Supporting information for: Knodle, a Support Vector Machines-based automatic perception of organic molecules from 3D coordinates

Maria Kadukova[†] and Sergei Grudinin[*,‡,¶,§]

† *Moscow Institute of Physics and Technology, Dolgoprudniy, Russia*

‡ *Univ. Grenoble Alpes, LJK, F-38000 Grenoble, France*

¶ *CNRS, LJK, F-38000 Grenoble, France*

§ *Inria, France*

E-mail: Sergei.Grudinin@inria.fr

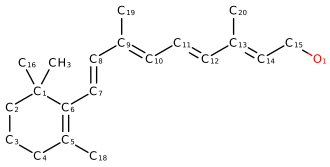
Contents

Table S1: Structure perception errors obtained on the Labute's benchmark set by the *Knodle*, fconv, NAOMI, and I-interpret algorithms.

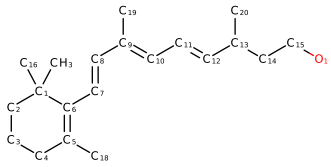
PDB ID(s)	Tool	Error description
1AAQ	*Knodle*	As C5 has a non-planar geometry, it is recognized as sp3 instead of sp2
	fconv	
	I-interpret	As C5 has a non-planar geometry, it is recognized as sp3 instead of sp2
1AQB	*Knodle*	Wrong C13 hybridization, thus missing C13-C14 double bond (Figs. S1a-S1b).
	fconv	
3FX2	*Knodle*	C2'-O2' bond is wrongly recognized as a double one due to a short length (1.37 Å) and a planar geometry of C2' atom.
	fconv	
	NAOMI	
	I-interpret	
8XIA	*Knodle*	O1-C1 bond length is 1.418 Å, it is not recognized as a double bond.
	fconv	
	NAOMI	
	I-interpret	
1NNB	*Knodle*	Wrong hybridization of the almost planar C2 atom (Figs. S1c-S1d).
1TLP, 2TMN	fconv	N should have the type of N.pl3, not N.3 because of a deprotonated PO_3 group nearby.
	NAOMI	
	I-interpret	
2R04	fconv	Wrong hybridizations of O1A, C4A, C5A: sp2 instead of sp3 (Figs. S1h-S1i).
	NAOMI	
	I-interpret	
1HTG	fconv	Pdb file contains two overlapping ligands with the same name, but different ids. *Knodle* and NAOMI separate them, while fconv shows it as a one ligand, which is chemically impossible.
4GR1 (RGS)	I-interpret	C4', a neighbour of an amide nitrogen N4' is sp2 instead of being sp3 and forms a double bond with N4' (Figs. S1j-S1k).
4FAB (FDS)	fconv	O3 has sp2 hybridization instead of sp3, consequently, aromatics is lost in C4-C5-C6-C7-C8-C9 ring (Figs. S1l-S1m).

	NAOMI	
1CPS	I-interpret	Aromatic properties of the ring CE2-CD2-CG-CD1-CE1-CZ are not recognized. It may happen due to long bonds in the ring (1.62 Å between CD1 and CE1).
1MNC	I-interpret	O1 should be sp3 instead of sp2, N1 should not form a double bond with O1.
5TLN	*Knodle*	The fast SVM kernels make the following error: almost planar CA2 has sp2 hybridization istead of sp3, which consequently causes sp2 hybrisization of CB2. (Fig. S1e)
	NAOMI	CA2 is sp3 instead of sp2, which causes changes in C1-O1 bond order (Fig. S1e). We do not consider it to be an error, although the most probable tautomeric form with sp3-hybridized O1 would be the one with C1-N1 double bond, not C1-CA2.
1BZM	NAOMI	Wrong bonds assignment: C2, N2 should form a double bond as a part of a heteroaromatic thiadiazole ring, there should be no double bond between C2 and N4 (Fig. S1n).
fconv also makes errors in all the HEM ligands (structures obtained by other packages are correct):		
7CAT	fconv	CHC in the porphyrin ring should be sp2 (Figs. S1q, S1p).
1PHF, 1PHE	fconv	CAB should be sp2 (Figs. S1s, S1p).
1MBI	fconv	C3A, C2A in the porphyrin ring should be sp2 (Figs. S1r, S1p).

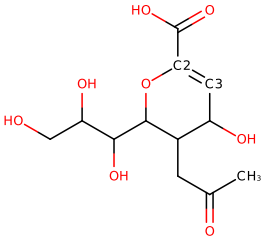
Figure S1: Structure perception errors in the Labute's benchmark set. Four methods are assessed, *Knodle*, fconv, NAOMI, and I-interpret. Chemical schemes from the RCSB PDB database are shown as the reference.



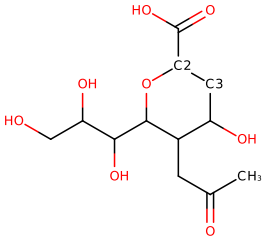
(a) 1aqb: reference



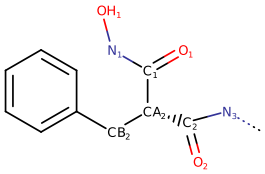
(b) 1aqb: *Knodle*



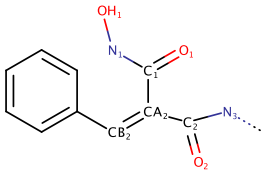
(c) 1nnb: reference



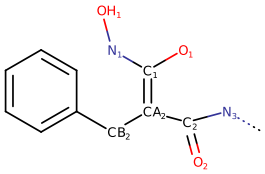
(d) 1nnb: *Knodle*



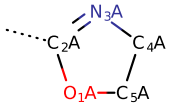
(e) 5tln: reference



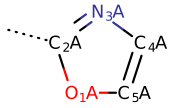
(f) 5tln: *Knodle (fast version)*



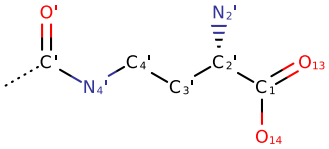
(g) 5tln: NAOMI



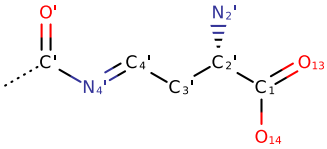
(h) 2r04: reference



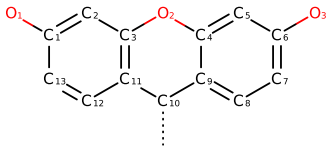
(i) 2r04: NAOMI, fconv, I-interpret



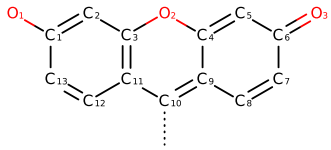
(j) 4gr1: reference



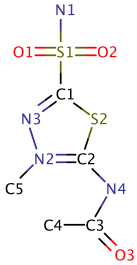
(k) 4gr1: I-interpret



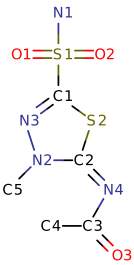
(l) 4fab: reference



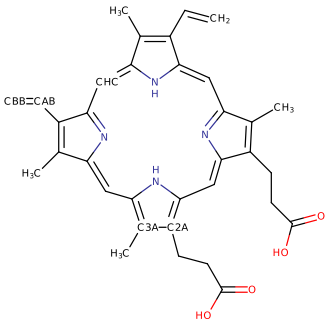
(m) 4fab: fconv and NAOMI



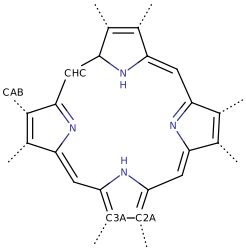
(n) 1bzm: reference



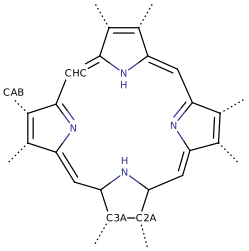
(o) 1bzm: NAOMI



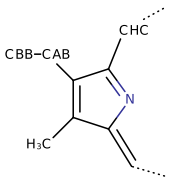
(p) HEM ligand: reference



(q) 7cat: fconv

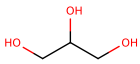


(r) 1mbi: fconv

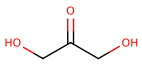


(s) 1phf, 1phe: fconv

Figure S2: Structure perception errors in the LigandExpo dataset. Chemical schemes of the 25 most frequent ligands in which *Knodle* perception errors occur are shown.



GOL (reference)



GOL (*Knodle*)

ACE (reference)

ACE (*Knodle*)

CMO (reference)

CMO (*Knodle*)

OXY (reference)

OXY (*Knodle*)

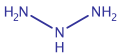
CYN (reference)

CYN (*Knodle*)

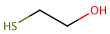
NO2 (reference)

NO2 (*Knodle*)

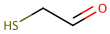
AZI (reference)



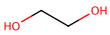
AZI (*Knodle*)



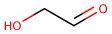
BME (reference)



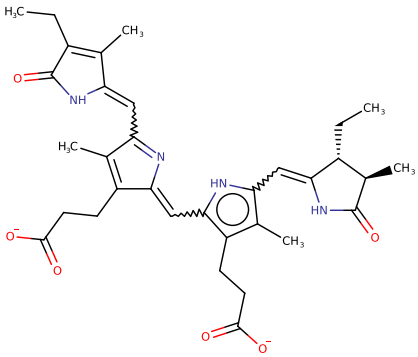
BME (*Knodle*)



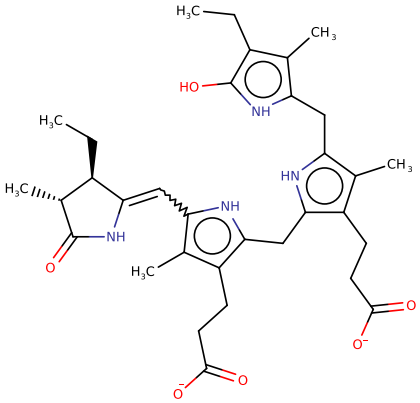
EDO (reference)



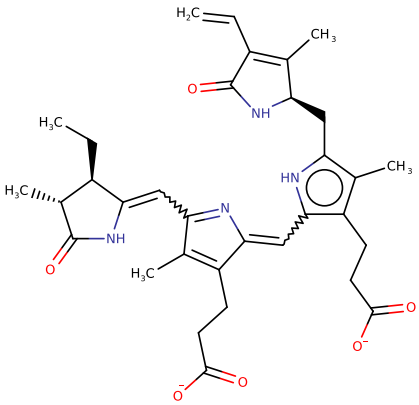
EDO (*Knodle*)



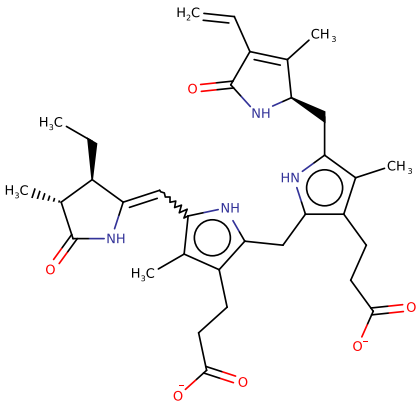
CYC (reference)



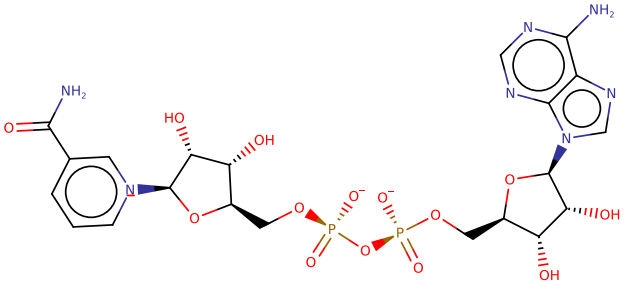
CYC (*Knodle*)



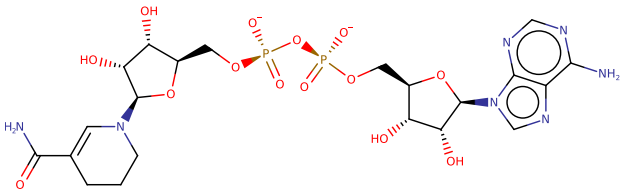
PEB (reference)



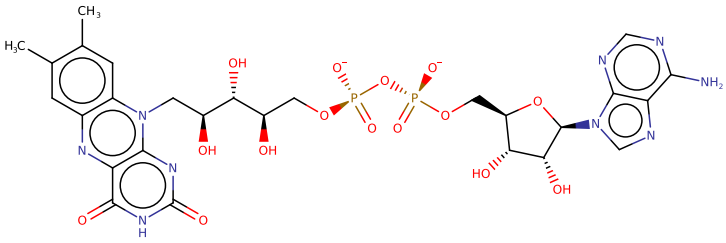
PEB (*Knodle*)



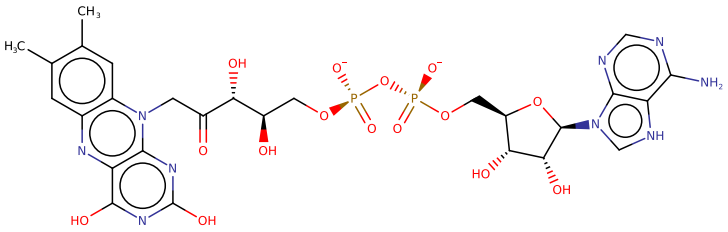
NAD (reference)



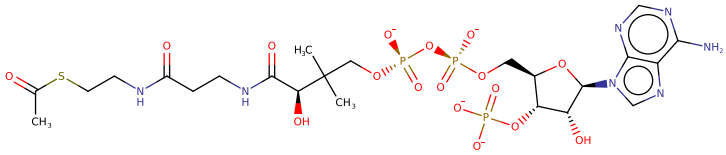
NAD (*Knodle*)



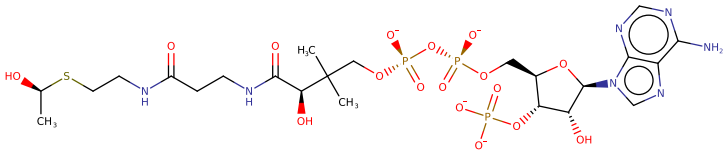
FAD (reference)



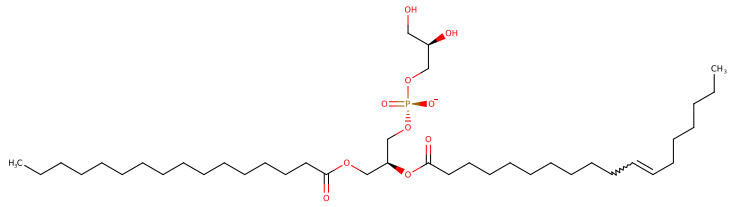
FAD (*Knodle*)



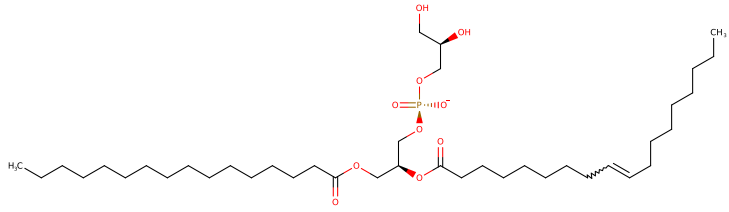
ACO (reference)



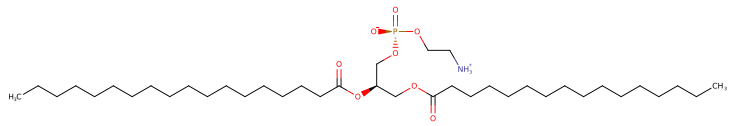
ACO (*Knodle*)



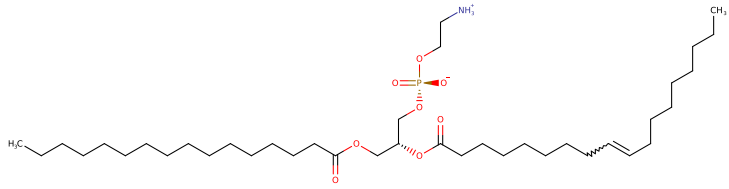
PGV (reference)



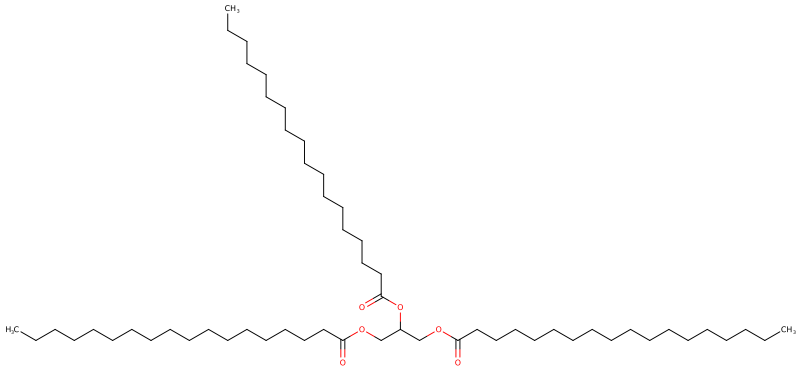
PGV (*Knodle*)



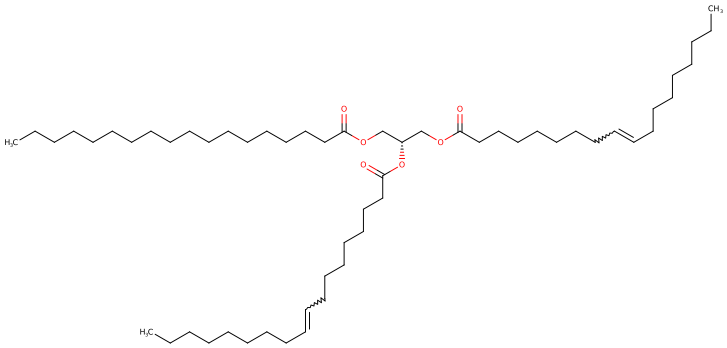
PEV (reference)



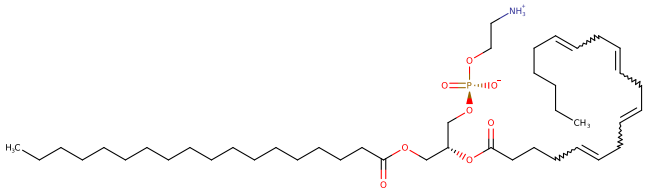
PEV (*Knodle*)



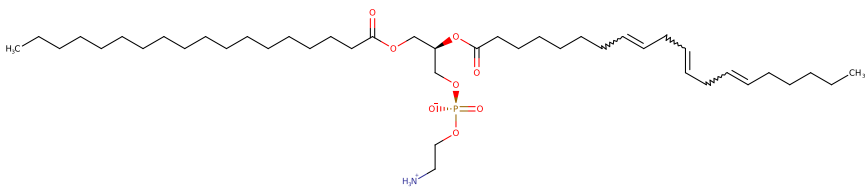
TGL (reference)



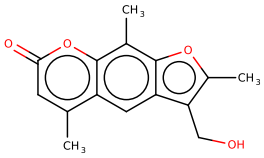
TGL (*Knodle*)



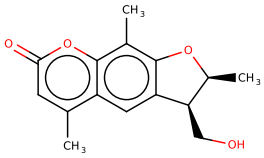
PEK (reference)



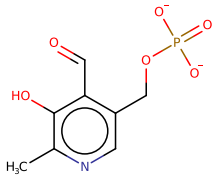
PEK (*Knodle*)



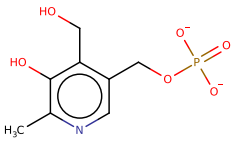
PSO (reference)



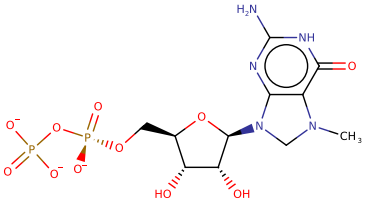
PSO (*Knodle*)



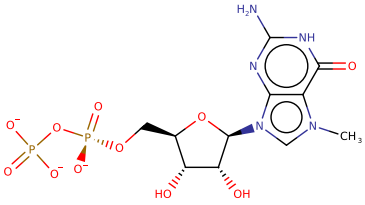
PLP (reference)



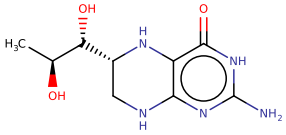
PLP (*Knodle*)



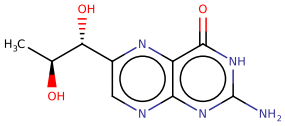
M7G (reference)



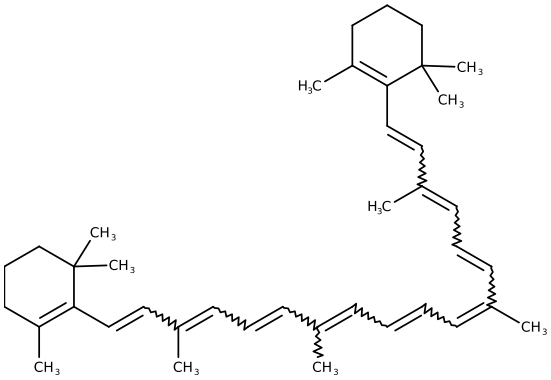
M7G (*Knodle*)



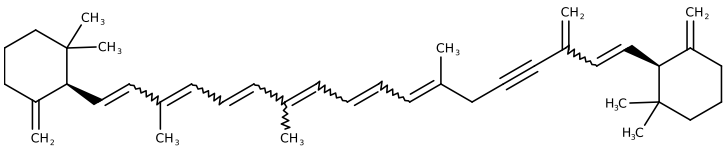
H4B (reference)



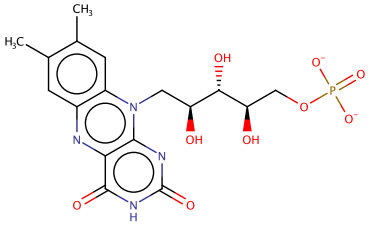
H4B (*Knodle*)



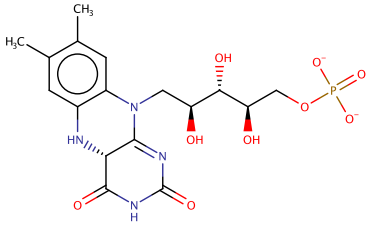
BCR (reference)



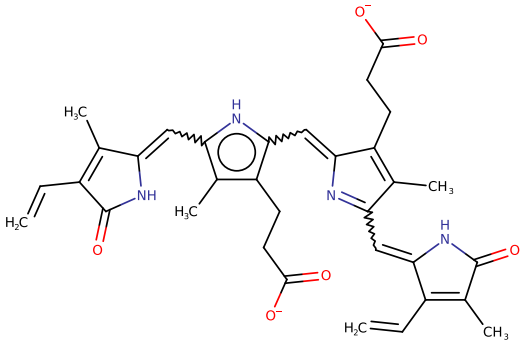
BCR (*Knodle*)



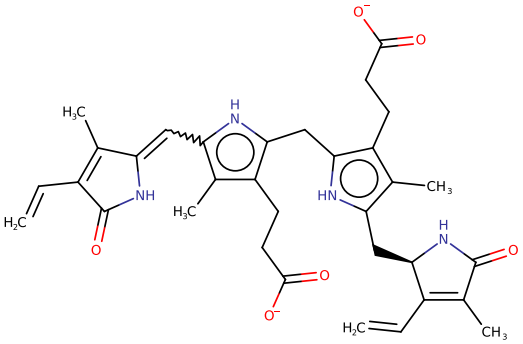
FMN (reference)



FMN (*Knodle*)



BLA (reference)



BLA (*Knodle*)

Table S2: Structure perception errors obtained on the PDBBindCN test set by the *Knodle*, fconv, NAOMI, and I-interpret algorithms.

PDB ID	type of inconsistency	*Knodle*	fconv	NAOMI	I-interpret
1ork	tautomer	different in C2-C3-O3			
1zz3	error	ok	ok	ok	error: OAB should be O.3
3dp1	error	ok	NAM should not form a db with an amide nitrogen NAN, CAG should be sp2 instead	NAM and CAG should be sp2	NAM should not form a db with an amide nitrogen NAN, CAG should be sp2 instead
3ds1	error	ok	C (94) should be sp2	ok	ok
3jrs	tautomer			C11-C10-O10	C11-C10-O10
3jrs	error	ok	C10 should be sp2	ok	ok
1k03	error	C5 should be aromatic, no db with oxygen	ok	ok	ok
1c7e	error	ok	ok	ok	C10 should be sp2
1c7e	tautomer	O2 is sp2 in the reference (exo oxygen)			
3h8c	error	C10,C17 should be sp3	ok	ok	ok
3msc	error	ok	CAQ,NAP,CAM,SAO should be sp2	CAQ,NAP should be sp2	CAQ,NAP should be sp2
3at3	error	c35 should be aromatic (ring is a little bit out of plane)	ok	ok	ok
1bky	error	O2 is sp2 in the reference (exo oxygen)	ok	ok	O2 is sp2 in the reference (exo oxygen)
3ms7	error	ok	N3,C8 should be sp2	N3,C8 should be sp2	C8 should be sp2, no db between N3 and amide N2

3qtw	error	ok	C5,O12 should be sp2	C5,O12 should be sp2	C5,O12 should be sp2
1rtl	error	C24 should be sp2	C24 should be sp2	ok	ok
3gz9	error	ok	C11,C19 should be sp1	C11,C19 should be sp1	C11,C19 should be sp1
3lp1	tautomer	O3 is sp2 in the reference (exocyclic oxygen)	O3 is sp2 in the reference (exocyclic oxygen)	O3 is sp2 in the reference (exocyclic oxygen)	O3 is sp2 in the reference (exocyclic oxygen)
2eas	error	ok	C2,O1 should be sp2	C2,O1 should be sp2	C2,O1 should be sp2
1tyn	error	O31 should be sp3 => missed aromatics	ok	ok	ok
3pcj	tautomer	O3 is sp3 in the reference (exocyclic oxygen)	O3 is sp3 in the reference (exocyclic oxygen)	ok	O3 is sp3 in the reference (exocyclic oxygen)
3hyf	tautomer	ok	O2 is sp3 in the reference (exo)	O2 is sp3 in the reference (exo)	O2 is sp3 in the reference (exo)
1nlt	error	C (58) should be C.2 (terminal CO in a peptide)	C (58) should be C.2 (terminal CO in a peptide)	ok	ok
3huc	error	lost db between CAC and CAA	ok	ok	ok
3tsd	tautomer	O2,O6 are sp2, not sp3 (exo oxygens near heteroaromatics)	ok	ok	O2,O6 are sp2, not sp3 (exo oxygens near heteroaromatics)
4det	error	C1-C2-C3-C4-C5-C6 ring should be aromatic	ok	ok	ok
3ti5	error	ok	C2,C3 should be sp2	C2,C3 should be sp2	C2,C3 should be sp2
3ggw	error	ok	C (99) should be sp2	ok	ok
1tg5	error	O1 should be sp3, ring should be aromatic	ok	ok	ok

3uh4	tautomer	ok	O6 is sp3 in reference (exo oxygen near heteroaromatics)	O6 is sp3 in reference (exo oxygen near heteroaromatics)	ok
3c79	error	C12,C13 should be sp2 (heteroaromatic)	ok	ok	ok
3abt	error	ok	CBK should be sp2	CBK should be sp2	CBK should be sp2
3r8v	error	ok	C5,O12 should be sp2	C5,O12 should be sp2	C5,O12 should be sp2
1ci7	tautomer	O4 is sp2 in the reference (exo oxygen near heteroaromatics)	ok	ok	O4 is sp2 in the reference (exo oxygen near heteroaromatics)
3zm4	error	ok	C15 should be sp2	C15 should be sp2	C15 should be sp2
1bkj	error	C2,O2 should be sp3 (but they are flat)	ok	ok	ok
2pqc	error	C1 should be sp2 insead of C3	ok	ok	ok
3d1v	tautomer			S1 is sp2 in the reference (exocyclic sulphur)	
3d1v	error	ok	ok	ok	C3 should be sp2
3qts	error	ok	O3,C11 should be sp2	O3,C11 should be sp2	O3,C11 should be sp2
1ke2	error	lost guanidinium	lost guanidinium	lost guanidinium	ok
1hv5	error	C36 should be sp2	ok	ok	C36 and all the other atoms from the ring should be sp2
4b5b	error	ok	CAF should be sp2	CAF should be sp2, NAP should be sp3	CAF should be sp2, NAP should be sp3
2j3q	error	ok	C8 should be sp2	C8 should be sp2	C8 should be sp2
1ftm	tautomer	ok	OE1 is sp3 in the reference	OE1 is sp3 in the reference	ok
2wyn	error	N4 should be sp3 here	ok	ok	ok

4ayy	error	ok	N37 should be sp2 here, lost guanidinium	ok	ok
3fdn	error	ok	CAJ should be sp2, NAH should be sp3	CAJ should be sp2, NAH should be sp3	CAJ should be sp2, NAH should be sp3
3rk7	error	ok	C9,O16 should be sp2	C9,O16 should be sp2	C9,O16 should be sp2
3ms4	errror	ok	N3,C8 should be sp2	N3,C8 should be sp2	N2 should be sp3, C8 should be sp2
2x7d	error	CAO,SAE should be sp2	CAO,SAE should be sp2	CAO,SAE should be sp2	CAO,SAE should be sp2
1gt5	error	ok	C7,O should be sp2	C7,O should be sp2	C7,O should be sp2
3uli	error	ok	ok	C27,C28 should be sp3, triangle	C27,C28 should be sp3, triangle
1a4h	error	C16,C17 should be sp2	O8,O9 should be sp2, the ring should not be aromatic	ok	ok
1oe0	tautomer	O4 is sp2 in the reference (exo oxygen near heteroaromatics)			
1oe0	error	ok	ok	ok	C27,C28 should be sp3, triangle
4gqq	error	C1',C2' should be sp2	C1',C2' should be sp2	ok	C1',C2' should be sp2
1qhr	error	ok	O8 should be sp2 (terminal oxy!)	ok	ok
4ef4	tautomer	O6, O61 are sp2 in the reference (exocyclic oxygens near heteroaromatics)	ok	ok	O6, O61 are sp2 in the reference (exocyclic oxygens near heteroaromatics)
3rt6	tautomer	O4 is sp2 in the reference (exo oxygen near heteroaromatics)			

3rt6	error	ok	ok	ok	C6 should be sp2
1ph0	error	O13 should be sp3 => lost aromatics	ok	ok	ok
2yiu	error	long chain: C15 should be sp2	ok	ok	ok
4m2v	error	C7,C8 should be sp2 (lost sulphur heteroaromatics)	ok	ok	ok
1xzx	error	C6 should be aromatic, O1 should be sp3	ok	ok	ok
3vbd	error	ok	NAP, CAS should be sp2	NAP, CAS should be sp2	NAP, CAS should be sp2
3ewz	tautomer	O2 is sp2 in the reference (exo oxygen)			
3ewz	error	ok	ok	ok	C2 should be sp2 (the ring should be aromatic)
3s0b	error	C9,C10 should be sp2 (bad ring)	C9,C10,C11,C12 should be sp2 (bad ring)	C9,C10 should be sp2 (bad ring)	ok
4bxn	error	CAI should be sp2	CAI,CAG,CAT,CAR should be sp2	ok	ok
1ol1	error	C2 should be sp2	ok	ok	ok
3st6	error	C5 should be sp2, O2 should be sp3 (+aromatics lost)	C3 should be sp2, O2 should be sp3 (+aromatics lost)	C3 should be sp2, O2 should be sp3 (+aromatics lost)	C3 should be sp2, O2 should be sp3 (+aromatics lost)
3kqa	error	ok	O3 should be sp2	O3 should be sp2	O3 should be sp2
3u6h	error	ok	C19 should be sp2	C19 should be sp2	C19 should be sp2
3doz	error	ok	CAG should be sp2, amide NAN sholdn't form a db with NAM	CAG,NAM should be sp2	CAG should be sp2, amide NAN sholdn't form a db with NAM
2oxx	error	S10 should be sp2	S10 should be sp2	ok	ok
3c1k	error	ok	C14 should be sp2	C14 should be sp2	C14 should be sp2
2q9m	error	ok	C2,C3 should be sp2	C2,C3 should be sp2	C2,C3 should be sp2

3as0	error	ok	C12 should be sp2 (lost aromatics)	C12 should be sp2 (lost aromatics)	C12 should be sp2 (lost aromatics)
3fcf	error	ok	C15 should be sp2 (aromatic, actually)	C15 should be sp2 (aromatic, actually)	C15 should be sp2 (aromatic, actually)
1hkm	error	C8 should be sp2 (no amidinium-like conjugation)	C8 should be sp2 (no amidinium-like conjugation)	ok	ok
2aie	error	ok	C23,O24 should be sp2	C23,O24 should be sp2	C23,O24 should be sp2
2h65	error	C (33) should be sp2 (terminal CO in a peptide)	C (33) should be sp2 (terminal CO in a peptide)	ok	C (33) should be sp2 (terminal CO in a peptide)
2kce	tautomer	O4 is sp2 in the reference (exo oxygen)	ok	ok	O4 is sp2 in the reference (exo oxygen)
3r4p	error	C8 should be sp2 (lost aromatics)	ok	ok	ok
4ewn	error	C5 should be sp2	C2 should be sp2	ok	ok
3s9y	tautomer	O2,O4 are sp2 in the reference (exocyclic oxygens)	ok	ok	ok
2vk2	error	C1,C2,C4,C5,O2,O1 should be sp3	C1,C2,C4,C5 should be sp3	C1,C2,C4,C5 should be sp3	C1,C2,C4,C5 should be sp3
3n86	error	ok	CAV,OAE should be sp2	CAV,OAE should be sp2	CAV,OAE should be sp2
2nqi	error	CAM should be sp2	ok	ok	ok
3so6	error	CB (54) should be sp2	CB (54) should be sp2	ok	ok
4hmq	metal				
2chm	error	C20, C21,C22 should be sp2 (terminal atoms)	C21,C22 should be sp2	C21,C22 should be sp2	C21,C22 should be sp2
1nu1	error	O41 should be sp3 (lost aromatics in cycle)	ok	ok	ok

1akv	tautomer	O2 is sp2 in the reference (exo oxygen)			
1akv	error	ok	ok	ok	C4 should be sp2
3t3y	error	ok	CA,N should be sp3	CA,N should be sp3	CA,N should be sp3
4bqs	error	ok	CAG should be sp2, CAI should be sp3	CAG should be sp2, CAI should be sp3	CAG should be sp2, CAI should be sp3
3lxl	error	ok	C13 should be sp2 (lost aromatics)	C13 should be sp2 (lost aromatics)	C13 should be sp2 (lost aromatics)
4ab9	error	ok	ok	CM,C5 should be sp3 (no aromatics here)	CM,C5 should be sp3 (no aromatics here)
4ahs	error	C1 should be sp2, lost aromatics	ok	ok	ok
3v0l	tautomer	OAA is sp2 in the reference			
3v0l	error	ok	ok	ok	CBC should be sp2
3v5q	error	N4 is planar here (sum of angles >350), aromatics is in two atoms from N, so it can be N.pl3	C29 should be sp2	ok	ok
3kvw	tautomer	CAT,NAL,OAC			
3kvw	error	CAU should be sp2, there should be no aromatics in the ring	CAU should be sp2, there should be no aromatics in the ring, missed aromatics in a 6-membered ring	no db with NAL, CAU should be sp2	CAU should be sp2, there should be no aromatics in the ring
3d9p	error	ok	terminal CO: C (62) should be sp2	ok	ok
2ql7	error	ok	ok	terminal C (27) should be sp3	ok
3fzn	error	ok	C02 should be sp2 (lost aromatics)	C02 should be sp2 (lost aromatics)	C02 should be sp2 (lost aromatics)
2bmc	error	C15,C16 should be sp2	C3,C14,C15,C16 should be sp2	C3 should be sp2	C3 should be sp2

1gag	error	ok	C (89) should be sp2 (terminal CO in a peptide)	ok	ok
1qpl	error	C64 should be sp2 (bad ring)	ok	ok	ok
3sud	error	ok	ok	CAK,CAL should be sp3 (triangle)	CAK,CAL should be sp3 (triangle)
3h23	error	ok	C4,C5 should be C2 (aromatic)	O24 should be sp2, C5 should be sp2 (sromatic)	C5 should be C2 (aromatic)
3r9n	error	ok	O3,C7 should be sp2	O3,C7 should be sp2	O3,C7 should be sp2
1xh4	error	O92 should be sp3 – > aromatics lost	ok	ok	ok
1a3e	error	ok	ok	CB,CG should be sp3	CB,CG should be sp3
3h2m	error	C4-N9-C10-C11-N12-C7 ring shouldn't be aromatic: C11 should be sp3!	C4,C7 should be sp2; thus the ring C1-N3-C5-N6-C7-C4 will be aromatic	C4 should be sp2	C4 should be sp2; thus the ring C1-N3-C5-N6-C7-C4 will be aromatic
3g8e	error	ok	CAO should be sp2, the ring is aromatic	CAO should be sp2, the ring is aromatic	CAO should be sp2, the ring is aromatic
4eop	error	C11,C12 should be sp2	the ring shouldn't be aromatic, C11 should be sp2 and form a db with C12	ok	ok
3qcj	error	C19,C20 should be sp2 (lost aromatics in a ring with sulphur)	ok	ok	ok
3iww	error	terminal S2 should be sp2 (terminal C-C)	ok	ok	ok
4imz	error	ok	ok	terminal CO: C (30) should be sp3	terminal CO: C (30) should be sp3

3rah	error	ok	C11,O3 should be sp2	C11,O3 should be sp2	C11,O3 should be sp2
2va5	tautomer	O5 is sp2 in the reference (exo oxygen)	ok	ok	O5 is sp2 in the reference (exo oxygen)
3zlo	error	ok	C8 should be sp2	C9,C8,C11 should be sp3, no aromaticity in that ring, rin with S1 should be aromatic	C9,C8,C11 should be sp3, no aromaticity in that ring
2qtn	error	ok	C2 should be sp2, the ring containing it should be aromatic	C2 should be sp2, the ring containing it should be aromatic	C2 should be sp2, the ring containing it should be aromatic
3el0	error	ok	O38 should be sp2, the ring should be aromatic	O38 should be sp2, the ring should be aromatic	O38 should be sp2, the ring should be aromatic
1fkf	error	ok	C39,C40 should be sp2	C38 should be sp3, C40 should be sp2	C38 should be sp3, C40 should be sp2
4b4g	error	ok	CAF,CAO should be sp2 (aromatics lost)	CAF should be sp2, NAP should be sp3 (aromatics lost)	CAF should be sp2, NAP should be sp3 (aromatics lost)
1ctu	error	terminal exoO near ring: C4 should be sp3	ok	ok	ok
3skf	error	ok	O43,C19 should be sp2	ok	O43,C19 should be sp2
3skf	tautomer	O43-C19-N39		O43-C19-N39	
3eax	error	ok	rings C1-C2-C3-C4-C5-C6 and C33-C34-C35-C36-C37-C38 should be aromatic	rings C1-C2-C3-C4-C5-C6 and C33-C34-C35-C36-C37-C38 should be aromatic	rings C1-C2-C3-C4-C5-C6 and C33-C34-C35-C36-C37-C38 should be aromatic
3mz3	error	ok	ok	ok	O4 should be sp3
2gbf	error	ok	ok	C8 should be sp3	C8 should be sp3
1tou	tautomer	O13 is sp2 in the reference	ok	ok	O13 is sp2 in the reference

1qnh	error	CE,CZ should be sp2	was not processed	was not processed	ok
3r5m	error	CAA,CAE,CAB,CAN should be sp2	CAB should be sp2, CAN should be sp3	CAB should be sp2, CAN should be sp3	CAB should be sp2, CAN should be sp3
2vvu	tautomer	O25 is sp2 in the reference (exo oxygen)			
2vvu	error	ok	ok	ok	C8 should be sp2
4asy	error	ok	CAF,CAO should be sp2	CAF should be sp2, NAP should be sp3 (lost aromaticity)	CAF should be sp2, NAP should be sp3 (lost aromaticity)
4mz4	error	ok	C23 should be sp2 (lost aromaticity in a bad ring)	C23 should be sp2 (lost aromaticity in a bad ring)	C23 should be sp2 (lost aromaticity in a bad ring)
3vp1	error	CAN,SAY should be sp3 (the ligand is truncated)	CAN,SAY should be sp3 (the ligand is truncated)	CAN,SAY should be sp3 (the ligand is truncated)	CAN,SAY should be sp3 (the ligand is truncated)
1nms	error	O23 should be sp2 (exo oxygen), thus C19 should be sp2 and the ring should be aromatic	ok	ok	ok
4bkz	error	ok	C3,C4 should be sp3 (not a planar ring!)	C3,C4 should be sp3 (not a planar ring!)	C3,C4 should be sp3 (not a planar ring!)
3qgy	error	C1-C2 should be sp2 (disrupted structure)	ok	ok	ok
3u4o	error	ok	C1 should be sp2 (aromatic ring)	C1 should be sp2 (aromatic ring)	C1 should be sp2 (aromatic ring)
3fy0	metal				
2wtj	error	ring should be aromatic	C12,S13,C14,C15,C20 ring should be aromatic	C11-C12 should not form a db	C12,S13,C14,C15,C20 ring should be aromatic
3cst	metal				
1br6	tautomer		O1 is sp3 in the reference (exo oxygen near heteroaromatics)	O1 is sp3 in the reference (exo oxygen near heteroaromatics)	

1br6	error	ok	ok	ok	C2 should be sp2
3ti1	error	O27 should be sp2, the ring shouldn't be aromatic => C12 should be sp2	O27 should be sp2, the ring shouldn't be aromatic => C12 should be sp2	ok	O27 should be sp2, the ring shouldn't be aromatic => C12 should be sp2
4euv	tautomer	O6 is sp2 in the reference (exo oxygen near heteroaromatics)	ok	ok	O6 is sp2 in the reference (exo oxygen near heteroaromatics)
2ihq	error	ok	ok	ok	C7 should be sp2
2ihq	tautomer	O8-C7-N9			
3t0t	error	ok	NAL should not form a db with an amide nitrogen NAN, CAP should be sp2 instead	was not processed	NAL should not form a db with an amide nitrogen NAN, CAP should be sp2 instead
4meo	error	ok	C7 should be sp2	C7 should be sp2	C7 should be sp2
4h3a	error	ok	CAU, CAF should be sp2	OAE or CAF should be sp2	CAU, CAF should be sp2
4h3a	tautomer	CAF,CAU,OAE			
2yac	error	ok	C7,C8 should be sp3	C7,C8 should be sp3	C7,C8 should be sp3
3zjv	not a full ligand: both double bonds are possible; a ring was here				
3ppr	error	ok	NAG,N,CAI form an amidinium group; they should be sp2	NAG,N,CAI form an amidinium group; they should be sp2	NAG,N,CAI form an amidinium group; they should be sp2
3s0o	error	ok	C5,O12 should be sp2	C5,O12 should be sp2	C5,O12 should be sp2
3m1k	error	ok	I'ts an aromatic pyridinium, C3 should be sp2	I'ts an aromatic pyridinium, C3 should be sp2	I'ts an aromatic pyridinium, C3 should be sp2

4jjq	error	C (21) should be sp2 (terminal CO in a peptide)	C (21) should be sp2 (terminal CO in a peptide)	ok	ok
1uy7	error	C17 should be sp3	ok	ok	ok
3ioi	error	CAQ,CBF,CAP,CBE should be sp2 (sulphur ring), CAO,OAB should be sp2 (terminal CO)	CAQ,CBF,CAP,CBE should be sp2 (sulphur ring), CAO,OAB should be sp2 (terminal CO)	was not processed	CAQ,CBF,CAP,CBE should be sp2 (sulphur ring), CAO,OAB should be sp2 (terminal CO)
1w2h	tautomer	O4 could be sp2 as in the reference (exo oxygen)	ok	ok	O4 could be sp2 as in the reference (exo oxygen)
4ie2	Boron has too long bonds, all the methods failed				
3ppo	error	ok	ok	C06,O07 should be sp3	C06,O07 should be sp3
2rfh	error	NAO should be sp2, it's NO2 group	NAO should be sp2, it's NO2 group	NAO should be sp2, it's NO2 group	NAO should be sp2, it's NO2 group
3hav	error	ok	CG2,OG2 should be sp2 (terminal CO)	ok	ok
3cyu	error	ok	ok	C46,C47 should be sp3	C46,C47 should be sp3
3udv	error	ok	C26 should be sp2	C25 should be sp2	C26 should be sp2
3exe	error	ok	C2,N3 should be sp2 (lost aromatics in a ring with sulphur)	C2,N3 should be sp2 (lost aromatics in a ring with sulphur)	C2,N3 should be sp2 (lost aromatics in a ring with sulphur)
2qbw	error	CB (12) should be sp2	CB (12) should be sp2	ok	ok
1lor	error	C5 should be sp2, O1 should be sp3 (lost aromaticity because of an exo oxygen)	ok	ok	ok

3te5	error	C4N should be sp3, no aromaticity in the corresponding cycle	C6N should be sp2, C4N should be sp3	C6N should be sp2, C4N should be sp3	C6N should be sp2, C4N should be sp3
3btj	error	ok	C9,C14 should be sp2 (both rings should be aromatic)	C9,C14 should be sp2 (both rings should be aromatic)	C9,C14 should be sp2 (both rings should be aromatic)
3kwb	error	C14,N15 should be sp2	N3,N4 should be sp2	C14,N15 should form a db, ring should be aromatic	C14,N15 should form a db, ring should be aromatic
2fxs	error	O6,O8 should be sp2, the ring should be aromatic	O6,O8 should be sp2, the ring should be aromatic	O6,O8 should be sp2, the ring should be aromatic	O6,O8 should be sp2, the ring should be aromatic
1kf6	error	O4,O1 should be sp3, the ring will be aromatic	ok	ok	O4,O1 should be sp3, the ring will be aromatic
1ih0	error	ok	N4,C5 should be sp2	N4,C5 should be sp2	N4,C5 should be sp2
1tyr	error	C7-C8 should be sp2 (chain)	C7-C8 should be sp2 (chain)	C7 should be sp2 (chain)	ok
3nhi	error	C7,C8 should be sp2 (long chain)	C7,C8,C9,C10,C11,C12 should be sp2 (long chain)	C7,C8,C9,C12 should be sp2 (long chain)	C7,C8,C9,C12 should be sp2 (long chain)
3su5	error	CBB should be sp2 (lost heteroaromatics)	ok	ok	ok
1uwu	error	ok	N1 should be sp2	N1 should be sp2	N1 should be sp2
2g5p	error	ok	ok	C21,N22 should be sp3	C21,N22 should be sp3
2xuf	error	ok	C6 should be sp2 (aromatics lost)	ok	C7 should be sp2 (aromatics lost)
3d9l	error	ok	C (80) should be sp2 (terminal CO in a peptide)	ok	ok
2g5t	error	ok	ok	C21,N22 should be sp3	C21,N22 should be sp3

4apr	error	C (42) should be sp2 (terminal CO in a peptide)	C (42) should be sp2 (terminal CO in a peptide)	ok	ok
2w10	error	ok	C (85) should be sp2 (terminal CO in a peptide)	ok	ok
4h2m	error	ok	CAG,CAE,CAF,CAH should be sp1	CAG,CAE should be sp1	CAG,CAE should be sp1
4e26	error	CAS-CAR-CAO-CAF-OAW ring should be aromatic	ok	ok	ok
1exv	error	C7 should be sp2, the ring should be aromatic	ok	ok	ok
1k2i	error	C2 should be sp2 (terminal CO)	ok	ok	ok
3ggj	tautomer	O6 is sp2 in the reference (exo oxygen near heteroaromatics)	ok	ok	ok
2q2z	error	ok	C24 should be sp2	C24 should be sp2	C24 should be sp2
4cc3	error	CG should be sp2	ok	ok	ok
2o3p	error	ok	C5 should be sp2 (aromatic)	C5 should be sp2 (aromatic)	C5 should be sp2 (aromatic)
1rev	error	S2 should be sp2	S2 should be sp2	ok	ok
3s1h	error	ok	O3,C11 should be sp2	O3,C11 should be sp2	O3,C11 should be sp2
4fli	error	ok	CO2,CO3 should be sp2 (stand-alone bond in a chain)	CO2,CO3 should be sp2 (stand-alone bond in a chain)	CO2,CO3 should be sp2 (stand-alone bond in a chain)
1ppw	error	ok	C11,C12 should be sp3 (stand-alone bond in a chain)	C11,C12 should be sp3 (stand-alone bond in a chain)	C11,C12 should be sp3 (stand-alone bond in a chain)

1il5	tautomer	O6 is sp3 in the reference (exo oxygen near heteroaromatic)	ok	ok	ok
4hxs	error	CAU should be sp2 (a non-aromatic heteroring with exo oxygen)	ok	ok	ok
4dn0	tautomer	O6 is sp2 in the reference (exo oxygen near heteroaromatics)	ok	ok	O6 is sp2 in the reference (exo oxygen near heteroaromatics)
3jzb	error	C1,C3,C5,C7,C9,C11 ring should be aromatic	C1,C3,C5,C7,C9,C11 ring shoild be aromatic	C1,C3,C5,C7,C9,C11 ring shoild be aromatic	C1,C3,C5,C7,C9,C11 ring shoild be aromatic
2w2u	error	ok	C (87) should be sp2 (terminal CO in a peptide)	ok	ok
3iae	error	ok	C03,C04,S1,C02,N3 should be sp2, ring should e aromatic	C03,C04,S1,C02,N3 should be sp2, ring should e aromatic	C03,C04,S1,C02,N3 should be sp2, ring should e aromatic
2xup	error	ok	N3,C11 should be sp2, ring should be aromatic	N3,C11 should be sp2, ring should be aromatic	N3,C11 should be sp2, ring should be aromatic
2y6o	error	C21 should be sp3 (terminal oxygen)	ok	ok	ok
1ekb	error	C (27) should be sp3 (terminal oxygen)	ok	ok	ok
2qlb	error	ok	ok	C (26) and O (27) should be sp3 (terminal CO in peptide)	ok
1uy8	error	C17,C16 should be sp3 (stand-alone bond in a chain)	ok	ok	ok

1gx0	error	ok	C5,C6 should be sp2, ring should be aromatic	C5,C6 should be sp2, ring should be aromatic	C5,C6 should be sp2, ring should be aromatic
3sjo	error	C2 should be sp2, ring should be aromatic	ok	ok	ok
3kig	error	C11 should be sp1	C11 should be sp1	C11 should be sp1	ok
2alv	error	C23 should be sp2 (stand-alone db in a chain)	C23 should be sp2 (stand-alone db in a chain)	ok	C23 should be sp2 (stand-alone db in a chain)
1x8b	error	ok	O3 should be sp3, ring should be aromatic	O3 should be sp3, ring should be aromatic	O3 should be sp3, ring should be aromatic
4gzt	error	ok	C2,C7 should be sp2 (stand-alone bond in a non-aromatic cycle)	C2,C7 should be sp2 (stand-alone bond in a non-aromatic cycle)	C2,C7 should be sp2 (stand-alone bond in a non-aromatic cycle)
3sut	error	ok	CAA,NAY should be sp2	CAA,NAY should be sp2	CAA,NAY should be sp2
4erw	error	O5 should be sp2, thus the ring cannot be aromatic	O5 should be sp2	O5 should be sp2	O5 should be sp2
1w6y	error	O1 should be sp3, the ring should be aromatic => C2 should be sp2	ok	ok	ok
3saz	tautomer	C06-C07-O08			
1vea	error	CA,CB,N should be sp3	ok	ok	ok
4bdj	error	ok	ok	C9,C10 should be sp3 (triangle)	C9,C10 should be sp3 (triangle)
1w2g	tautomer	O4 is sp2 in the reference (exo oxygen near heteroaromatics)	ok	ok	ok

3mtd	error	ok	NAP,CAQ should be sp2	NAP,CAQ should be sp2	NAP,CAQ should form a db
4ijq	tautomer	O6 is sp2 in the reference (exo oxygen near heteroaromatics)	ok	ok	O6 is sp2 in the reference (exo oxygen near heteroaromatics)
2wzx	error	ok	C11 should be sp2 (terminal CO)	ok	ok
3fbr	error	ok	C (40) should be sp2 (terminal CO)	ok	ok
2p59	error	ok	ok	C7,C28 should be sp3 (triangle)	C7,C28 should be sp3 (triangle)
1h24	error	C (72) should be sp2 (terminal CO in a peptide)	C (72) should be sp2 (terminal CO in a peptide)	ok	ok
4i0z	error	ok	C13 should be sp2 (lost aromatics)	C13 should be sp2 (lost aromatics)	C13 should be sp2 (lost aromatics)
3sug	error	ok	ok	CAK,CAL should be sp3 (triangle)	CAK,CAL should be sp3 (triangle)
2r6n	error	C09,N10 should be sp3	C09,N10 should be sp3	C09 should be sp3, ring should be aromatic	C09,N10 should be sp3
3v4j	error	O4,O5 should be sp3, ring should be aromatic	ok	N3,O3 or N3,N2 should be sp2	O4,O5 should be sp3, ring should be aromatic
3gl6	error	C (59) should be sp2 (terminal CO in a peptide)	C (59) should be sp2 (terminal CO in a peptide)	ok	ok
3sad	tautomer	O05,C04,O06			
1ths	error	CG (34) should be sp3 (no double bond in a non-aomatic heteroring)	ok	ok	ok

4d9p	tautomer	O1 ,O3 are sp2 in reference (exo oxy in heteroaromatic ring)	O1 is sp2 in reference (exo oxy in heteroaromatic ring)	O1 is sp2 in reference (exo oxy in heteroaromatic ring)	
4d9p	error	ok	ok	ok	C10 should be sp2 (aromatics lost)
2z3h	tautomer			O1 is be sp3 in the reference (exo oxy near heteroaromatics)	
2z3h	error	C1' should be sp3 (but it is planar), C3' should be sp2	C2',C3' should be sp2	ok	C4 should be sp2
3vb4	error	C (28) should be sp2 (terminal CN)	ok	ok	ok
4fse	error	ok	C01-C07-C02-C04-C11-C03 should be aromatic (sp2 hybridizations)	C01-C07-C02-C04-C11-C03 should be aromatic (sp2 hybridizations)	C01-C07-C02-C04-C11-C03 should be aromatic (sp2 hybridizations)
1rdt	error	C9 should be sp2, C,C1,C2,C3 should be sp3	C9 should be sp2, C,C1,C2,C3 should be sp3	C9 should be sp2, C,C1,C2,C3 should be sp3	C9 should be sp2, C,C1,C2,C3 should be sp3
1d6n	tautomer		O6 is sp3 in reference (exo oxygen near heteroaromatics)		
1d6n	error	C9 should be sp3	ok	O6 should be sp3 (exo oxygen near heteroaromatics), N7,N8should be sp2	C4 should be sp2, O6 should be sp3, N8 should be sp2
4mep	error	ok	C7 should be sp2 (aromatics lost)	C7 should be sp2 (aromatics lost)	C7 should be sp2 (aromatics lost)
1oky	error	ok	C20 should be sp3 (non-aromatic heterocycle)	C20 should be sp3 (non-aromatic heterocycle)	C20 should be sp3 (non-aromatic heterocycle)

4fz3	error	C6 should be sp2 (stand-alone double bond in a non-aromatic cycle)	ok	ok	ok
2clo	error	C5-C6-C7-C8-C9-C10 ring should be sp2	C5-C6-C7-C8-C9-C10 ring should be sp2	ok	C5-C6-C7-C8-C9-C10 ring should be sp2
2v10	error	ok	ok	C32,C33 should be sp3 (stand-alone bond in a chain)	C32,C33 should be sp3 (stand-alone bond in a chain)
3eyh	error	ok	C13 should be sp2 (lost aromatics)	C13 should be sp2 (lost aromatics)	C13 should be sp2 (lost aromatics)
2q2y	error	ok	C8 should be sp2	C8,N12 should be sp2	C8 should be sp2
3vw0	error	ok	C5-C6-C7-C8-C9-N1 and C10-C11-N2-C13-C14 should be sp2 (and aromatic)	C5-C6-C7-C8-C9-N1 and C10-C11-N2-C13-C14 should be sp2 (and aromatic)	C5-C6-C7-C8-C9-N1 and C10-C11-N2-C13-C14 should be sp2 (and aromatic)
3ddu	error	ok	C12 should be sp2 (lost aromatics)	C12 should be sp2 (lost aromatics)	C12 should be sp2 (lost aromatics)
1ove	error	ok	ok	C23,C24,C25,C26 hould be sp3	ok
2fl2	error	C7,N1 should be sp3 (terminal NO)	ok	ok	ok
4hxq	Boron				
3n76	error	C3,C8 should be sp3	C2,C3,C8,C9 should be sp3	C2,C3,C8,C9 should be sp3	C2,C3,C8,C9 should be sp3
1nf8	error	C2',C3' should be sp2 (terminal CC)	ok	ok	ok
2fky	error	C10 should be sp2	C10 should be sp2	ok	C10 should be sp2
1unh	error	C14 should be sp2 (and one of the rings shouldn't be aromatic, and more..)	O23 should be sp2, both 5-membered rings should be non-aromatic	ok	C14 should be sp2 (and one of the rings shouldn't be aromatic, and more..)

3zze	error	N17,C15 should be sp3 (C15 is planar)	C7-C8-C9-C12-C11-C10 ring should be aromatic, no db with N13	ok	ok
1nc6	error	NW,NH1,NH2 should be sp2, missed guanidinium	NW,NH1,NH2 should be sp2, missed guanidinium	ok	ok
3nc4	error	ok	CAQ,NAP,SAO,CAM should be sp2	CAQ,NAP should be sp2	CAQ should be sp2, there should be no db betseen NAN,NAP
2xyu	error	ok	ok	C08,C09,C10,C11 should be sp3 (non-planar ring)	C08,C09,C10,C11 should be sp3 (non-planar ring)
1rmz	error	C1-C2-C3-C4-C5-C6 ring should be aromatic (non-planar ring)	C1-C2-C3-C4-C5-C6 ring should be aromatic (non-planar ring)	C1-C2-C3-C4-C5-C6 ring should be aromatic (non-planar ring)	ok
1thz	tautomer			O17 is sp3 in the reference (exocyclic oxygen)	
1thz	error	ok	N20,N19 should be sp2 (stand-alone double bond)	ok	C15 should be sp2
2jb6	error	wrong chain of sp2 carbons, and some other mistakes as a concequence	ok	ok	ok
1e02	error	ok	C1,O1 should be sp2 (terminal CO)	ok	ok
1oz0	error	ok	C4P should be sp2 (aromatics in a heteroring (not very planar as it should be))	C4P should be sp2 (aromatics in a heteroring (not very planar as it should be))	ok
3p44	metal				

1ngw	error	porphyrin: C15,C25,C35,C45 should be sp2 (aromatic)	porphyrin: C15,C25,C35,C45 should be sp2 (aromatic)	porphyrin: C15,C25,C35,C44,C45 should be sp2 (aromatic)	porphyrin: C15,C25,C35,C45 should be sp2 (aromatic)
3gm0	error	ok	ok	CE1 should be sp2	ok
2g01	error	ok	O2 should be sp2 (exo oxygen)	O2 should be sp2 (exo oxygen)	ok
4m0e	error	ok	exo O21,O20 should be sp2, the ring shouldn't be aromatic	exo O21,O20 should be sp2, the ring shouldn't be aromatic	ok
1bjr	error	terminal CO: C (54) should be sp2 (terminal CO in peptide)	terminal CO: C (54) should be sp2 (terminal CO in peptide)	ok	ok
3lkj	error	C7 should be sp2, the ring should be aromatic	ok	ok	ok
2l65	error	C2,C3 and C7,C6 should be sp1	ok	C2,C3 and C7,C6 should be sp1	ok
2p99	error	ok	NAL,NAM should be sp3 (terminal NN)	ok	NAL,NAM should be sp3 (terminal NN)
4n5d	error	ok	C4 should be sp2	C4,O should be sp2	C4 should be sp2
3hig	error	ok	ok	N should be sp2	ok
3gy2	error	ok	ok	N should be sp2	ok
3v9b	error	ok	ok	ok	O17 should be sp3
3vw9	tautomer	ok	ok	O21 is sp2 in the reference (exo oxygen near heteroring)	ok
4ge4	tautomer	ok	ok	O7 is sp2 in the reference (exo)	ok
4kn7	error	ok	O1 (exo) should be sp2, the corresponding ring shouldn't be aromatic	O1 (exo) should be sp2, the corresponding ring shouldn't be aromatic	ok
3p3h	metal				

2r02	error	ok	C (85) should be sp2 (it is COOH); oxygen bond is too long to be recognized by fconv	ok	C (85) should be sp2 (it is COOH); oxygen bond is too long to be recognized by iintr
2h2d	error	terminal CO: C (102) should be sp2 (terminal CO in peptide)	terminal CO: C (102) should be sp2 (terminal CO in peptide)	ok	ok
4nkt	tautomer	O4 is sp2 in the reference(exo oxygen near heteroaromatics)	ok	ok	O4 is sp2 in the reference(exo oxygen near heteroaromatics)
1nhz	error	C31 should be sp1 (big angle, not too short bond)	C31 should be sp1	ok	ok
1pip	error	ok	ok	ok	C4 should be sp2 (amide in a peptide)
1l5s	error	ok	C5 should be sp2 (aromatic)	ok	C5 should be sp2 (aromatic)
1i1e	error	ok	ok	ok	C1 should be sp2 (CO)
3hl8	error	ok	C10 should be sp2	ok	ok
3hl8	tautomer				O1,C8,N2
4j8r	error	ok	ok	ok	CG should be sp2 (heteroaromatic ring)
1n5r	error	ok	ok	ok	C22 should be sp2 (aromatic)
2y36	error	ok	ok	ok	O (20) should be sp2 (amide in a peptide)
1uu3	error	ok	ok	ok	CAH should be sp2 (non-aomatic 5-membered planar ring)
2a5b	error	ok	ok	ok	C5 should be sp2 (5-membered heteroring)

3t2w	tautomer				N2,C3,O3
4gk4	error	ok	ok	ok	C3 should be sp2
1d3q	error	ok	ok	ok	C1 should be sp2 (aromatic ring)
1v11	error	ok	ok	ok	C2 should be sp2
2cle	error	ok	C12 should be sp2 (amide)	ok	C12 should be sp2 (amide)
2c9d	error	ok	ok	ok	C9 should be sp2 (heteroring)
6std	error	ok	ok	ok	O9B should be sp2 (SO)
1g05	error	ok	ok	ok	C15,O18 should be sp2
1qk4	error	ok	ok	ok	C2 should be sp2 (heteroaromatics)
2gm1	error	ok	ok	ok	CG should be sp2 (aromatic)
1bqm	error	ok	ok	ok	C2,C1 should be sp2
1bqm	tautomer		S1,C2,N2		
1ilh	error	ok	C7 should be sp2	ok	C7 should be sp2
1q84	error	ok	ok	ok	C7 should be sp2 (aromatic)
3q2m	error	ok	ok	ok	O25 should be sp2 (SO2)
3uwl	error	ok	ok	ok	C4A should be sp2 (heterocycle)
1yyy	error	ok	C (13) should be sp2 (amide)	ok	C (13) should be sp2 (amide)
1qs4	error	ok	ok	ok	C8 should be sp2 (CO)
1g9a	error	ok	ok	ok	C8 should be sp2 (aromatic 5-membered ring)
1u0h	error	ok	ok	ok	CA5 should be sp2 (aromatic)

1onz	error	ok	C18 should be sp2 (CO)	ok	C18 should be sp2 (CO)
1v1m	error	ok	ok	ok	C2 should be sp2 (heteroaromatics)
1pmn	1	ok	ok	ok	C4 should be sp2 (heteroaromatics)
3bjc	tautomer	ok	O16 is sp2 in the reference (exo oxygen near heteroaromaics)	ok	ok
2ans	error	ok	N should be sp2 (N.pl3): aromatics is nearby	ok	ok
2uzl	error	ok	C6 should be sp2, because O1 should be sp2, and the ring shouldn't be aromatic	ok	ok
2vgo	error	ok	N6 should be sp2 (N.pl3): aromatics is nearby	ok	ok
3unj	error	ok	N04 should be sp2 (N.pl3): aromatics is nearby	ok	ok
3qzq	error	ok	C01,O23 should be sp2 (CO)	ok	ok
2o9j	error	ok	C3 should be sp2 (non-aromatic heteroring)	ok	ok
1hlf	error	ok	S8 should be sp2 (exo suphur near 5-membered ring)	ok	ok
2ieh	error	ok	S should be sp2 (near heteroaromatic ring)	ok	ok

3jsi	tautomer	ok	O15 is sp2 in the reference (exo oxygen near aromatic ring)	ok	ok
3d1g	tautomer		S1-C8-S2		
2zjw	error	ok	O21,O24 should be sp2 (exo oxygens)	ok	ok
1iem	error	ok	N16 should be sp2	ok	ok
3ppj	error	C21,C22,C23,C24 should be sp2	C21,C22,C23,C24 should be sp2	C21,C22,C23,C24 should be sp2	C21,C22,C23,C24 should be sp2
2f14	error	ok	S2 should be sp2 (amide)	ok	ok
1iky	error	ok	S10 should be sp2 (amide)	ok	ok
1q0b	error	ok	S20 should be sp2 (exo sulfur near heteroaromatics)	ok	ok
1m2p	error	ok	O16 should be sp2 (exo oxygen)	ok	ok
4azg	error	ok	CAA should be sp2	ok	ok
4a4o	error	ok	C29 should be sp2 (5-membered heteroaromatics)	ok	ok
2oa0	error	ok	C3 should be sp2 (planar carbon)	ok	ok
1o4j	error	ok	C12 should be sp2 (amide)	ok	ok
1m2r	error	ok	O16 should be sp2 (exo oxygen)	ok	ok
2ph9	error	ok	C13 should be sp2 (the ring it belongs to should be aromatic)	ok	ok
4jin	error	ok	C10 should be sp2	ok	ok

2uzn	error	ok	O1 should be sp3, ring wouldn't be aromatic => C6 should be sp2	ok	ok
1fw0	error	ok	CD1 should be sp2 (terminal CC)	ok	ok
3wd1	error	ok	N11 should be sp2	ok	ok
1agm	error	ok	C5A,C7A should be sp2 (stand-alone bond in a non-aromatic ring)	ok	ok
2xu3	error	ok	C19 should be sp2 (terminal CN)	ok	ok
1h46	error	ok	C13 should be sp2 (stand-alone CN bond in a chain)	ok	ok
2o7e	error	ok	C2' should be sp2 (aromatic)	ok	ok
2b55	error	ok	O54 should be sp2 (exo oxygen of non-aromatc 5-membered ring)	ok	ok
3pb7	error	ok	SAC should be sp2	ok	ok
3q4c	error	ok	O (1) should be sp2 (exo oxygen near 5-membered heteroring)	ok	ok
3ft8	error	ok	N24 should be sp2	ok	ok
1bji	error	ok	C2 should be sp2 (stand-alone bond in a non-aromatic ring)	ok	ok
1ke5	error	ok	O11 should be sp2	ok	ok
1q3w	error	ok	C5 should be sp2 (aromatic), thus no C6-N1 double bond	ok	ok

4iaw	error	ok	S1 should be sp2	ok	ok
4hy1	error	ok	C22 should be sp2 (and aromatic)	ok	ok
3ao2	error	ok	C5 should be sp2 (aromatic)	ok	ok
1bug	error	ok	S1 should be sp2 (terminal CS)	ok	ok
2e9n	error	ok	C18 should be sp2 (and aromatic)	ok	ok
3py0	error	ok	C6' should be sp2, the should be not aromatic because of the oxygen	ok	ok
2h42	tautomer	ok	O27 is sp2 in the reference (exo oxygen near heteroaromatic ring)	ok	ok
4aj4	error	ok	C1 should be sp2 (terminal CC)	ok	ok
4gbd	error	ok	C9 should be sp2	ok	ok
3pbb	tautomer		NAO,CAR,SAC		
2vcb	error	ok	NAY should be sp2	ok	ok
4kot	error	ok	N2 should be sp2	ok	ok
3r2a	error	ok	OAB should be sp2 (exo oxygen near non-aromatic ring)	ok	ok
3eft	error	ok	SAT should be sp2	ok	ok
1gx8	error	ok	C5 should be sp2 (stand-alone db in a ring)	ok	ok
3hec	error	ok	N48 should be sp3; it is two atoms far from aromatics and is not absolutely planar	ok	ok
4j8t	error	ok	C22 should be sp2		

1dtq	error	ok	S9 should be sp2	ok	ok
2w92	error	ok	N2 should be sp2 (and aromatic)	ok	ok
3zs1	error	ok	CHA,CHB,CHC,CHD should be aromatic (HEM)	ok	ok
1nym	error	ok	N16 should be sp2	ok	ok
2g79	error	ok	C11 should be sp2 (long sp2 chain)	ok	ok
3cde	error	ok	N22 should be sp2 (in a non-aromatic heteroring)	ok	ok
3ozp	error	ok	CAA should be sp2	ok	ok
3ogx	error	ok	C19 should be sp2 (stand-alone double bond in a non-aromatic ring)	ok	ok
1y2j	error	ok	either O21 or O20 should be sp2 (NO2)	ok	ok
2yjx	error	ok	C2 should be sp2 (stand-alone db in a non-aromatic ring) while C1-C3-C7-C9-N4 ring should be non-aromatic because of the oxygen	ok	ok
2wca	error	ok	C10 should be sp2	ok	ok
1ur9	error	ok	N2 should be sp2 (we are not sure, because it is actually, not one ligand, but several joined ligands)	ok	ok
3dpk	tautomer	ok	O16 is sp2 in the reference (exo oxygen near heteroaromaics)	ok	ok

3gk2	error	C12 should be sp3, a db should be between N3,N4	a db should be formed between N3,N4; C13 should be sp2	C12 should be sp3, a db should be between N3,N4	C12 should be sp3, a db should be between N3,N4
1u65	error	ok	C14 should be sp2	ok	ok
1u33	error	ok	NA1 should be sp2 (amidinium)	ok	ok
4enx	error	ok	C13 should be sp2 while O21 should be sp2 and the ring shouldn't be aromatic	ok	ok
2w4x	error	ok	NAO should be sp2 (terminal NO)	ok	ok
4eny	error	ok	C15, N8 should be sp2, the ring shouldn't be aromatic	ok	ok
4hbp	tautomer		O1,C2,N10		
3s3o	error	ok	OAD should be sp2 (exo oxygen near heterocycle)	ok	ok
2oxn	error	CAF,OAM should be sp3 (terminal oxygen)	NAY should be sp2	ok	ok
2vvp	error	ok	O2 should be sp2 (terminal CO)	ok	ok
2i72	tautomer		O18-C17-N22		
2i1r	error	ok	C8 should be sp2, the adjacent ring should be non-aromatic because of an sp2 exo oxygen	ok	ok
3vru	error	ok	CAV should be sp2	ok	ok

1ke8	error	ok	O5 should be sp2, the ring shouldn't be aromatic	ok	ok
3s3m	error	ok	OAD should be sp2 (exo oxygen near heterocycle)	ok	ok
3s6t	error	ok	NAY should be sp2	ok	ok
2wly	error	ok	S7 should be sp2 (amide)	ok	ok
3zlr	error	ok	C7 should be sp2 (planar carbon)	ok	ok
2yki	error	ok	C19 should be sp2 (aromatic)	ok	ok
2x2l	error	ok	CAF should be sp2, the adjacent ring should be non-aromatic because of an sp2 exo oxygen	ok	ok
2clk	error	ok	O1 should be sp2 (terminal CO)	ok	ok
4f9v	error	ok	SAC should be sp2	ok	ok
2nv7	error	ok	N1 should be sp2	ok	ok
3lbl	error	ok	C7 should be sp2 (aromatics lost)	ok	ok
3opp	error	C5,N4 should not form a db (stand-alone bond in a chain)	ok	C5,N4 should not form a db (stand-alone bond in a chain)	C5,N4 should not form a db (stand-alone bond in a chain)
1d7x	error	C17,N21 should not form a db (stand-alone bond in a chain)	ok	C17,N21 should not form a db (stand-alone bond in a chain)	C17,N21 should not form a db (stand-alone bond in a chain)
1rw8	error	ok	C5 should be sp2 (aromatic)	ok	ok

1pxk	error	ok	N9 should be sp2 (amidinium)	ok	ok
1p1n	error	ok	CD1 should be sp2 (terminal CC)	ok	ok
4bjc	error	ok	C10 should be sp2 (and aromatic)	ok	ok
2xg5	error	ok	C22-C27-C28-C29-S30 cycle should be non-aromatic	ok	ok
3g90	error	ok	N2 should be sp2, the adjacent ring should be non-aromatic because of an sp2 exo oxygen	ok	ok
1tt1	error	ok	CD1 should be sp2 (terminal CC)	ok	ok
2vvs	error	ok	CAA should be sp2	ok	ok
2o5d	error	ok	C8 should be sp2, the adjacent ring should be non-aromatic because of an sp2 exo oxygen	ok	ok
3bki	error	ok	NAF, NAP should be sp2 (and aromatic)	ok	ok
3zw3	error	ok	C10,O17 should be sp2	ok	ok
3psd	error	ok	C25 should be sp2 (and aromatic)	ok	ok
3ao4	error	ok	C5 should be sp2 (aromatic)	ok	ok
3ewc	error	ok	C9 should be sp2 (and aromatic)	ok	ok

2qrh	error	ok	N1 should be sp2 (stand-alone double bond in a non-aromatic ring)	ok	ok
4ie7	error	ok	OAB should be sp2 (exo oxygen near non-aromatic ring)	ok	ok
3fnm	error	ok	N2 should be sp2 (5-membered heteroaromatics)	ok	ok
2a4z	error	ok	S4 should be sp3 (a very strange error)	ok	ok
1pmv	error	ok	O22 should be sp2 (exo oxygen near conjugated aromatics)	ok	ok
3dej	error	ok	OAF should be sp2, the heterocycle should be not aromatic	ok	ok
1ke9	error	ok	O11 should be sp2 (exo oxygen near non-aromatic heteroring)	ok	ok
2yof	error	ok	SAE should be sp2	ok	ok
1f0q	error	ok	O6,O19 should be sp2 (exo oxygens near non-aromatic ring)	ok	ok
SOME OTHER ISSUES					
Fconv experience big problems with N.pl3 nitrogens : there are 62 more ligands that are not listed here, in which a nitrogen atom (mainly located near aromatics) should be sp2 and N.pl3, not N.3					

Table S3: Running times of the *Knodle*, fconv, NAOMI, and I-interpret methods for the PDB-Bind 2014 dataset. Timings for two series of tests are given, the first for 10,605 protein-ligand complexes, and the second for 10,605 isolated ligands. Running times for *Knodle* are given for the two versions of its SVM kernels, the exact one and the approximate one.

Package	Total time of extraction from the PDB format (protein-ligand complexes, 10,605 instances), s		Total time of extraction from the PDB format (isolated ligands, 10,605 instances), s	
	Linux	Mac OS	Linux	Mac OS
Knodle (exact slow kernels)	107	129.7	80	97.8
Knodle (approximate fast kernels)	87	112.8	59	79.5
Fconv	212	370.4	30	65.7
NAOMI	232	-[a]	74	-[a]
I-interpret	-[b]	-[a]	78	-[a]

[a] NAOMI and I-interpret executables for Mac OS are not available. [b] Unfortunately, the desktop version of I-interpret often hanged on protein complexes with more than one ligand and thus we excluded I-interpret from this comparison.