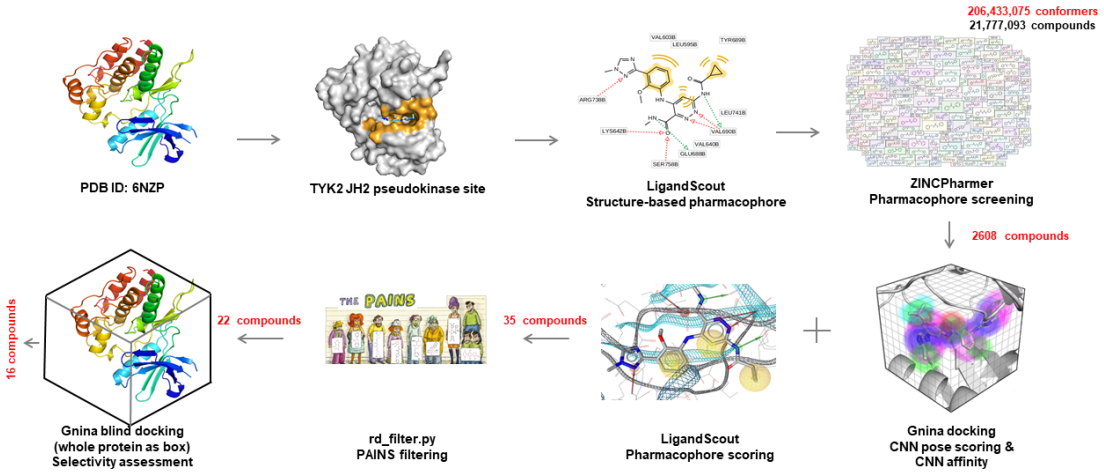
6_TYK2_JH2_Potential_Selective_Allosteric_Inhibitors

April 8, 2022

Pipeline Map



Results from pharmacophore- and docking-based virtual screening targeting TYK2 Pseudokinase domain (JH2) to identify selective inhibitors through the pipeline illustrated below

[1]:
```
import pandas as pd
from rdkit import Chem
from rdkit.Chem.Draw import MolsToGridImage
from rdkit.Chem import PandasTools
```

[2]:
```
hits_df = pd.read_csv('../results/16.smi', sep=' ', header=None)
hits_df.shape
```

[2]: (16, 2)

[3]:
```
hits_df
```

[3]:

	0	1
0	O(c1ccc(cc1)[C@H]4n2nc(nc2NC(c3cc(OC)ccc3)C4)N)CC	ZINC13131776
1	Clc1c(ccc(Cl)c1)[C@H]3C[C@H](n2nc(nc2N3)N)c4cc…	ZINC13131298
2	O=C(c3cc2n1c(nnc1)c(nc2cc3)NC4CCCC4)c5ccccc5	ZINC06901595
3	O=C(c3cc2n1c(nnc1C)c(nc2cc3)NC[C@@H]4N(CCC4)CC…	ZINC20939758

```
4       O=C(N(C)C)c3ccc(c1nc(nc(n1)N)Nc2ccc(cc2)C)cc3  ZINC32540567
5         O(C(c3ccc(c1nc(nc(n1)N)Nc2ccc(cc2)C)cc3)C)C  ZINC32541042
6         Fc1c(c(F)ccc1)CNc4n2ncnc2nc(c3ccc(OC)cc3)c4  ZINC23432014
7   O(c1ccc(cc1)[C@H]4n2nc(nc2NC(c3ccc(OC)cc3)C4)N…  ZINC02472318
8   O(c1ccc(cc1)[C@H]4n2nc(nc2NC(c3cc(OC)ccc3)C4)N…  ZINC02472304
9          S=C(Nc2nc(c1c(O)cccc1)cc(n2)c3ccc(OC)cc3)N  ZINC01084111
10  Clc1ccc(cc1)C(=O)c3sc2n(c(=O)[nH]c(=O)c2c3N)c4…  ZINC02638281
11  Fc1ccc(cc1)C3Nc2n(nc(n2)N)[C@@H](C3)c5ccc(OCc4…  ZINC02472295
12    [nH]2c1nc(NCCN(C)C)cc(c1cc2)c4cc3c(cnc(N)c3)cc4  ZINC72143513
13  O(c1c(cccc1)CNc3nc4c(n2c3nnc2C)cc(cc4)C(=O)c5c…  ZINC09224703
14  o1c(nnc1)c5ccc(c4c2c([nH]cc2)nc(NCc3ccncc3)c4)cc5  ZINC72454514
15     n1(ncc4c1ccc(c3c2c([nH]cc2)nc(NCCN(C)C)c3)c4)C  ZINC72168947
```

[4]:
```
hits_df.columns = ['SMILES', 'ZINC_ID']
hits_df
```

[4]:
```
                                               SMILES       ZINC_ID
0   O(c1ccc(cc1)[C@H]4n2nc(nc2NC(c3cc(OC)ccc3)C4)N)CC  ZINC13131776
1   Clc1c(ccc(Cl)c1)[C@H]3C[C@H](n2nc(nc2N3)N)c4cc…  ZINC13131298
2         O=C(c3cc2n1c(nnc1)c(nc2cc3)NC4CCCC4)c5ccccc5  ZINC06901595
3   O=C(c3cc2n1c(nnc1C)c(nc2cc3)NC[C@@H]4N(CCC4)CC…  ZINC20939758
4       O=C(N(C)C)c3ccc(c1nc(nc(n1)N)Nc2ccc(cc2)C)cc3  ZINC32540567
5         O(C(c3ccc(c1nc(nc(n1)N)Nc2ccc(cc2)C)cc3)C)C  ZINC32541042
6         Fc1c(c(F)ccc1)CNc4n2ncnc2nc(c3ccc(OC)cc3)c4  ZINC23432014
7   O(c1ccc(cc1)[C@H]4n2nc(nc2NC(c3ccc(OC)cc3)C4)N…  ZINC02472318
8   O(c1ccc(cc1)[C@H]4n2nc(nc2NC(c3cc(OC)ccc3)C4)N…  ZINC02472304
9          S=C(Nc2nc(c1c(O)cccc1)cc(n2)c3ccc(OC)cc3)N  ZINC01084111
10  Clc1ccc(cc1)C(=O)c3sc2n(c(=O)[nH]c(=O)c2c3N)c4…  ZINC02638281
11  Fc1ccc(cc1)C3Nc2n(nc(n2)N)[C@@H](C3)c5ccc(OCc4…  ZINC02472295
12    [nH]2c1nc(NCCN(C)C)cc(c1cc2)c4cc3c(cnc(N)c3)cc4  ZINC72143513
13  O(c1c(cccc1)CNc3nc4c(n2c3nnc2C)cc(cc4)C(=O)c5c…  ZINC09224703
14  o1c(nnc1)c5ccc(c4c2c([nH]cc2)nc(NCc3ccncc3)c4)cc5  ZINC72454514
15     n1(ncc4c1ccc(c3c2c([nH]cc2)nc(NCCN(C)C)c3)c4)C  ZINC72168947
```

[5]:
```
PandasTools.AddMoleculeColumnToFrame(hits_df, 'SMILES', 'Mol')
hits_df
```

[5]:
```
                                               SMILES       ZINC_ID  \
0   O(c1ccc(cc1)[C@H]4n2nc(nc2NC(c3cc(OC)ccc3)C4)N)CC  ZINC13131776
1   Clc1c(ccc(Cl)c1)[C@H]3C[C@H](n2nc(nc2N3)N)c4cc…  ZINC13131298
2         O=C(c3cc2n1c(nnc1)c(nc2cc3)NC4CCCC4)c5ccccc5  ZINC06901595
3   O=C(c3cc2n1c(nnc1C)c(nc2cc3)NC[C@@H]4N(CCC4)CC…  ZINC20939758
4       O=C(N(C)C)c3ccc(c1nc(nc(n1)N)Nc2ccc(cc2)C)cc3  ZINC32540567
5         O(C(c3ccc(c1nc(nc(n1)N)Nc2ccc(cc2)C)cc3)C)C  ZINC32541042
6         Fc1c(c(F)ccc1)CNc4n2ncnc2nc(c3ccc(OC)cc3)c4  ZINC23432014
7   O(c1ccc(cc1)[C@H]4n2nc(nc2NC(c3ccc(OC)cc3)C4)N…  ZINC02472318
8   O(c1ccc(cc1)[C@H]4n2nc(nc2NC(c3cc(OC)ccc3)C4)N…  ZINC02472304
```

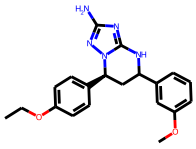
```
9          S=C(Nc2nc(c1c(O)cccc1)cc(n2)c3ccc(OC)cc3)N  ZINC01084111
10  Clc1ccc(cc1)C(=O)c3sc2n(c(=O)[nH]c(=O)c2c3N)c4…  ZINC02638281
11  Fc1ccc(cc1)C3Nc2n(nc(n2)N)[C@@H](C3)c5ccc(OCc4…  ZINC02472295
12    [nH]2c1nc(NCCN(C)C)cc(c1cc2)c4cc3c(cnc(N)c3)cc4  ZINC72143513
13  O(c1c(cccc1)CNc3nc4c(n2c3nnc2C)cc(cc4)C(=O)c5c…  ZINC09224703
14  o1c(nnc1)c5ccc(c4c2c([nH]cc2)nc(NCc3ccncc3)c4)cc5  ZINC72454514
15     n1(ncc4c1ccc(c3c2c([nH]cc2)nc(NCCN(C)C)c3)c4)C  ZINC72168947

                                                  Mol
0   <img data-content="rdkit/molecule" src="data:i…
1   <img data-content="rdkit/molecule" src="data:i…
2   <img data-content="rdkit/molecule" src="data:i…
3   <img data-content="rdkit/molecule" src="data:i…
4   <img data-content="rdkit/molecule" src="data:i…
5   <img data-content="rdkit/molecule" src="data:i…
6   <img data-content="rdkit/molecule" src="data:i…
7   <img data-content="rdkit/molecule" src="data:i…
8   <img data-content="rdkit/molecule" src="data:i…
9   <img data-content="rdkit/molecule" src="data:i…
10  <img data-content="rdkit/molecule" src="data:i…
11  <img data-content="rdkit/molecule" src="data:i…
12  <img data-content="rdkit/molecule" src="data:i…
13  <img data-content="rdkit/molecule" src="data:i…
14  <img data-content="rdkit/molecule" src="data:i…
15  <img data-content="rdkit/molecule" src="data:i…
```

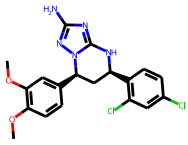
[6]:
```
molecules = [Chem.MolFromSmiles(m) for m in hits_df['SMILES']]
```

[7]:
```
MolsToGridImage([x for x in molecules],
                useSVG=True, molsPerRow=4,
                subImgSize=(200,200),
                legends=[f"{x}" for x in hits_df['ZINC_ID']])
```

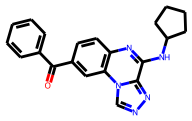
[7]:



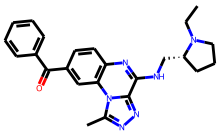
ZINC13131776



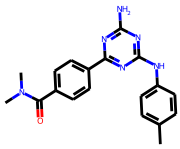
ZINC13131298



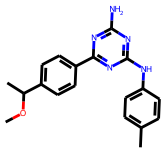
ZINC06901595



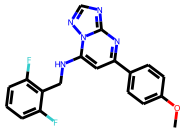
ZINC20939758



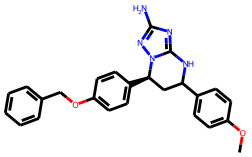
ZINC32540567



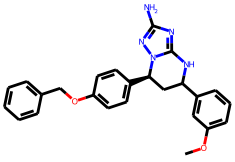
ZINC32541042



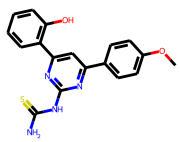
ZINC23432014



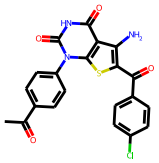
ZINC02472318



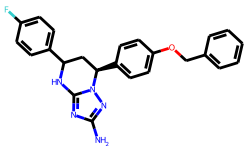
ZINC02472304



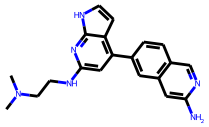
ZINC01084111



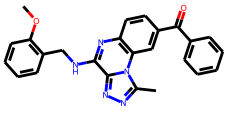
ZINC02638281



ZINC02472295

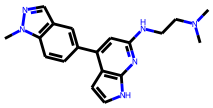


ZINC72143513



ZINC09224703

ZINC72454514



ZINC72168947

[8]:
```
from rdkit.Chem import QED

qed_list = [QED.qed(m) for m in molecules]
qed_list
```

[8]:
```
[0.7203864633232515,
 0.653518332862722,
 0.5615565560749445,
 0.4849126795348229,
 0.7523483858687771,
 0.7357045017110975,
 0.5828318093121403,
 0.4678482661435929,
```

```
 0.4678482661435929,
 0.6209545271886019,
 0.4726425674450768,
 0.4929576995484239,
 0.516186638351186,
 0.40685772684468846,
 0.4807601278517356,
 0.588671762802852]
```

[9]:
```
from rdkit import DataStructs                          #fingerprint handling
from rdkit.ML.Cluster import Butina                    #cluster molecules
from rdkit.Chem import rdMolDescriptors as rdmd        #descriptors

def butina_cluster(mol_list, cutoff=0.35):
    fp_list = [rdmd.GetMorganFingerprintAsBitVect(m, 3, nBits=2048)
               for m in mol_list]
    dists = []
    nfps = len(fp_list)
    for i in range(1, nfps):
        sims = DataStructs.BulkTanimotoSimilarity(fp_list[i], fp_list[:i])
        dists.extend([1-x for x in sims])
    mol_clusters = Butina.ClusterData(dists, nfps, cutoff,
                                      isDistData=True)

    cluster_id_list = [0] * nfps
    for idx, cluster in enumerate(mol_clusters, 1):
        for member in cluster:
            cluster_id_list[member] = idx

    return cluster_id_list

cluster_df = hits_df.copy()

cluster_df['Cluster'] = butina_cluster(cluster_df.Mol)
cluster_df
```

[9]:

	SMILES	ZINC_ID	\
0	O(c1ccc(cc1)[C@H]4n2nc(nc2NC(c3cc(OC)ccc3)C4)N)CC	ZINC13131776	
1	Clc1c(ccc(Cl)c1)[C@H]3C[C@H](n2nc(nc2N3)N)c4cc…	ZINC13131298	
2	O=C(c3cc2n1c(nnc1)c(nc2cc3)NC4CCCC4)c5ccccc5	ZINC06901595	
3	O=C(c3cc2n1c(nnc1C)c(nc2cc3)NC[C@@H]4N(CCC4)CC…	ZINC20939758	
4	O=C(N(C)C)c3ccc(c1nc(nc(n1)N)Nc2ccc(cc2)C)cc3	ZINC32540567	
5	O(C(c3ccc(c1nc(nc(n1)N)Nc2ccc(cc2)C)cc3)C)C	ZINC32541042	
6	Fc1c(c(F)ccc1)CNc4n2ncnc2nc(c3ccc(OC)cc3)c4	ZINC23432014	
7	O(c1ccc(cc1)[C@H]4n2nc(nc2NC(c3ccc(OC)cc3)C4)N…	ZINC02472318	
8	O(c1ccc(cc1)[C@H]4n2nc(nc2NC(c3cc(OC)ccc3)C4)N…	ZINC02472304	
9	S=C(Nc2nc(c1c(O)cccc1)cc(n2)c3ccc(OC)cc3)N	ZINC01084111	

10	Clc1ccc(cc1)C(=O)c3sc2n(c(=O)[nH]c(=O)c2c3N)c4…	ZINC02638281	
11	Fc1ccc(cc1)C3Nc2n(nc(n2)N)[C@@H](C3)c5ccc(OCc4…	ZINC02472295	
12	[nH]2c1nc(NCCN(C)C)cc(c1cc2)c4cc3c(cnc(N)c3)cc4	ZINC72143513	
13	O(c1c(cccc1)CNc3nc4c(n2c3nnc2C)cc(cc4)C(=O)c5c…	ZINC09224703	
14	o1c(nnc1)c5ccc(c4c2c([nH]cc2)nc(NCc3ccncc3)c4)cc5	ZINC72454514	
15	n1(ncc4c1ccc(c3c2c([nH]cc2)nc(NCCN(C)C)c3)c4)C	ZINC72168947	

	Mol	Cluster
0	<img data-content="rdkit/molecule" src="data:i…	1
1	<img data-content="rdkit/molecule" src="data:i…	13
2	<img data-content="rdkit/molecule" src="data:i…	12
3	<img data-content="rdkit/molecule" src="data:i…	11
4	<img data-content="rdkit/molecule" src="data:i…	10
5	<img data-content="rdkit/molecule" src="data:i…	9
6	<img data-content="rdkit/molecule" src="data:i…	8
7	<img data-content="rdkit/molecule" src="data:i…	1
8	<img data-content="rdkit/molecule" src="data:i…	1
9	<img data-content="rdkit/molecule" src="data:i…	7
10	<img data-content="rdkit/molecule" src="data:i…	6
11	<img data-content="rdkit/molecule" src="data:i…	1
12	<img data-content="rdkit/molecule" src="data:i…	5
13	<img data-content="rdkit/molecule" src="data:i…	4
14	<img data-content="rdkit/molecule" src="data:i…	3
15	<img data-content="rdkit/molecule" src="data:i…	2

[10]:
```
len(cluster_df.Cluster.unique())
```

[10]: 13

[11]:
```
scores_df = pd.read_csv('../results/hits_scores.csv')
scores_df
```

[11]:

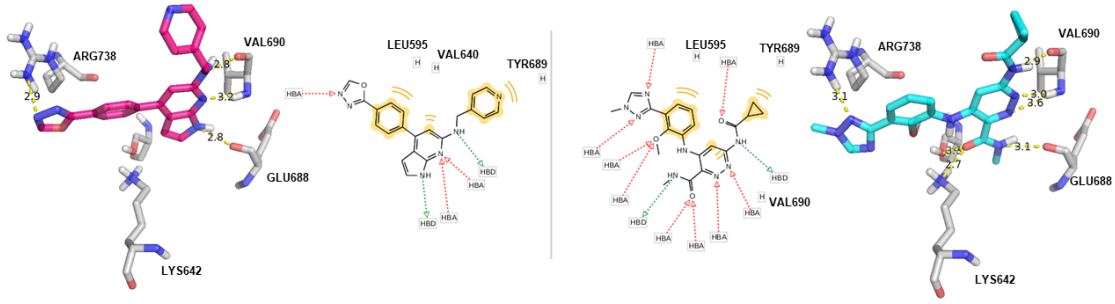
	Name	Smiles	\
0	BMS-986165	O(c3c(Nc2c(nnc(NC(=O)C1CC1)c2)C(=O)NC)cccc3c4n…	
1	ZINC72454514	o1c(nnc1)c5ccc(c4c2c([nH]cc2)nc(NCc3ccncc3)c4)cc5	
2	ZINC72168947	n1(ncc4c1ccc(c3c2c([nH]cc2)nc(NCCN(C)C)c3)c4)C	
3	ZINC72143513	[nH]2c1nc(NCCN(C)C)cc(c1cc2)c4cc3c(cnc(N)c3)cc4	
4	ZINC20939758	O=C(c3cc2n1c(nnc1C)c(nc2cc3)NCC4N(CCC4)CC)c5cc…	
5	ZINC23432014	Fc1c(c(F)ccc1)CNc4n2ncnc2nc(c3ccc(OC)cc3)c4	
6	ZINC32540567	O=C(N(C)C)c3ccc(c1nc(nc(n1)N)Nc2ccc(cc2)C)cc3	
7	ZINC01084111	S=C(Nc2nc(c1c(O)cccc1)cc(n2)c3ccc(OC)cc3)N	
8	ZINC02638281	Clc1ccc(cc1)C(=O)c3sc2n(c(=O)[nH]c(=O)c2c3N)c4…	
9	ZINC13131776	O(c1ccc(cc1)C4n2nc(nc2NC(c3cc(OC)ccc3)C4)N)CC	
10	ZINC06901595	O=C(c3cc2n1c(nnc1)c(nc2cc3)NC4CCCC4)c5ccccc5	
11	ZINC13131298	Clc1c(ccc(Cl)c1)C4Nc2n(nc(n2)N)C(c3cc(OC)c(OC)…	
12	ZINC02472304	O(c1ccc(cc1)C4n2nc(nc2NC(c3cc(OC)ccc3)C4)N)Cc5…	
13	ZINC02472318	O(c1ccc(cc1)C4n2nc(nc2NC(c3ccc(OC)cc3)C4)N)Cc5…	

```
14  ZINC02472295  Fc1ccc(cc1)C5Nc2n(nc(n2)N)C(c4ccc(OCc3ccccc3)c…

    CNNaffinity  CNNaffinity_variance  CNNscore  minimizedAffinity  \
0          7.85                 0.26      1.00             -10.33
1          7.38                 0.19      0.99             -10.15
2          7.07                 0.15      0.99              -9.75
3          7.16                 0.12      0.99             -10.56
4          7.57                 0.06      0.98              -9.55
5          7.35                 0.08      0.98              -9.34
6          7.11                 0.03      0.97              -9.34
7          7.09                 0.08      0.97              -9.59
8          7.29                 0.10      0.96              -9.52
9          7.01                 0.03      0.96              -9.02
10         7.12                 0.09      0.93              -9.39
11         7.25                 0.09      0.92              -9.07
12         7.36                 0.05      0.91              -9.50
13         7.14                 0.07      0.89              -9.39
14         7.07                 0.09      0.88              -9.66

    Pharm. Score  rmsd
0            NaN   NaN
1          86.01  0.98
2          86.60  0.84
3          76.65  0.70
4          67.66  0.96
5          65.42  0.89
6          46.58  1.08
7          45.81  1.00
8          76.25  0.96
9          56.40  0.94
10         55.76  1.09
11         36.13  0.94
12         45.86  0.94
13         56.98  0.94
14         56.98  0.94
```

ZINC72454514 Selective Allosteric TYK2 Inhibitor



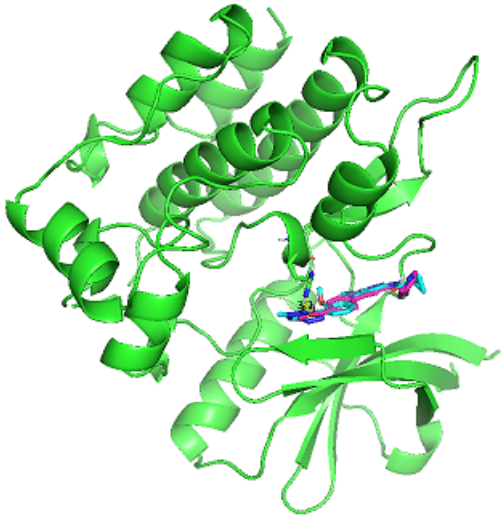
ZINC72454514

CNNaffinity = 7.34

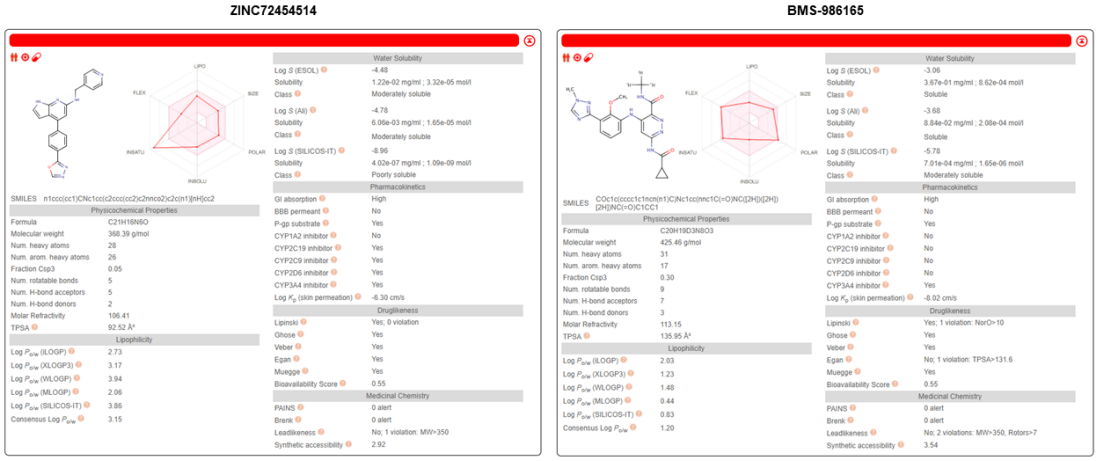
BMS-986165

CNNaffinity = 7.87
pChEMBL Value = 7.89

6

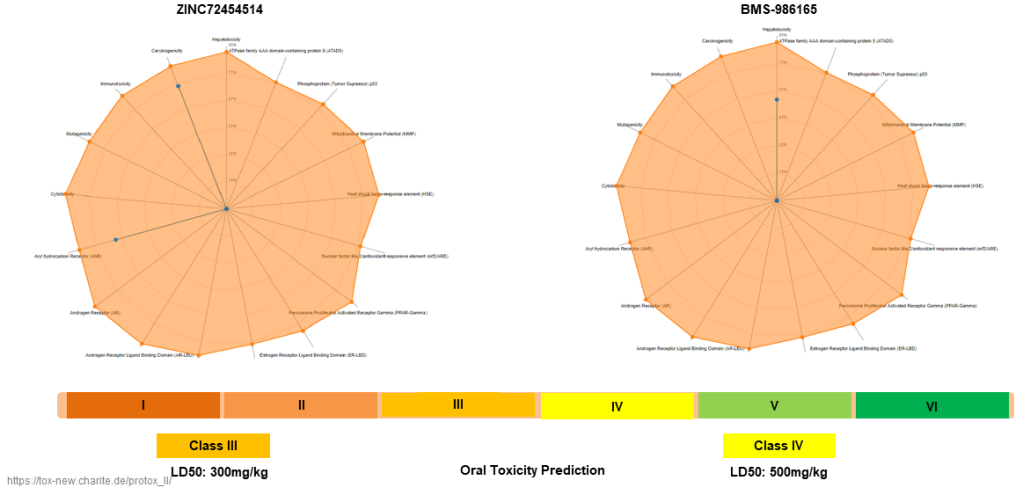


ZINC72454514 'Viable' TYK2 Inhibitor? ADMET



http://www.swissadme.ch/

ZINC72454514 'Viable' TYK2 Inhibitor? ADMET



https://tox-new.charite.de/protox_II/

[ ]: