Supporting Information

Assessment of the commercially available chemical space for the using in ^{19}F NMR FAXS method: A Enamine Ltd. case

O. V. Oksiuta[1,2], Ya. I. Filatov[2,3]

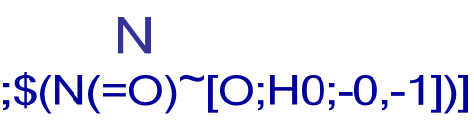
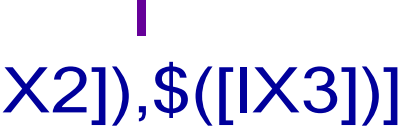
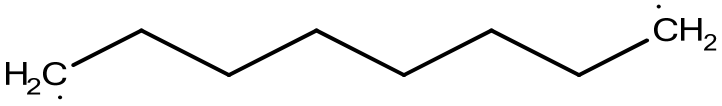
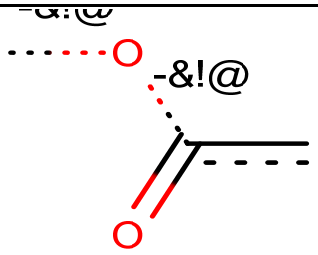
[1]Institute of Organic Chemistry of the National Academy of Sciences of Ukraine, 5, Academician Kukhar str., 02660 Kyiv, Ukraine

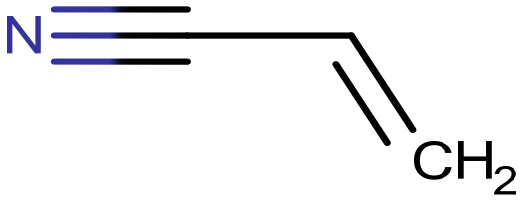
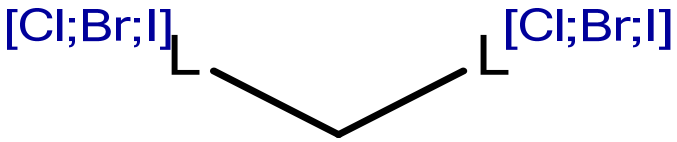
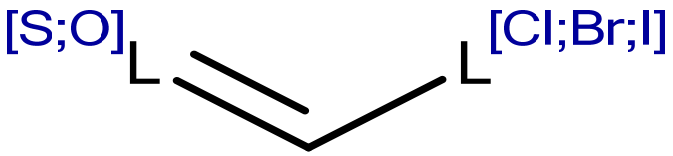
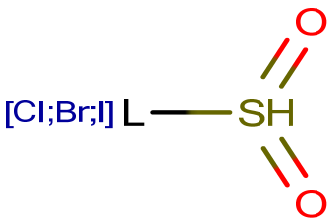
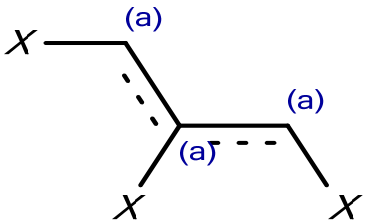
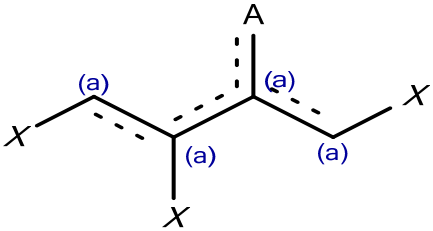
[2]Chemspace LLC, 85, Winston Churchill str., 02094 Kyiv, Ukraine

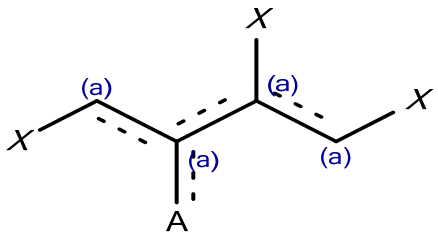
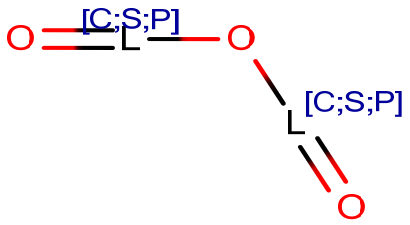
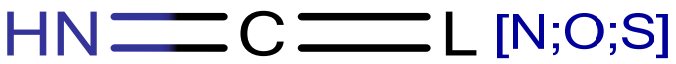
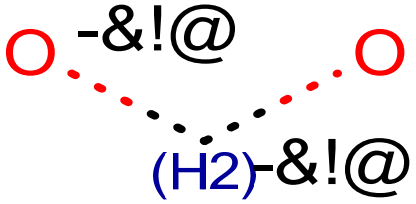
[3]V.N. Karazin Kharkiv National University, 4, Svobody Sq., 61022 Kharkiv, Ukraine

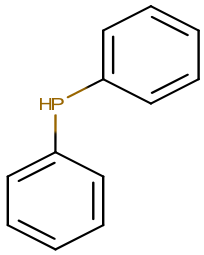
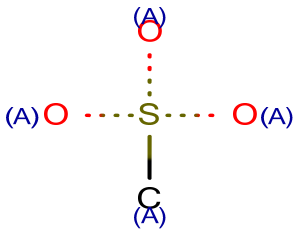
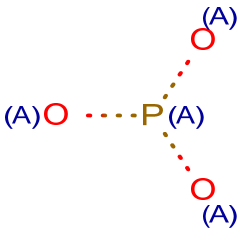
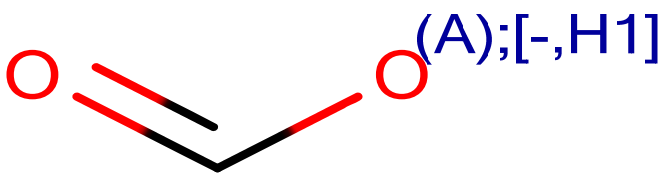
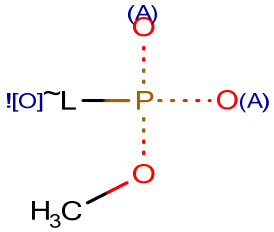
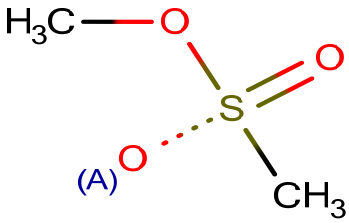
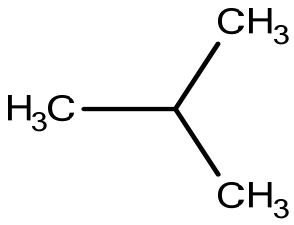
Table of Contents

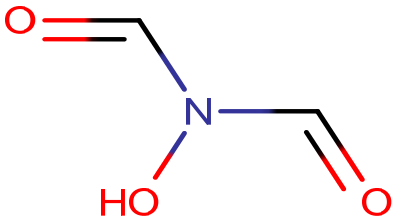
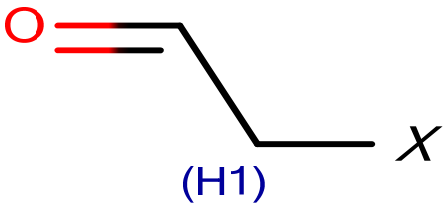
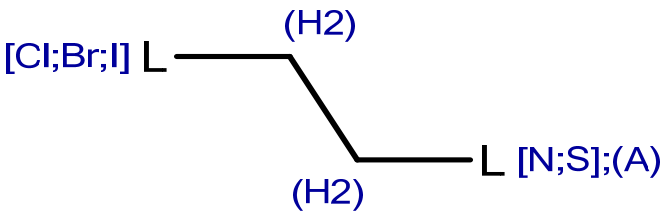
Table S1. List of SMARTS used in REOS filters

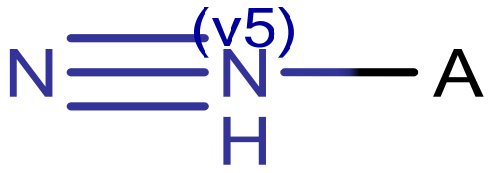
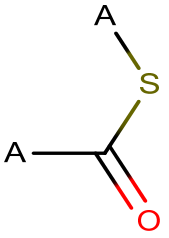
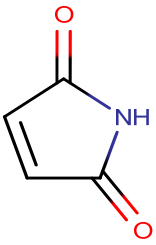
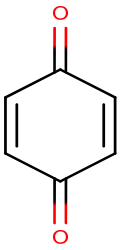
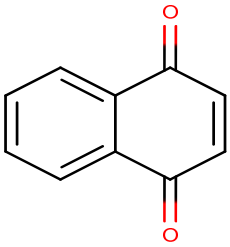
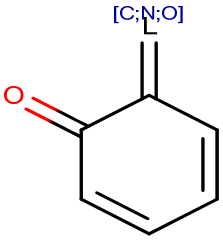
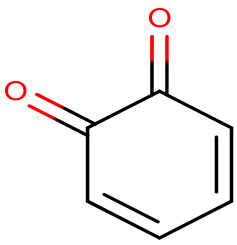
#	SMARTS	Structure	Name
1	[Cl,Br,I]		non-Fluorine halogens
2	[N;+0,+1;$(N(=O)~[O;H0;-0,-1])]		nitro groups
3	[I;$([IX2]),$([IX3])]		multivalent iodine
4	[CH2][CH2][CH2][CH2][CH2][CH2][CH2][CH2]		long aliphatic
5	[!#1;!#3;!#6;!#7;!#8;!#9;!#11;!#12;!#15;!#16;!#17;!#18;!#20;!#35;!#53]		atom types other than H,C,O,N,S,P,F,Cl,Br,I,Na,K,Mg,Ca,Li
6	[Cl,Br,I][CH2]		primary alkyl halides
7	[#6][#6](=O)-&!@O-&!@[#6]		esters

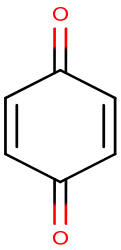
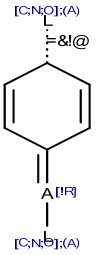
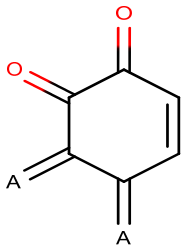
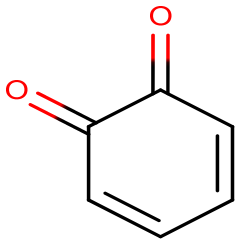
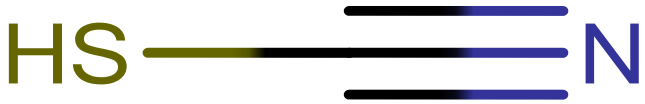
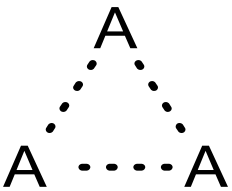
8	C=CC#N		alpha beta unsaturated nitriles
9	[C;H1]=O		aldehydes
10	[Cl,Br,I]-C-[Cl,Br,I]		gem-dihalo compounds
11	[S,O]=C-[Cl,Br,I]		acid halides
12	S(=O)(=O)-[Cl,Br,I]		sulfonyl halides
13	c([F,Cl,Br,I])c([F,Cl,Br,I])c([F,Cl,Br,I])		perhalo compounds
14	c([F,Cl,Br,I])c([F,Cl,Br,I])c(*)c([F,Cl,Br,I])		perhalo compounds

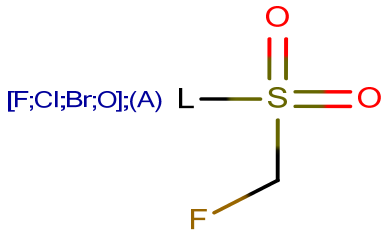
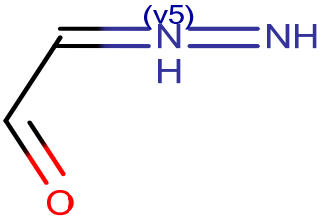
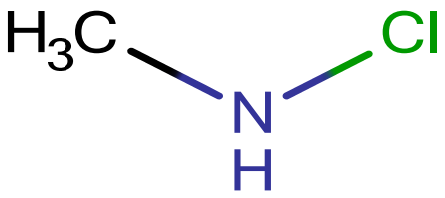
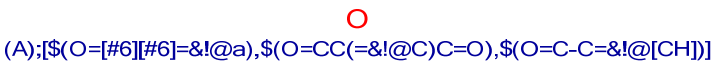
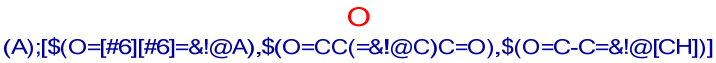
15	c([F,Cl,Br,I])c(*)c([F,Cl,Br,I]) c([F,Cl,Br,I])		perhalo compounds
16	O=[C,S,P]-O-[C,S,P]=O		anhydrides
17	C#N		nitriles
18	N=C=[N,O,S]		diimides isocyanates isothiocyanates
19	[Cl,Br,I]~[S,P,Si]		halogen bonded to S,P,or Si
20	F		fluorine
21	O-&!@[CH2]-&!@O		acetals

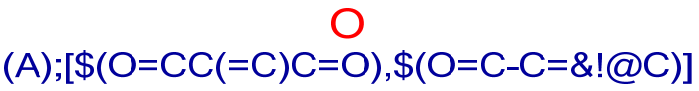
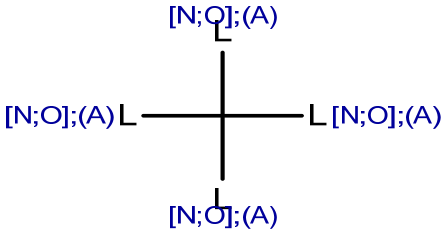
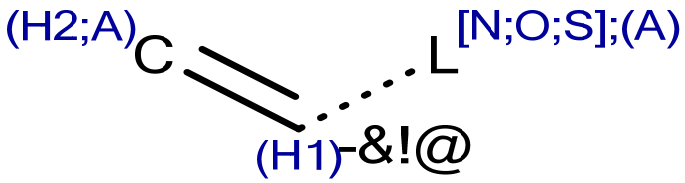
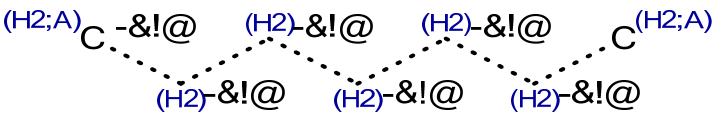
22	c1c(cccc1)Pc2ccccc2		gem-diphenyl phosphorous
23	S(C)(~O)(~O)~O		SO3
24	P(~O)(~O)~O		PO3
25	C(=O)[O-,OH]		carboxylic acids
26	P(~O[#6])(~O)(~O)[!#8]		Alkyl esters of phosphonic acids
27	[#6]S(=O)(~O)O[#6]		Alkyl esters of sulphonic acids
28	C([CH3])([CH3])([CH3])		t-butyl groups

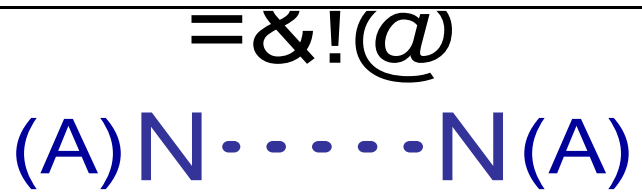
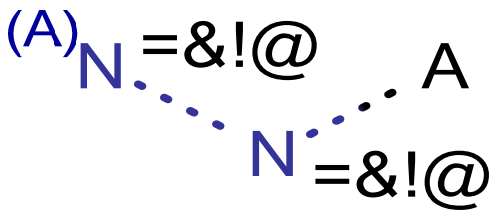
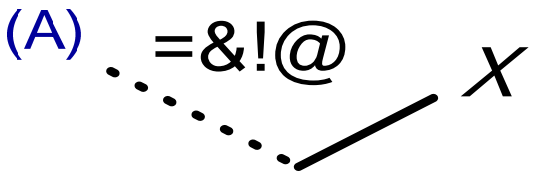
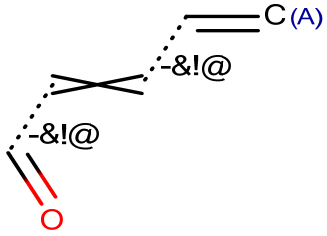
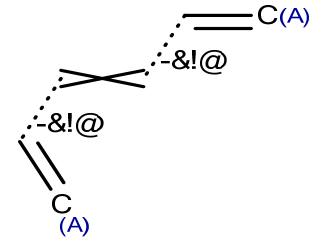
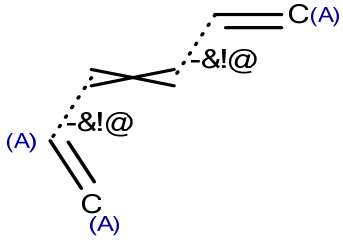
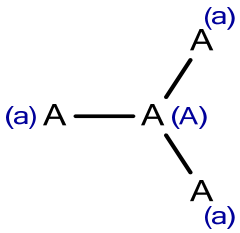
29	[#16]-[#16]		S-S
30	S=P~*		P=S
31	[F,Cl,Br,I]-[N,O,S]		Halogen-N,O,S
32	ON(C=O)C=O		hydroxy imines & imide esters
33	[F,Cl,Br,I]-[CH]-C=O		alpha halo ketones
34	S		sulfurs
35	[N,S][CH2][CH2][Cl,Br,I]		N and S mustards

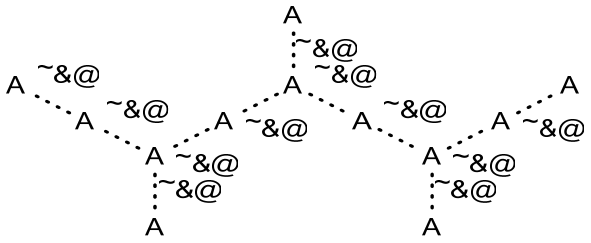
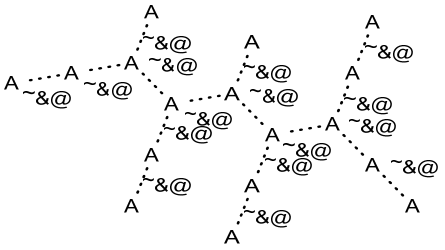
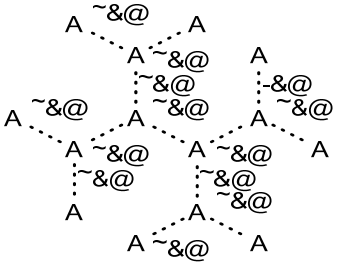
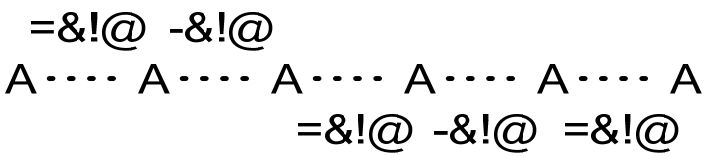
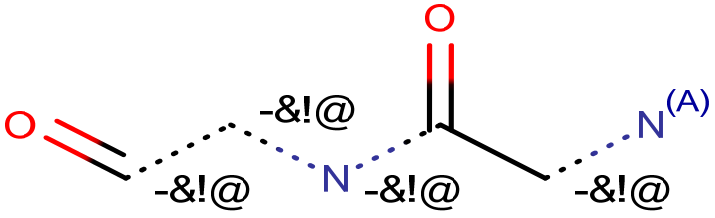
36	N#N*		diazo compounds
37	*-S-C(=O)-*		thioesters
38	O=C1NC(=O)C=C1		maleimides
39	C1C(C=CC(C=1)=O)=O		quinones
40	c21ccccc1C(=O)C=CC2=O		quinones
41	C1(C=CC=CC1=O)=[C,N,O]		quinones
42	C1=CC=CC(C1=O)=O		quinones

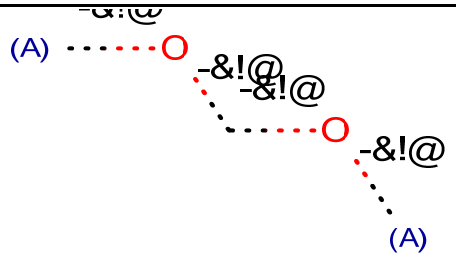
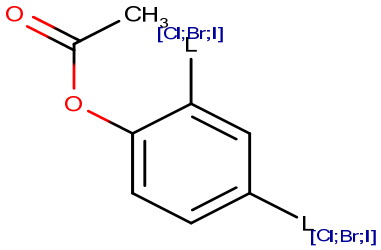
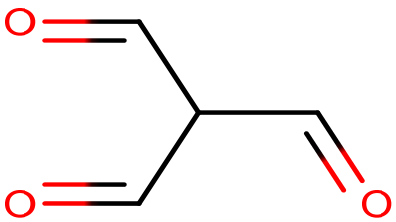
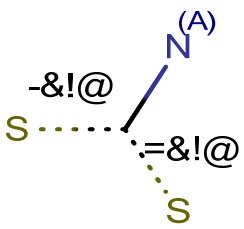
43	C1C(C:CC(C:1)=O)=O		quinones
44	[O;$(O=C1c2ccccc2C(=[C,O,N])c3ccccc13),$(O=C1C=CC(=[O,C,N])C=C1)]		quinones
45	C1=CC(C=CC1=&!@[C,N,O])=[!R][C,N,O]		quinones
46	C1=CC(C(C(C1=O)=O)=*)=*		quinones
47	C1:CC:CC(C1=O)=O		quinones
48	S-C#N		thiocyanates
49	*1~*~*~1		three membered rings

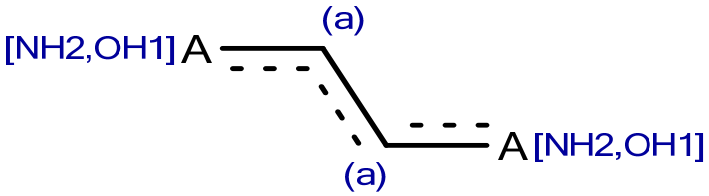
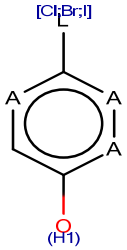
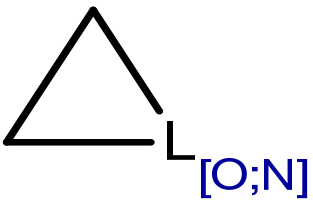
50	O-O		peroxides
51	FCS(=O)(=O)[F,Cl,Br,O]		triflates
52	N-[F,Cl,Br,I]		N-haloamines
53	N=N=C-C=O		diazomethylketones
54	C-N-Cl		N-chloramines
55	[O;$(O=[#6][#6]=&!@a),$(O=CC(=&!@C)C=O),$(O=C-C=&!@[CH])]		alpha beta unsat ketones aromatic
56	[O;$(O=[#6][#6]=&!@A),$(O=CC(=&!@C)C=O),$(O=C-C=&!@[CH])]		alpha beta unsat ketones aliphatic

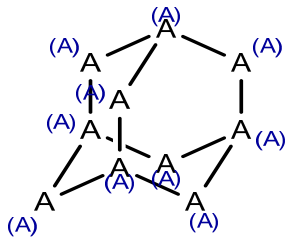
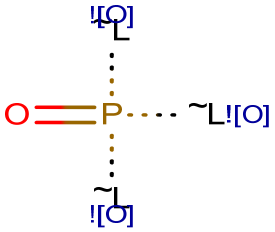
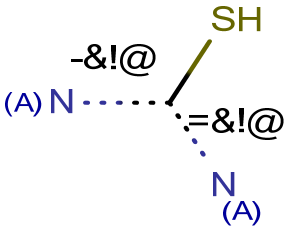
57	[O;$(O=CC(=C)C=O),$(O=C-C=&!@C)]		Michael acceptors
58	C=C=C		allenes
59	C([N,O])([N,O])([N,O])[N,O]		carbon attached to 4 N or O
60	c[OH]		aromatic hydroxyls
61	[CH2]=[CH]-&!@[N,O,S]		terminal vinyl groups
62	[CH2]-&!@[CH2]-&!@[CH2]-&!@[CH2]-&!@[CH2]-&!@[CH2]		long aliphatic
63	[Si]O		siloxanes

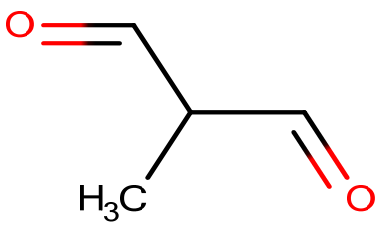
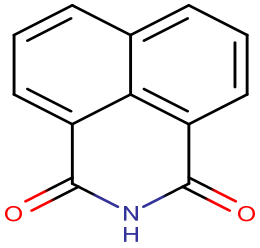
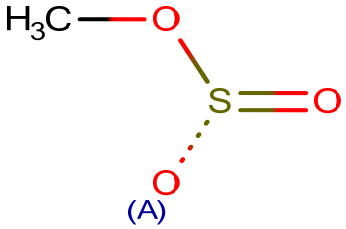
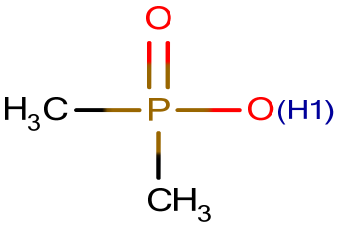
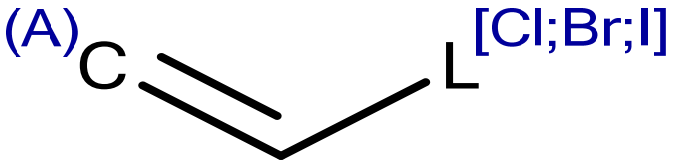
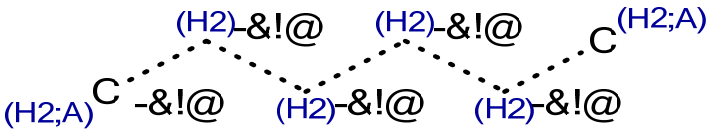
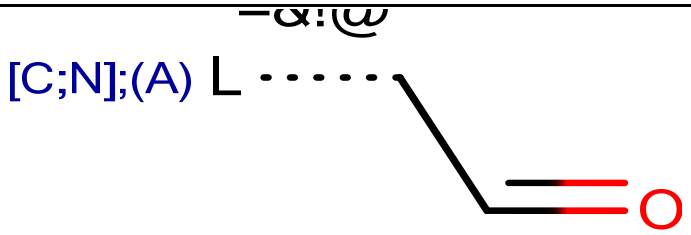
64	N=&!@N		azides
65	N=&!@N=&!@*		diazonium
66	C=&!@C[F,Cl,Br,I]		halogen alpha to a double bond
67	C=C-&!@C=C-&!@C=O		conjugated double bonds
68	C=C-&!@C=C-&!@C=C		conjugated double bonds
69	C=C-&!@C=C-&!@C:C		conjugated double bonds
70	A(a)(a)a		triphenyl anything

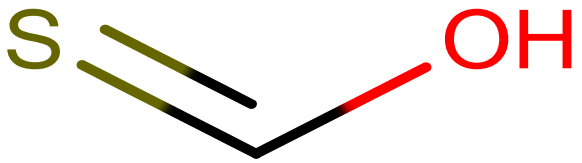
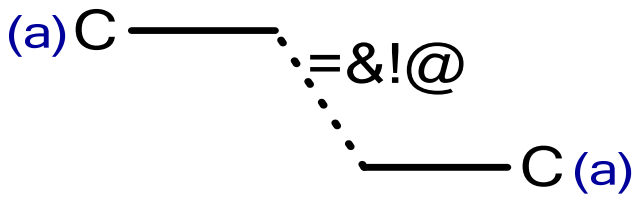
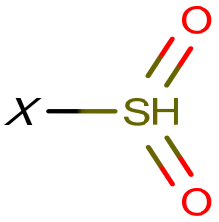
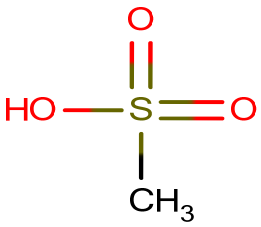
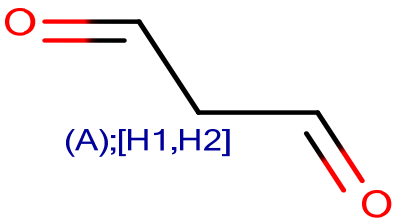
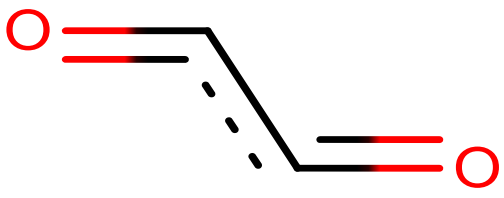
71	*~&@*(~&@*~&@*(~&@*)~&@*~&@*)~&@*~&@*(~&@*)~&@*~&@*		systems with 4 fused rings
72	*(~&@*)~&@*(~&@*(~&@*~&@*)~&@*(~&@*)~&@*(~&@*~&@*)~&@*(~&@*~&@*)~&@*)~&@*~&@*		systems with 4 fused rings
73	*(~&@*)(~&@*(~&@*(~&@*(~&@*)~&@*)~&@*(~&@*)~&@*)~&@*(-&@*)~&@*)~&@*		systems with 4 fused rings
74	*=&!@*-&!@*=&!@*-&!@*=&!@*		3 conjugated acyclic double bonds
75	[SX3](=O)		sulfoxides
76	[SH]		thiols
77	O=C-&!@C-&!@N-&!@C(=O)C-&!@N		dipeptides

78	C-&!@O-&!@C-&!@O-&!@C		acyclic ketals
79	c1(:c:c(:c:c:c:1OC(=O)[#6])[Cl,Br,I])[Cl,Br,I]		active_esters
80	N-C-[F,Cl,Br,I]		N-C-haloamines
81	cBr		aromatic bromide
82	n-O		pyridine n-oxide
83	C(C=O)(C=O)(C=O)		carbon alpha to 3 carbonyls
84	NC(=&!@S)(-&!@S)		thioguanadine

85	C([OH])C#C		hydroxy alpha to alkyne
86	C#C		more than one alykne
87	[NH2,OH]cc[NH2,OH]		catechol
88	c1([Cl,Br,I])aac([OH])ca1		halogen para to phenol
89	[r8,r9,r10,r11,r12,r13,r14,r15,r16,r17,r18,r19,r20]	A [r8,r9,r10,r11,r12,r13,r14,r15,r16,r17,r18,r19,r20]	ring size larger than 7
90	[a;$(aa[R3](a)a[R3](a)a),$(aa[R3](a)aa[R3](a)a)]	A (a);[$(aa[R3](a)a[R3](a)a),$(aa[R3](a)aa[R3](a)a)]	three fused aromatic rings
91	C1[O,N]C1		epoxides

92	A1A2AA3AA1AA(A2)A3		adamantane
93	Br	HBr	bromine
94	[O+,o+]	A^+ [O,o]	oxonium
95	P(~[!#8])(~[!#8])(~[!#8])=O		phosphine oxides and related non-phosphates
96	N-[OH]		N-OH
97	S-C(=&!@N)-&!@N		thioguanadine
98	C=&!@C	=&!@ (A) (A)	acyclic olefins

99	C(=O)C([#6])C(=O)		tertiary carbon between two carbonyls
100	O=C1NC(=O)c2cccc3cccc1c23		1,8-Naphthalimide
101	S(=O)(~O)O[#6]		sulphonate esters
102	[#6]P(=O)([OH])[#6]		CP(=O)(OH)C
103	C=C-[Cl,Br,I]		vinyl halides
104	[CH2]-&!@[CH2]-&!@[CH2]-&!@[CH2]-&!@[CH2]		hexyl chain
105	C(=O)C=&!@[C,N]		acyclic double bond alpha to carbonyl

106	C(=S)O		thioester
107	cC=&!@Cc		stilbene
108	N-N=O		nitrosoamines
109	S(=O)(=O)[F,Cl,Br,I]		sulfonyl halides
110	[#6]S(=O)(=O)O		sulfonates
111	C(=O)[CH,CH2]C(=O)		1,3 diketones
112	[#6](=O)[#6](=O)		alpha diketones

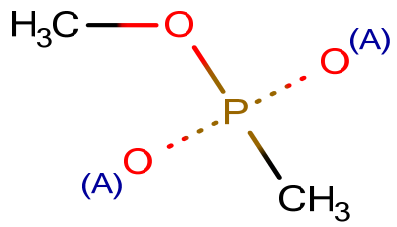
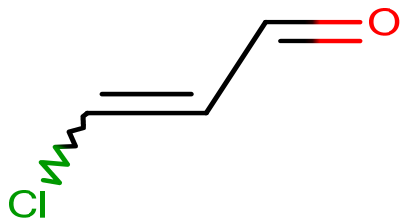
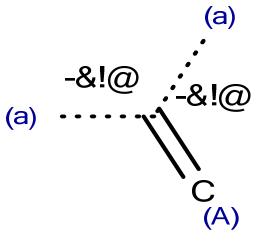
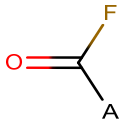
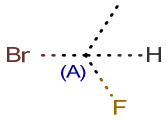
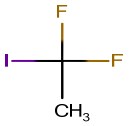
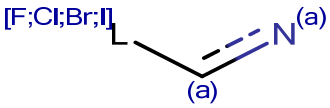
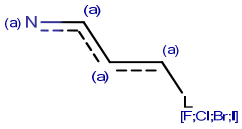
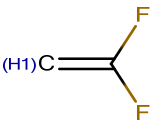
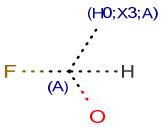
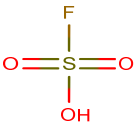
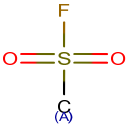
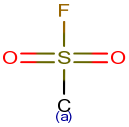
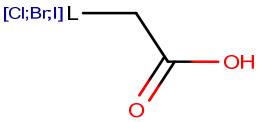
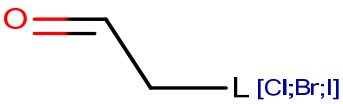
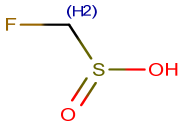
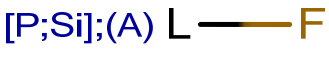
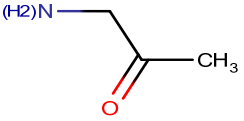
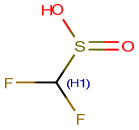
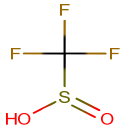
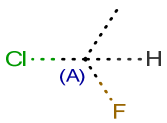
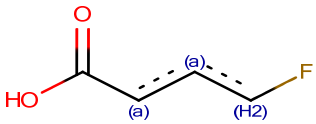
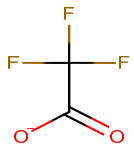
113	[#6]P(~O)(~O)O[#6]		phosphonate esters
114	C(=O)C=CCl		vinylagous acid chlorides
115	[#6]=[NH]		primary imines
116	[OH]cccc[OH]		para phenols
117	c-&!@C(=C)-&!@c		vinyl group between two aromatic carbons

Table S2. List of SMARTS used for in-house Custom Filtering

#	SMARTS	Comment	Structure
1	[#6]F.[#6]F	FCFC	
2	F[#6](-*)=O	C(=O)F	
3	[H]~[#6;A](~[#6])(~F)~Br	CCHFBr	
4	[#6]S(F)(=O)=O	SO2F	
5	[#6]C(F)(F)I	CCF2I	
6	[#9,#17,#35,#53]-[#6;a]:[#7;a]	ortho-Arylators	
7	[#9,#17,#35,#53]-[#6;a]:[#6;a]:[#6;a]:[#7;a]	para-Arylators	
8	F[#6](F)=[#6H1]	CH=CF2	
9	[H]~[#6;A](~[#6;A;H0X3])(~[#8])~F	C-CHF-O	
10	[#8]S(F)(=O)=O	FluoroSulfates	
11	[#6;A]S(F)(=O)=O	sulfonyl_fluoride_(Alyph)	
12	[#6;a]S(F)(=O)=O	sulfonyl_fluoride_(Aryl)	

13	[F-]	F-	
14	F[H]	HF	
15	[#8]-[#6](=O)-[#6]-[#17,#35,#53]	alpha-Hal-Carb acid	
16	[#17,#35,#53]-[#6]-[#6]=O	alpha-Hal-Carbonyls	
17	[#8]S(=O)[#6H2]F	O-S(=O)CH2F	
18	F[#15,#14;A]	(P\Si)-F	
19	[#6]-[#6](=O)-[#6]-[#7H2]	alpha-amino-Ketons	
20	[#8]S(=O)[#6H1](F)F	O-S(=O)CF2	
21	[#8]S(=O)[#6](F)(F)F	O-S(=O)CF3	
22	[H]~[#6;A](~[#6])(~F)~Cl	CCHClF	
23	[#8]-[#6](=O)-[#6;a][#6;a][#6H2]F	Ortho-Carboxylic-CH2F	
24	[N+]	Quaternary and pyridinium salts	

25	[#8-]-[#6](=O)C(F)(F)F	TFA salts	
26	[CX1-]#[NX2+]	Isonitriles