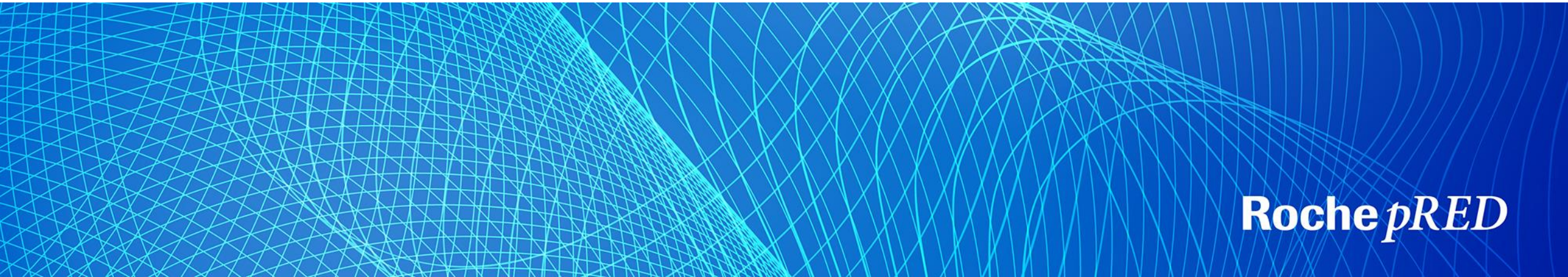
The playbooks of medicinal chemistry design moves

Mahendra Awale

RDKit UGM Oct 6-8, 2020



Hero's of Small Molecule Drug Design (Medicinal Chemists)

1910

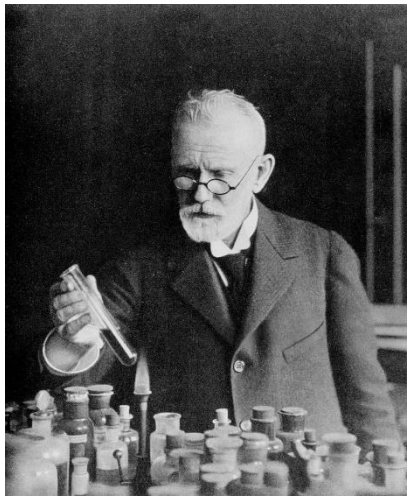
Paul Ehrlich (Physicist)

Alfred Bertheim (Chemist)

Salvarsan,

Syphilis

First rational Design



1960

Leo Sternbach (Chemist)

Roche

Benzodiazepines

Anxiety, seizures, muscle spasms

2020

Hassane Ratni (Chemist)

Roche

Risdiplam

Spinal muscular atrophy

Compound optimization, now and then

Roche

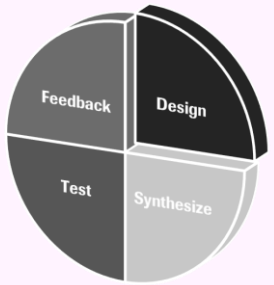
1910s

Paul Ehrlich (Physicist)

Alfred Bertheim (Chemist)

Salvarsan, Syphilis

First rational Design



~900 designs

Intuitions of chemist/chemistry

>100 years

QSAR analysis

Molecular Docking

Generative Deep neural Network

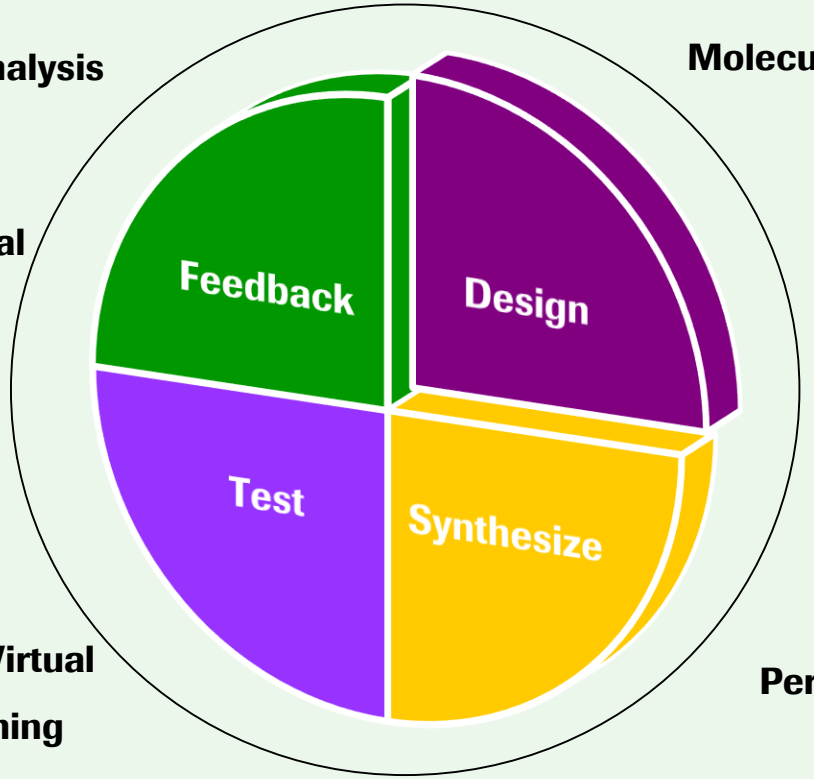
3D-Modelling

Machine learning models

Molecular Dynamics

2D/3D Virtual Screening

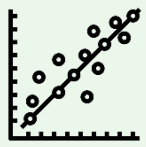
Free-Energy Perturbation study

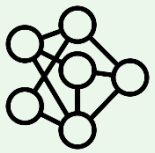


Databases

Compounds/bioactivity/target/structure/reactions

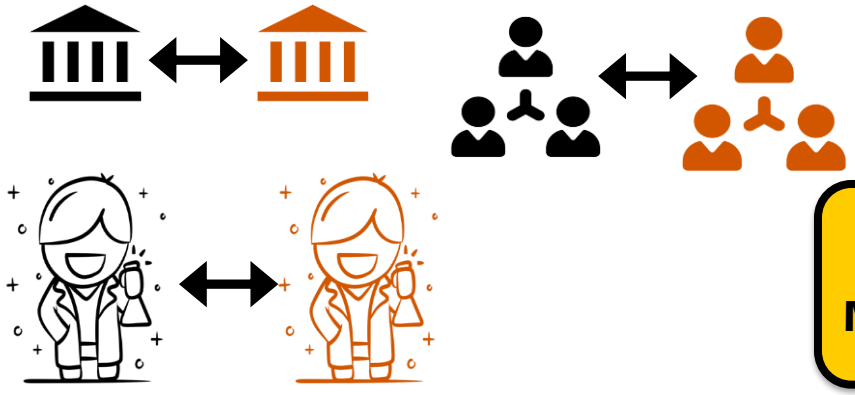
Computers

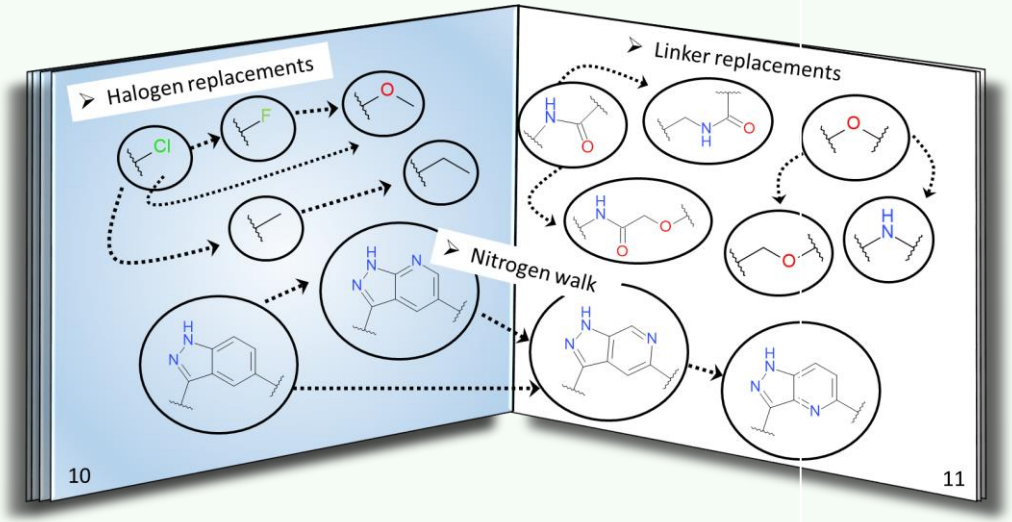




Access/Sharing, Search, Analyze

Wouldn't it be great to have a way to capture MedChem Design Moves?





Systematic

Comprehensive

Automated idea generation

Explainable

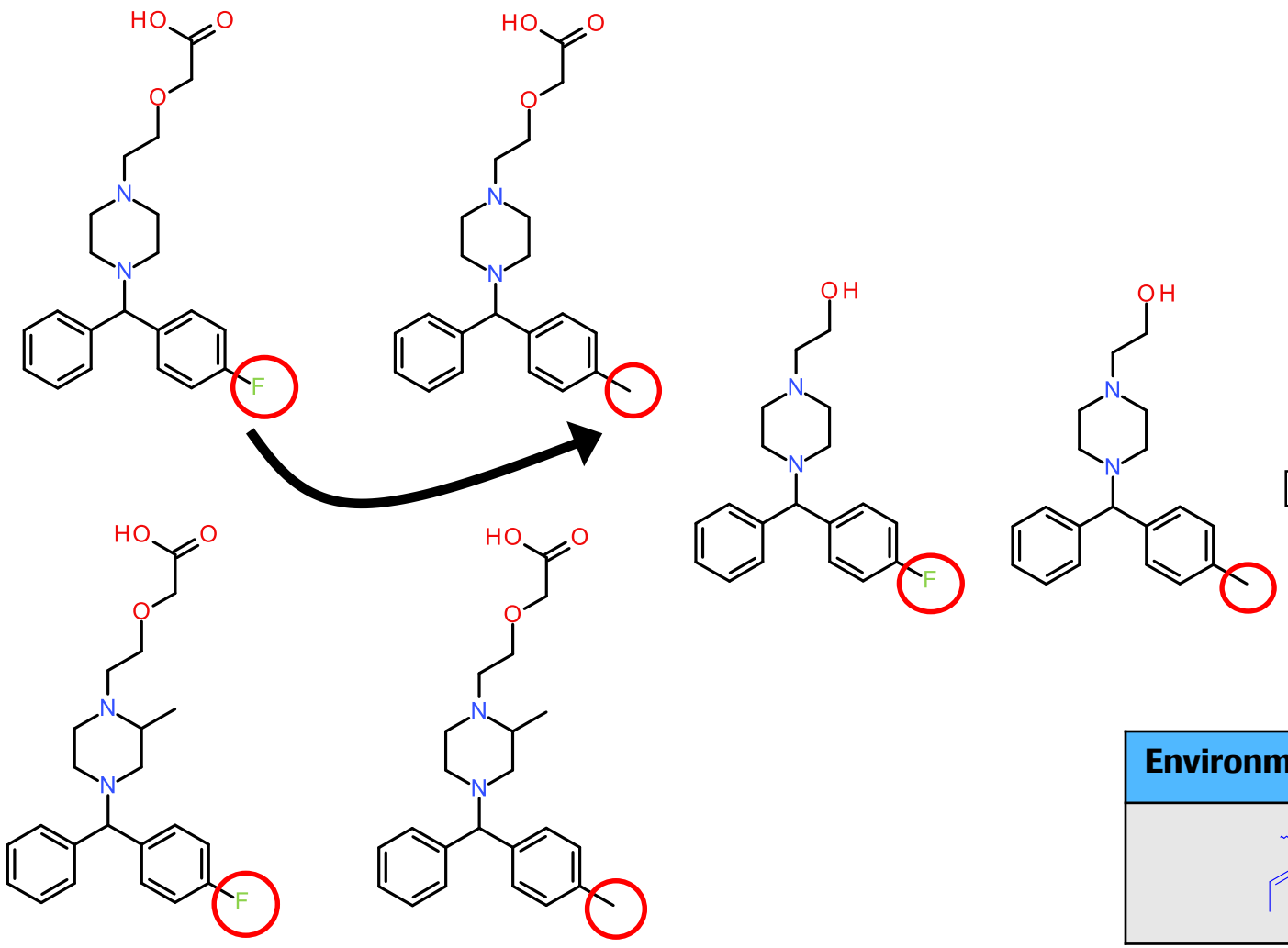
Synthetically tractable solutions

Simulate MedChem Thinking

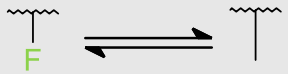
Democratizing Medchem. moves

Design Intuition

Matched Molecular Pairs (MMP): Generating Med Chem Knowledge



Molecule A	Molecule B	Δ Property
ID01	ID04	-0.2
ID02	ID05	
ID03	ID06	-0.4
…		

Environment	Rule	~~Mean Δ~~	~~Std. Dev.~~	Frequency
		~~-0.4~~	~~0.2~~	3

Increase scope of compounds

Intentional design moves

unintentional design moves

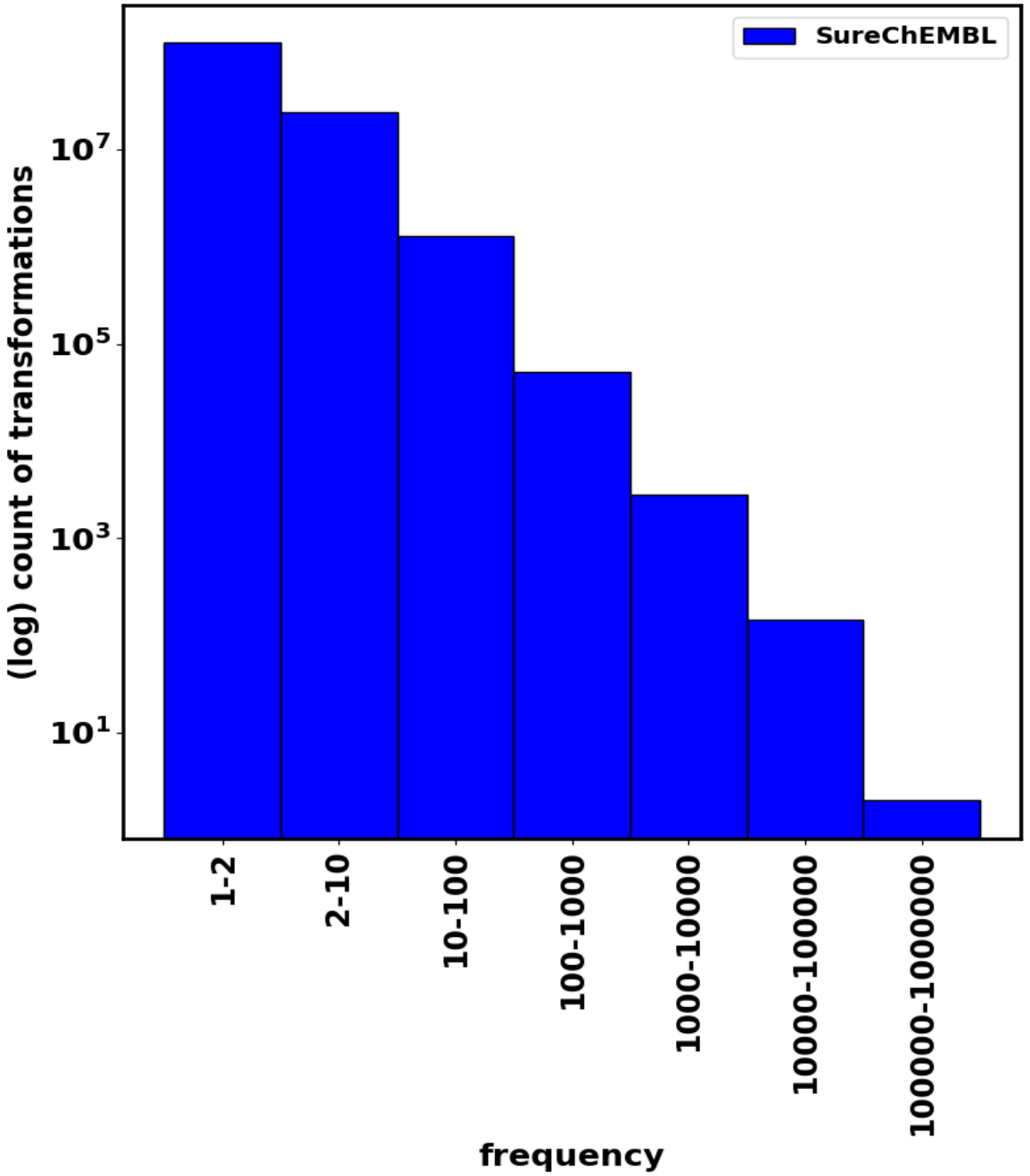
Compiling the Playbooks

Chemical Environment at Radius 3

ChEMBL: ~10 Million

Roche: ~29 Million

SureChEMBL: ~162 Million



Inside the Playbooks

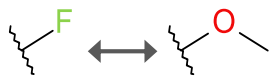
a) Frequent Single cut transformations

274,779/98,305/28,764

256,153/117,394/28,609

155,563/94,255/25,118

191,133/99,566/23,758

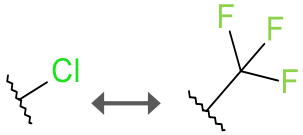


142,704/87,788/21,813

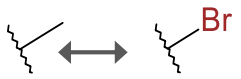
190,340/98,441/19,879

118,568/42,100/15,239

123,088/56,003/11,828



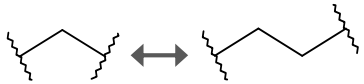
119,714/33,659/10,504



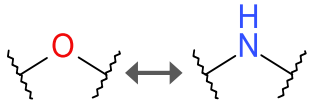
90,392/33,873/9,664

b) Frequent Double cut , acyclic-acyclic transformations

SureChEMBL

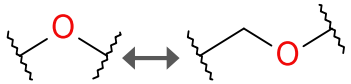


14,749/21,028/3,448

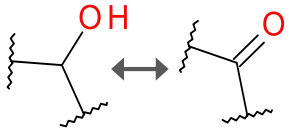


13,888/1,786/951

9,905/2,455/993

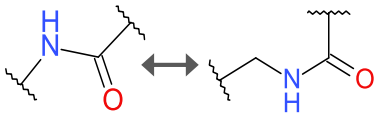


9,464/1,613/839

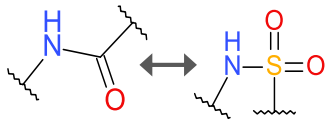


6,883/762/408

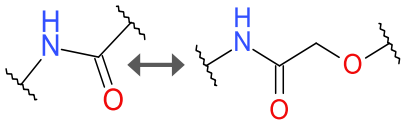
RocheDB



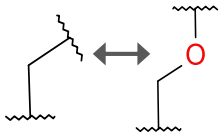
1,226/10,110/953



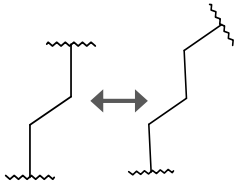
2,817/4,465/459



411/4,137/283



2,646/4,110/520

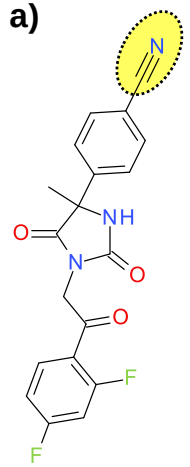


3,811/4,054/961

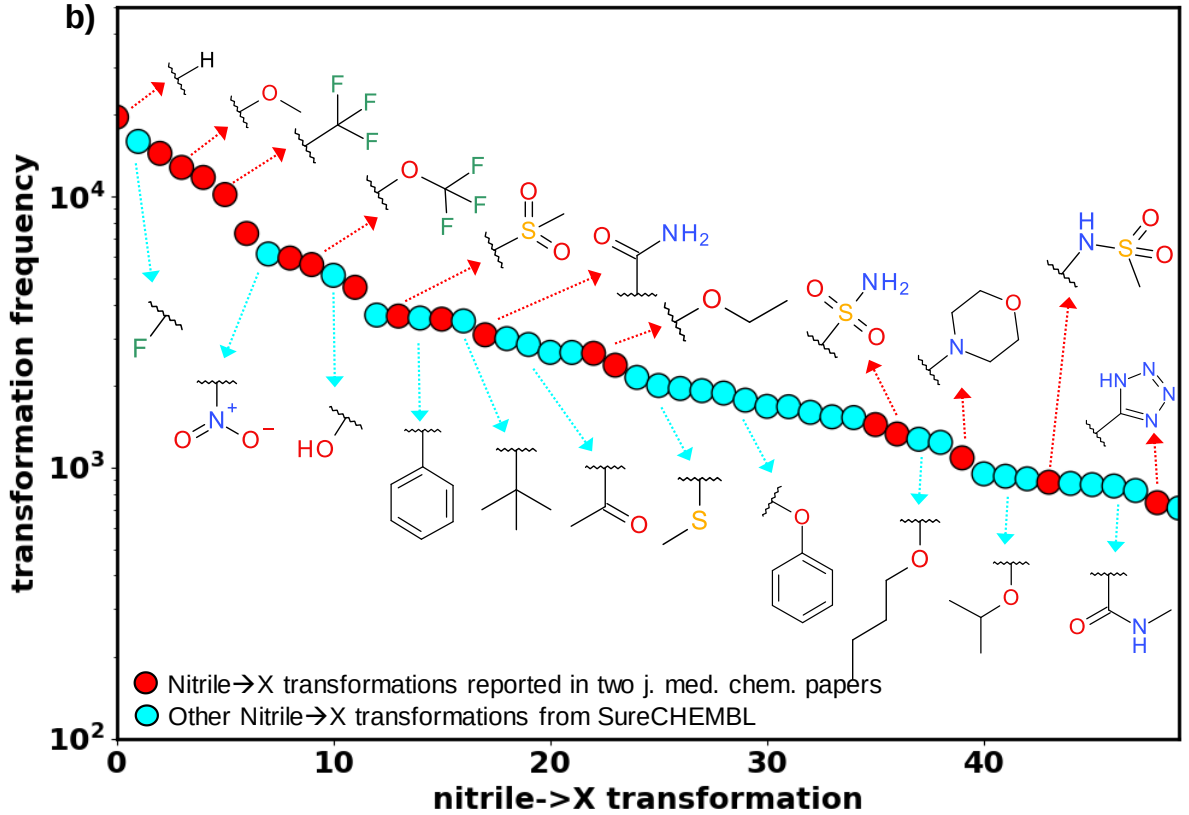
Validation Study

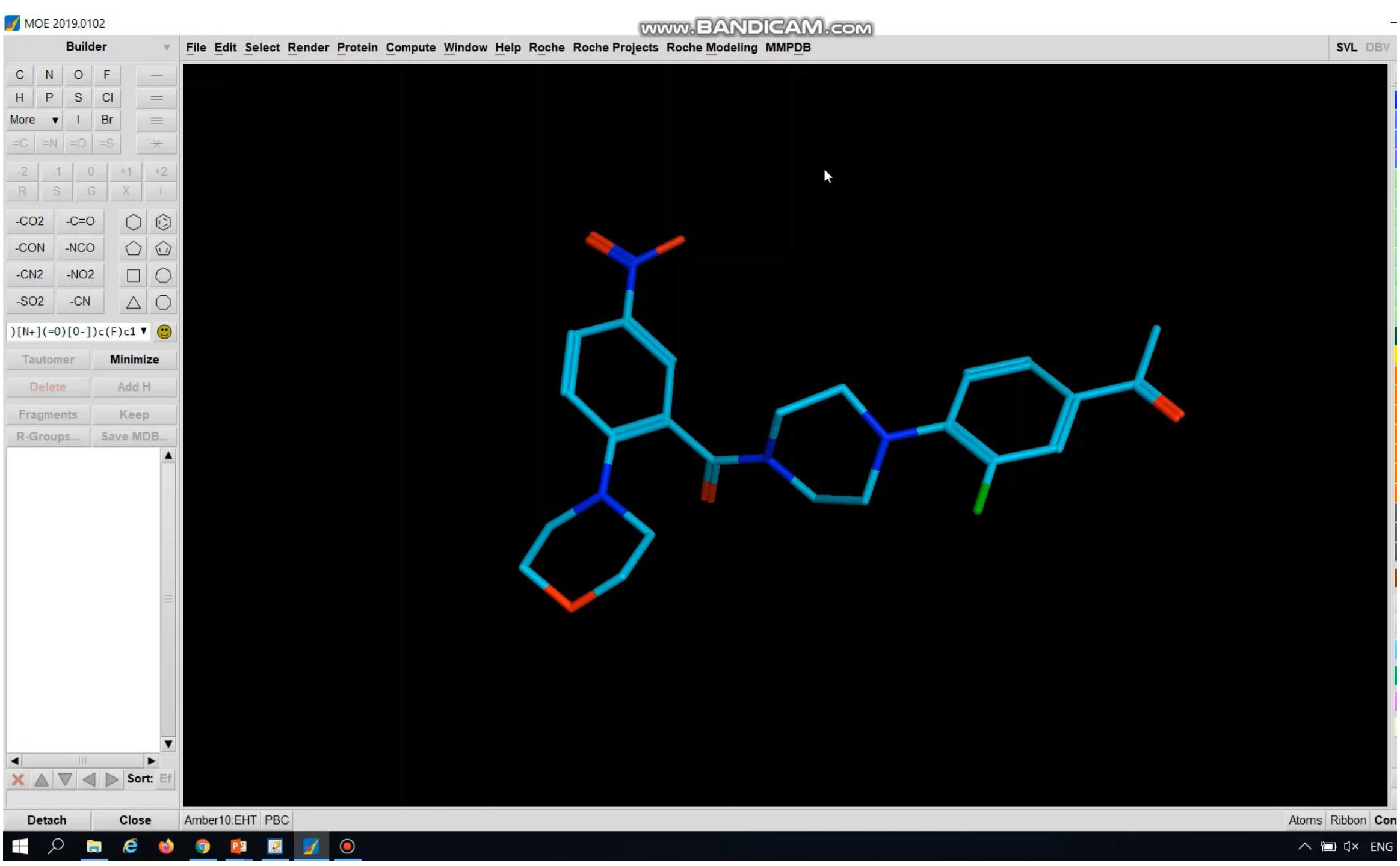
transformation frequency	no. of transformations from SureChEMBL	no. of recovered transformations
>=1 and <10	11,852	1
>=10 and <100	1,457	5
>=100 and <1000	206	7
>=1000	40	17
Total	**13,555**	**30/30**

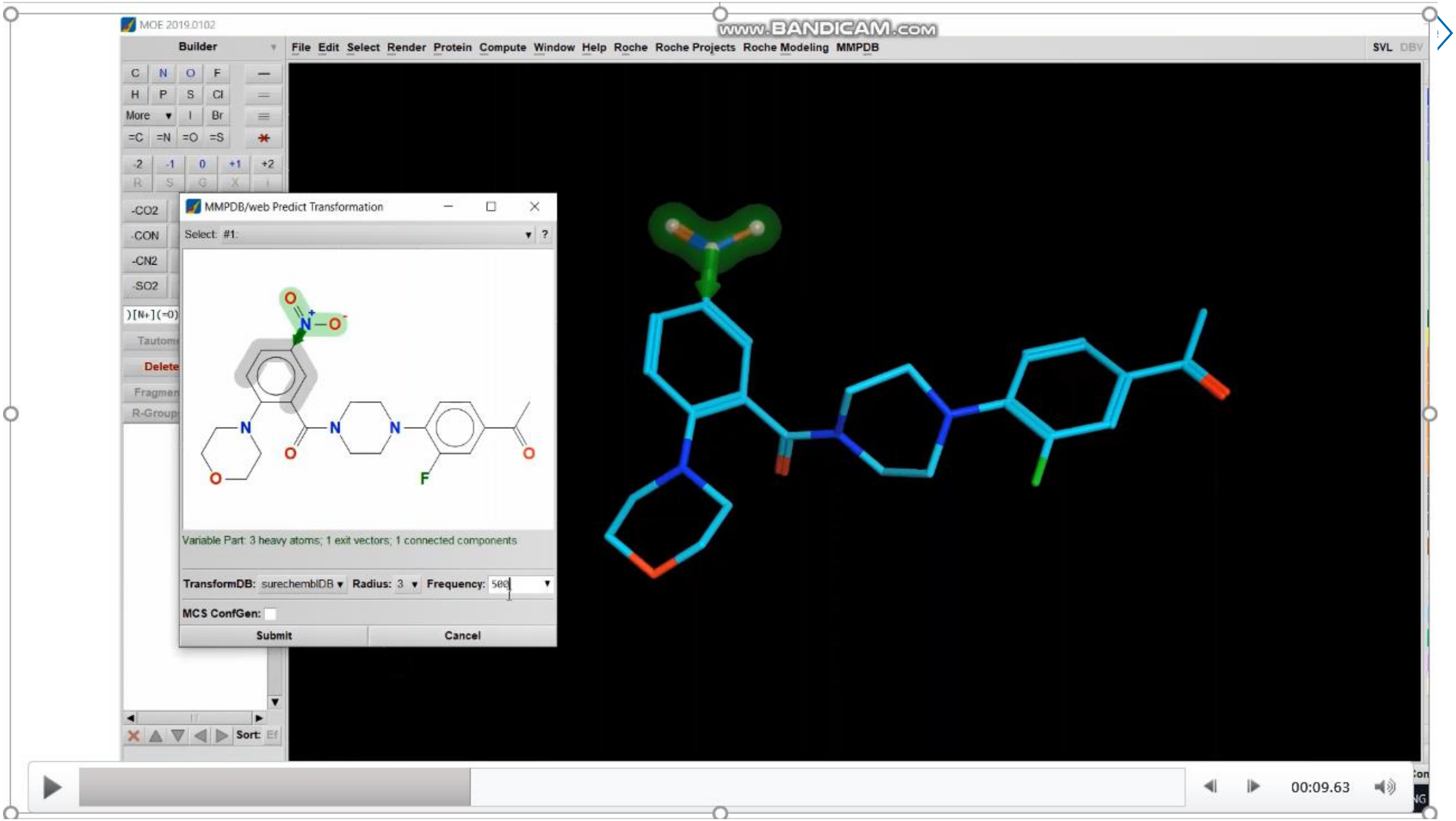
a)



b)

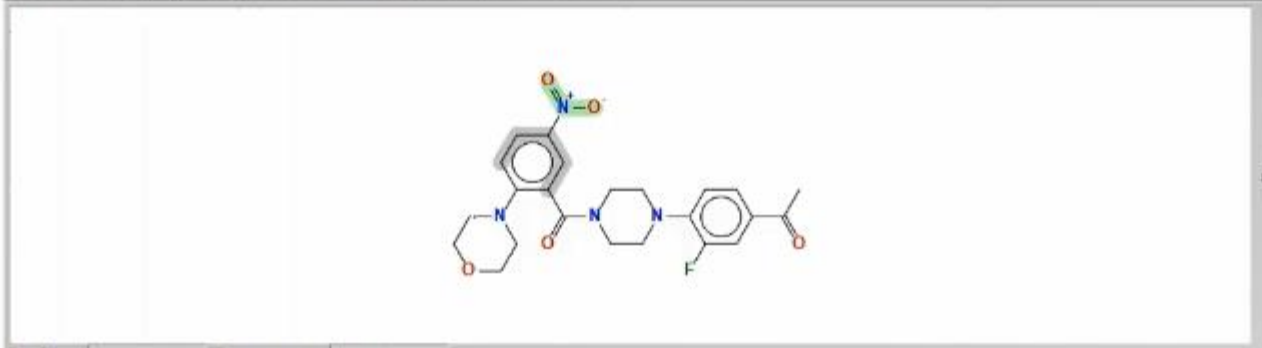


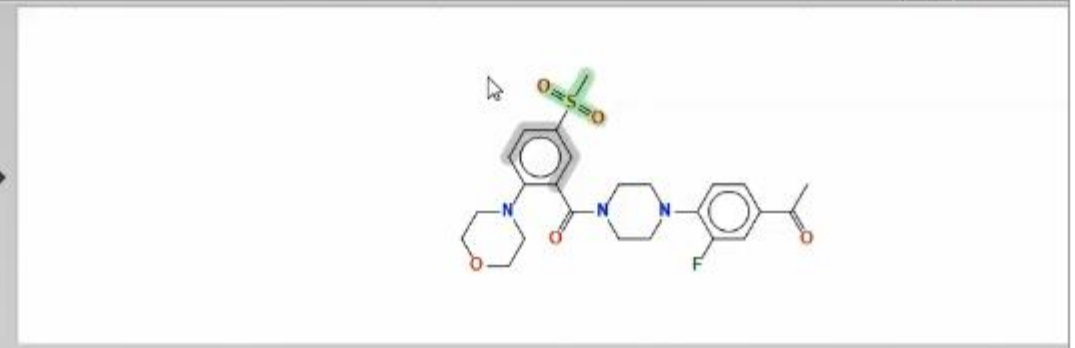

MOE 2019.0102
www.BANDICAM.com
Builder
File Edit Select Render Protein Compute Window Help Roche Roche Projects Roche Modeling MMPDB
SVL DBV
C N O F
H P S Cl
More I Br
=C =N =O =S
-2 -1 0 +1 +2
R S G X i
-CO2 -C=O
-CON -NCO
-CN2 -NO2
-SO2 -CN
)[N+](=O)[O-])c(F)c1
Tautomer Minimize
Delete Add H
Fragments Keep
R-Groups... Save MDB...
Sort: Ef
Detach Close
Amber10:EHT PBC
Atoms Ribbon Con
ENG


MOE 2019.0102
www.BANDICAM.com
Builder
File Edit Select Render Protein Compute Window Help Roche Roche Projects Roche Modeling MMPDB
SVL DBV
MMPDB/web Predict Transformation
Select: #1:
Variable Part: 3 heavy atoms; 1 exit vectors; 1 connected components
TransformDB: surechemblDB
Radius: 3
Frequency: 500
MCS ConfGen:
Submit
Cancel
00:09.63

Database Viewer : $TMP/mmpdb_mct_44836.mdb

File Edit Display Compute Window Help ScorpionHS MMPDB SVL DBV





Radius: 3 Frequency: 1681

	mol	example_lhs	example_rhs
12	COC(=O)c1ccc(N2	SCHEMBL947293	SCHEMBL14645833
13	CC(=O)Nc1ccc(N2	SCHEMBL1930466	SCHEMBL1930375
14	CC(=O)c1ccc(N2C	SCHEMBL947293	SCHEMBL16156066
15	CCc1ccc(N2CCOCC	SCHEMBL947293	SCHEMBL667823
16	CC(=O)c1ccc(N2C	SCHEMBL947293	SCHEMBL1602589
17	CC(=O)c1ccc(N2C	SCHEMBL2657985	SCHEMBL7037977
18	CC(=O)c1ccc(N2C	SCHEMBL947293	SCHEMBL18130731
19	CC(=O)c1ccc(N2C	SCHEMBL947293	SCHEMBL16398416
20	CC(=O)c1ccc(N2C	SCHEMBL2657985	SCHEMBL16775717
21	CC(=O)c1ccc(N2C	SCHEMBL1930466	SCHEMBL3862704
22	CC(=O)c1ccc(N2C	SCHEMBL20443239	SCHEMBL20443297
23	CC(=O)c1ccc(N2C	SCHEMBL947293	SCHEMBL4864533
24	CC(=O)c1ccc(N2C	SCHEMBL947293	SCHEMBL9015450
25	CC(=O)c1ccc(N2C	SCHEMBL947293	SCHEMBL8114283
26	CC(=O)c1ccc(N2C	SCHEMBL4817737	SCHEMBL14044709
27	CC(=O)c1ccc(N2C	SCHEMBL1930466	SCHEMBL1929695
28	CC(=O)c1ccc(N2C	SCHEMBL1930466	SCHEMBL1930420
29	CCOc1ccc(N2CCOC	SCHEMBL2310583	SCHEMBL2357454
30	CCOC(=O)c1ccc(N	SCHEMBL4030661	SCHEMBL20362242
31	CSc1ccc(N2CCOCC	SCHEMBL6750268	SCHEMBL8490674

Acknowledgement

- Dr. Jerome Hert
- Dr. Christian Kramer
- Dr. Laura Guasch
- Prof. Dr. Sereina Riniker
- Dr. Patrizio Mattei
- Dr. Giuseppe Cecere

Doing now what patients need next