
	5
	46
	60

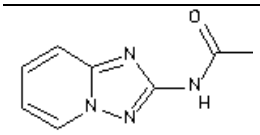
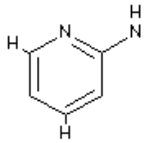
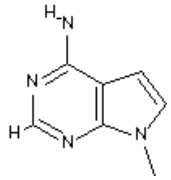
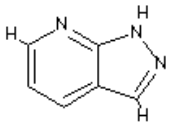
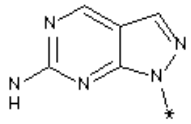
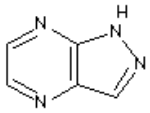
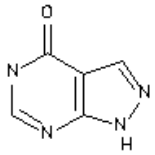
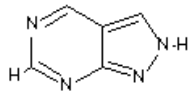
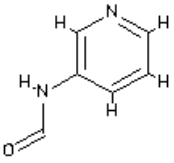
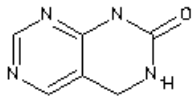
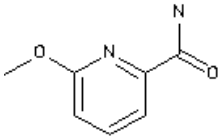
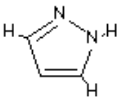
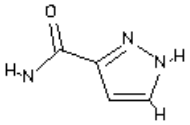
Core fragment	Hits
	6699
	2
	71
	2
	2219
	104
[a]	
	91
	3
	13
	1
	600

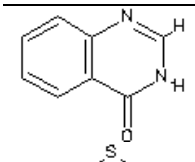
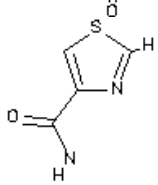
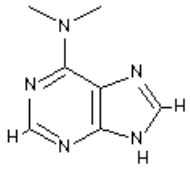
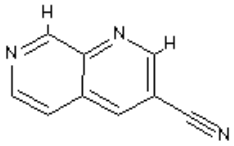
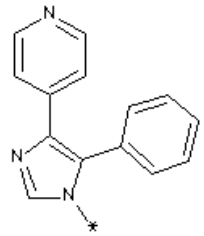
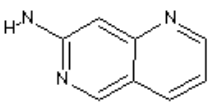
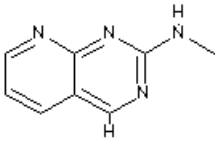
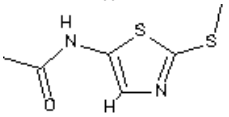
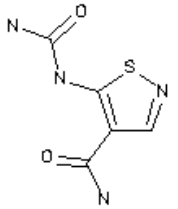
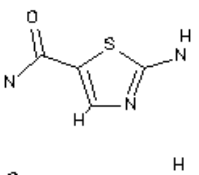
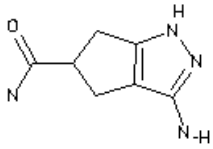
Core fragment	Hits
	87
	490
	4
	878
	107
	1
	5
	2
	7
	115
	1518

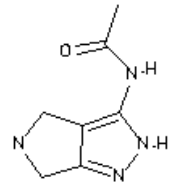
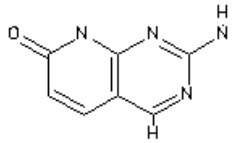
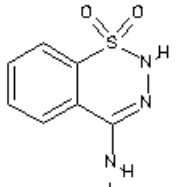
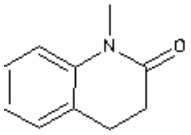
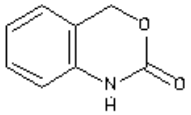
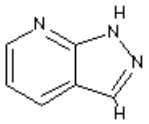
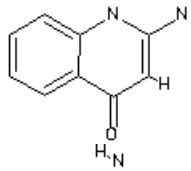
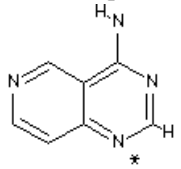
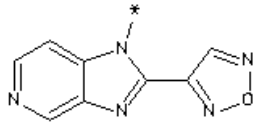
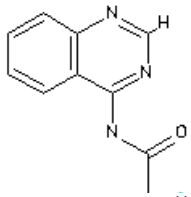
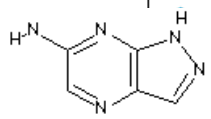
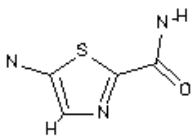
Core fragment	Hits
	334
	3
	457
	7
	107
	1
	36
	2
	20
	110
	38

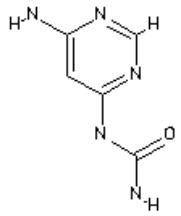
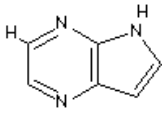
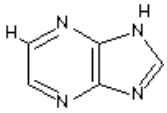
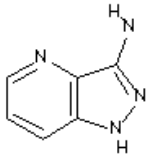
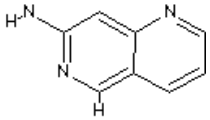
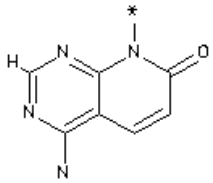
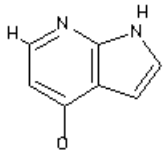
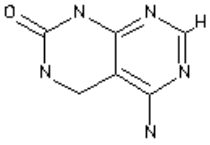
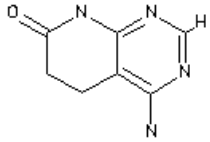
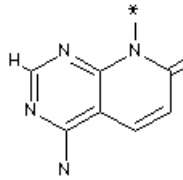
Core fragment	Hits
	2
	42
	302
	263
	1
	13
	165
	187
	203
	442
	25

Core fragment	Hits
	136
	68
	173
	35
	6
	129
	40
	35
	1245
	62
	1

Core fragment	Hits
	69
	3403
	118
	1
	48
	1
	115
	41
	876
	4
	1
	50
	27

Core fragment	Hits
	14
	5
	43
	0
	0
	0
	0
	0
	0
	0
	0

Core fragment	Hits
	0
	0
	0
	0
	0
	0
	0
	0
	0
	0
	0
	0

Core fragment	Hits
	0
	0
	0
	0
	0
	0
	0
	0
	0
	0

[a] * indicates any atom, including H