

SUPPORTING INFORMATION

Decomposition of Small Molecules for Fragment-Based Drug Design

Nikita N. Ivanov, Dmitry A. Shulga and Vladimir A. Palyulin *

Department of Chemistry, Lomonosov Moscow State University, 119991 Moscow, Russia

* Correspondence: vap@qsar.chem.msu.ru

Table of contents

Tables	1
Usages.....	22

Tables

Table S1. SMARTS expressions

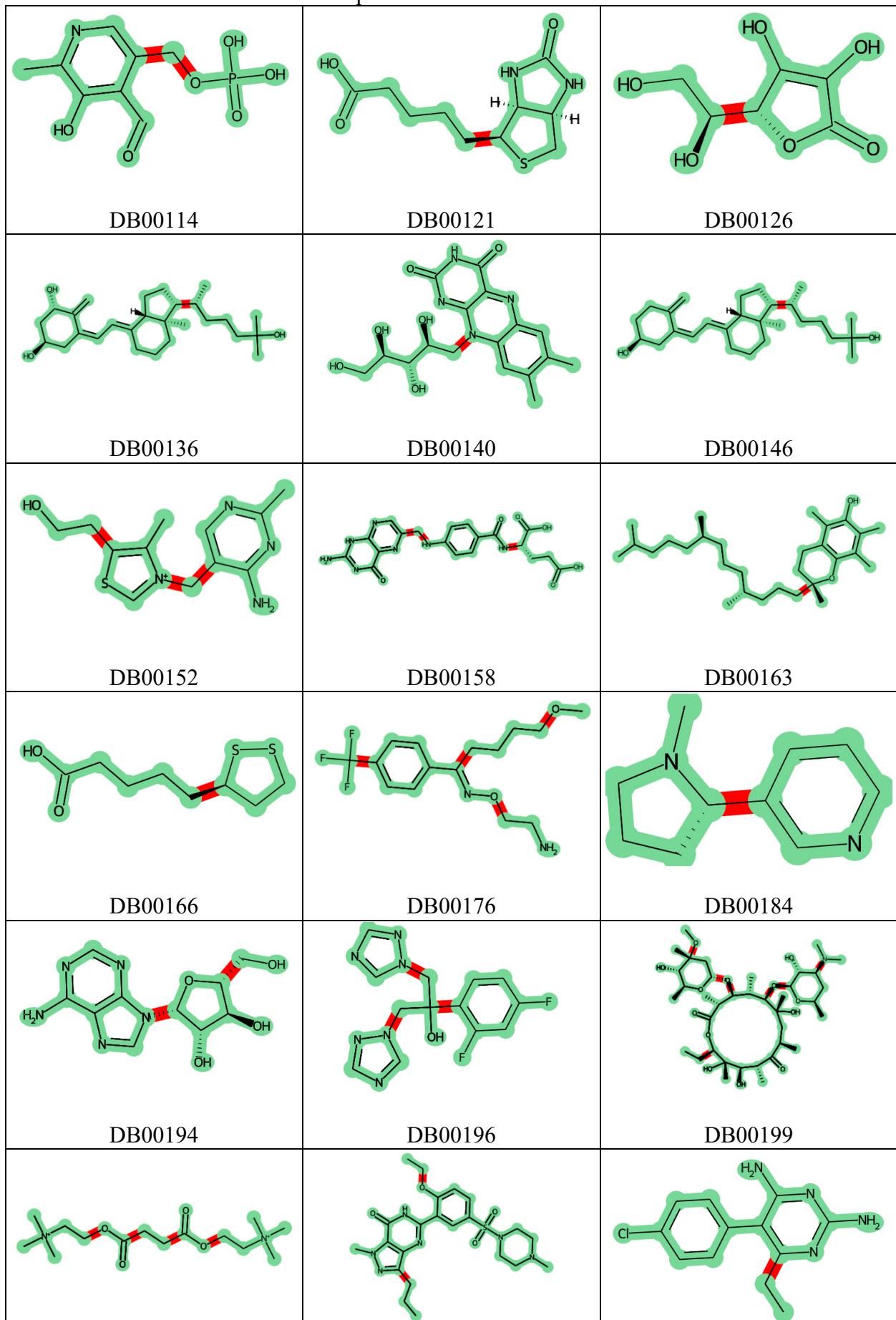
Nº	SMARTS expressions**	Indices*	Rule, №
1	[R] - !@[\$([CX4 ; H2 , H1 , H0])]	(0, 1)	1
2	[a] - !@[\$([NX3 ; H1 , H0]), \$([OX2 ; H0]), \$([SX2 ; H0])] - !@[\$([C ; H2 , H1 , H0]) ; !\$([CX3] = [OX1])]	(1, 2)	2
3	[a] - !@[\$([NX3 ; H1 , H0]), \$([OX2 ; H0]), \$([SX2 ; H0])] - !@[a]	(0, 1), (1, 2)	3
4	[a] - !@[\$([CX3] = [OX1 , NX2 , SX1 , CX3])] - !@[\$([CX4 ; H2 , H1 , H0])]	(1, 2)	2
5	[c] - !@[\$([CX3] = [OX1 , NX2 , SX1 , \$([CX3 ; H2])])] - !@[c]	(0, 1), (1, 2)	3
6	[n] - !@[\$([CX3] = [OX1 , NX2 , SX1 , \$([CX3 ; H2])])] - !@[c]	(1, 2)	2
7	[\$([CX4 ; H2 , H1 , H0])] - !@[CX3] (= [OX1]) [OX2 ; H0]	(0, 1)	4
8	[\$([CX4 ; H2 , H1 , H0])] - !@[OX2 ; H0] [CX3] (= [OX1])	(0, 1)	4
9	[a] - !@[CX3] (= [OX1]) O - !@[\$([CX4 ; H2 , H1 , H0])]	(3, 4)	2
10	[a] - !@[CX3] (= [OX1]) O - !@[a]	(0, 1), (3, 4)	3
11	[a] - !@[NX2 ; H0] = [NX2 ; H0] - !@[\$([CX4 ; H2 , H1 , H0])]	(2, 3)	2
12	[a] - !@[NX2 ; H0] = [NX2 ; H0] - !@[a]	(0, 1), (2, 3)	3
13	[a] - !@[NX3 ; H1] - !@[\$([CX3 ; H0] (= [OX1]))] - !@[\$([CX4 ; H2 , H1 , H0])]	(2, 3)	2
14	[a] - !@[\$([CX3 ; H0] (= [OX1]))] - !@[NX3 ; H1] - !@[\$([CX4 ; H2 , H1 , H0])]	(2, 3)	2
15	[a] - !@[NX3 ; H1] - !@[\$([CX3 ; H0] (= [OX1]))] - !@[a]	(0, 1), (2, 3)	3
16	[a] - !@[\$([CX3 ; H0] (= [OX1]))] - !@[NX3 ; H1] - !@[a]	(0, 1), (2, 3)	3
17	[a] - !@[SX4] (= [OX1]) (= [OX1]) [NX3 ; H1] - !@[\$([CX4 ; H2 , H1 , H0])]	(4, 5)	2
18	[a] - !@[NX3 ; H1] [SX4] (= [OX1]) (= [OX1]) - !@[\$([CX4 ; H2 , H1 , H0])]	(4, 5)	2
19	[a] - !@[SX4] (= [OX1]) (= [OX1]) [NX3 ; H1] - !@[a]	(0, 1), (4, 5)	3
20	[a] - !@[NX3 ; H1] [SX4] (= [OX1]) (= [OX1]) - !@[NX3 ; H1] - !@[a]	(0, 1), (4, 5)	3
21	[\$([CX4 ; H2 , H1 , H0])] - !@[NX3] [CX3] (= [OX1])	(0, 1)	4
22	[\$([CX4 ; H2 , H1 , H0])] - !@[CX3] (= [OX1]) [NX3]	(0, 1)	4
23	[\$([CX4 ; H2 , H1 , H0])] - !@[\$([NX3 ; H1 , H0])]	(0, 1)	4
24	[\$([CX4 ; H2 , H1 , H0])] - !@[\$([OX2 ; H0])]	(0, 1)	4
25	[\$([CX4 ; H2 , H1 , H0])] - !@[\$([SX2 ; H0])]	(0, 1)	4

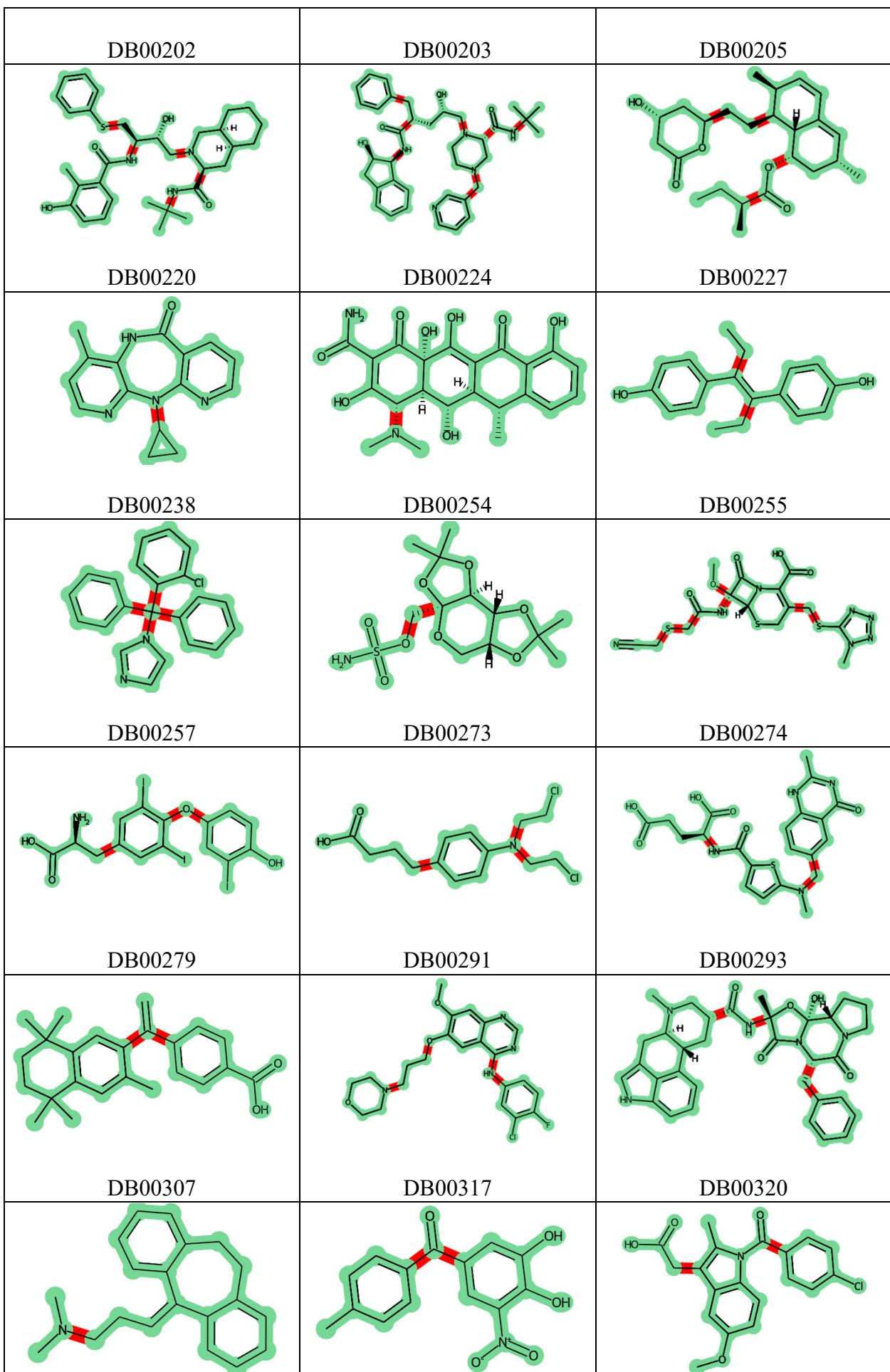
No	SMARTS expressions**	Indices*	Rule, №
26	<chem>[\$([CX4;H2,H1,H0])]-!@[SX4](=[OX1])(=[OX1])[NX3;H1]</chem>	(0, 1)	4
27	<chem>[\$([CX4;H2,H1,H0])]-!@[NX3;H1][SX4](=[OX1])(=[OX1])</chem>	(0, 1)	4

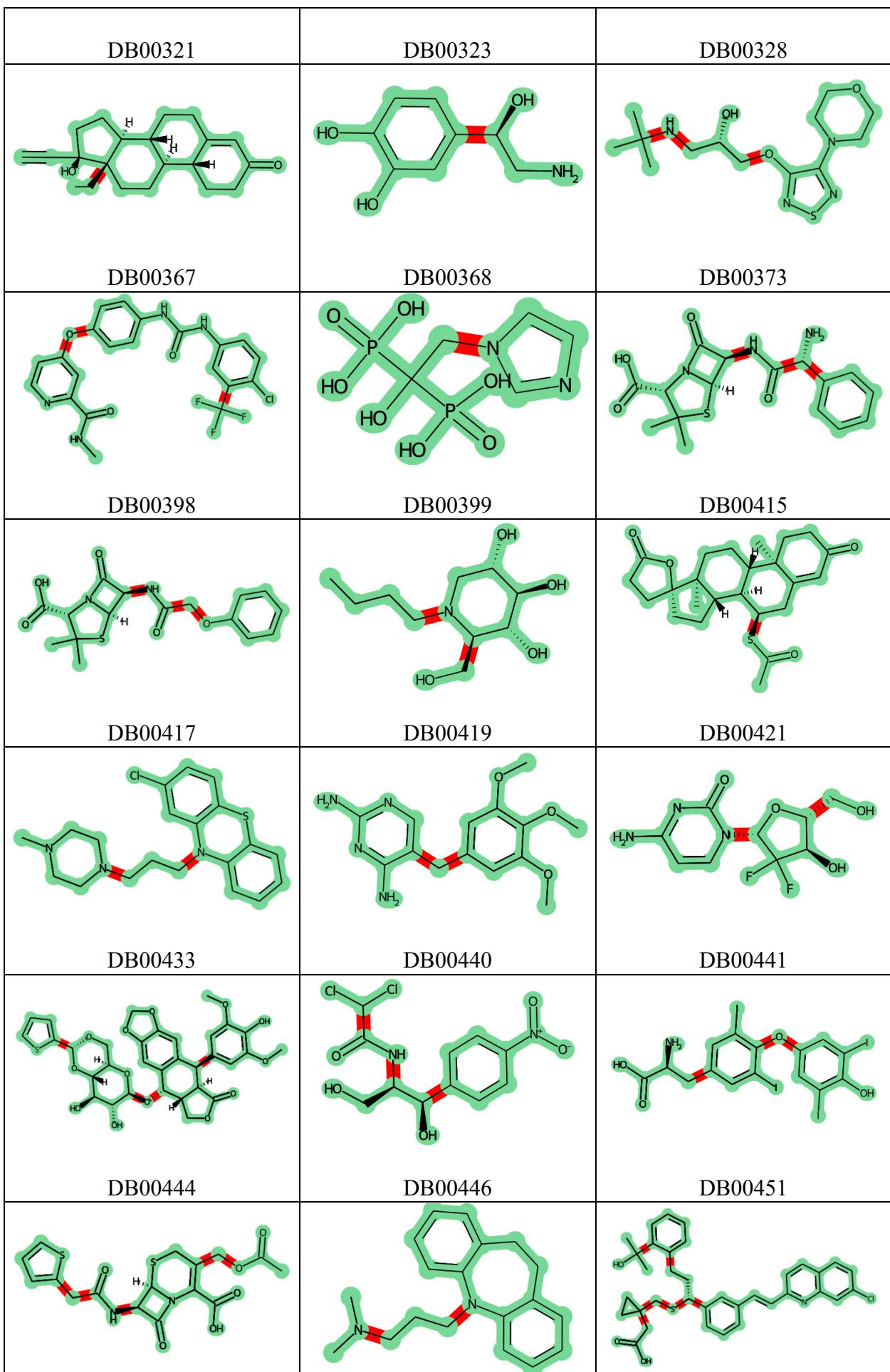
*-the indices of atomic pairs between which the bond should be removed; indexing starts from 0.
 ** - the bonds that need to be broken are marked by red color.

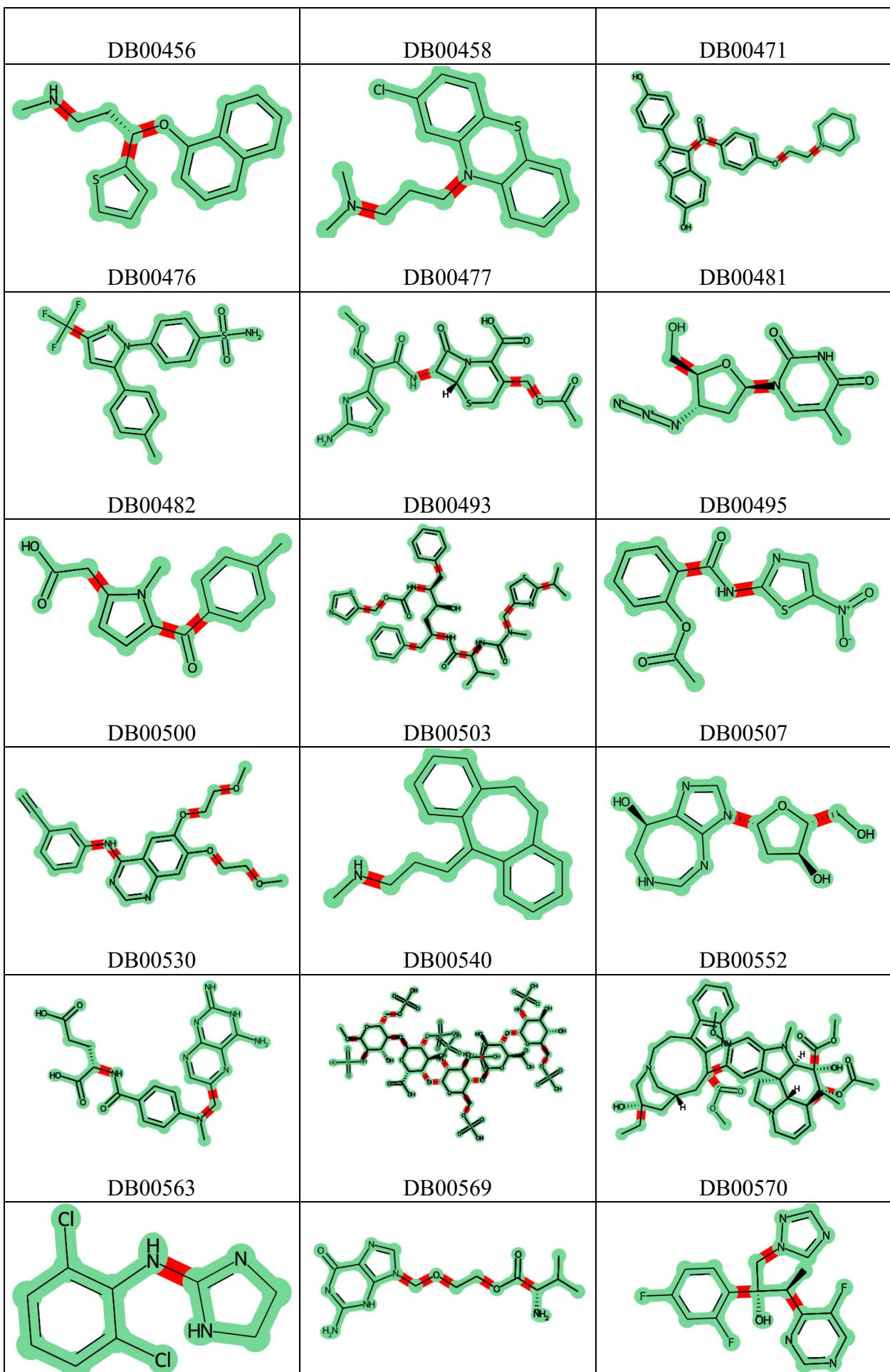
The set of SMARTS and the script for decomposition of molecules into fragments is freely accessible on the web (http://molmodel.com/hg/medchem_fragment_splitter).

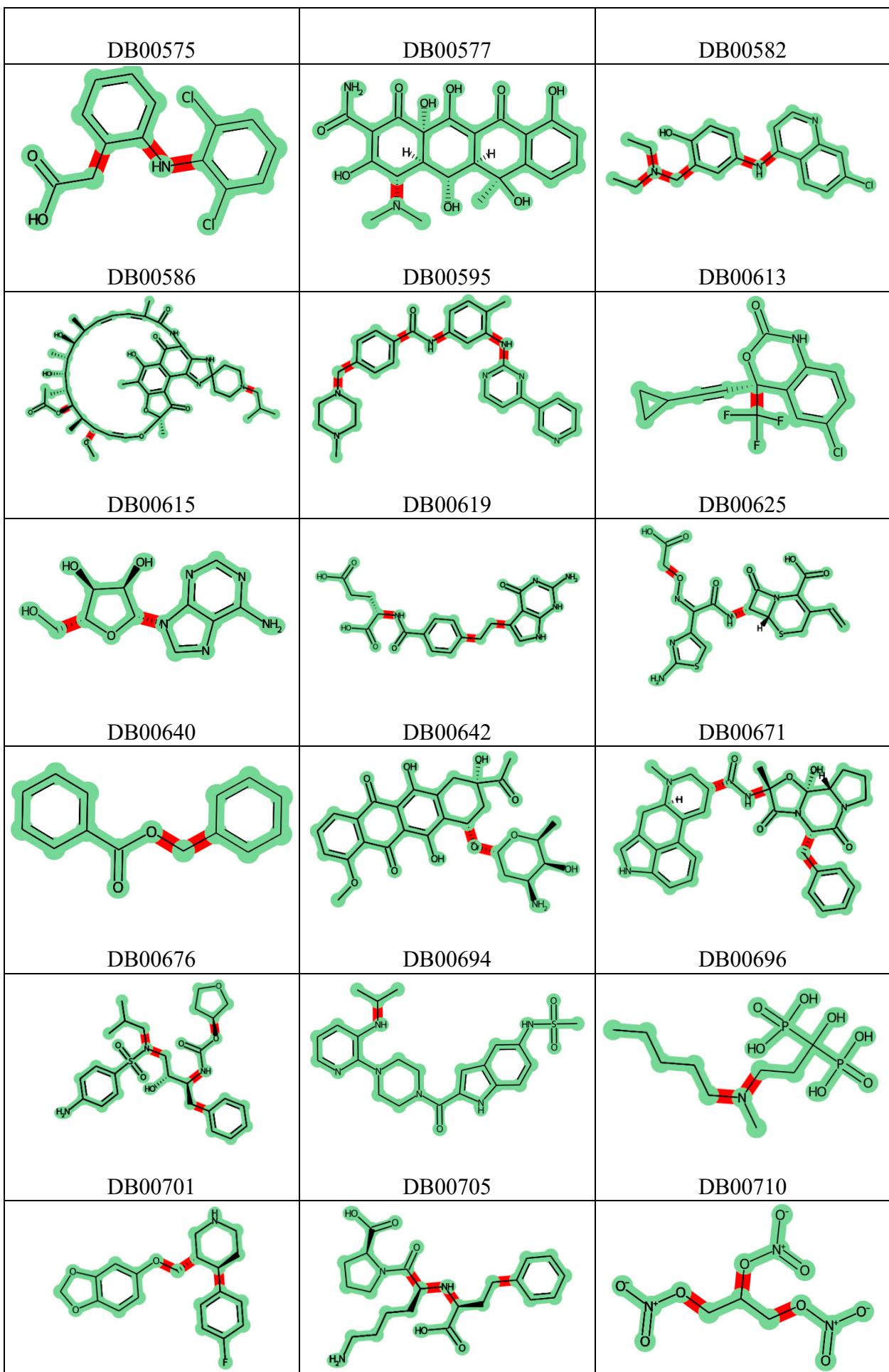
Table S2. Results of molecules decomposition

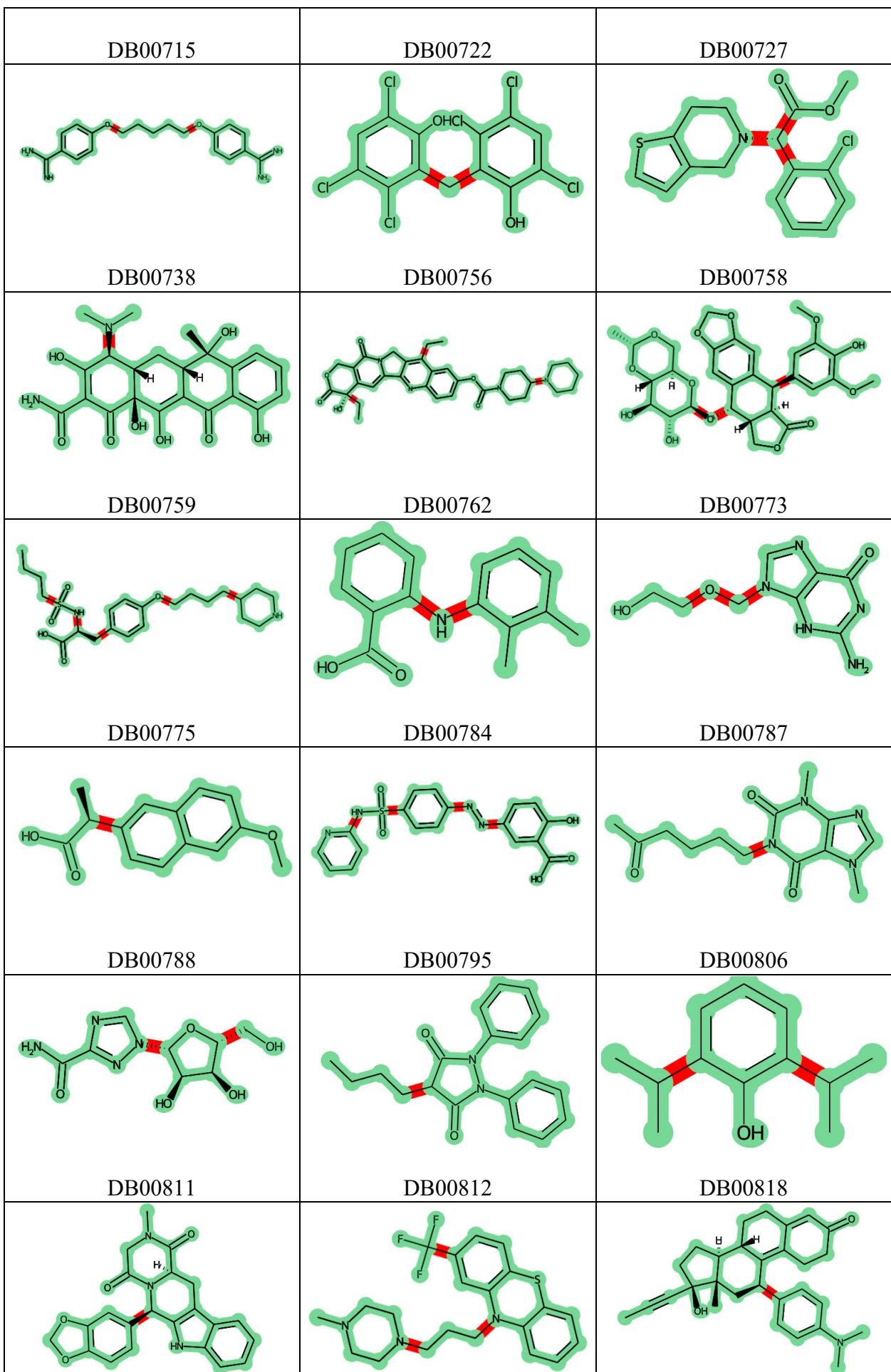


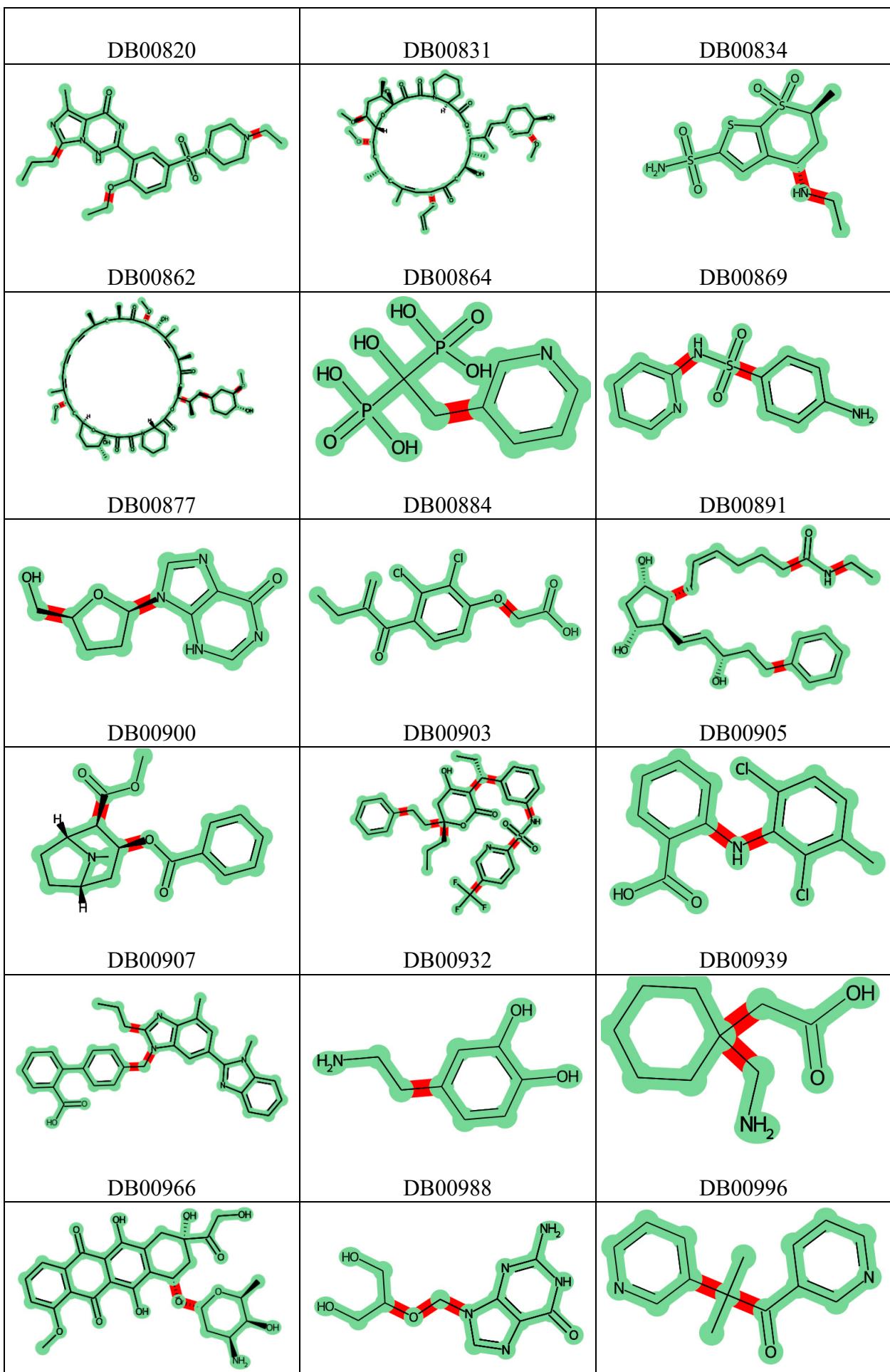


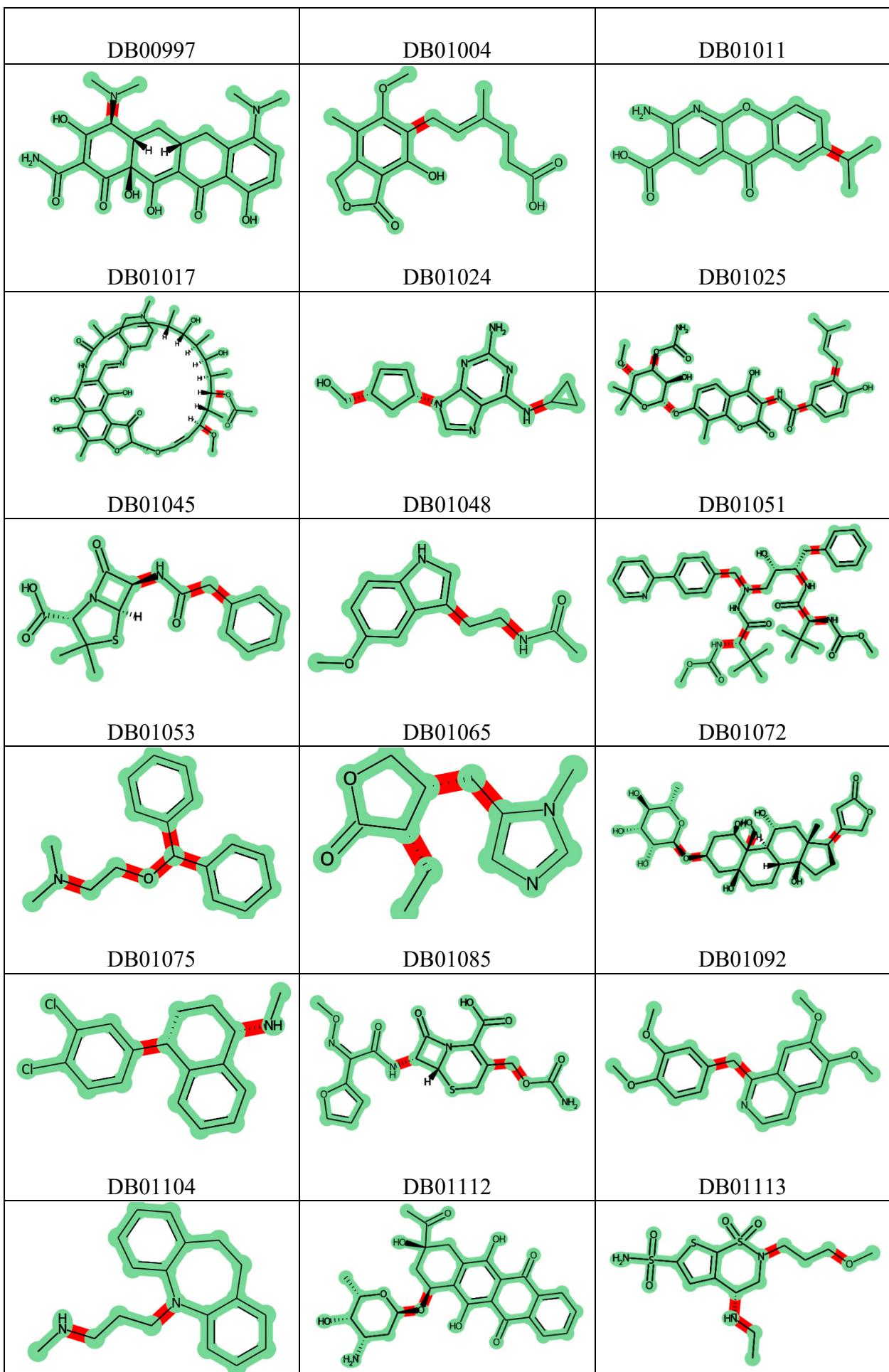


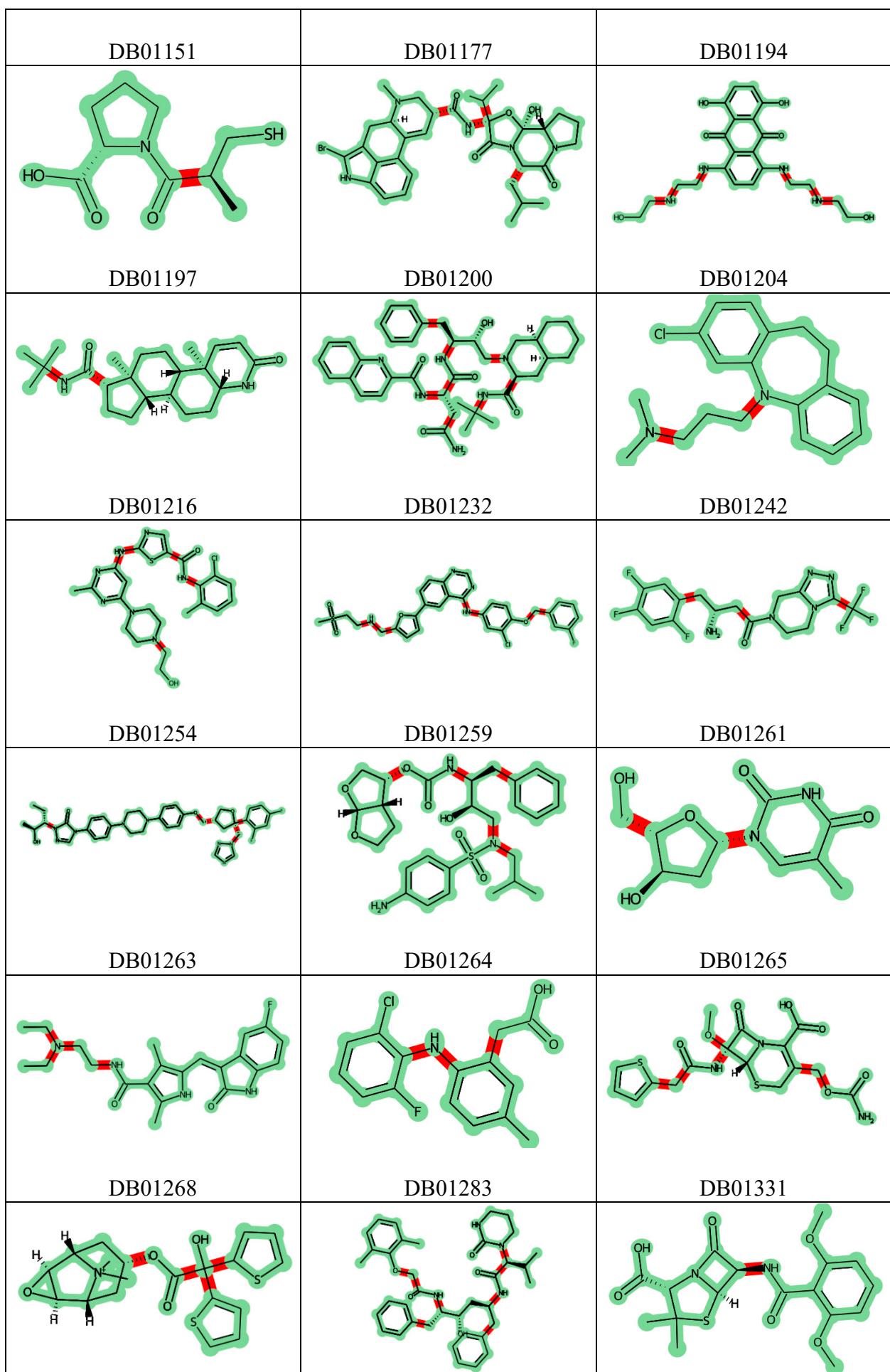


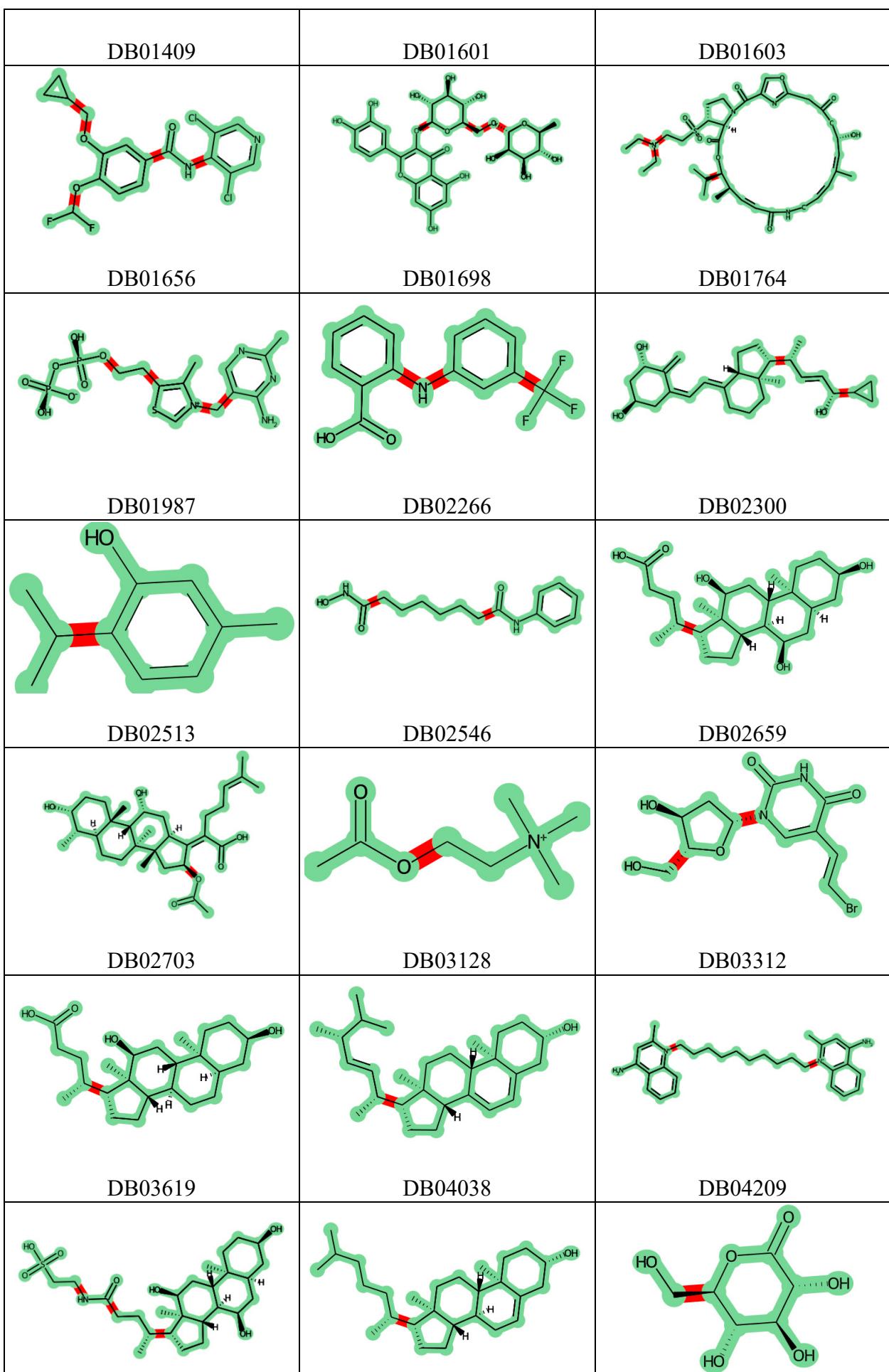


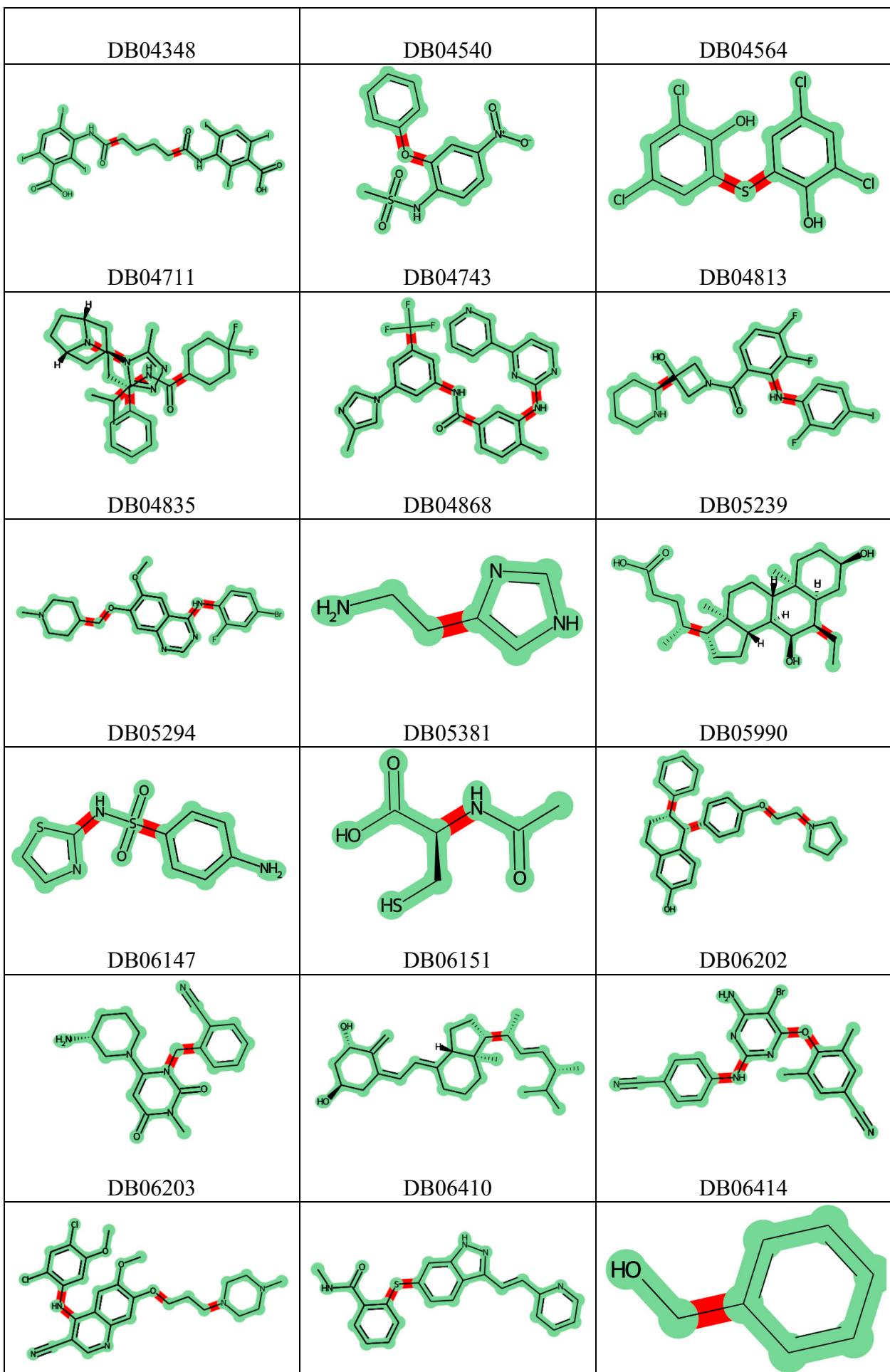


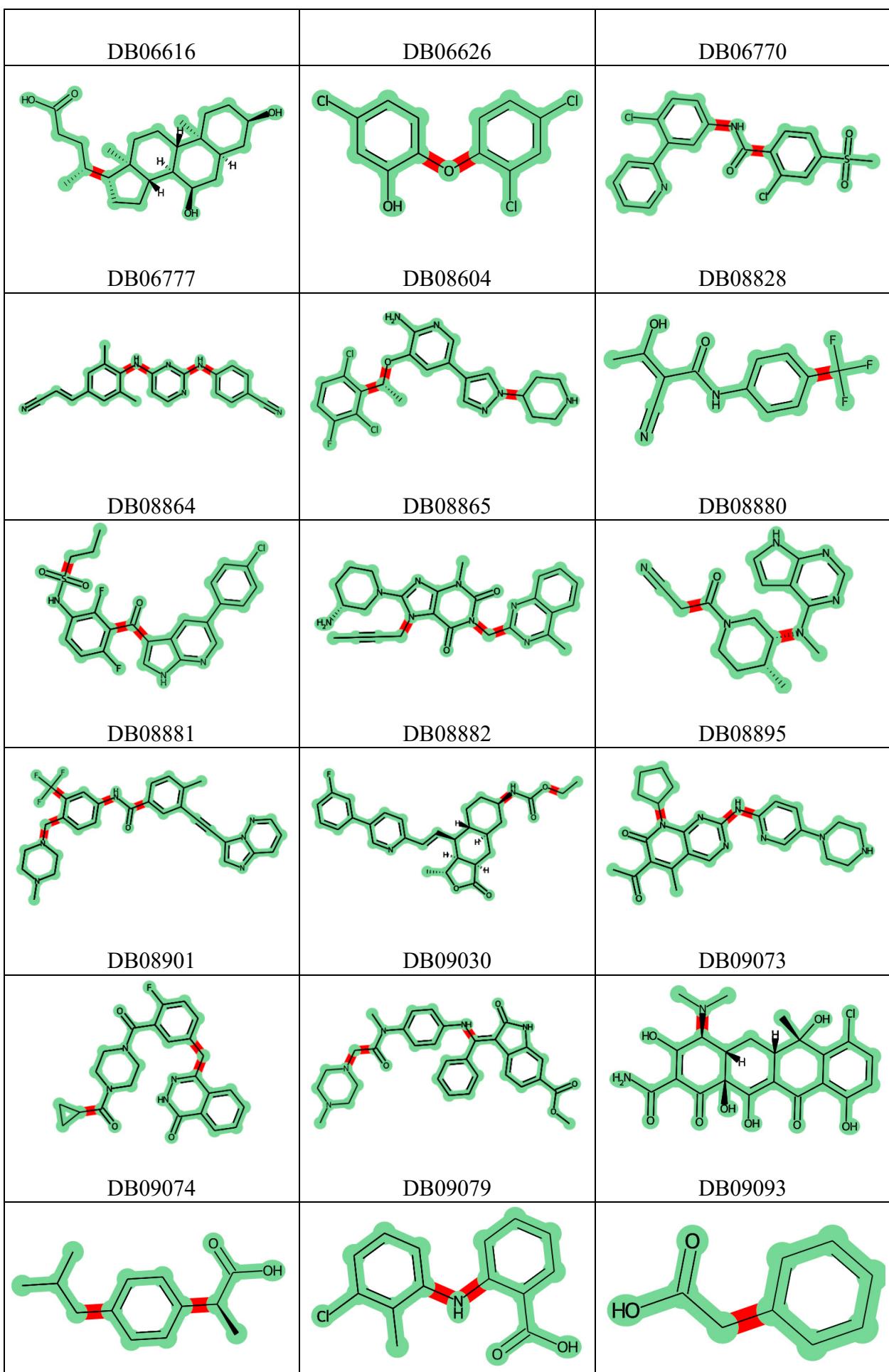












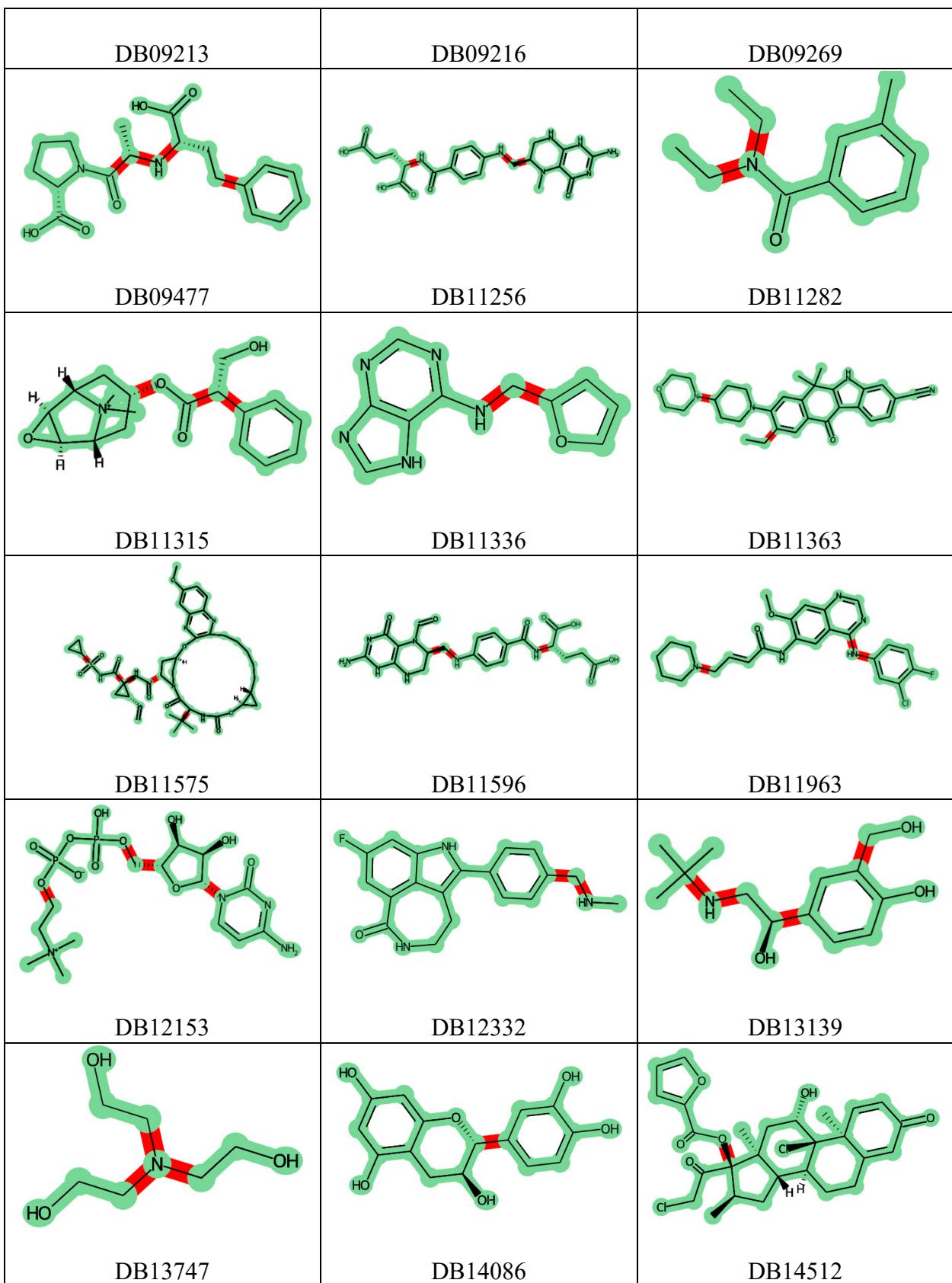
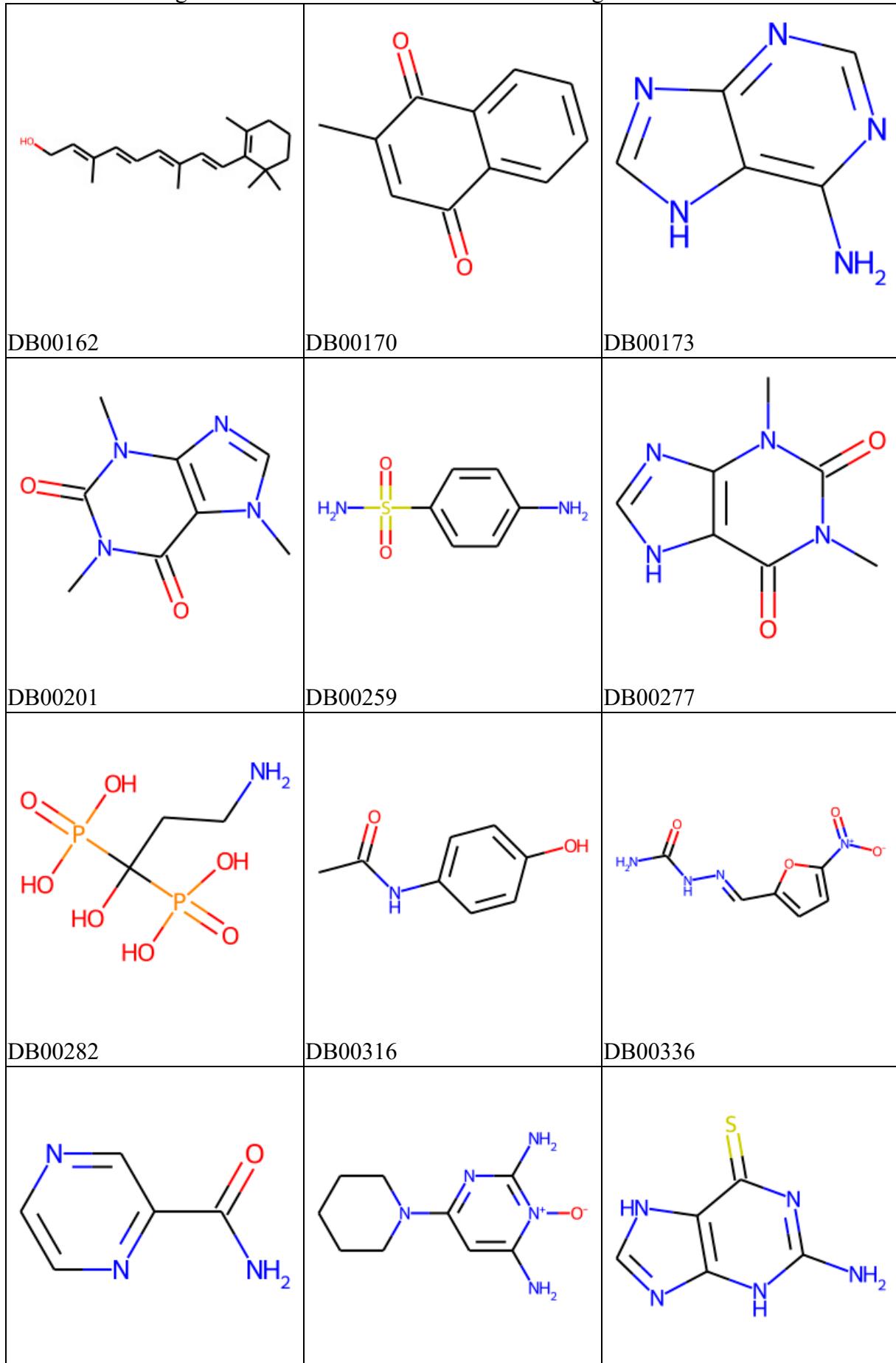
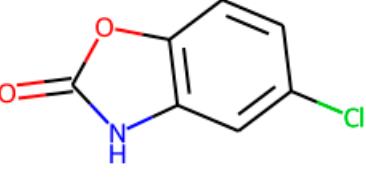
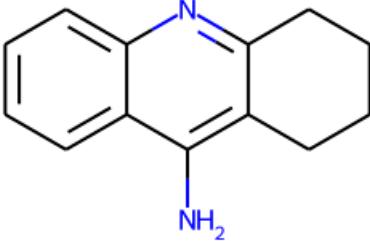
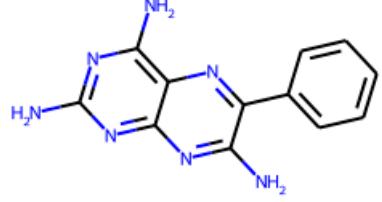
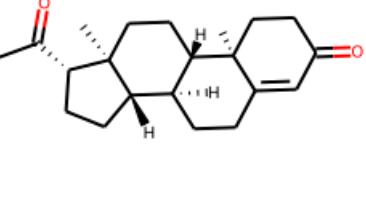
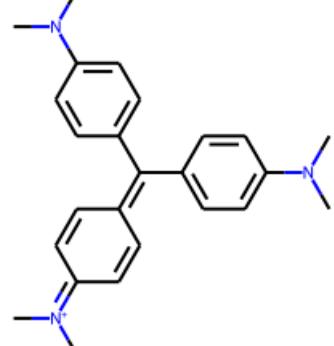
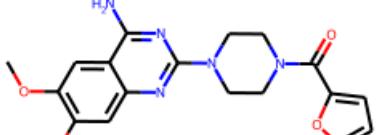
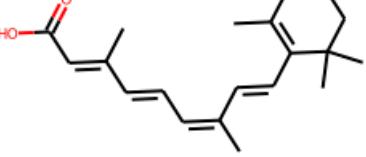
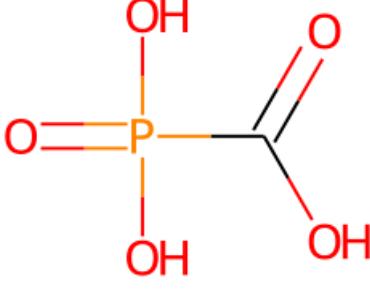
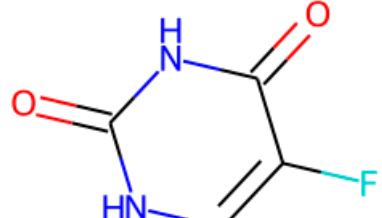
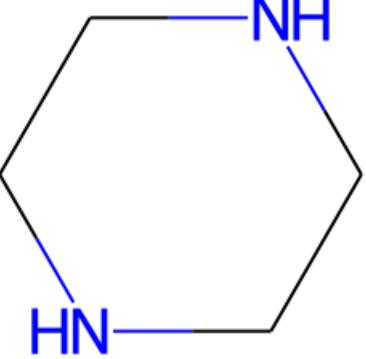
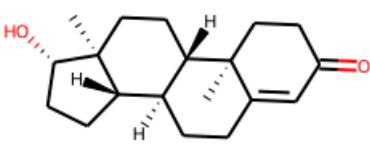
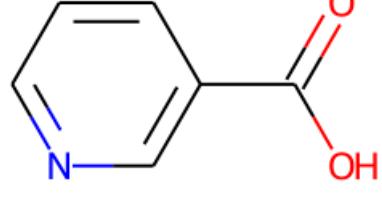
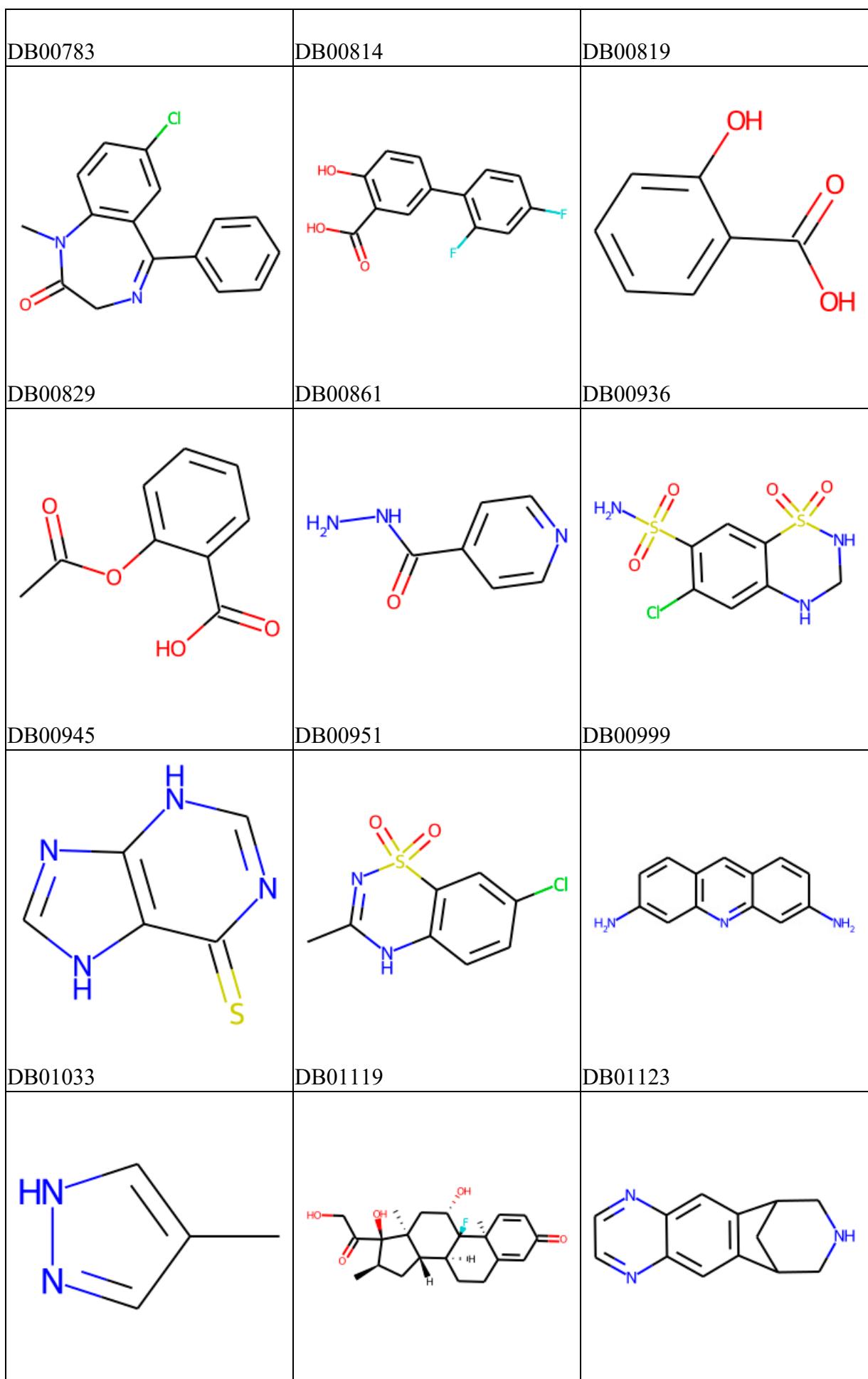


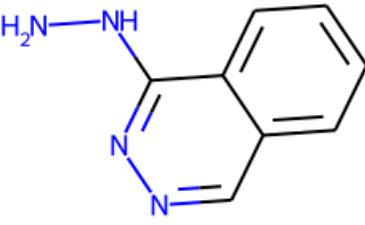
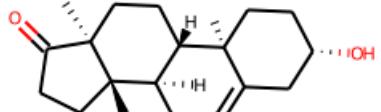
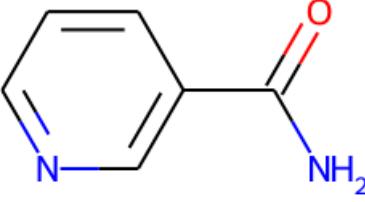
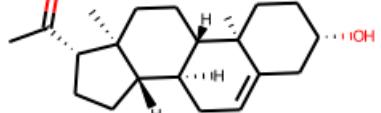
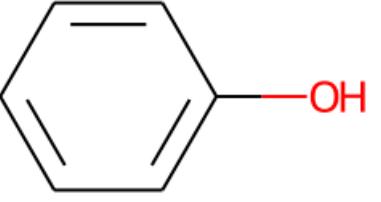
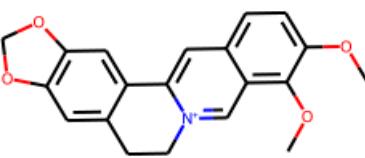
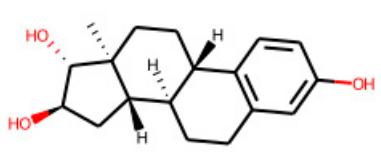
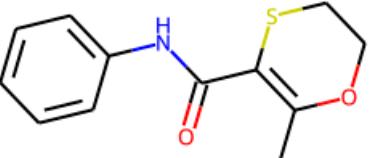
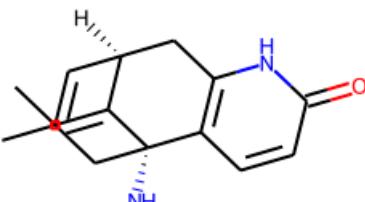
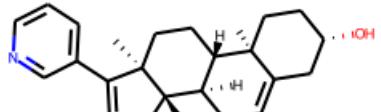
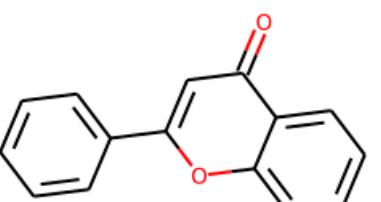
Table S3. The drug molecules which were not divided into fragments

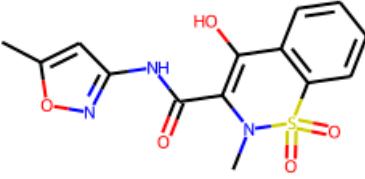
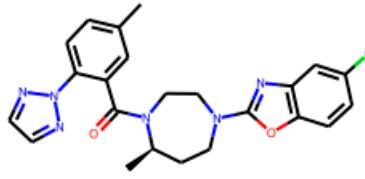
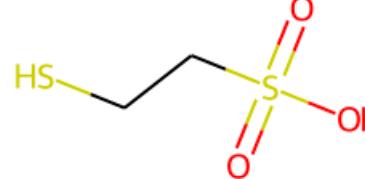
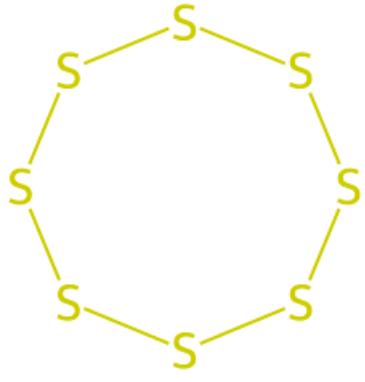
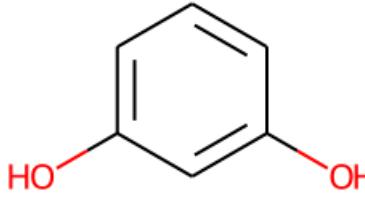
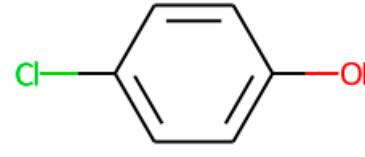
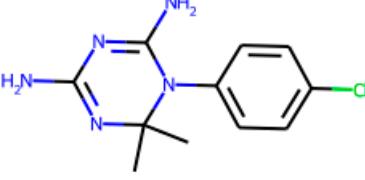


DB00339	DB00350	DB00352
		
DB00356	DB00382	DB00384
		
DB00396	DB00406	DB00457
		
DB00523	DB00529	DB00544
		

DB00592	DB00624	DB00627
DB00655	DB00674	DB00687
DB00717	DB00730	DB00741
DB00755	DB00763	DB00766



DB01213	DB01234	DB01273
		
DB01275	DB01708	DB02362
		
DB02701	DB02789	DB03255
		
DB04115	DB04573	DB04657
		

DB04864	DB05812	DB07776
		
DB08942	DB09034	DB09110
		
DB09353	DB11085	DB13154
		
DB14763		

Usages

The application consists of python scripts and SMARTS expression data. The following options are available in script:

- l - path to ligand file;
- w - path to working directory;
- s - path to file with SMARTS expression;
- p - no parameter, option if needed a picture;
- x - no parameter, option if needed to save fragment-files
- a - no parameter, option if needed to be decomposed by inner rules

The application can work in several modes. The first mode decomposes molecules by the built-in set of SMARTS-expressions. It is achieved using parameter ‘-s’ and a file with defined SMARTS-expressions:

```
./defragmentation.py -l ..path-to-ligand -w ..path-to-work-
directory -s ..SMARTS-expressions -x -p
```

What is more, these expressions must describe non-overlaid fragments. The second mode lets to decompose molecules by the rules described in current article. It is achieved using parameter ‘-a’:

```
./defragmentation.py -l ..path-to-ligand -w ..path-to-work-
directory -x -p -a
```

The parameter ‘-x’ is used to write files of fragments:

```
./defragmentation.py -l ..path-to-ligand -w ..path-to-work-
directory -x -a
```

The parameter ‘-p’ is used to generate a picture of a molecule's structure with colored fragments:

```
./defragmentation.py -l ..path-to-ligand -w ..path-to-work-
directory -p -a
```

To produce the output at least one parameter ‘-x’ or ‘-p’ should be used.