FEBRUARY 2024

MEET-EU

Amina ALIOUA
Adel BIBI TRIKI
Késia DETHELOT-DELAG
Tiziri TERKMANI

TABLE
of contents

Introduction

Find inhibitor for NSP13

NSP13 : vital protein of SRAS-coV-2 virus

Pipeline

Database of compounds

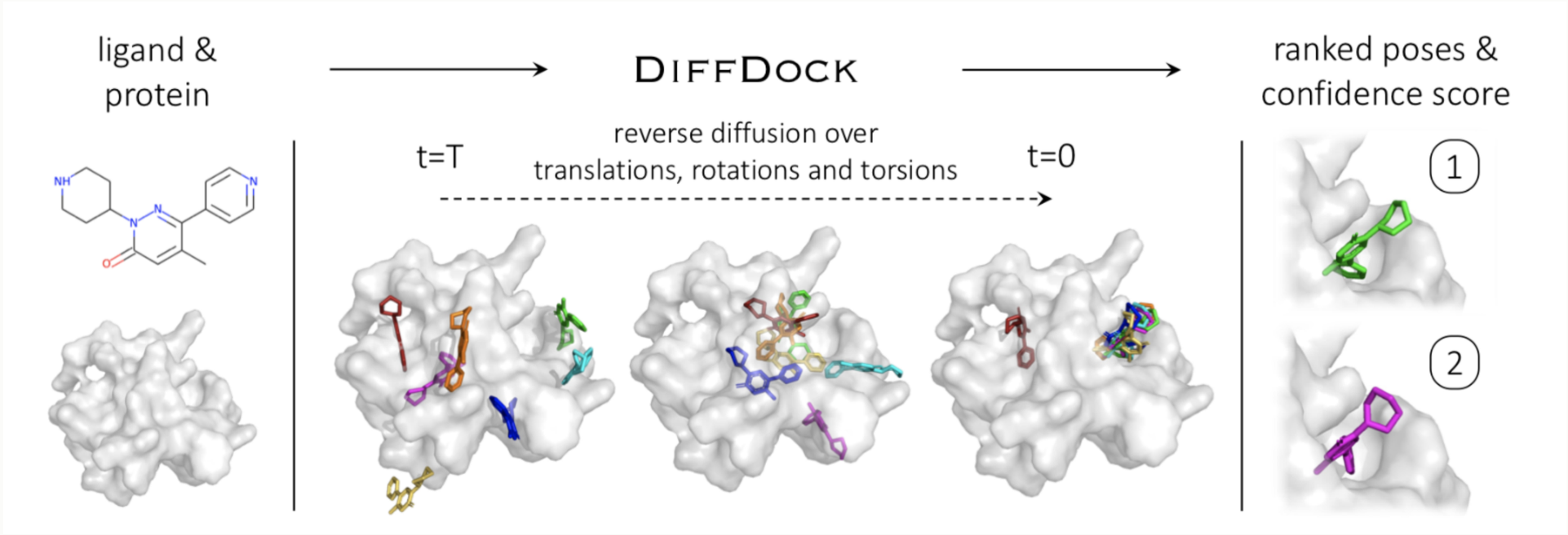
Clustering

Estimator

Prediction

Diffdock

- Molecular docking
- Slow method



Fingerprints

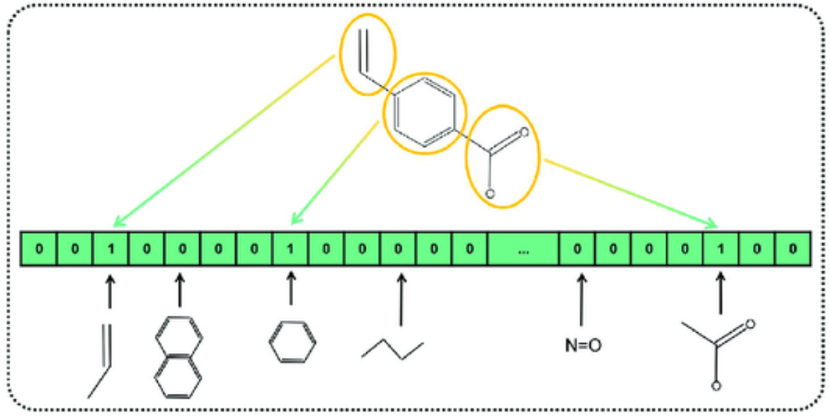
Fingerprints

Morgan

- bit vector, unfixed size
- radius parameter adjustment
- distinguish stereoisomers

MACCS

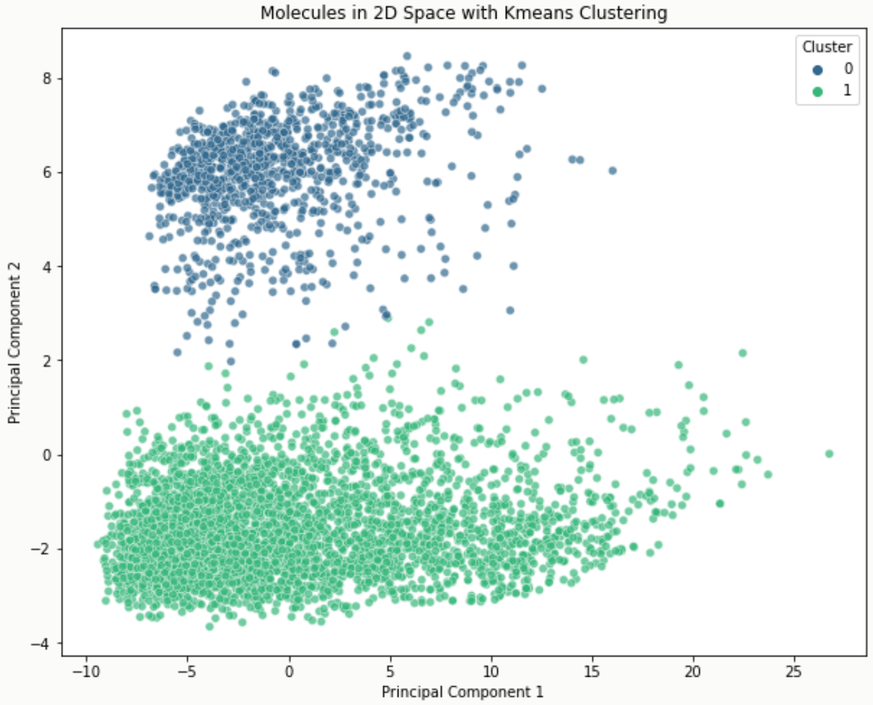
- bit vector, fixed size of 166
- computationally lighter



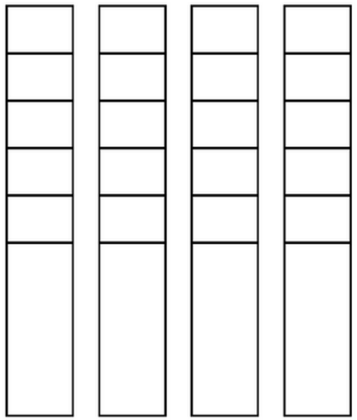
Results Clusters

Tanimoto distance → Kmeans → 2 clusters with MACCS

Kmeans → No segregation with Morgan



Prediction score



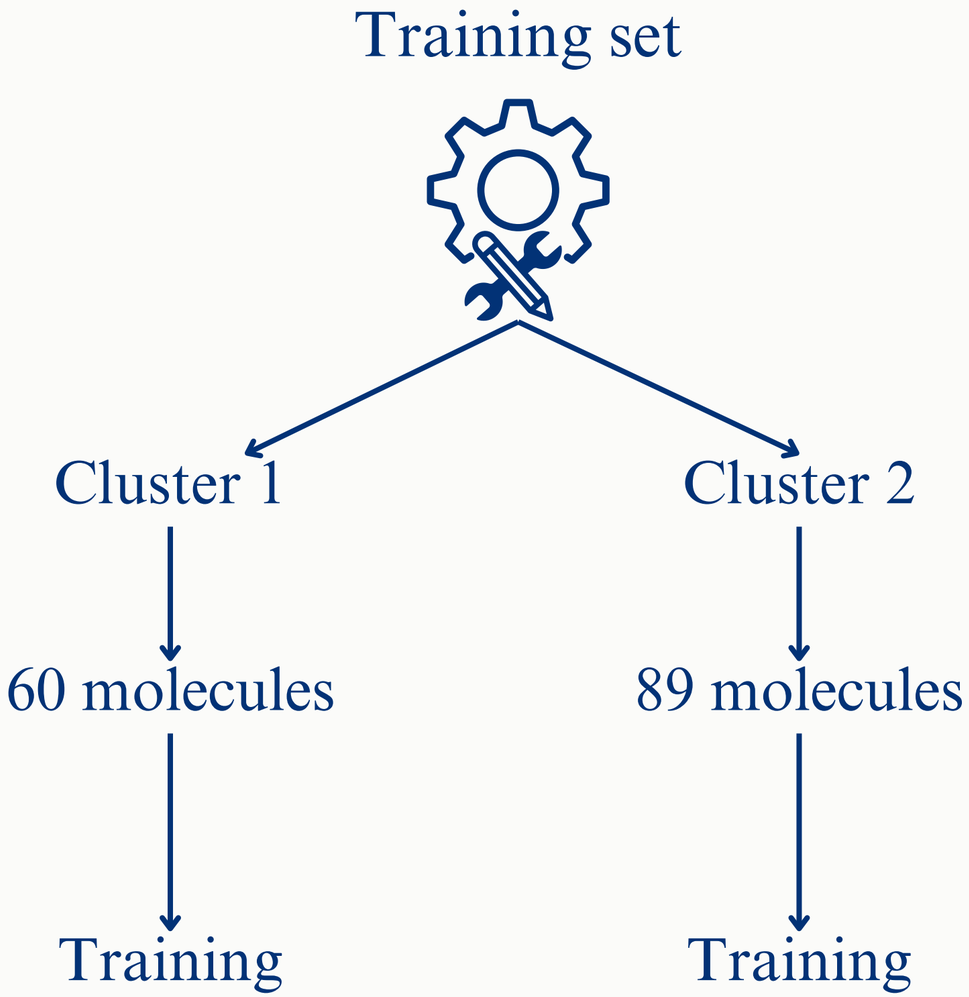
Bit vector fingerprints

Features fot the prediction

Random Forest

Neural Network

Neural Network



Validation

Cluster	MAE	MSE
Cluster 1	0.3415	0.1368
Cluster 2	0.3193	0.1154

Results Best Candidates

Cluster 1

	eos	smiles	Conf	Conf_Average
0	EOS100346	N=C(N)NOCC[C@H](N)C(=O)O.O=S(=O)(O)O	1.095006	-0.339066
1	EOS100001	CC[C@]1(O)C[C@@H]2CN(CCc3c([nH]c4ccccc34)[C@@]...	0.187590	-0.480802
2	EOS100611	CCOC(=O)c1ccc(OC(=O)CCCCCNC(=N)N)cc1.CS(=O)(=O)O	2.083342	-0.204106
3	EOS100734	CN1C2CCC1CC(OC(c1ccccc1)c1ccccc1)C2.CS(=O)(=O)O	-1.180006	-0.427780
4	EOS101103	CC[C@]1(O)C[C@@H]2CN(CCc3c([nH]c4ccccc34)[C@@]...	0.581058	-0.365956
5	EOS101129	CC(C)C[C@H]1C(=O)N2CCC[C@H]2[C@]2(O)O[C@](NC(=...	-0.045090	-0.261124
6	EOS101317	COc1cc2nc(N3CCN(C(=O)C4COc5ccccc5O4)CC3)nc(N)c...	-0.735330	-0.459855
7	EOS101424	COC(=O)[C@H](c1ccccc1Cl)N1CCc2sccc2C1.O=S(=O)(O)O	-0.410535	-0.234433
8	EOS100994	CO/N=C1\CN(c2nc3c(cc2F)c(=O)c(C(=O)O)cn3C2CC2)...	-1.306605	-0.417591
9	EOS101450	O=C(Nc1cnn2ccc(N3CCC[C@@H]3c3cc(F)ccc3F)nc12)N...	0.872005	-0.471151
10	EOS101000	N#Cc1c(Oc2cccc(OS(=O)(=O)CCCC(F)(F)F)c2)cccc1C...	-0.235021	-0.445547
11	EOS101001	CS(=O)(=O)O.Cn1cc(C2=C(c3cn(CCCSC(=N)N)c4ccccc...	0.109194	-0.119145
12	EOS101653	CS(=O)(=O)O.Oc1ccc(C2CNCCc3c2cc(O)c(O)c3Cl)cc1	0.090173	-0.249529
13	Acetylcysteine	CC(=O)NC(CS)C(=O)O	-0.245664	-0.479596
14	Clavulanic Acid	C1C2N(C1=O)C(C(=CCO)O2)C(=O)O	-0.436168	-0.237415

Known inhibitors

14 molecules

Cluster 2

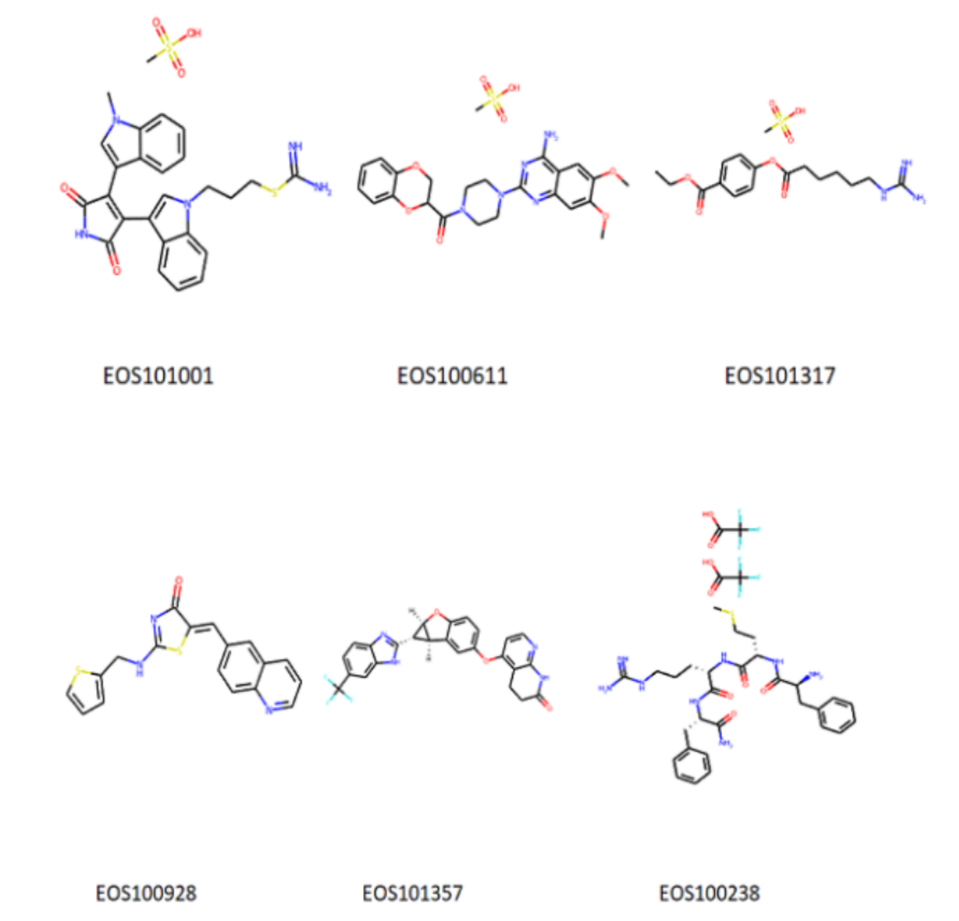
16	EOS100964	CCC(CC)(NC(=O)C(C)C)C(=O)N[C@@H](CCCNC(=N)N)C(...	-0.938951	-0.493857
17	EOS101357	O=C1CCc2c(Oc3ccc4c(c3)[C@@H]3[C@H](O4)[C@H]3c3...	-0.495913	-0.361723
18	EOS101461	N=C(N)NCCC[C@H](NC(=O)COCC(c1ccccc1)c1ccccc1)C...	-0.213281	-0.227900
19	EOS101898	O=C(Cc1ccccc1F)Nc1nc2ccccc2[nH]1	-1.831513	-0.488446
20	EOS101916	COc1cccc(C(=O)Nc2ncc(Cc3cccc(C(F)(F)F)c3)s2)c1	-0.567059	-0.470855
21	EOS101923	CSc1nc2ccc(NC(=O)c3cc4ccccc4o3)cc2s1	-0.908828	-0.420800
22	EOS101925	Cc1cccc(C(=O)NC2=NCCS2)c1	-0.824673	-0.343452
23	EOS101934	Cc1ccc2nc(NC(=O)CCCc3nc4ccccc4s3)sc2c1	-1.138363	-0.420967
24	EOS102089	O=C(Cc1cccs1)Nc1nc2ccccc2s1	-1.527312	-0.489443
25	EOS102139	C=CCn1c2ccccc2c2nnc(SCCCn3c(=O)[nH]c4ccccc43)nc21	-0.272600	-0.468202
26	EOS102330	CSc1nc2ccc3nc(NC(=O)C(c4ccccc4)c4ccccc4)sc3c2s1	-1.435418	-0.457659
27	EOS102371	COc1ccc(/C=C/C(=O)Nc2ncc(Cc3cccc(C(F)(F)F)c3)s...	-0.348160	-0.282739
28	EOS102233	O=C(Nc1ccccc1)Nc1cnns1	-2.051171	-0.447467
29	EOS102433	Br.Br.N=C(N)SCCc1ccc(CCSC(=N)N)cc1	-0.360232	-0.459541

29 molecules

Threshold of -0.5

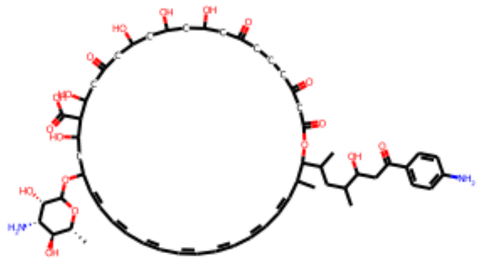
Analysis of best hits

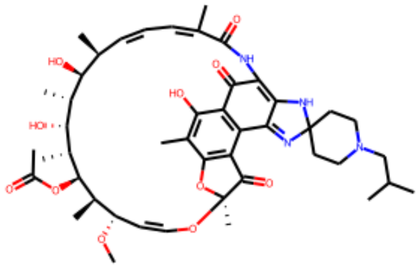
- Our best prediction with Neural Network
  - Cyclic molecules
  - Contain sulfur

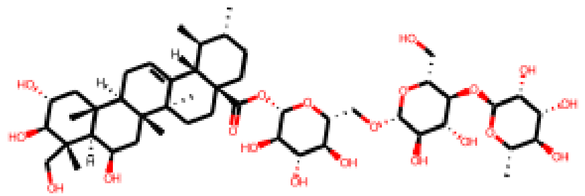


Collaboration with Charles team

Diffdock and Autodock Vina







Bigger moleclules !

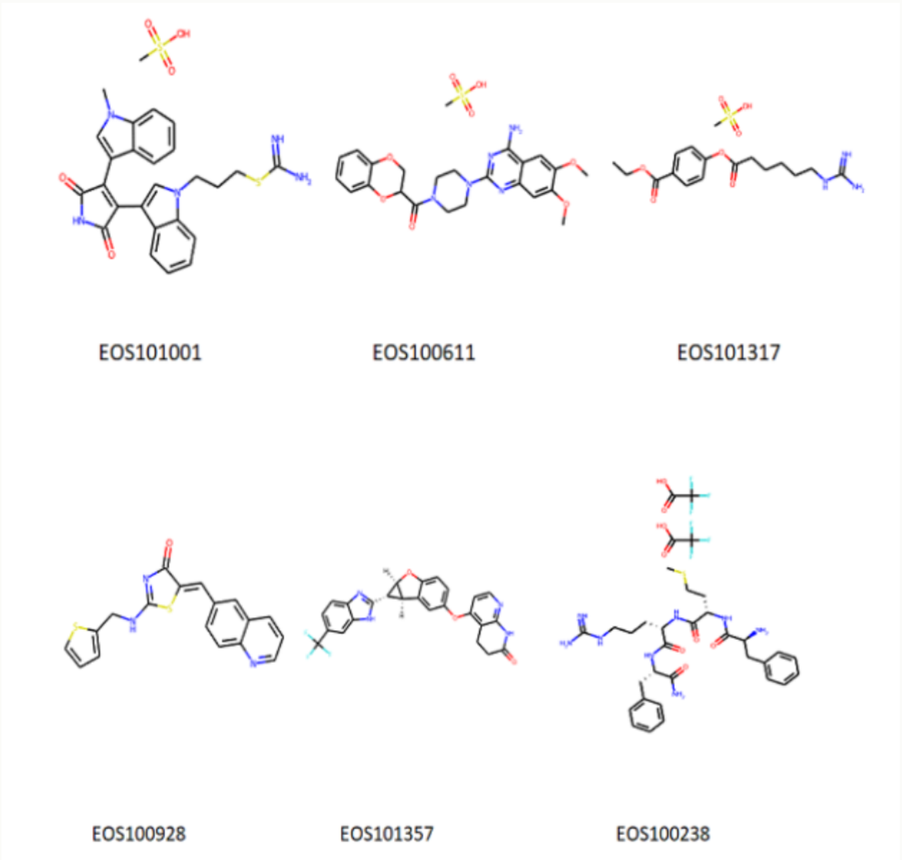
Testing with our model

Metric	Value
Mean Absolute Error (MAE)	3.1642
Mean Squared Error (MSE)	15.4131

Conclusion

- Conclusive results.
- Would we suggest DiffDock ?
    - Yes, but....

- Major issues:
    - Computationnal
    - Unxepected results



Perspectives

- Other methods:
  - Fingerprints: MAP4, MHFP6
  - Docking engines: Equibind
- Other features in the database.
- Using the pocket.
- “Reverse screening”

THANK YOU FOR YOU ATTENTION