

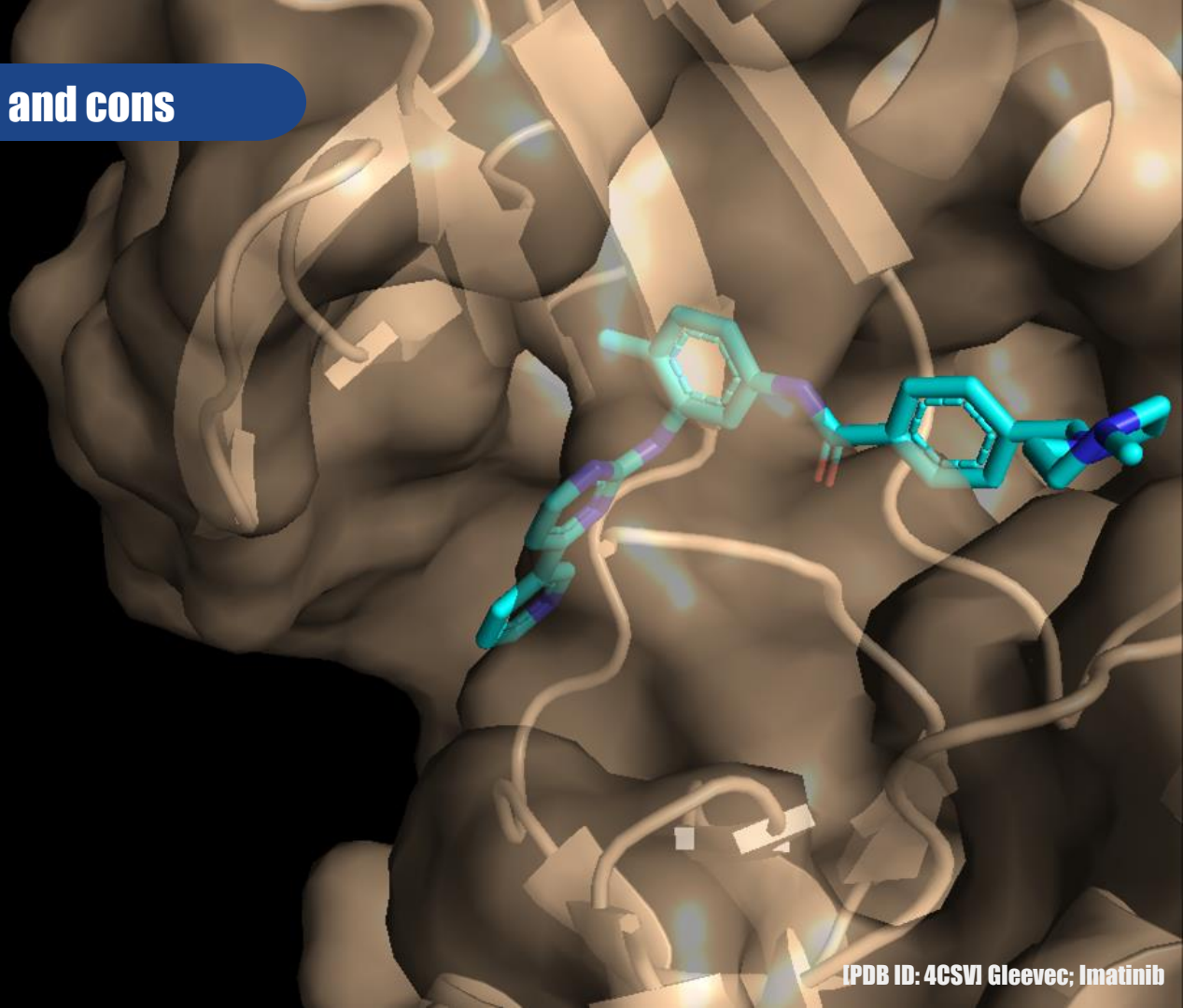
*AI-Driven Medical Breakthrough*

***Virtual Screening:***  
**Undruggable Target Breakthrough**

[Triphos]

박채현 - 응용화학생명공학과  
박민서 - 응용화학생명공학과  
이상욱 - 응용화학생명공학과

## Small Molecules in drug; pros and cons



IPDB ID: 4CSV1 Gleevec; Imatinib

## Small Molecules in drug; pros and cons

### PROS

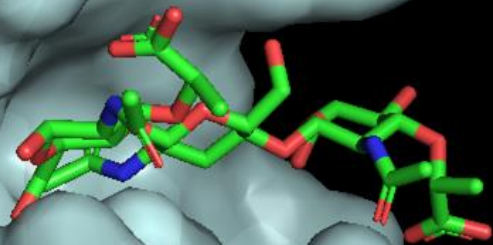
작은 크기로 세포 내 침투 용이.  
생산 과정이 상대적으로 간단함.

### CONS

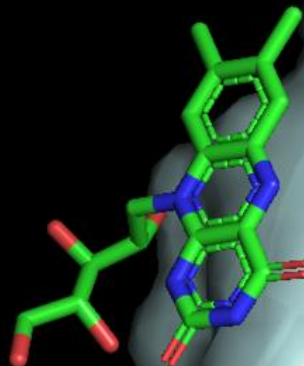
Undruggable-Target의 한계점 존재  
Target protein에 대한 Specificity ↓



**Binding Pocket 0**



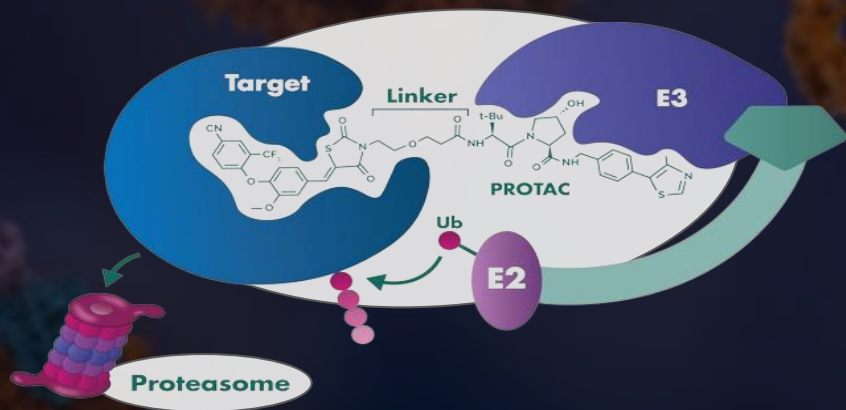
***Undruggable  
Target***



**Binding Pocket X**

# PROTAC (PROteolysis-Targeting Chimera)

## ***PROTAC***



## TPD (Target Protein Degradation)

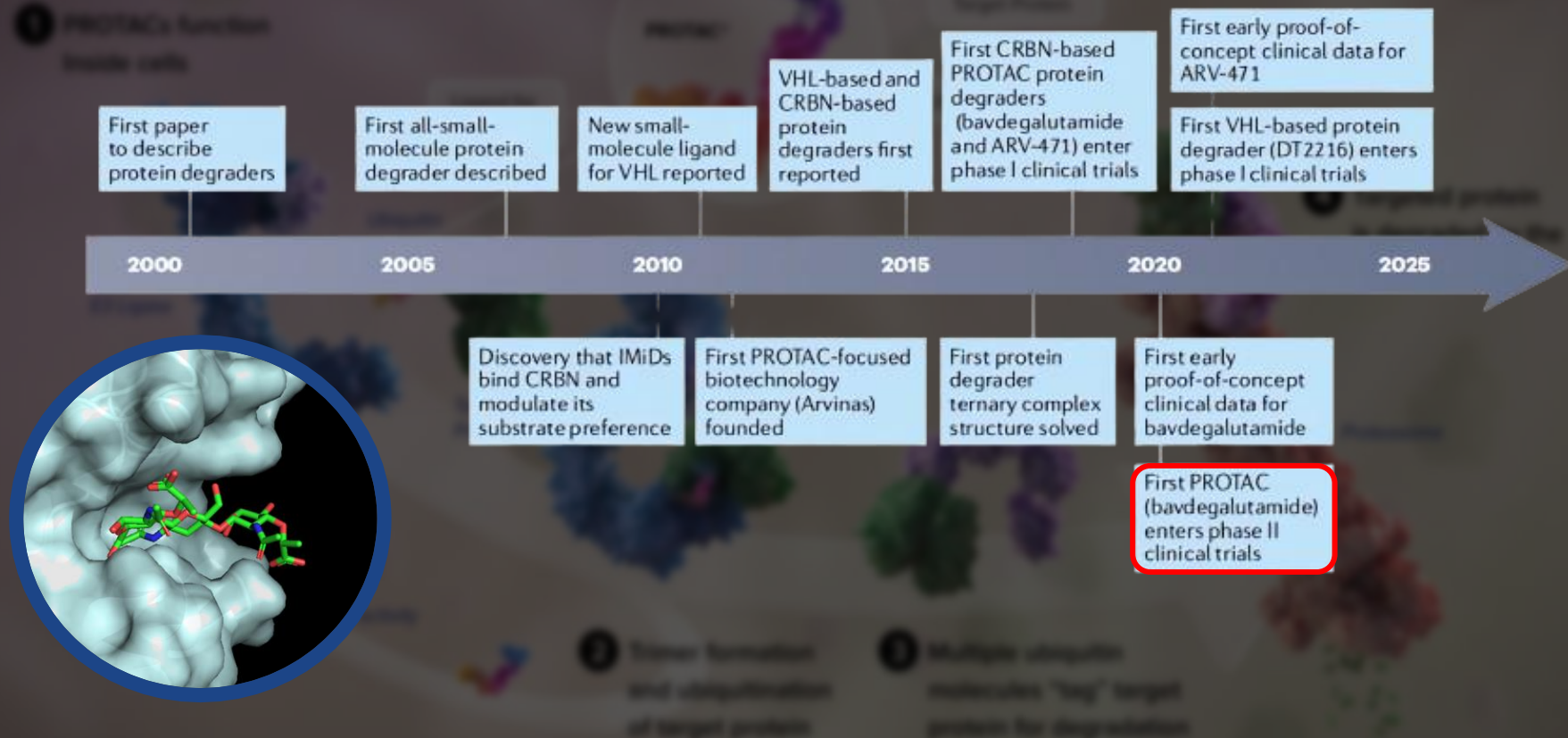
세포 안 Protein Target 가능



High-Selectivity

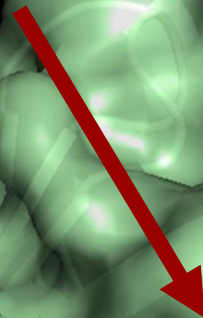
TPD Modality ⇒ 질병 치료

# History of PROTAC





***Shallow Pocket***



# HTS: High-Throughput Screening

```
#####
```

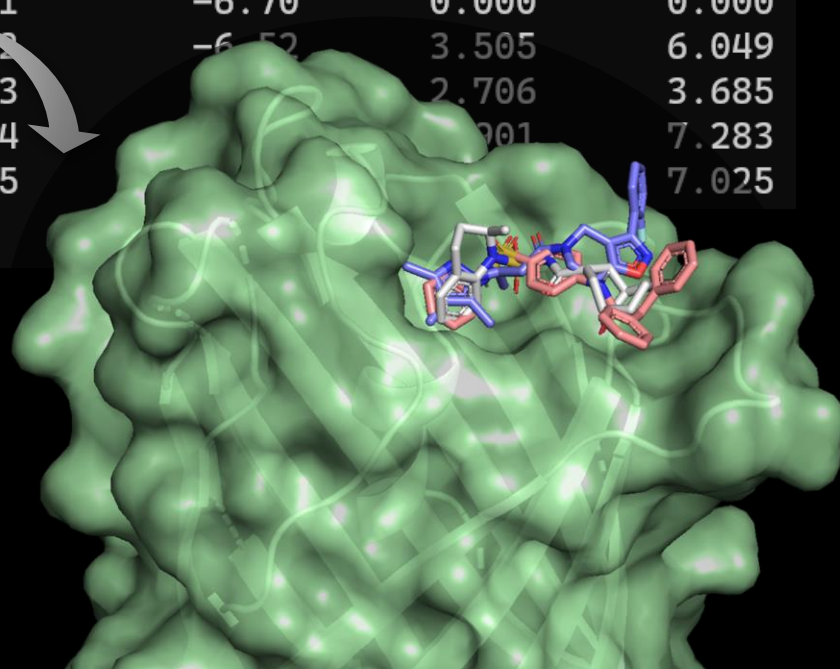
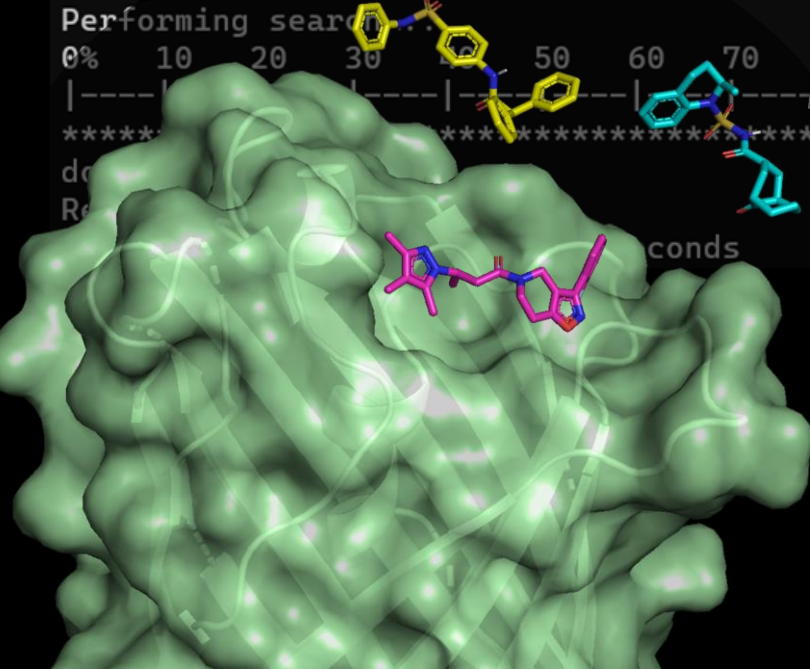
```
Reading input ... done.
Setting up the scoring function ... done.
Analyzing the binding site ... done.
Using random seed: 1001325748
```

```
Performing search...
```

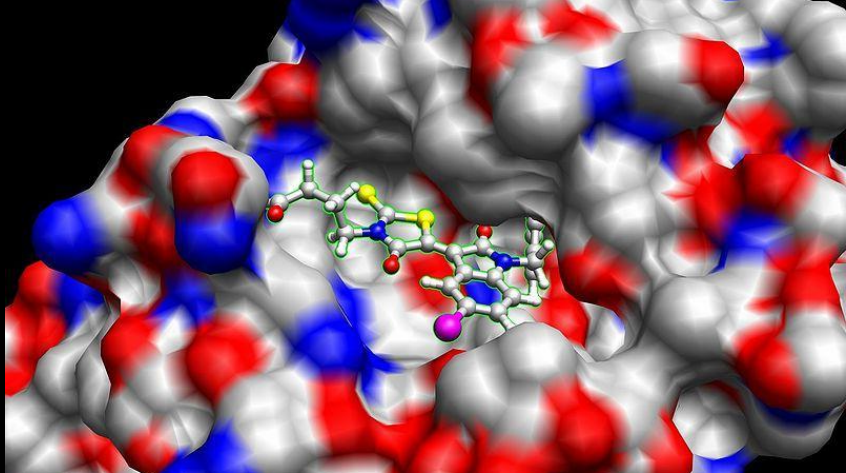
```
0% 10 20 30 40 50 60 70 80 90
|----|-----|-----|-----|-----|-----|-----|-----|
****|-----|-----|-----|-----|-----|-----|-----|
do|-----|-----|-----|-----|-----|-----|-----|
Re|-----|-----|-----|-----|-----|-----|-----|
```

conds

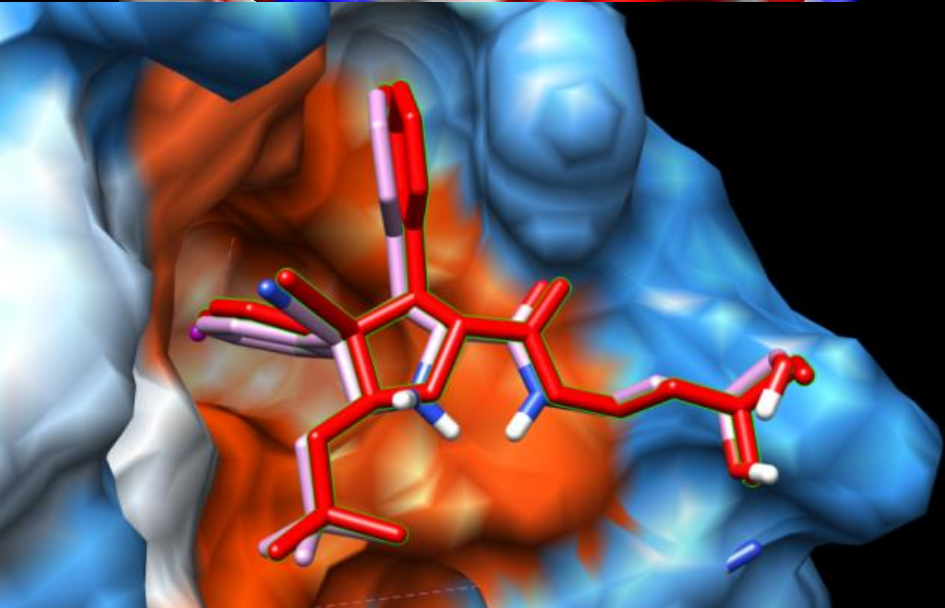
mode	affinity (kcal/mol)	dist from best mode rmsd l.b.	rmsd u.b.
1	-6.70	0.000	0.000
2	-6.52	3.505	6.049
3		2.706	3.685
4		3.901	7.283
5			7.025



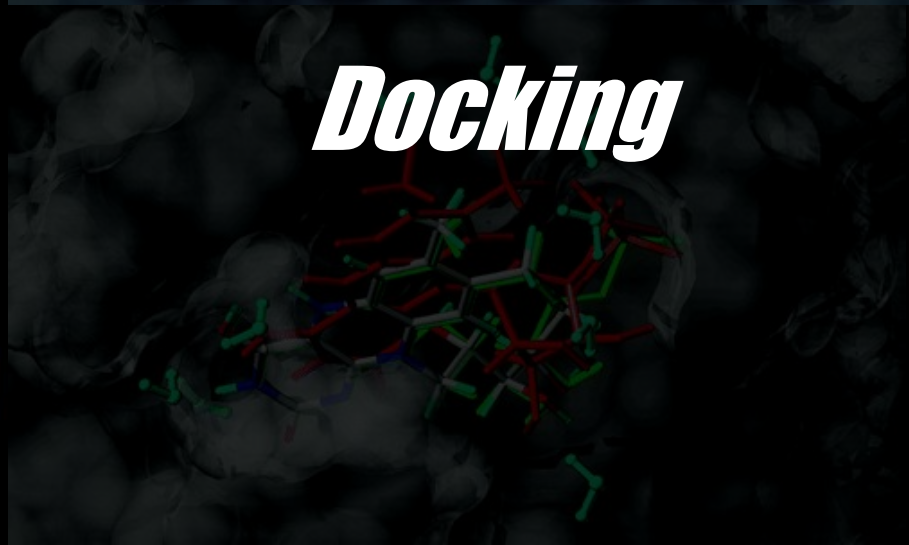


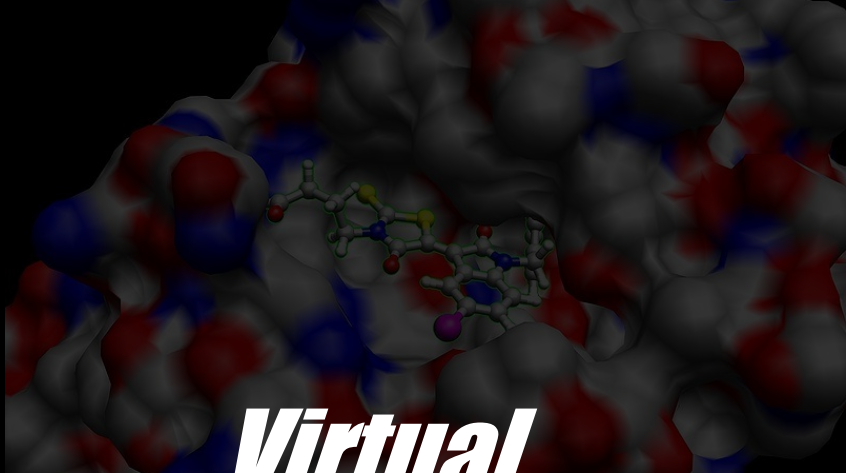


***Molecular***

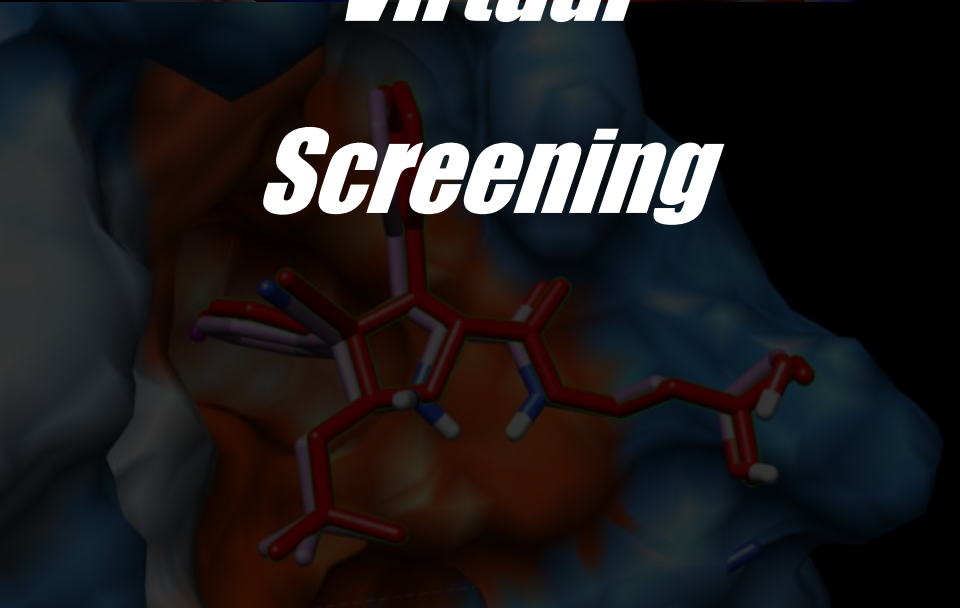
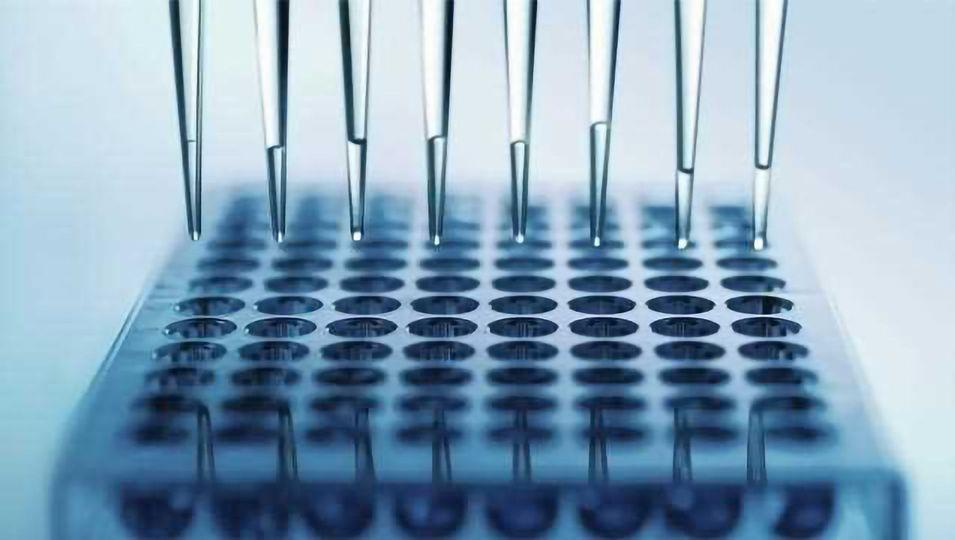


***Docking***

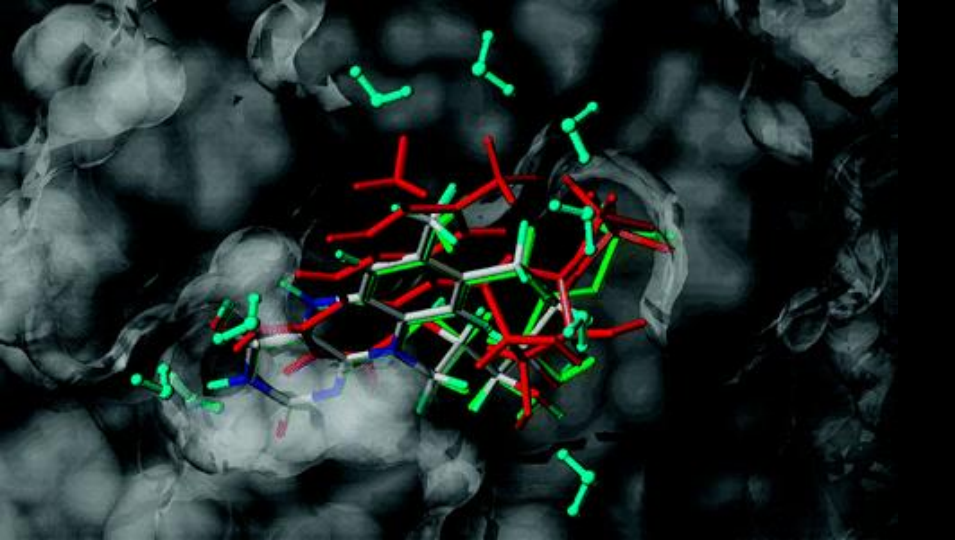




***Virtual***

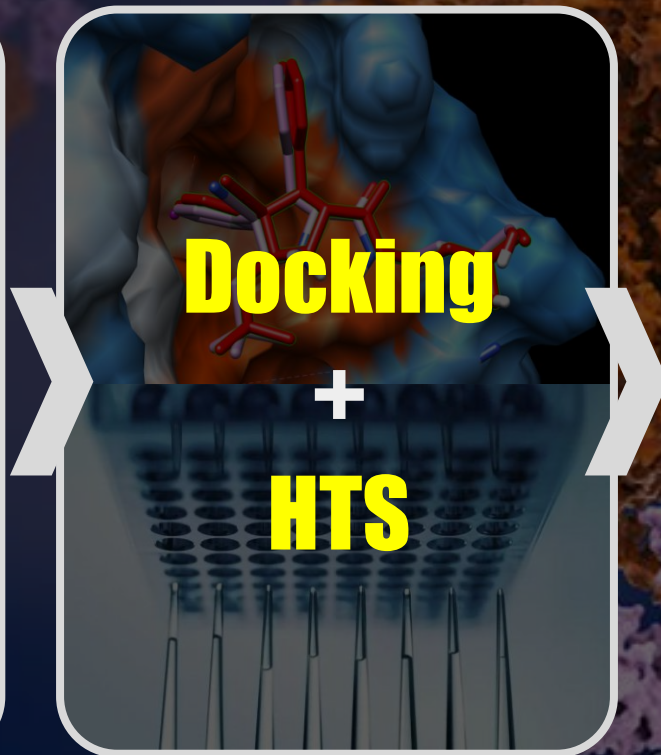
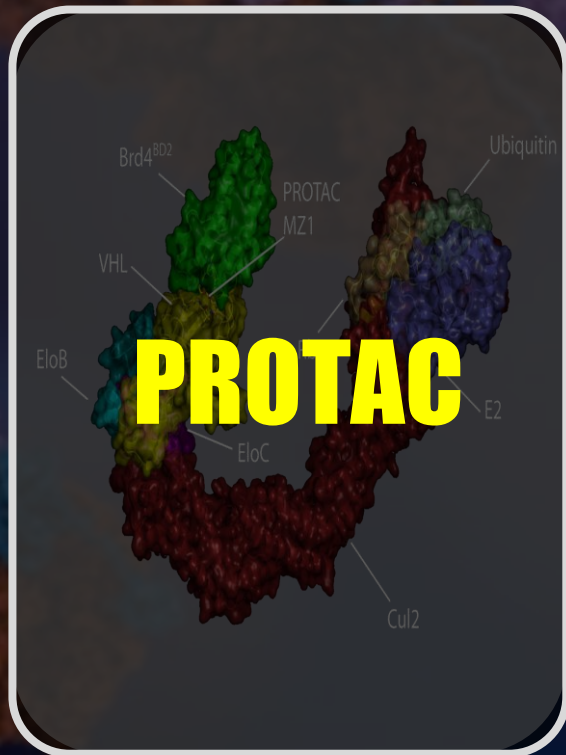


***Screening***





# ***Our Goal***





## Main Project - Timeline

Sep.

[Developing Programming Skills] - Python and Linux using Visual Studio & Jupyter Notebook  
[Proficiency in visualization tools] - Pymol, Discovery Studio, Chimera, and Chemdraw  
[Studying Cheminformatics Packages] - RDKit & OpenBabel  
[Exploring & Applying Docking Tools] - LeDock, MolAIcal, and AutoDock.Vina  
[Validation of Ligand 3D Conformation] - RDKit & OpenBabel  
[Validating docking tools] - MolAIcal (3rd Attempts)

Oct.

Generation and preparation of 3D ligand structure  
Large-Scale Library Docking on Naver & Google Servers (Linux / Window)  $\Rightarrow$  '1,000,000 Compounds'

Nov.

[MolAIcal Validation] - Measurement of  $K_d$ -value via MST(MicroScale Thermophoresis) assay

Dec.

(What's next?)  $\Rightarrow$  AI-based ligand modification

# PROCESS

**Environment  
SETUP**



**Method  
Validation**



**Library  
preparation**



**Virtual-Screening  
Docking**



**MST Assay  
Experiment &  
Data analysis**



**Environment  
SETUP**



**Method  
Validation**

**Library  
preparation**

**Virtual-Screening  
Docking**

**MST Assay  
Experiment &  
Data analysis**



```
rdkit
from rdkit.Chem import AllChem
from rdkit.Chem.Draw import IPythonConsole
IPythonConsole.ipython_3d = True

In [ ]: name = "PMG"

In [ ]: f = open(f"/MY_Library/{name}.txt")
file = f.readlines()
f.close()

In [ ]: for i in range(len(file)):
file[i] = file[i].replace("\n", " ")

In [ ]: j = 1 #처음 한번만 돌리고 지우기

In [ ]: w = open(f"/MY_Library/{name}/{j}.txt", "w")
for i in range(len(file)):
if file[i] != "END":
w.write(file[i])
j+=1
w.close()
else:
w.write(file[i])
w.write("\n")
w.close()

100% [██████████] 14307128/14307128 10s
```

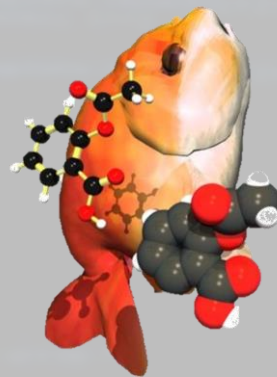


python™

Linux™



Open-Source Cheminformatics  
and Machine Learning



***Python***

***Linux***

***RDKit***

***OpenBabel***



**Environment  
SETUP**



**Method  
Validation**

**Library  
preparation**

**Virtual-Screening  
Docking**

**MST Assay  
Experiment &  
Data analysis**

***Servers***



**Google Cloud**



***Subtools***



**PuTTY**



**GitHub**



## Environment SETUP



## Method Validation

## Library preparation

## Virtual-Screening Docking

## MST Assay Experiment & Data analysis

RUN

```
python mdock_vina.py -v molaicald -c config.txt -l pdb_list.txt -d dock -o dock.txt -p 16
```

```
spider@DESKTOP-2SL4271: ~  
0[|||||100.0%] 4[|||||100.0%] 8[|||||100.0%] 12[|||||100.0%]  
1[|||||100.0%] 5[|||||100.0%] 9[|||||100.0%] 13[|||||100.0%]  
2[|||||100.0%] 6[|||||100.0%] 10[|||||100.0%] 14[|||||100.0%]  
3[|||||100.0%] 7[|||||100.0%] 11[|||||100.0%] 15[|||||100.0%]  
Mem[|||||1.33G/15.5G] Tasks: 91, 353 thr; 16 running  
Swp[|||||0K/4.00G] Load average: 103.39 103.31 103.47  
Uptime: 09:07:29  
  
T RES SHR S CPU% MEM% TIME+ Command  
M 31092 2676 S 136.0 0.2 0:10.73 molaicald --config config.txt --ligand dock/532339/532339.pdbqt --out dock/532339/dock_532339.pdbq  
M 31096 2676 S 135.0 0.2 0:19.98 molaicald --config config.txt --ligand dock/531901/531901.pdbqt --out dock/531901/dock_531901.pdbq  
M 29508 2676 S 134.0 0.2 0:25.25 molaicald --config config.txt --ligand dock/531675/531675.pdbqt --out dock/531675/dock_531675.pdbq  
M 32680 2676 S 132.0 0.2 0:51.42 python 2.py -v molaicald -c config.txt -l pdb_list.txt -d dock -o dock.txt -p 16  
M 29508 2676 S 132.0 0.2 0:16.26 molaicald --config config.txt --ligand dock/531948/531948.pdbqt --out dock/531948/dock_531948.pdbq  
M 31092 2676 S 132.0 0.2 0:14.73 molaicald --config config.txt --ligand dock/532028/532028.pdbqt --out dock/532028/dock_532028.pdbq  
M 31092 2676 S 131.0 0.2 0:03.41 molaicald --config config.txt --ligand dock/532756/532756.pdbqt --out dock/532756/dock_532756.pdbq  
M 31092 2676 S 130.0 0.2 0:11.35 molaicald --config config.txt --ligand dock/532209/532209.pdbqt --out dock/532209/dock_532209.pdbq  
M 31560 2676 S 130.0 0.2 0:06.83 molaicald --config config.txt --ligand dock/532687/532687.pdbqt --out dock/532687/dock_532687.pdbq  
M 32680 2676 S 129.0 0.2 0:11.86 molaicald --config config.txt --ligand dock/532321/532321.pdbqt --out dock/532321/dock_532321.pdbq  
M 29508 2676 S 128.0 0.2 0:09.46 molaicald --config config.txt --ligand dock/532514/532514.pdbqt --out dock/532514/dock_532514.pdbq  
M 32676 2676 S 62.0 0.2 0:45.37 molaicald --config config.txt --ligand dock/531080/531080.pdbqt --out dock/531080/dock_531080.pdbq  
M 28188 2676 S 36.0 0.2 0:37.11 molaicald --config config.txt --ligand dock/531199/531199.pdbqt --out dock/531199/dock_531199.pdbq  
M 39948 16700 R 32.7 0.2 0:00.50 /home/spider/anaconda3/envs/pdb/bin/python /home/spider/anaconda3/envs/pdb/bin/prepare_ligand4.py  
M 32676 2676 R 18.7 0.2 0:05.71 molaicald --config config.txt --ligand dock/531080/531080.pdbqt --out dock/531080/dock_531080.pdbq  
M 31092 2676 R 18.7 0.2 0:01.26 molaicald --config config.txt --ligand dock/532339/532339.pdbqt --out dock/532339/dock_532339.pdbq  
M 32680 2676 R 18.0 0.2 0:06.33 molaicald --config config.txt --ligand dock/531043/531043.pdbqt --out dock/531043/dock_531043.pdbq  
M 29508 2676 R 18.0 0.2 0:03.11 molaicald --config config.txt --ligand dock/531675/531675.pdbqt --out dock/531675/dock_531675.pdbq  
M 29508 2676 R 18.0 0.2 0:03.08 molaicald --config config.txt --ligand dock/531675/531675.pdbqt --out dock/531675/dock_531675.pdbq  
M 31096 2676 R 18.0 0.2 0:02.40 molaicald --config config.txt --ligand dock/531901/531901.pdbqt --out dock/531901/dock_531901.pdbq  
M 31092 2676 R 18.0 0.2 0:00.37 molaicald --config config.txt --ligand dock/532756/532756.pdbqt --out dock/532756/dock_532756.pdbq  
8 28452 2676 R 18.0 0.2 0:00.27 molaicald --config config.txt --ligand dock/533008/533008.pdbqt --out dock/533008/dock_533008.pdbq  
F1Help F2Setup F3Search F4Filter F5Tree F6SortBy F7Nice F8Nice F9Kill F10Quit
```

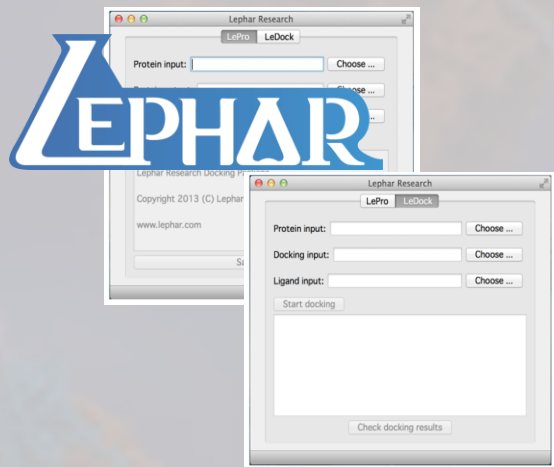
## Environment SETUP

## Method Validation

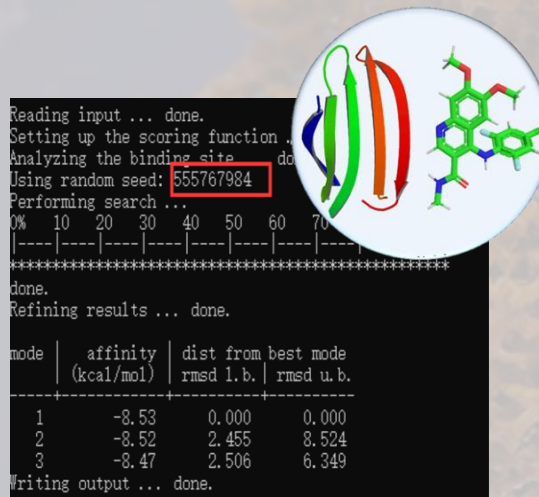
## Library preparation

## Virtual-Screening Docking

## MST Assay Experiment & Data analysis



***LeDock***



***MolAI/Cal***



***AutoDock.Vina***



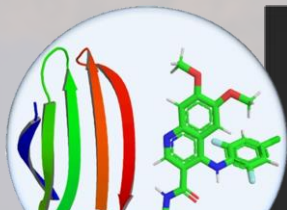
Environment  
SETUP

Method  
Validation

Library  
preparation

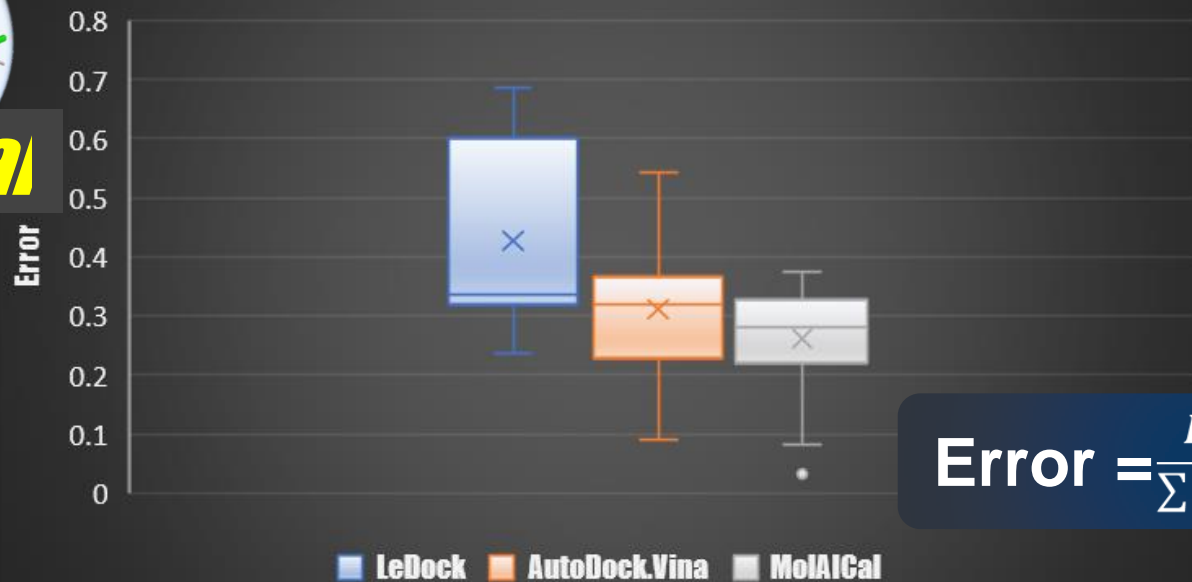
Virtual-Screening  
Docking

MST Assay  
Experiment &  
Data analysis



**MolAI-Cal**

## Accuracy



$$\text{Error} = \frac{\text{Error factor}}{\sum \text{Error factor}_i}$$



# Environment SETUP

# Method Validation

# Library preparation

# Virtual-Screening Docking

# MST Assay Experiment & Data analysis



```
from rdkit.Chem import Chem
from rdkit.Chem import SDmolSupplier, SDWriter
from rdkit.Chem import Descriptors
from tqdm import tqdm

input_sdf_file = 'LSW.sdf'
output_sdf_file = 'output.sdf'
sdf_supplier = SDmolSupplier(input_sdf_file)
writer = SDWriter(output_sdf_file)

total_molecules = 0
removed_fragments = 0

for molecule in tqdm(sdf_supplier):
    if molecule is not None:
        total_molecules += 1

        is_multifragment = len(Chem.GetMolFrags(
            molecule)) > 1

        if is_multifragment:
            fragments = Chem.GetMolFrags(molecule)

            max_fragment_weight = 0.0
            max_fragment = None

            for fragment in fragments:
                fragment_molecular_weight = Chem
                fragment_molecular_weight = Chem

                if fragment_molecular_weight
                    max_fragment_weight = frv
                    max_fragment = fragment

            writer.write(max_fragment)

        for fragment in fragments:
            if fragment != max_fragment:
                removed_fragments += 1
        else:
            writer.write(molecule)
```

```
from openbabel import openbabel
import os
from tqdm import tqdm

# Directory containing the input PDB files
input_directory = "LSW/"

# Output directory for the modified PDB files
output_directory = "LSW1/"
```

```
# Create the output directory if it doesn't exist
os.makedirs(output_directory, exist_ok=True)
```

```
# Loop through PDB files numbered from 1 to 20
for pdb_number in tqdm(range(1, 302738)):
    input_pdb_file = os.path.join(input_directory,
    output_pdb_file = os.path.join(output_dir,
```

```
# Initialize the Open Babel molecule
mol = openbabel.OBMol()
```

```
# Read the input PDB file
obconversion = openbabel.OBConversion()
obconversion.SetInAndOutFormats("pdb", "pdb")
obconversion.ReadFile(mol, input_pdb_file)
```

```
# Add hydrogen atoms
mol.AddHydrogens()
```

```
# Write the output PDB file
obconversion.WriteFile(mol, output_pdb_file)
```

```
print("Hydrogens added to all PDB files.")
```

```
def SdfToPdb_3D(file, path): #sdf 파일을 pc
    mols_E = []
    mols = []
    suppl = Chem.SDMolSupplier(file)
    print("File is loaded successfully.")
    for mol in tqdm(suppl):
        mols_E.append(mol) #읽어들이기
    Err = 0 #변수 선언
    mols = mols_E.copy() #복사본 생성
    print("Sdf molecules(", len(mols_E), "
    for x in range(len(mols_E)):
        if mols_E[x] is None:
            print("Error : molecule No.",
            del mols[x-Err]
            Err += 1
    print("Number of unexpected molecules:
    print("Number of molecules without err
    for m in tqdm(range(len(mols))):
        AllChem.EmbedMolecule(mols[m]) #3D
    print("Molecules are converted to 3D
    with Chem.PDBWriter(path) as w:
        for m in mols:
            w.write(m) #파일로 저장
    print("3D file is saved successfully.")
    return mols #반환
```

```
import os
from rdkit.Chem import Chem
from rdkit.Chem import Descriptors3D
import numpy as np
from tqdm import tqdm
```

```
data_dir = "./RESULT/10-100"
gyration_values = []
```

```
# Iterate over all files in the directory
for filename in tqdm(os.listdir(data_dir)):
    if filename.endswith(".pdb"):
        pdb_file = os.path.join(data_dir, filename)
        mol = Chem.MolFromPDBFile(pdb_file)
        if mol is not None:
            radius_of_gyration = Descriptors3D.RadiusOfGyration(mol)
            gyration_values.append(radius_of_gyration)
```

```
mean_gyration = np.mean(gyration_values)
normalized_mean_gyration = mean_gyration / 0.35 * 2
```

```
print(f"Mean Radius of Gyration: {mean_gyration}")
print(f"Normalized Mean Radius of Gyration: {normalized_mean_gyration}")
```

```
for filename in os.listdir(pdb_directory):
    if filename.endswith(".pdb"):
        pdb_path = os.path.join(pdb_directory, filename)
        mol = Chem.MolFromPDBFile(pdb_path)

        if mol is not None:
            mol = Chem.AddHs(mol)
            rg = Descriptors3D.RadiusOfGyration(mol)

            for min_range, max_range in radius_ranges:
                if min_range <= rg < max_range:
                    result_subdirectory = os.path.join(result_directory, f'{min_range}-{max_range}')
                    os.makedirs(result_subdirectory, exist_ok=True)
                    result_path = os.path.join(result_subdirectory, filename)
                    shutil.copy(pdb_path, result_path)
                    print(f'Moved: {filename} to {min_range}-{max_range} folder')
                    break
```



**Environment  
SETUP**

**Method  
Validation**

**Library  
preparation**

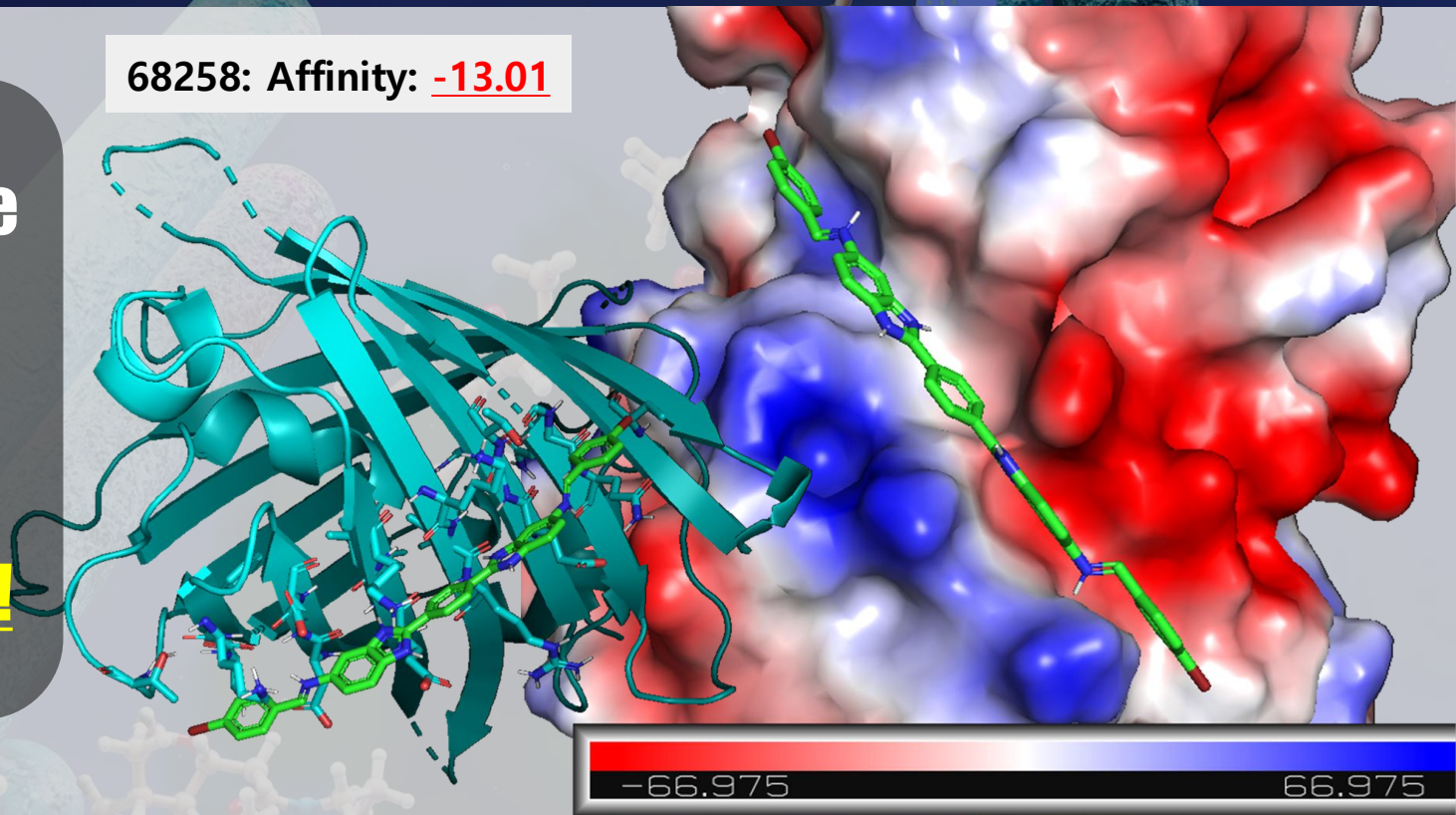
**Virtual-Screening  
Docking**

**MST Assay  
Experiment &  
Data analysis**

**Docking Pose  
of 1,000,000**

**Top 2800  
Compounds !**

68258: Affinity: **-13.01**



**Environment  
SETUP**

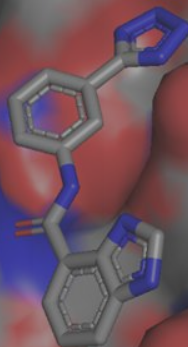
**Method  
Validation**

**Library  
preparation**

**Virtual-Screening  
Docking**

**MST Assay  
Experiment &  
Data analysis**

***1,000,000 >> 2,800 compounds***



***30-Compounds  
of Hit Candidates!!***

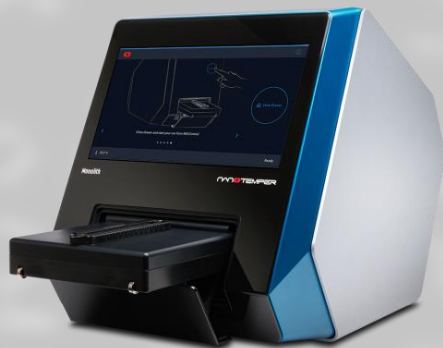
**Environment  
SETUP**

**Method  
Validation**

**Library  
preparation**

**Virtual-Screening  
Docking**

**MST Assay  
Experiment &  
Data analysis**



***MST Assay***



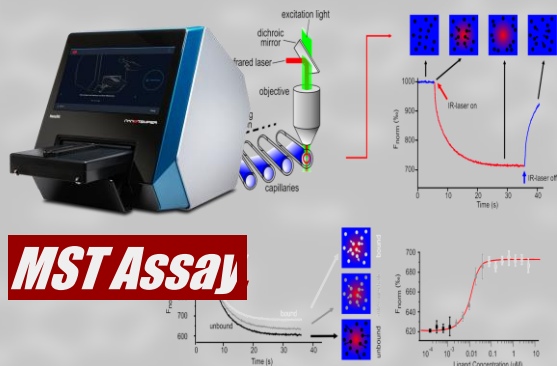
# Environment SETUP

# Method Validation

# Library preparation

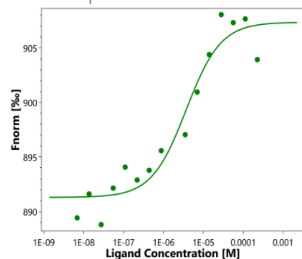
# Virtual-Screening Docking

# MST Assay Experiment & Data analysis



## Compound 20 - $6.72 \pm 4.28 \mu\text{M}$

### Dose Response



Response Evaluation: On Time 15s

Kd model

Unbound 891.3

Bound 907.3

Kd 3.69  $\mu\text{M}$

TargetConc 50 nM

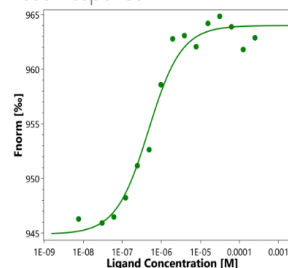
Response Amplitude: 16.0

Noise: 1.7

Signal to Noise Ratio: 9.2 ✓

## Compound 24 - 308nM

### Dose Response



Response Evaluation: On Time 1.5s

Kd model

Unbound 944.9

Bound 964

Kd 308 nM

TargetConc 50 nM

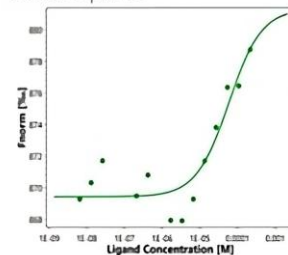
Response Amplitude: 19.1

Noise: 1.2

Signal to Noise Ratio: 15.5 ✓

## Compound 22 - $63.2 \pm 6.08 \mu\text{M}$

### Dose Response



Response Evaluation: On Time 20s

Kd model

Unbound 869.4

Bound 881.3

Kd 58.9  $\mu\text{M}$

TargetConc 50 nM

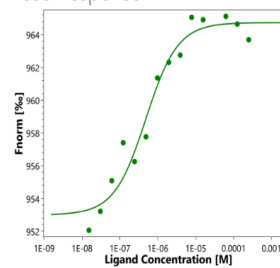
Response Amplitude: 11.9

Noise: 1.3

Signal to Noise Ratio: 9.2 ✓

## Compound 28 - 324nM

### Dose Response



Response Evaluation: On Time 1.5s

Kd model

Unbound 953

Bound 964.7

Kd 324 nM

TargetConc 50 nM

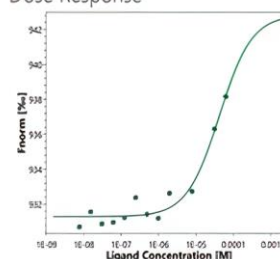
Response Amplitude: 11.7

Noise: 1.0

Signal to Noise Ratio: 12.2 ✓

## Compound 36 - $42.95 \pm 1.34 \mu\text{M}$

### Dose Response



Response Evaluation: On Time 2.5s

Kd model

Unbound 931.3

Bound 942.8

Kd 42  $\mu\text{M}$

TargetConc 50 nM

Response Amplitude: 11.5

Noise: 0.5

Signal to Noise Ratio: 23.2 ✓

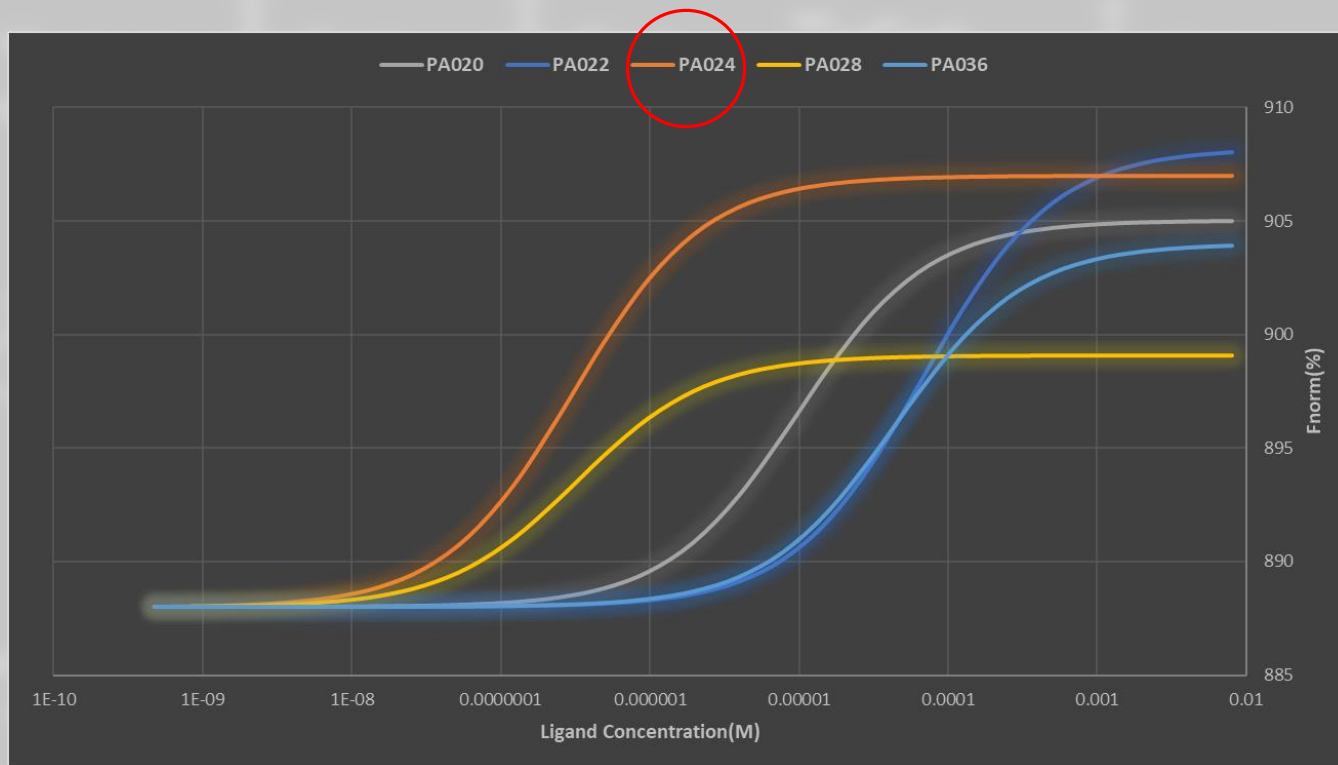
**Environment  
SETUP**

**Method  
Validation**

**Library  
preparation**

**Virtual-Screening  
Docking**

**MST Assay  
Experiment &  
Data analysis**



# ***Future Goal - Hit to Lead***

## ***Hit Compounds***

**PA020**

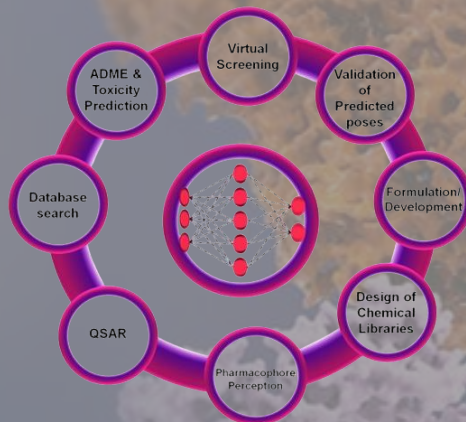
**PA022**

**PA024**

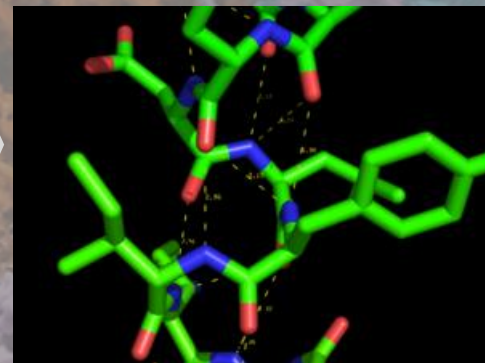
**PA028**

**PA036**

## ***AI-Based Modification***

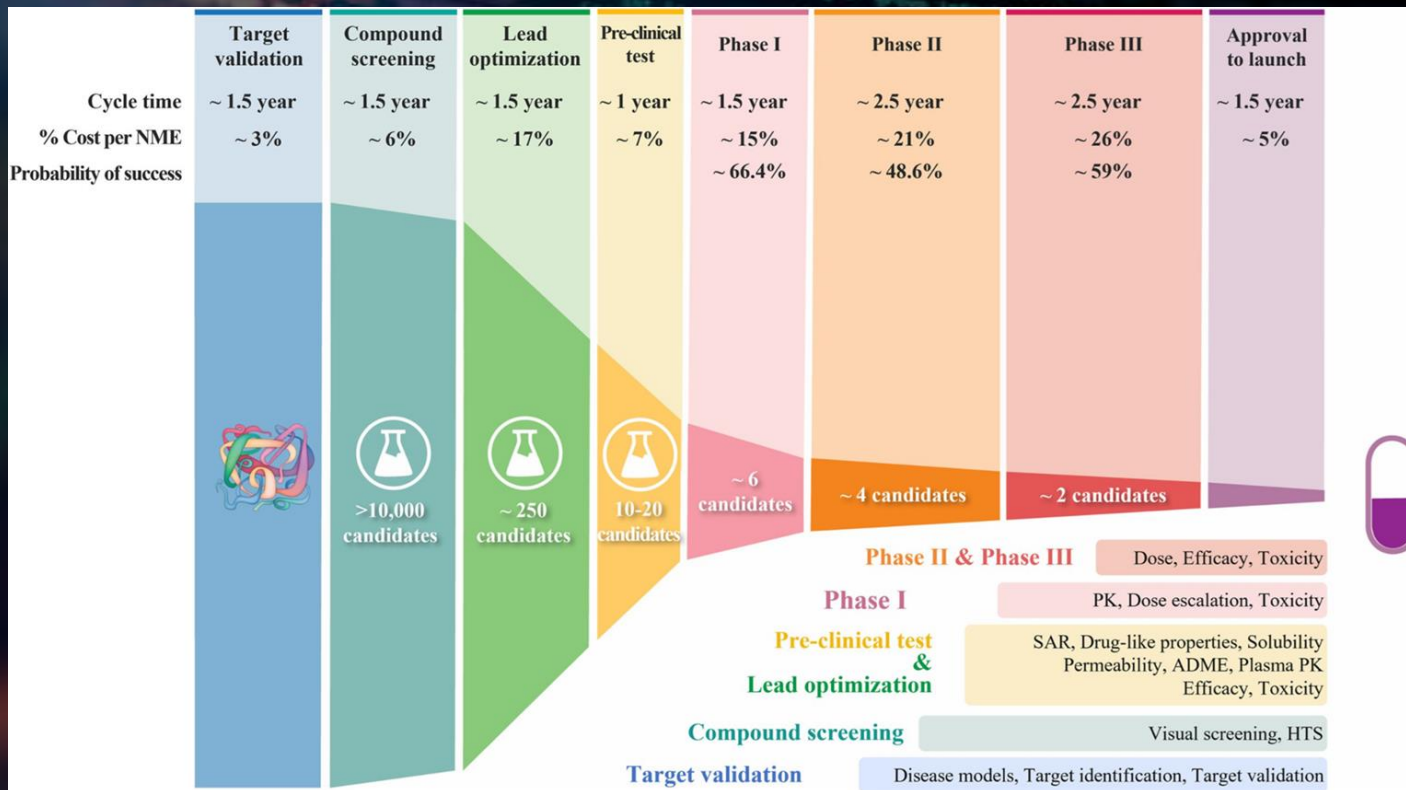


## ***Synthetic Analysis***





# Conclusion





**QNA**