

# Fragment-based Drug Discovery of Pan-Coronavirus Antivirals Assisted by Generative Machine Learning:

## Targeting the SARS-CoV-2 Helicase (nsp13)



Follow us on GitHub :

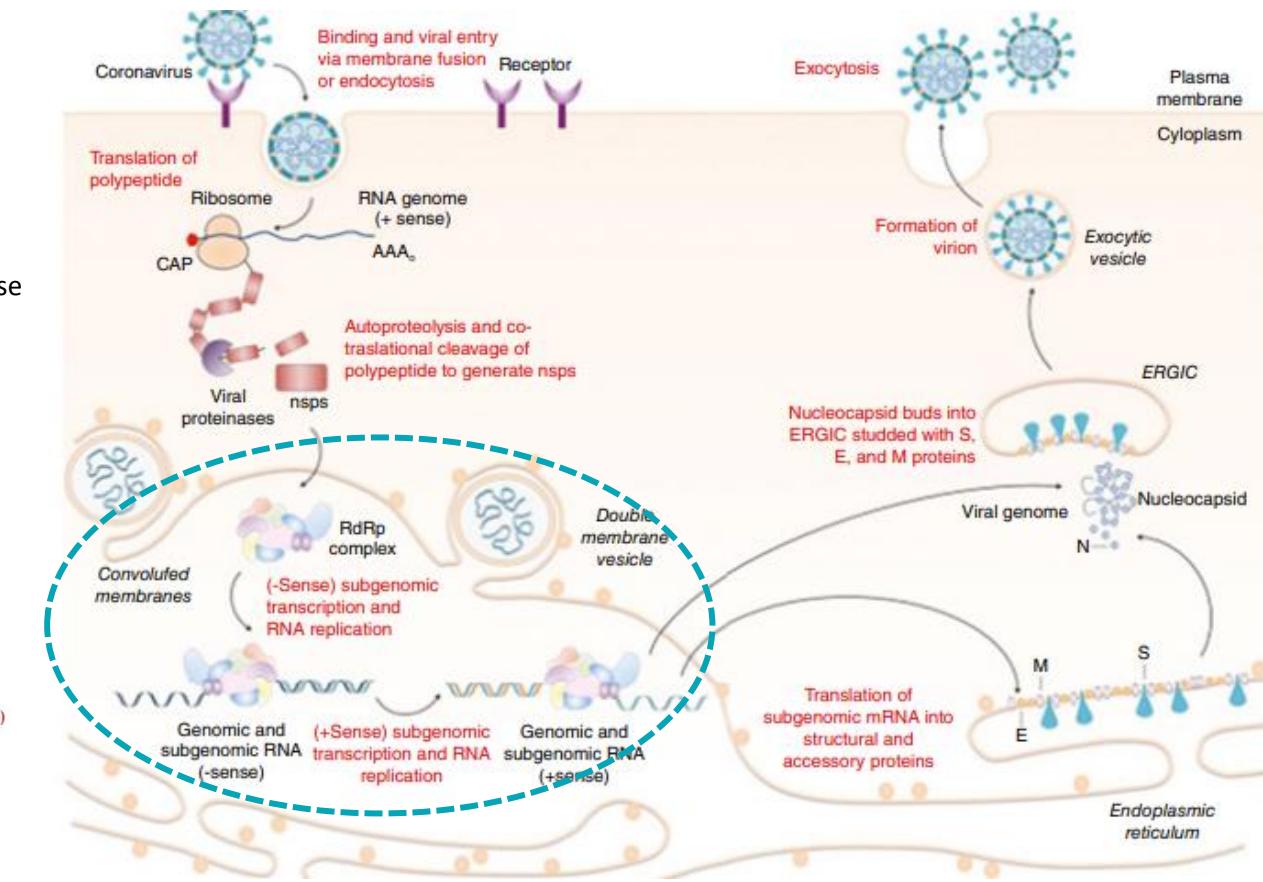
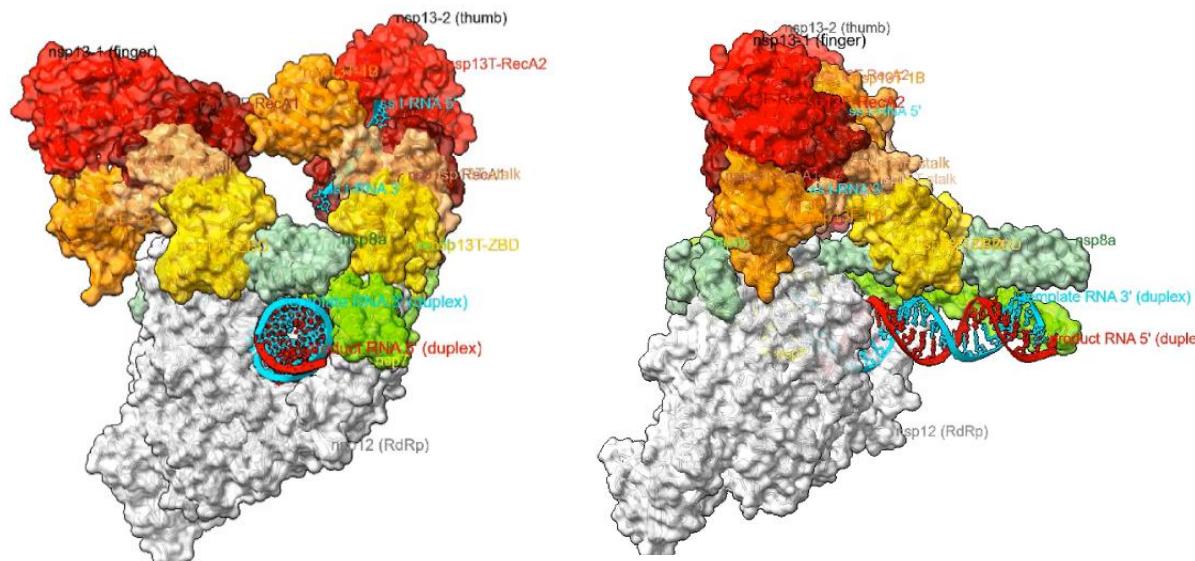


**Tom Knight, Ph.D. Candidate in Medicinal Chemistry  
Prof. Matthew Todd Research Group  
Department of Pharmaceutical and Biological Chemistry  
UCL School of Pharmacy**

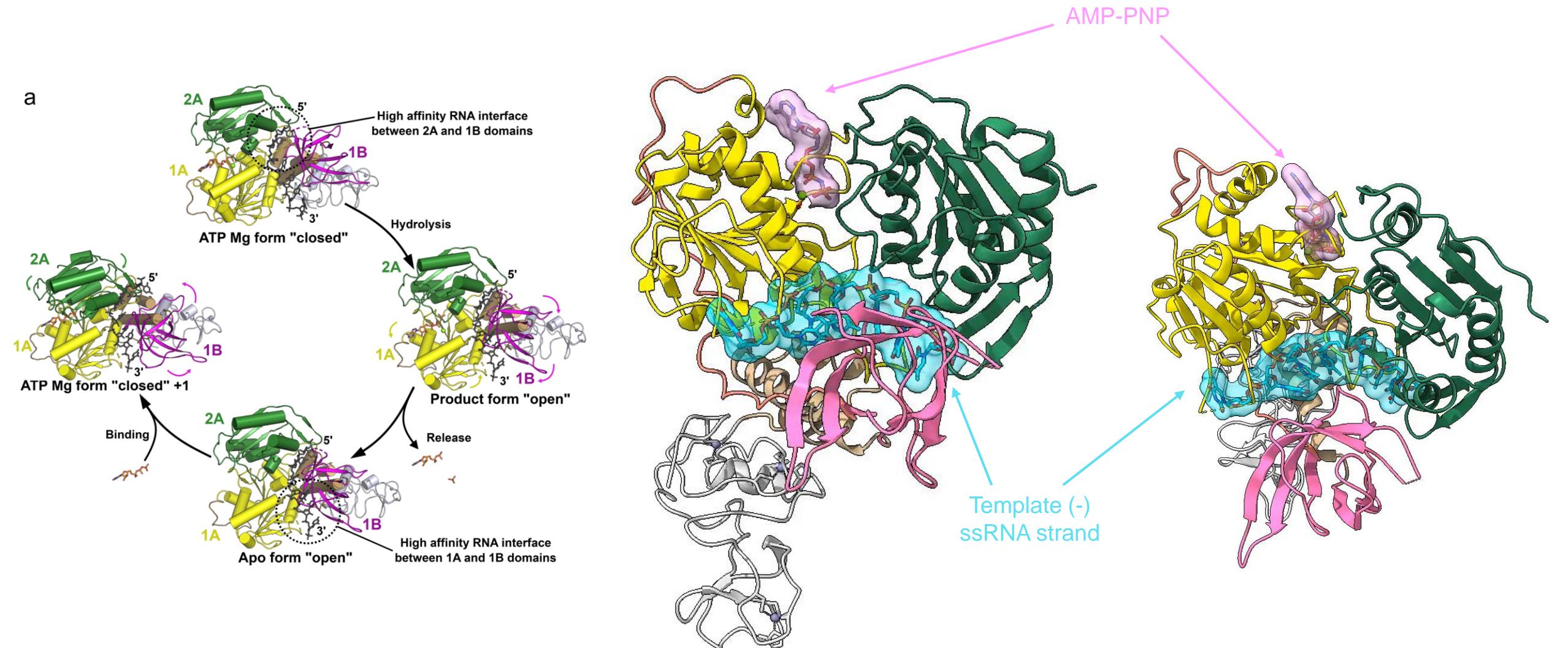
# SARS-CoV-2 Replication Transcription Complex (RTC): The (+)ssRNA replicative machinery of the coronavirus (PDB 7CXN)

## SARS-CoV-2 Helicase (non-structural protein-13, nsp13) enzymatic activity:

- CryoEM has revealed two copies of the helicase in the replicase-transcriptase complex; one active in RNA unwinding (thumb) and one passive (finger).
- ATP hydrolysis (ATPase), RNA-translocation (translocase) and RNA unwinding (helicase).
- Helicase is activated by binding to ATP and a 3' ssRNA overhang.
- Template (-)ssRNA is threaded through an RNA-binding channel between the Rec1A and Rec2A domains.
- Unwinding of RNA duplex is coupled to ATP hydrolysis and processive translocation of the helicase by one nucleotide.



# Nsp13 Helicase RNA Translocation Mechanism



# READDI-AC MedChem Core D

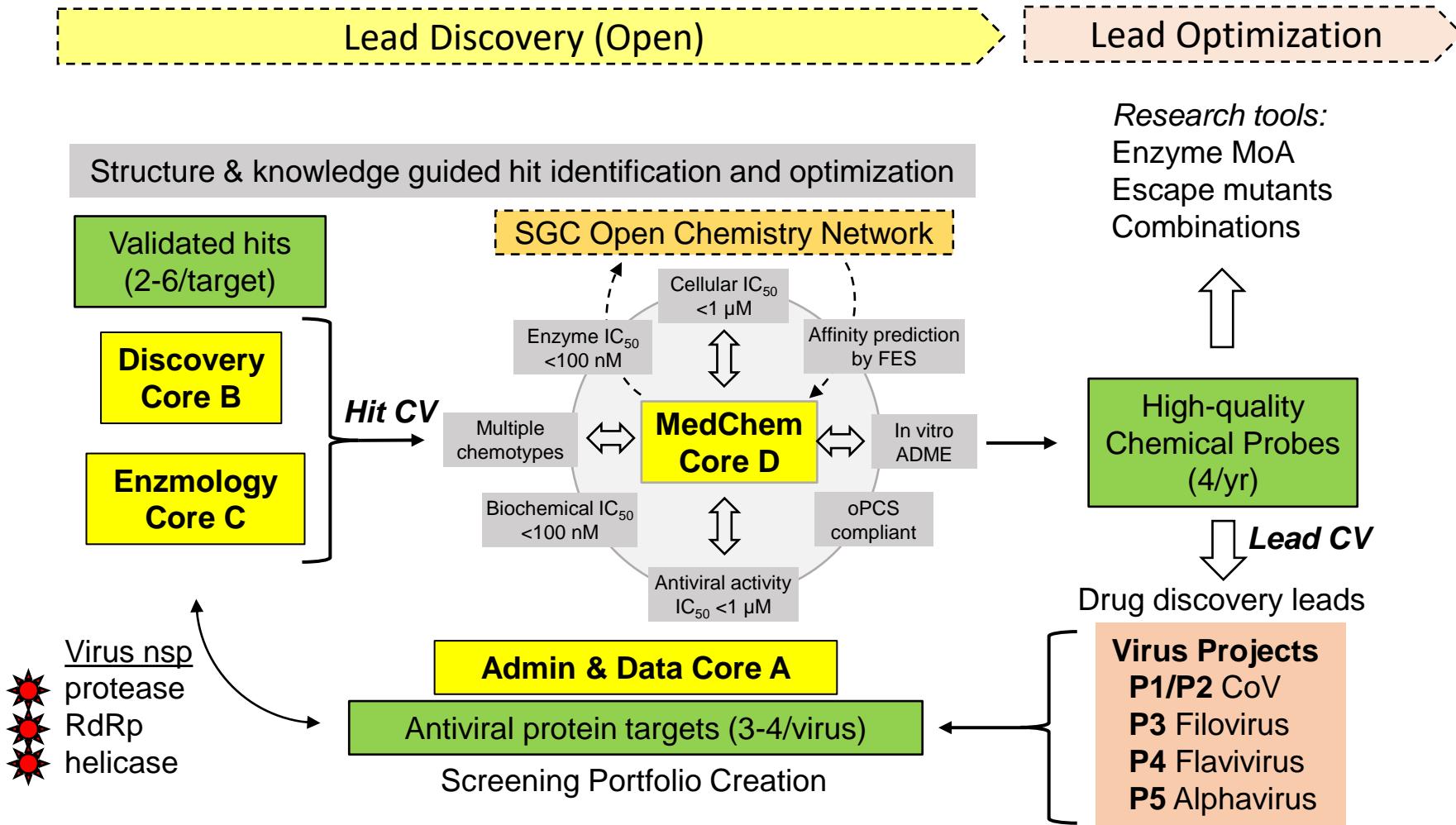
Goal:

- In partnership with Cores B & C, deliver 4 leads/yr to the Projects
- Focus on non-nucleoside inhibitors of viral proteases, replicases, and helicases

Will will use:

- Use open science to support quality and quantity of leads
- Use computational chemistry to accelerate hit generation

# READDI-AC Cores and Projects



Mat Todd



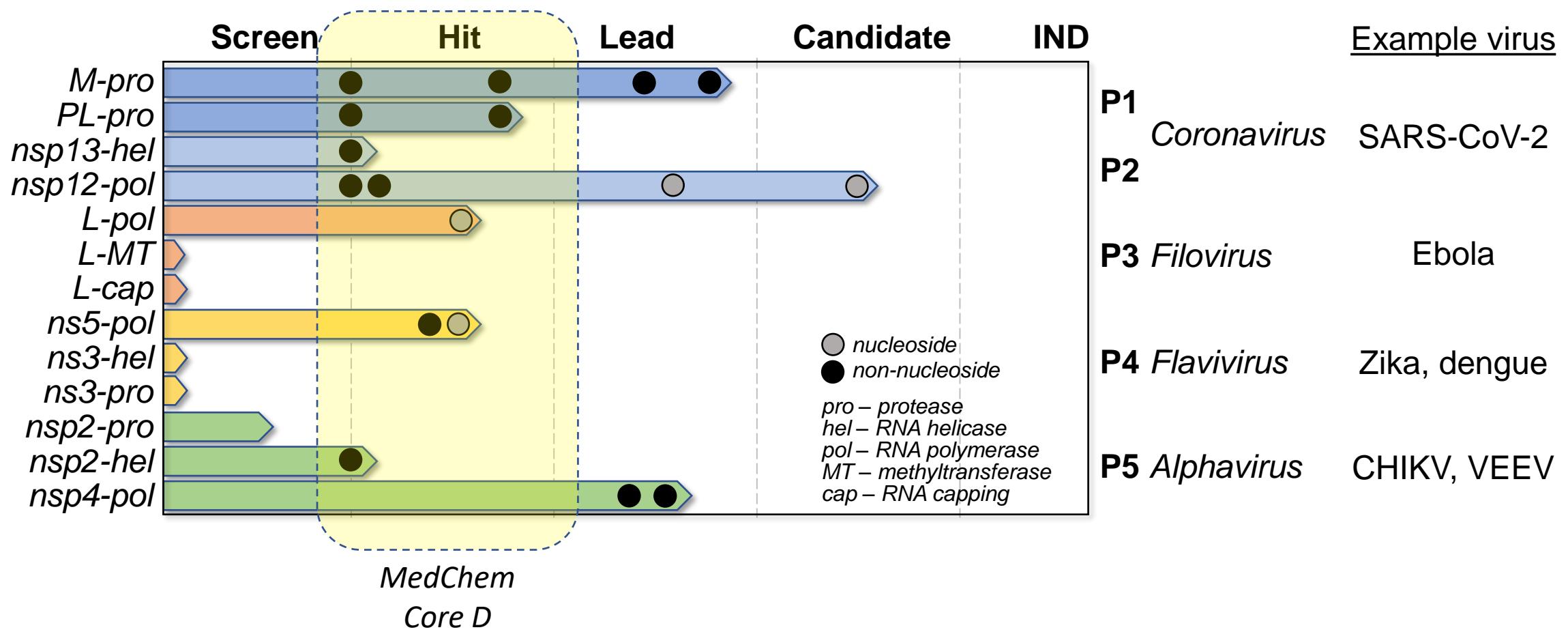
Peter Brown



Anwar Hossain



# READDI-AC Portfolio (Aug 2022)



MENU

# FRAGALYSIS: NSP13



SAVE



SHARE



DOWNLOAD STRUCTURES



TIMELINE



diamond



janssen

COVID Moonshot

CONTRIBUTORS

**Tag Details**

Tag name	Category	Creator	Date
A - Nucleotide Site	Sites	SELECT HITS	10/27/2021
B - RNA-3' Site	Sites	SELECT HITS	10/27/2021
B2 - RNA-3' Site 2	Sites	SELECT HITS	10/27/2021
C1 - RNA-5' Site	Sites	SELECT HITS	10/27/2021
C2 - RNA-5' Proximal	Sites	SELECT HITS	10/27/2021
D1 - RNA-central	Sites	SELECT HITS	10/27/2021

**Hit List Filter**

Union

Show untagged hits   Show all tags

Sites	Series	Discussion	Other
A - Nucleotide Site			
B - RNA-3' Site			
B2 - RNA-3' Site 2			
C1 - RNA-5' Site			
C2 - RNA-5' Proximal			
D1 - RNA-central			
E - Stalk			

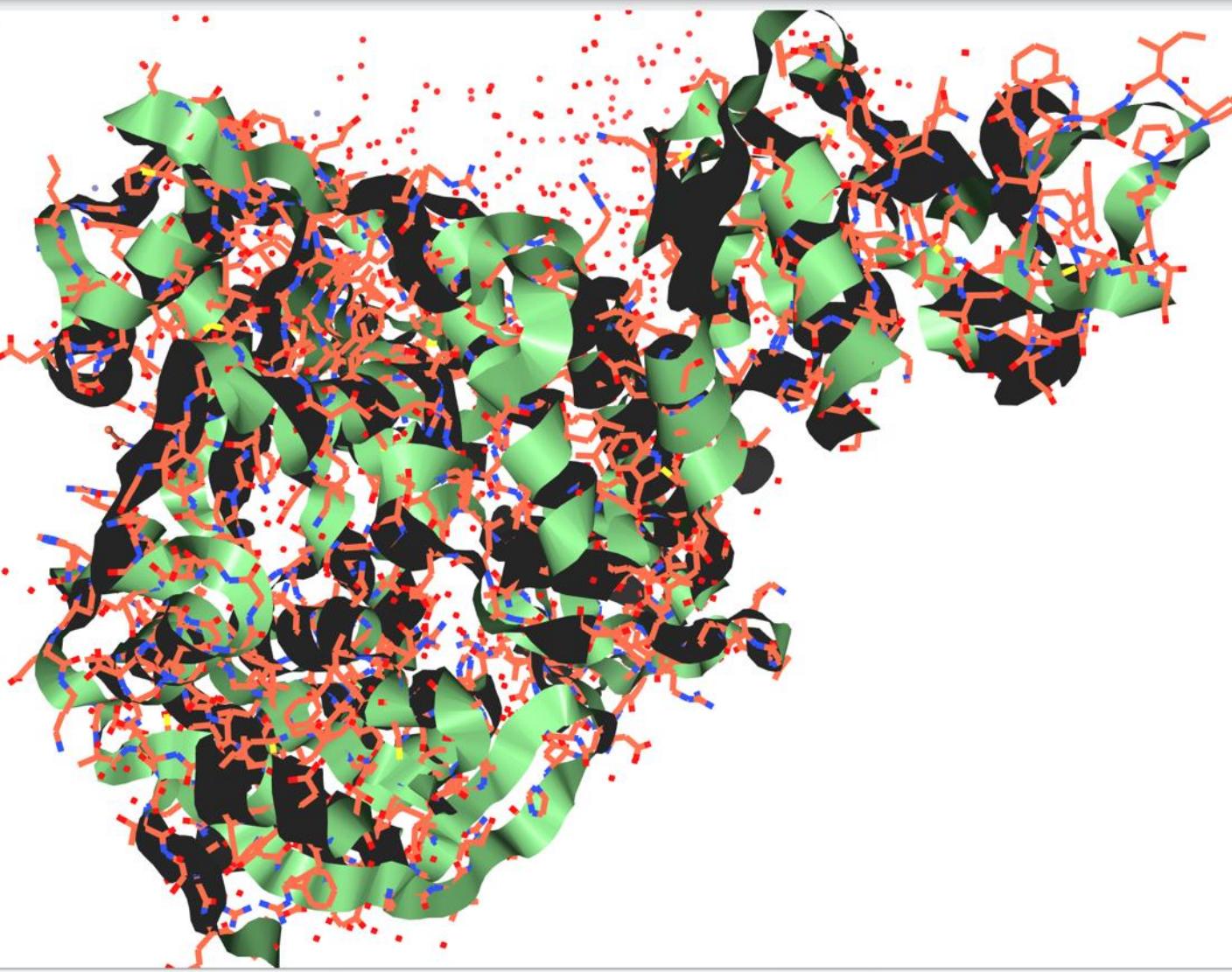
**Hit navigator**

Search

SELECT ALL HITS Selected: 0

MW	logP	TPSA	HA	Hacc	Hdon	Rots	Rings	Velec		
X0029_OA	197	1	66	14	2	2	2	1	74	
X0034_OB	203	1	60	13	2	1	2	1	72	
X0176_OB	199	1	46	13	2	1	4	1	72	
X0183_OB	199	1	60	13	2	1	3	1	72	

TOTAL 13   LOAD NEXT 30   LOAD NEXT 100   LOAD FULL LIST

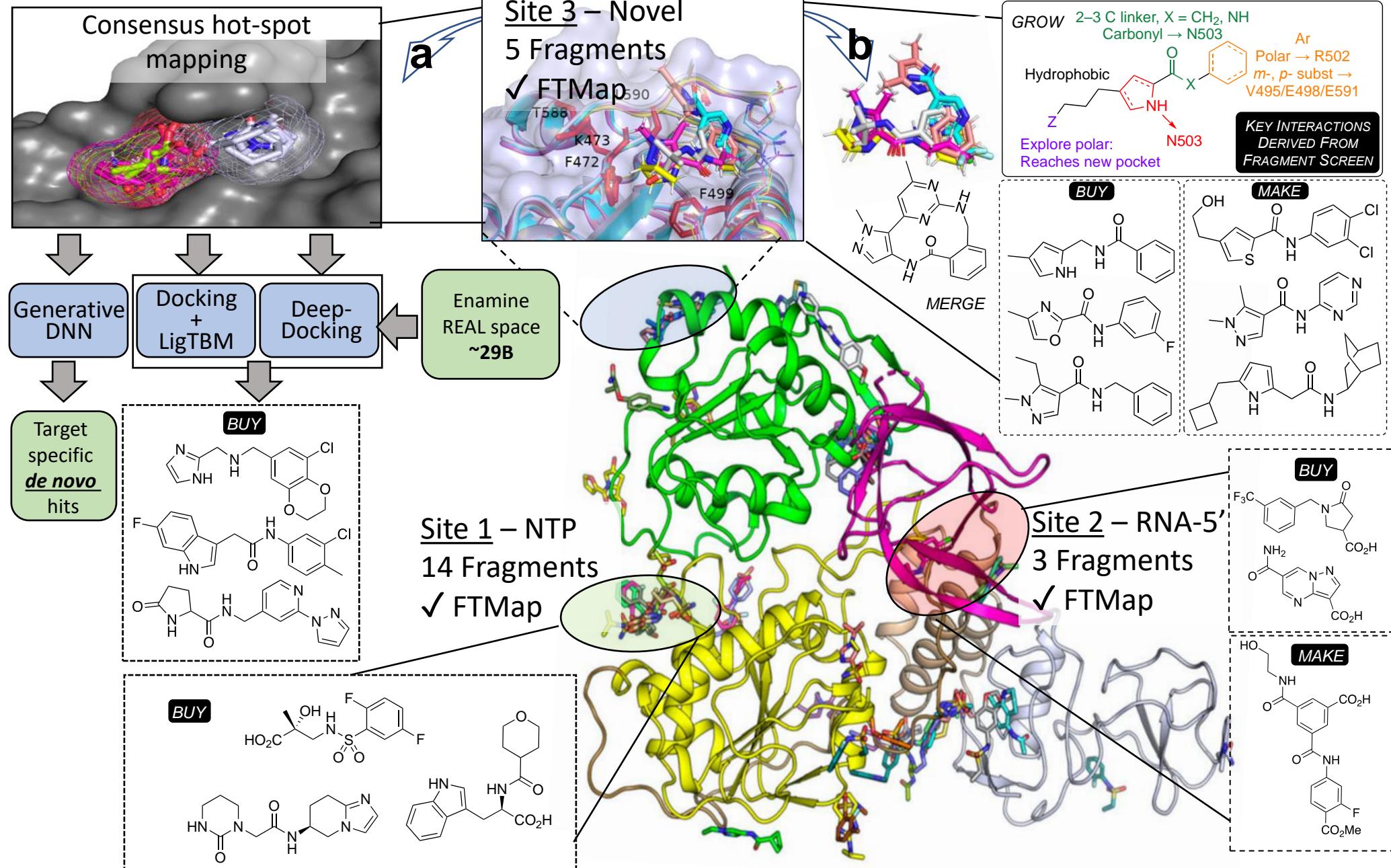


LHS

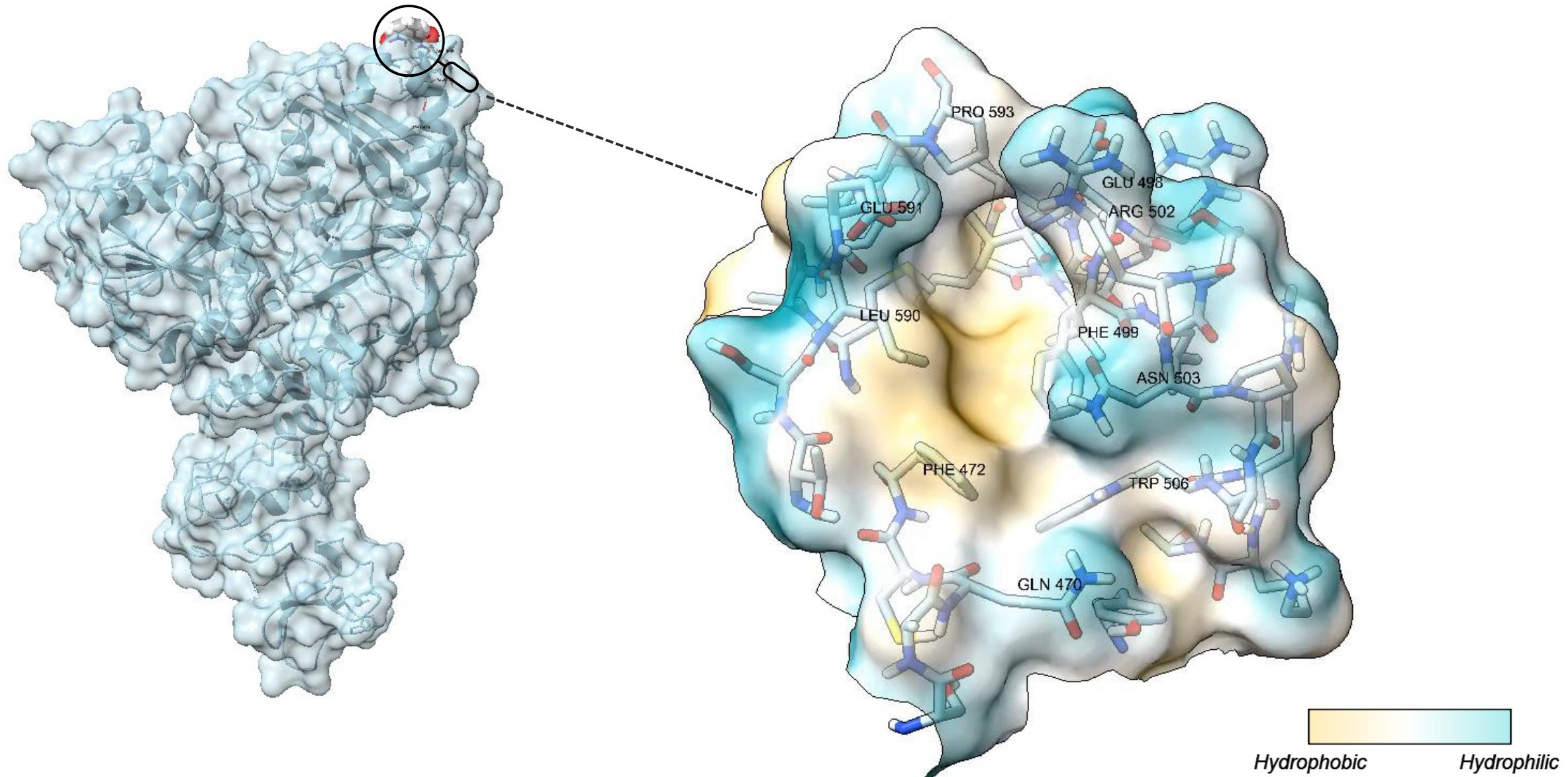


RHS

nsp13: Diamond X-ray Fragment Screen (Joe Newman and Frank von Delft)



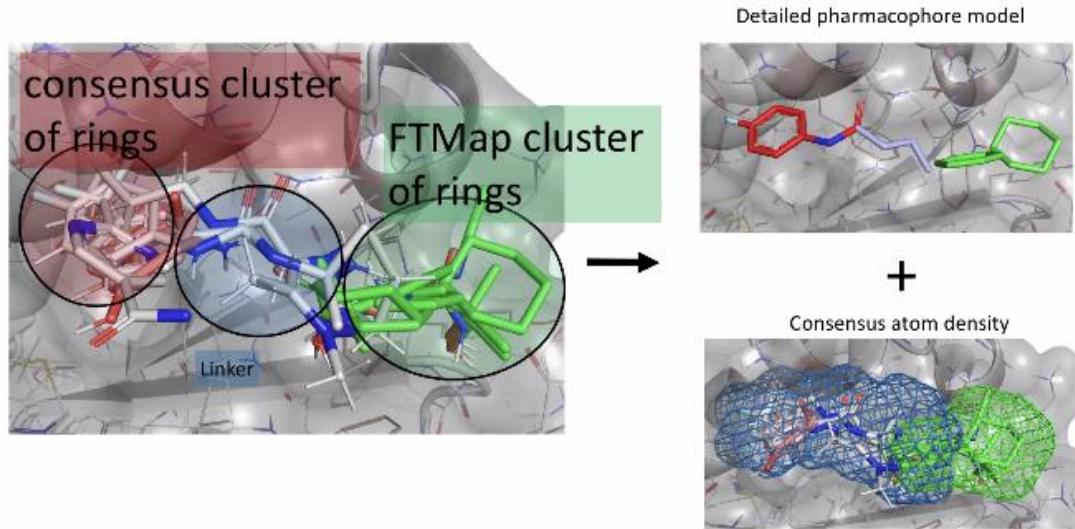
# Nsp13 Helicase C-terminus-B Pocket (Site 3)



# Crystal Fragment Hits to Pharmacophore Identification by Kostya



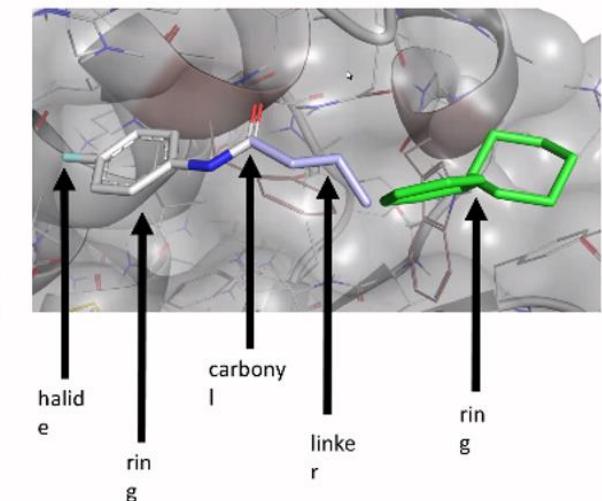
## Pharmacophore identification



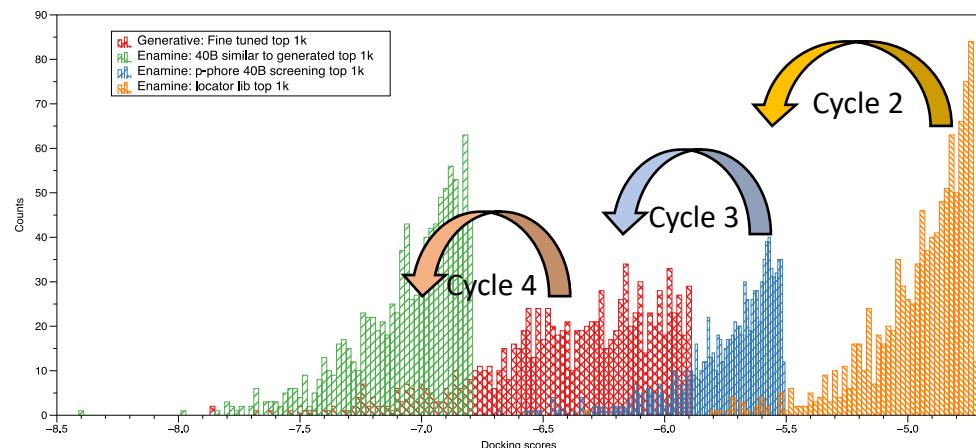
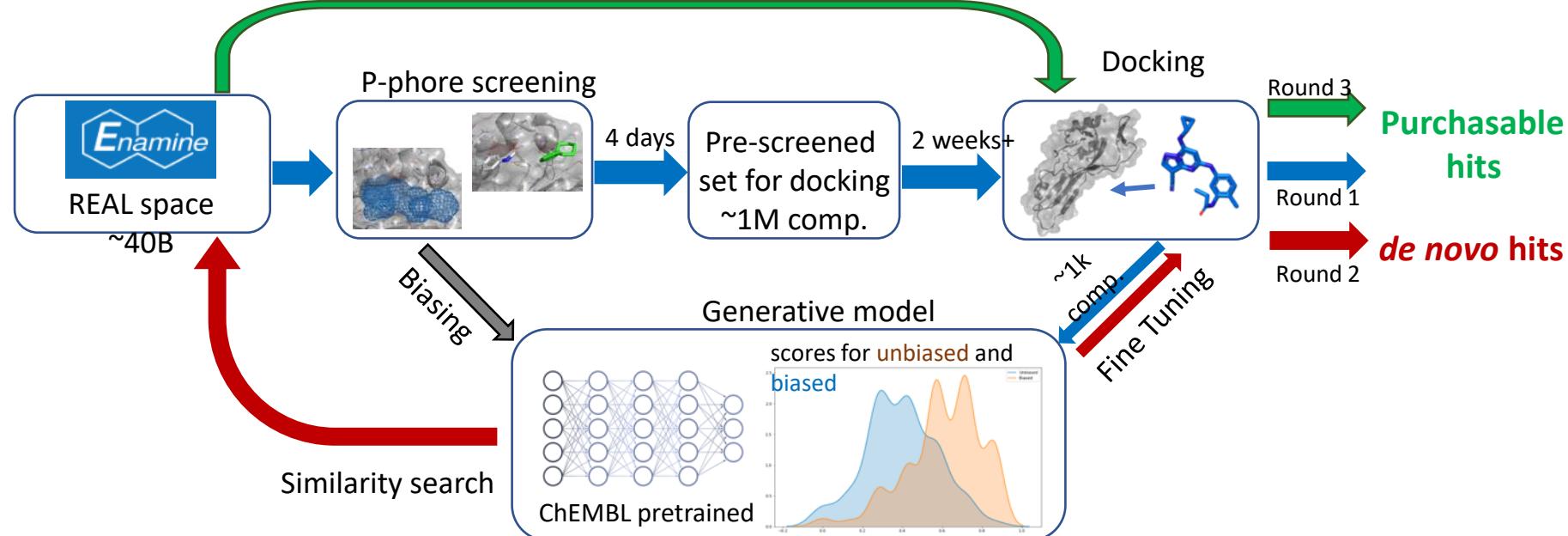
## Pharmacophore identification

### Score points:

- 1 point: Has only a halide
- 1 point: Has only a ring
- 1 point: Has only a carbonyl
- 2 points: Has a halide and a ring, unconnected
- 2 points: Has a carbonyl and a ring, unconnected
- 2 points: Has a halide and a carbonyl, unconnected
- 3 points: Has a halide, a carbonyl, and a ring, unconnected
- 3 points: Has a halide and ring connected correctly
- 3 points: Has a carbonyl and a ring connected correctly
- 4 points: Has a halide and a ring connected correctly, and a carbonyl
- 4 points: Has a carbonyl and a ring connected correctly, and a halide
- 5 points: Has full pharmacophore match



# Virtual Screening Workflow

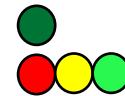
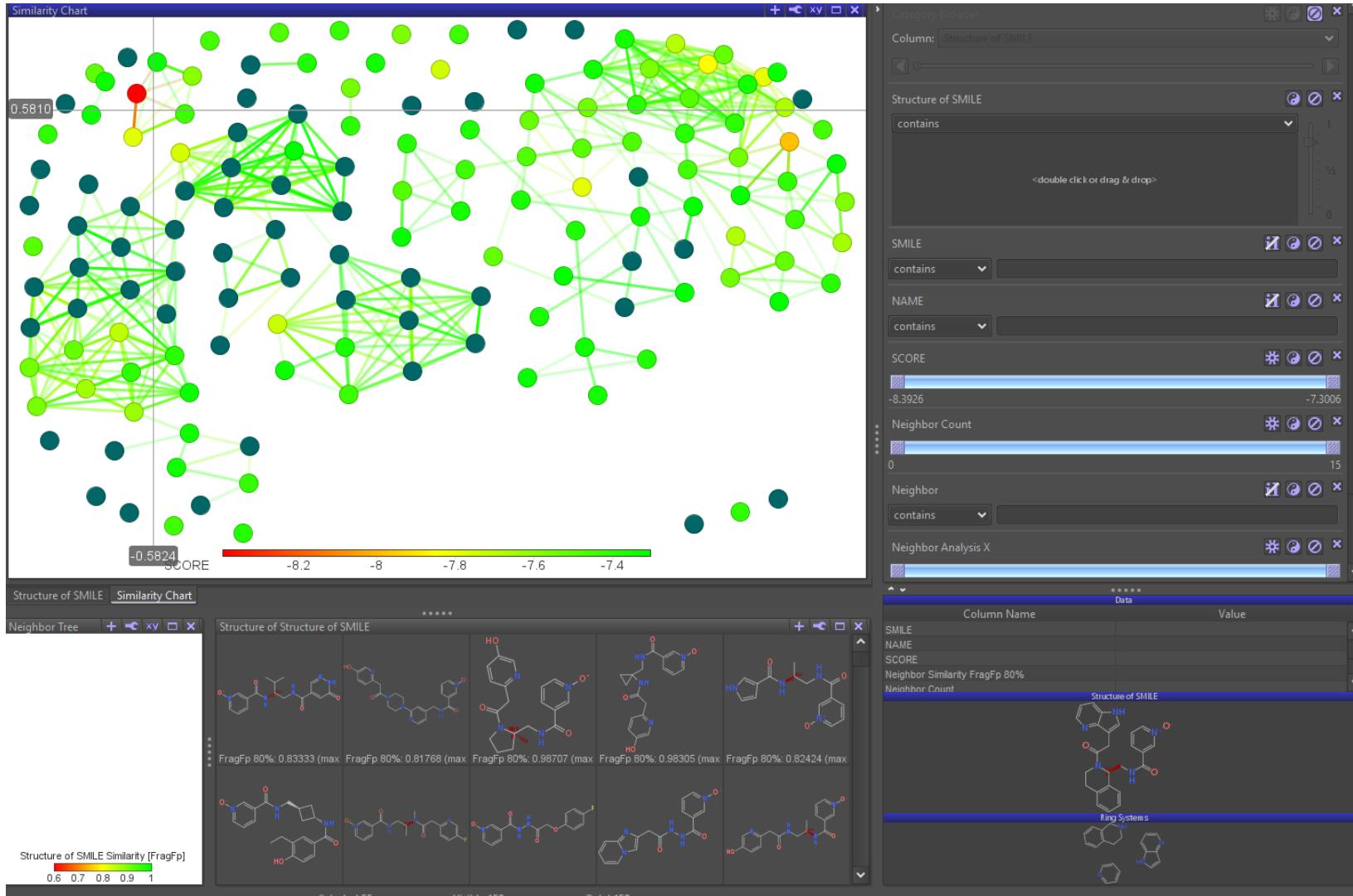


Pharmacophore screening of 40 Billion analogs  
Each cycle seeds the next generative model  
Iterative improvement in docking scores

# Enamine REAL Space (43 Billion Library) Top 150 List



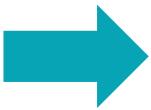
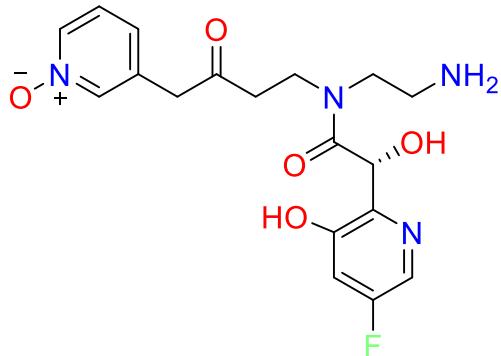
DataWarrior zip files can be found on GitHub: <https://github.com/StructuralGenomicsConsortium/CNP4-Nsp13-C-terminus-B/issues/23#issuecomment-1181839847>



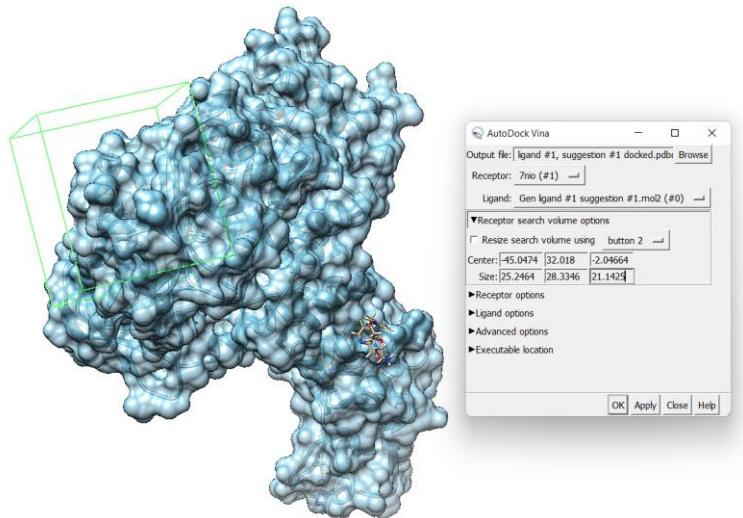
Enamine MOD quote provided.  
Glide score (lowest to highest).

# Generated #1: Following Suggested Modifications

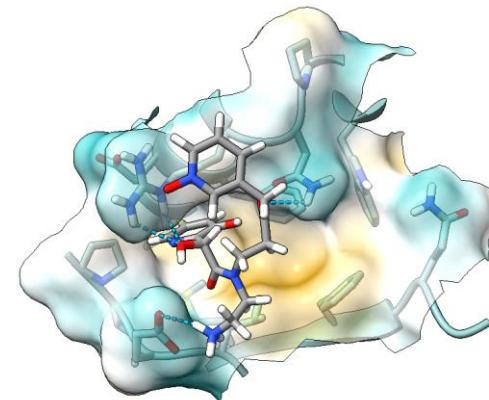
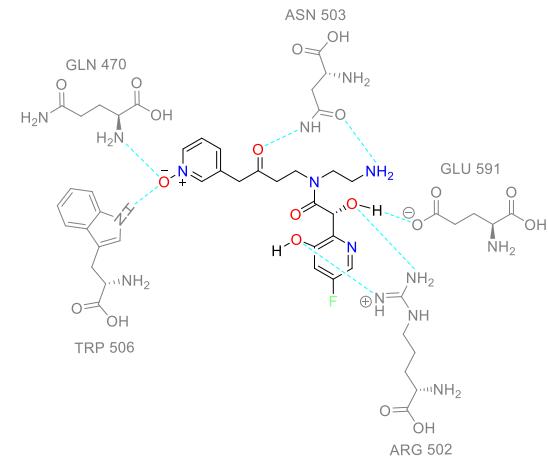
Generative Machine  
Learning  
Suggestions  
(Kostya, Glide)



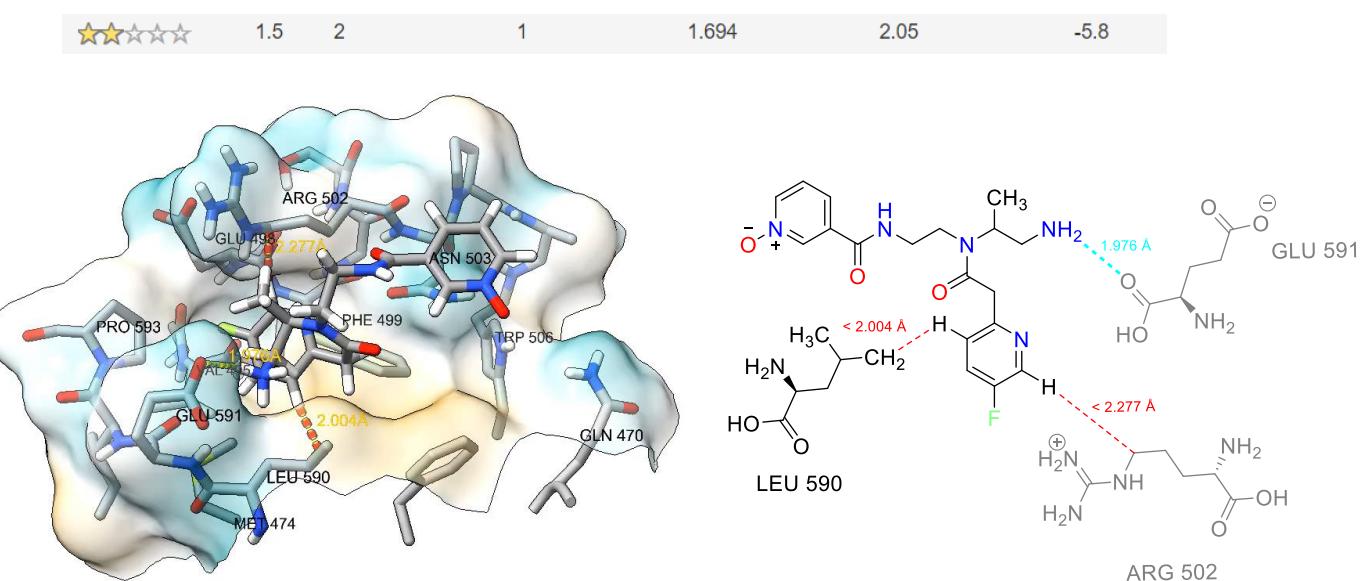
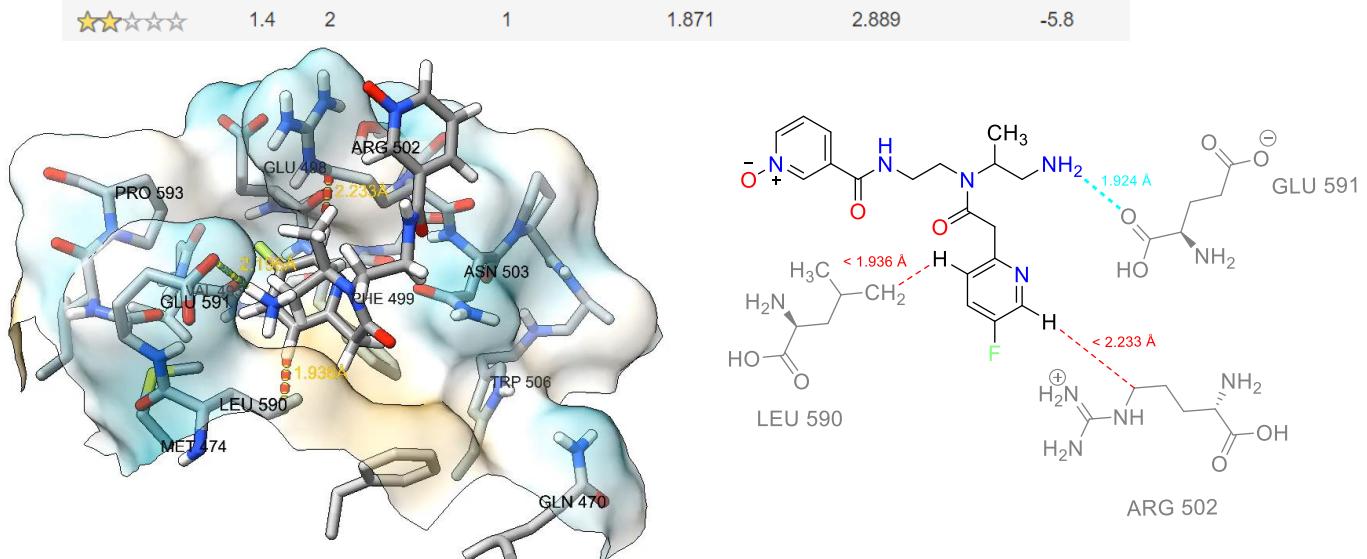
Molecular Docking  
(AutoDock Vina)



2D / 3D SAR Analysis  
(ChimeraX)



# Generated #1: AutoDock Vina Scores at C-terminus-B Pocket



RATING	ID	CLASHES	HBONDS	RMSD L.B.	RMSD U.B.	SCORE
★★★★★	1.1	2	1	0	0	-6.5
★★★★☆	1.2	2	0	3.562	5.003	-6.1
★★★★☆	1.3	2	3	3.489	4.648	-5.9
★★★★☆	1.4	2	1	1.871	2.889	-5.8
★★★★☆	1.5	2	1	1.694	2.05	-5.8

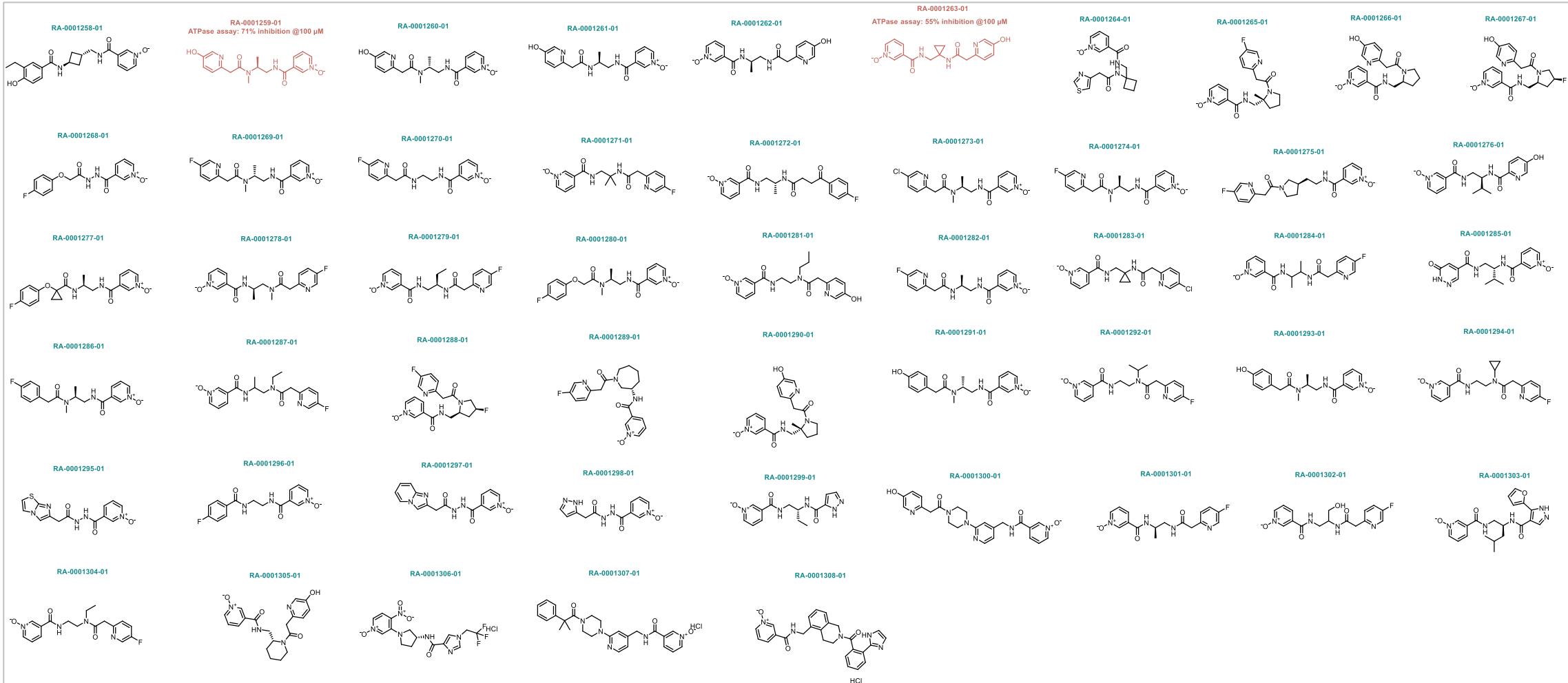
## Conclusions:

- Terminal -NH<sub>2</sub> forms H-bond interactions with GLU591 and ASN503.
- Steric clashes** with 3-fluoropyridine are common with ARG502, MET474, and LEU590 – **consider reducing ring size to improve binding**.
- CH<sub>3</sub> and Ar-H clashes with ARG502 – H-bond acceptor in this position should increase binding.
- Elongated conformer allows up to 3x H-bonds (2 with pyridine 1-oxide and GLN470/TRP506); maintain interaction by reducing 3-fluoropyridine ring size?
- Aromatic ring may allow pi-stacking with PHE499.
- GLU591 and ARG502 form a H-bond ‘gateway’ to the pocket (see slide 12)** – exploiting ring planarity with stereogenic centre might increase binding significantly!

# Enamine REAL Space-Trained 43 Billion Library, Top 150 List:

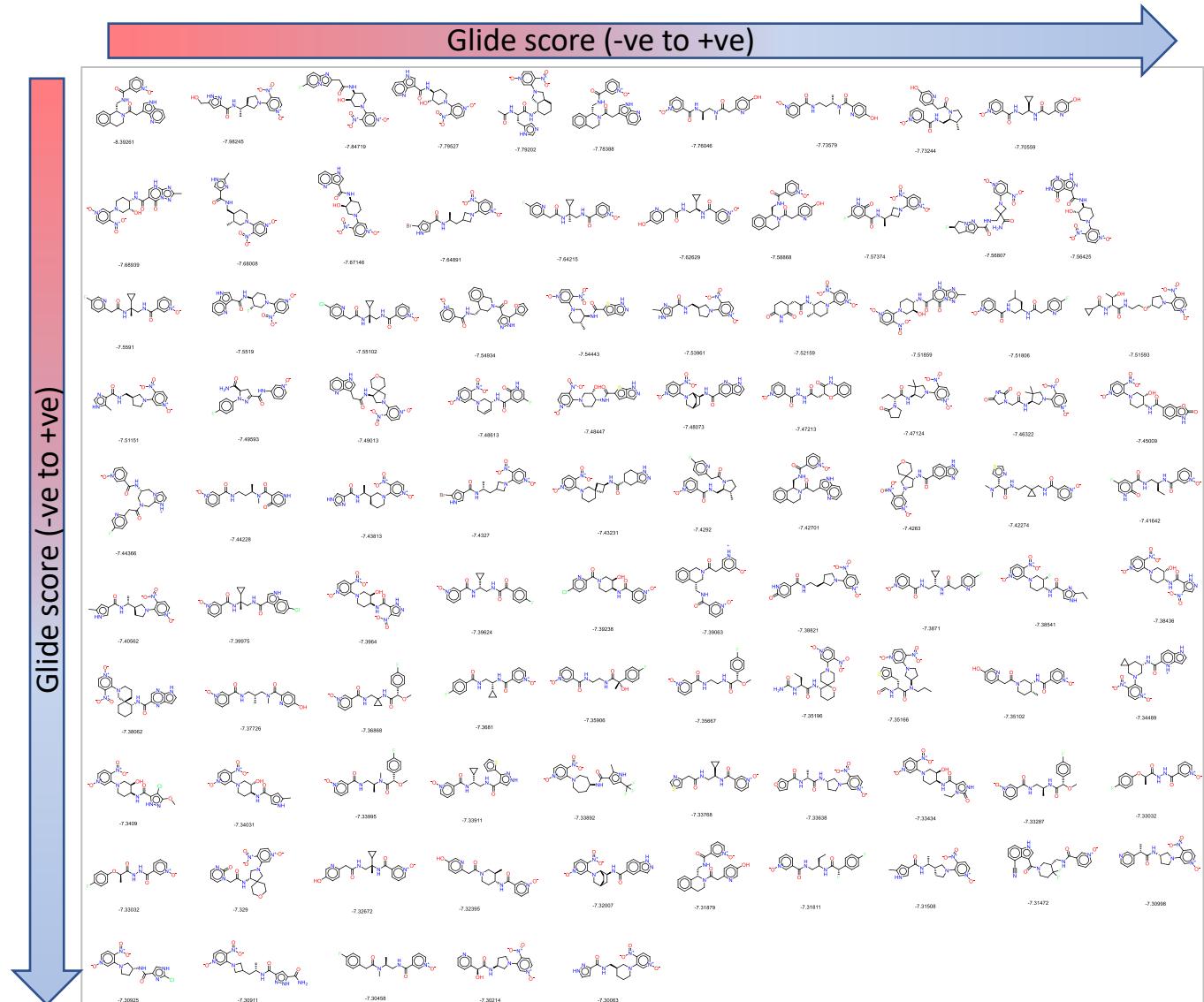


- 55 / 150 Enamine provided quote.
- **51 / 150 (1/3)** compounds shipped from Enamine (Anwar).
- **2 compounds of interest** based on ATPase assay (Sumera).

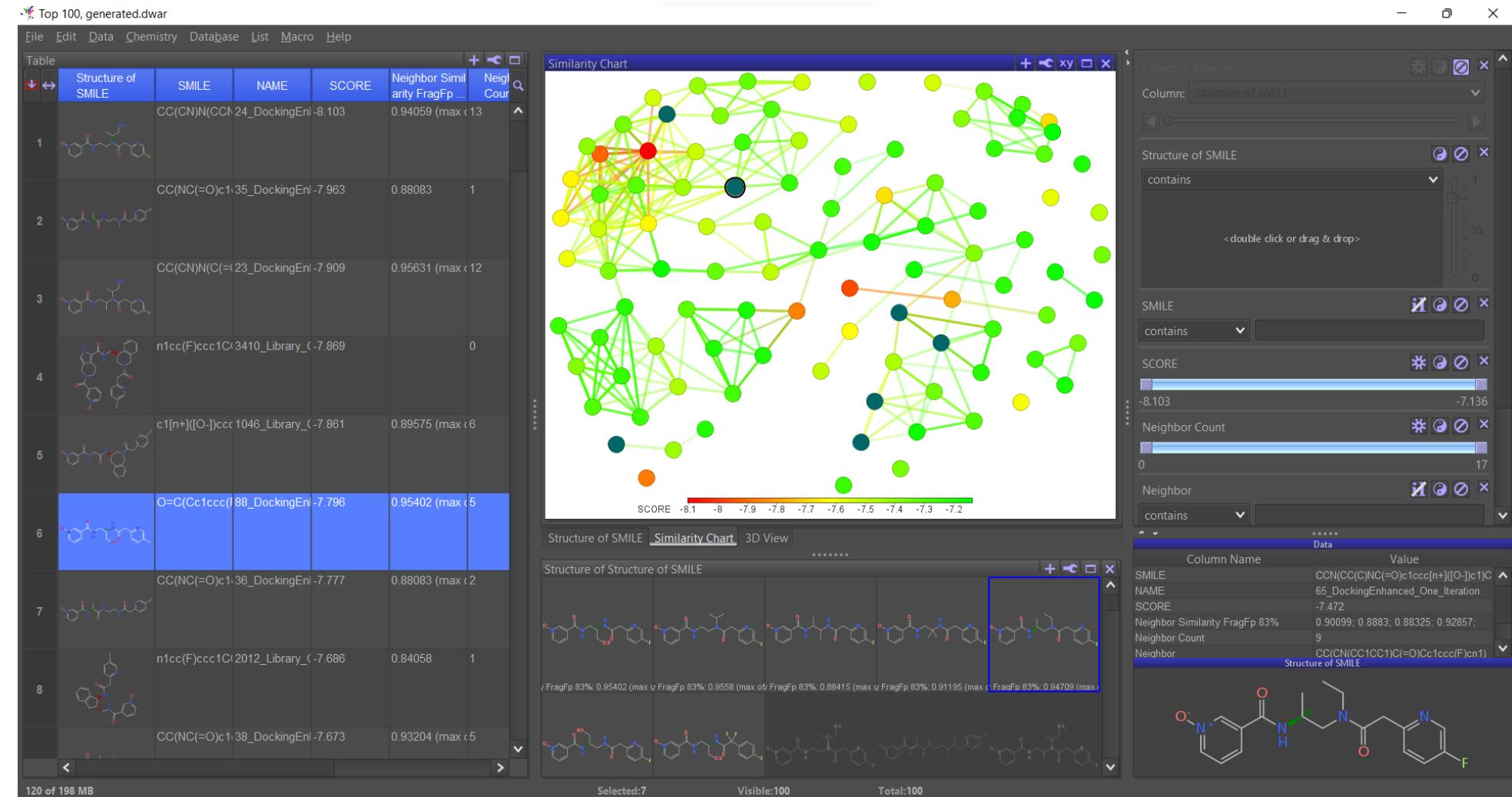


## Enamine REAL space-trained 43 billion library: Top 150 compound list

- 95 / 150 (1/3) compounds remaining to be made...

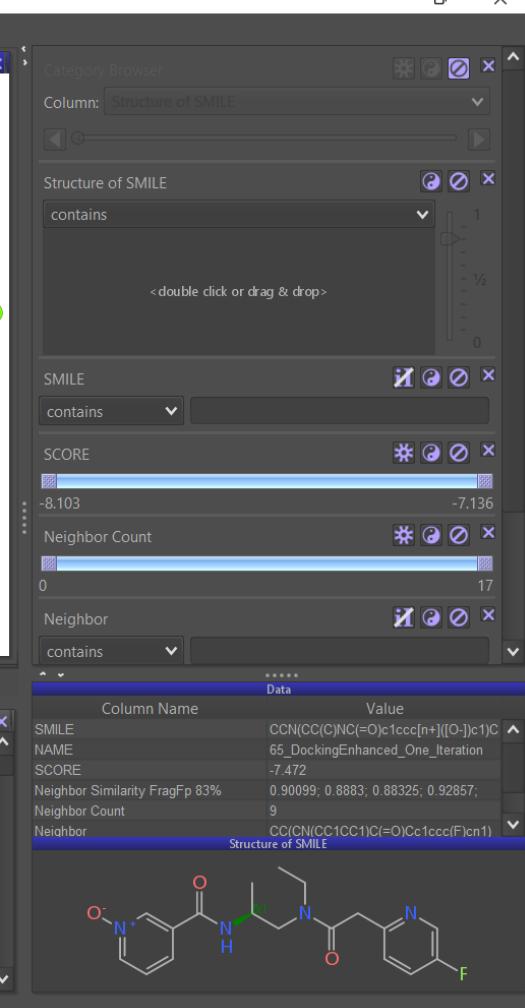


# De-novo generated compounds list: Glide top 100-scoring compounds



● Enamine MOD

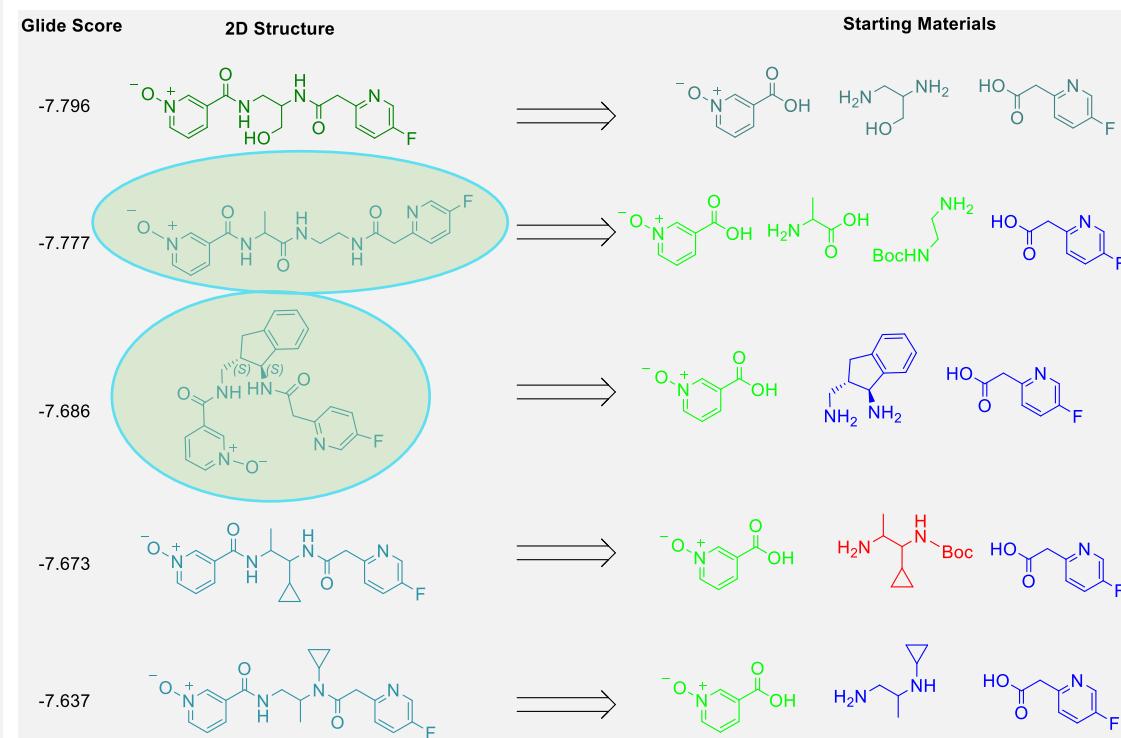
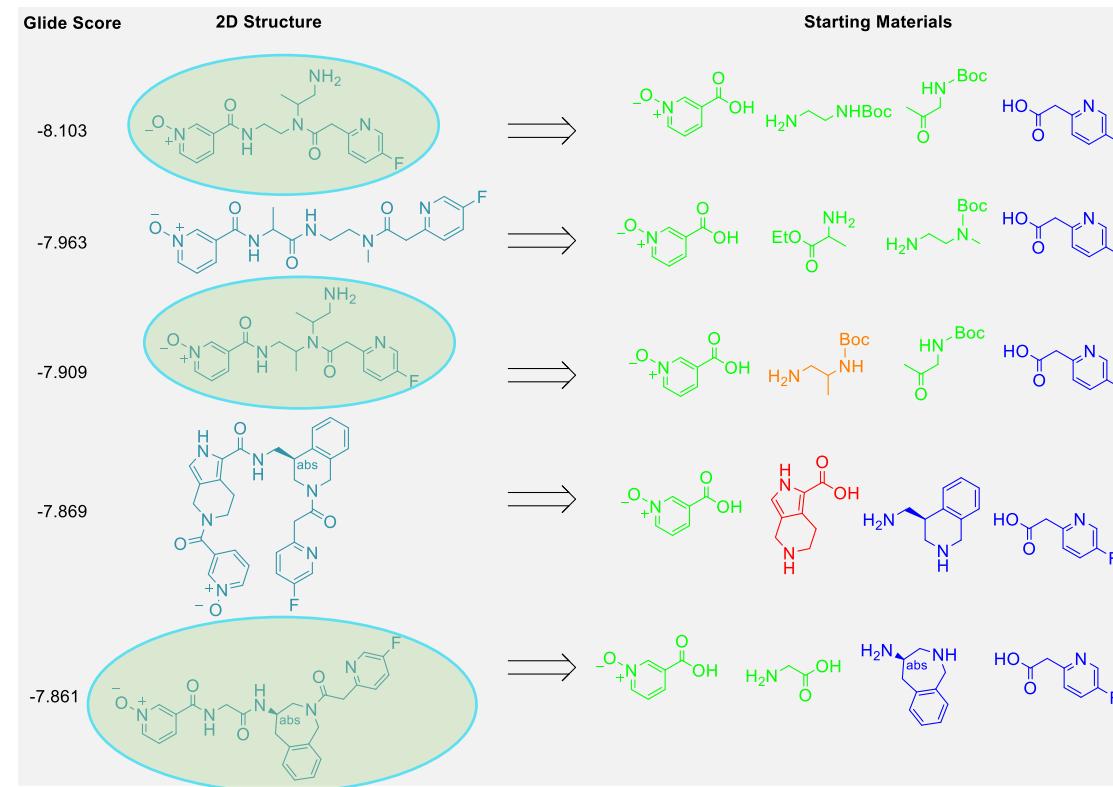
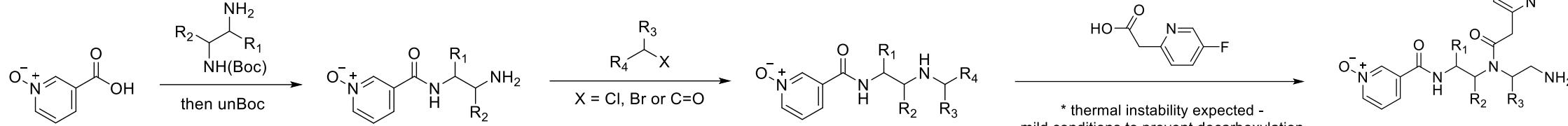
● ● ● Glide score (lowest to highest)



# De-Novo Generative Compounds: Top 10-scoring (Glide)

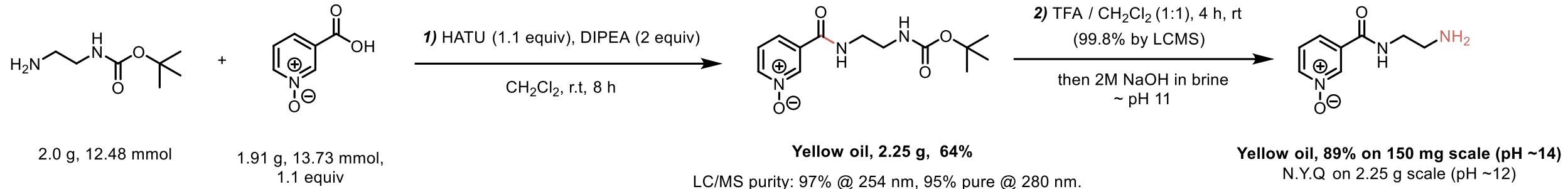


Purchasable vs unavailable (to-make) building blocks ...

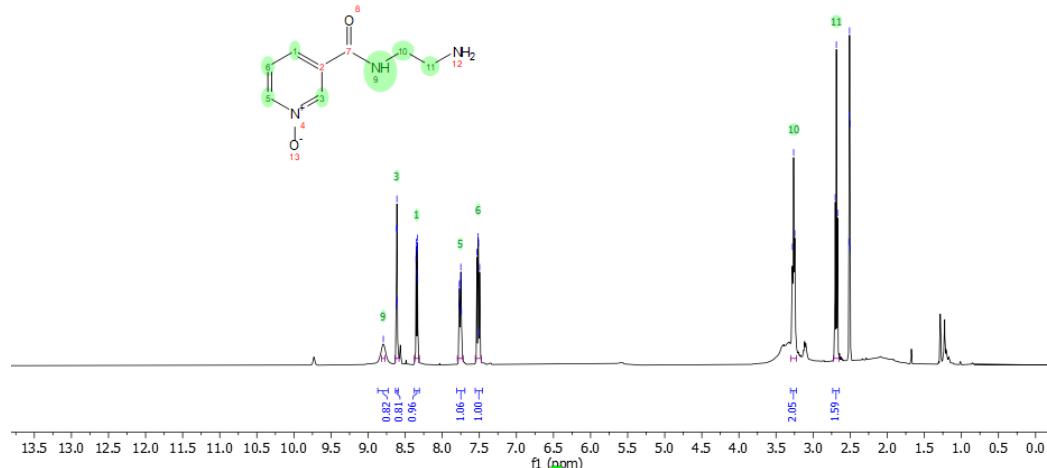


- Purchased
- Purchasable
- To make
- Unavailable
- Enamine MOD
- Current targets

# Chemistry Update:

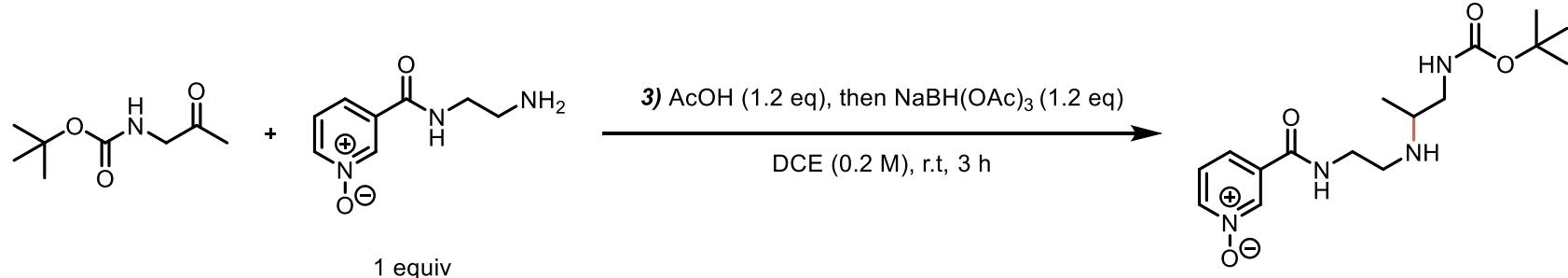


<sup>1</sup>H NMR (400 MHz, DMSO) δ 8.79 (s, 1H), 8.61 (q, *J* = 1.5 Hz, 1H), 8.34 (ddd, *J* = 6.5, 1.8, 1.0 Hz, 1H), 7.80 – 7.72 (m, 1H), 7.51 (dd, *J* = 8.0, 6.4 Hz, 1H), 3.25 (d, *J* = 6.5 Hz, 2H), 2.69 (t, *J* = 6.4 Hz, 2H).

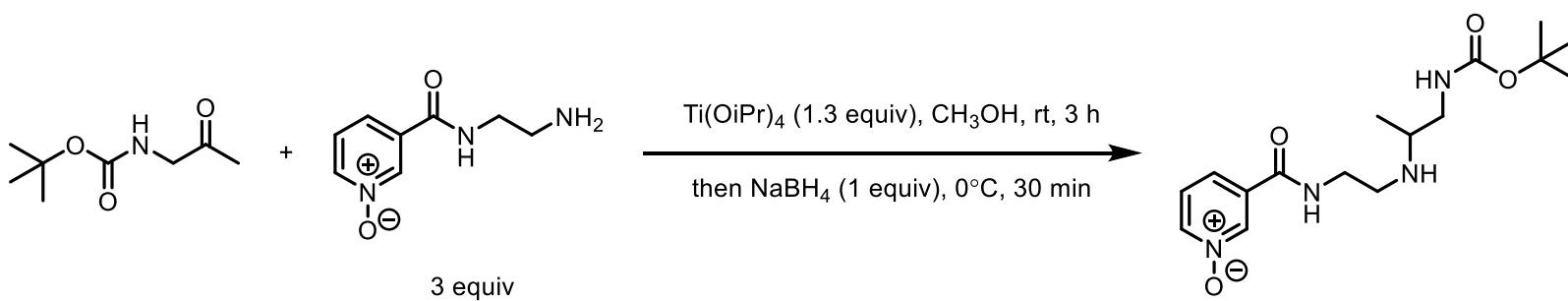


# Chemistry Update:

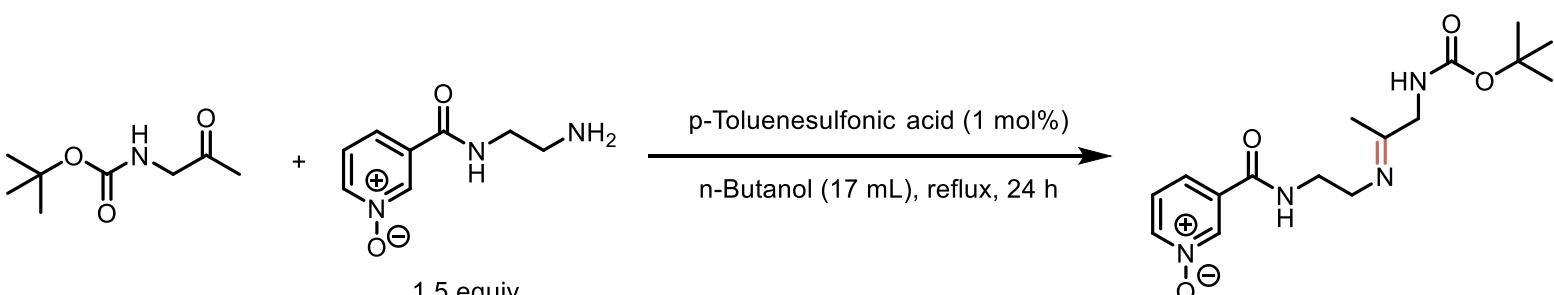
Next...



- US20090227560 A1 2009-09-10



- Eur. J. Org. Chem. 2020, 2745–2753. doi.org/10.1002/ejoc.202000322.



- Peng, Dongjie et al, Journal of the American Chemical Society (2013), 135(51), 19154-19166.

Amine solubility:

- ✓ MeOH, iPrOH (polar)
- ✗ DCE, DCM (non-polar)

No conversion

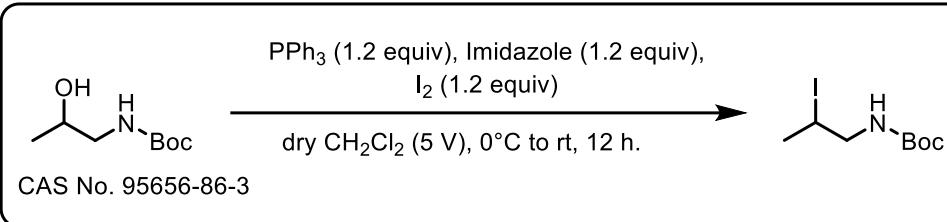
Issues:

- Chelation to Ti<sup>4+</sup> LA?
- Reduction of N-oxide.

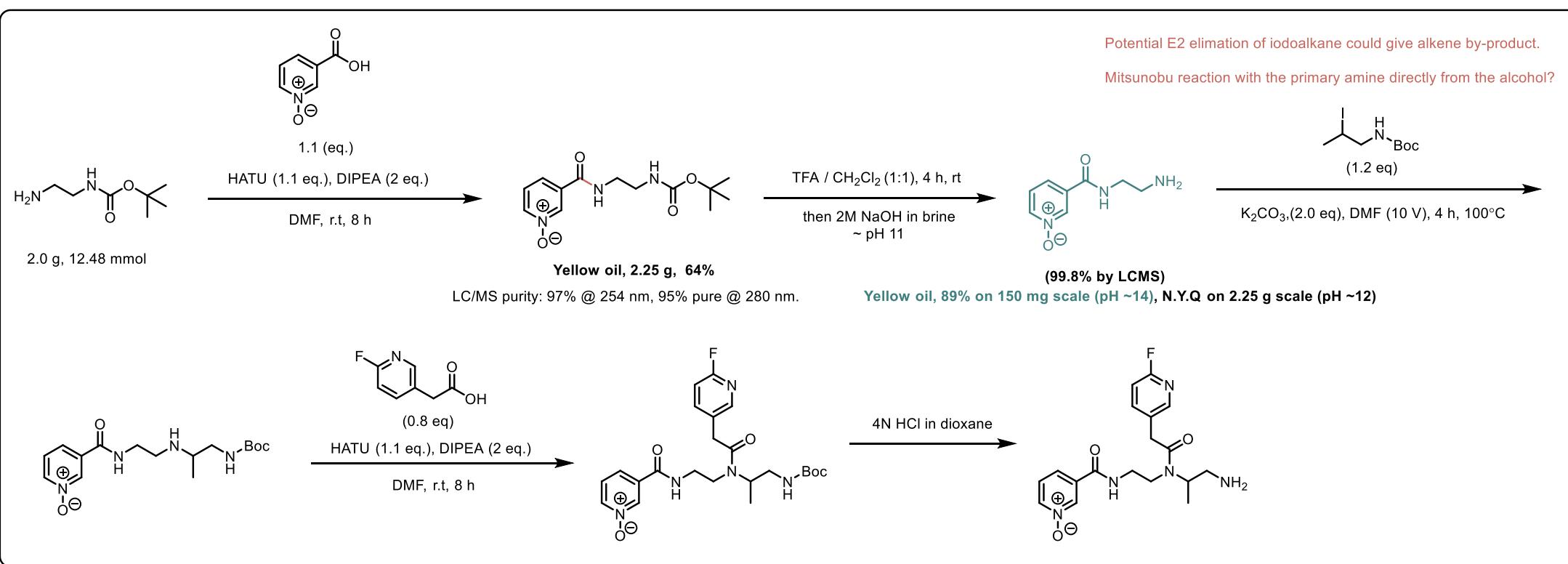
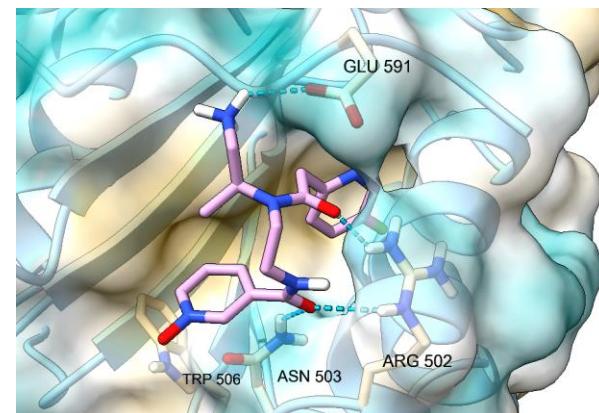
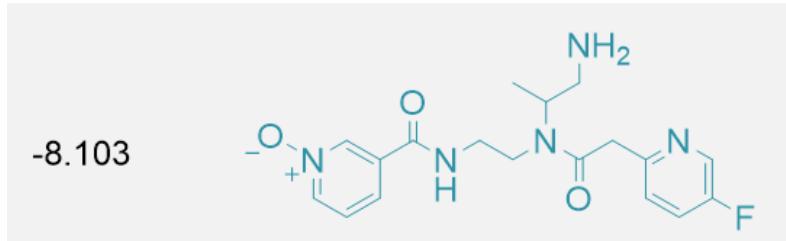
To attempt:

- Isolation of imine

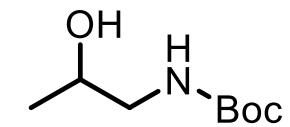
# De-Novo Generative Compounds: Top 10 Scoring (Glide)



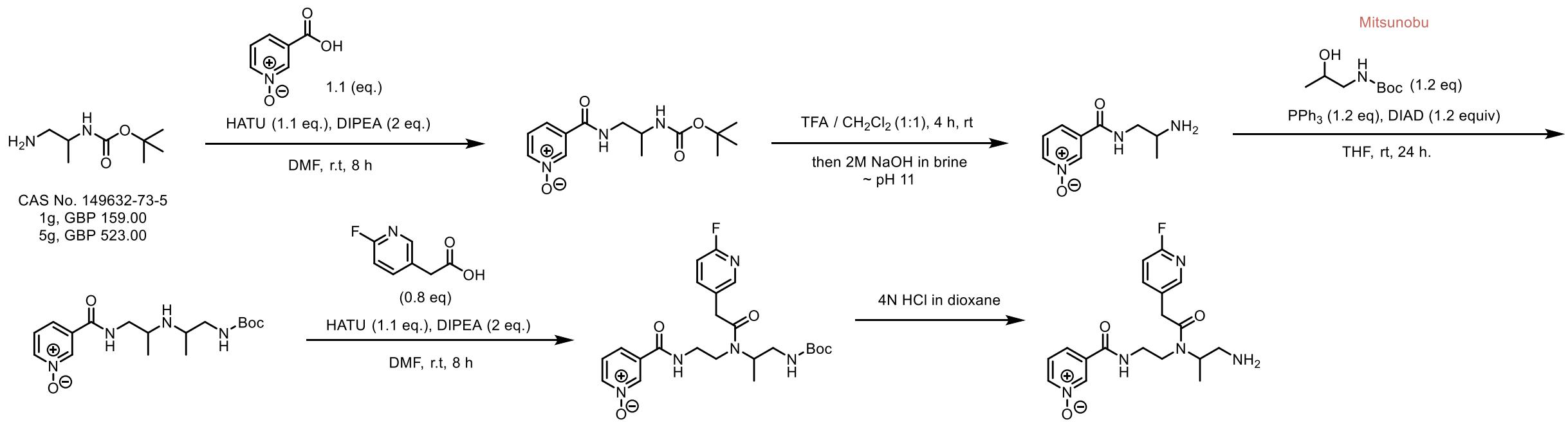
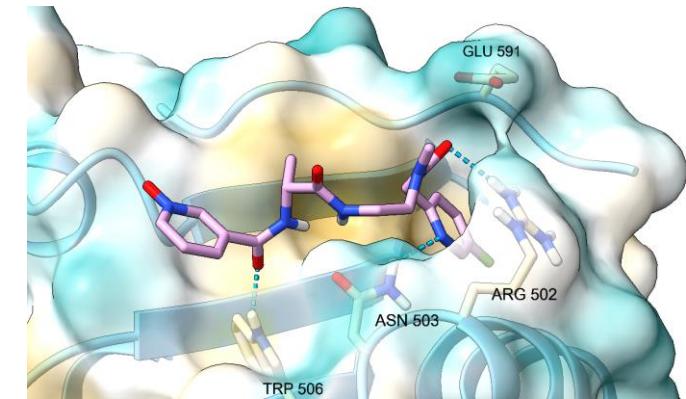
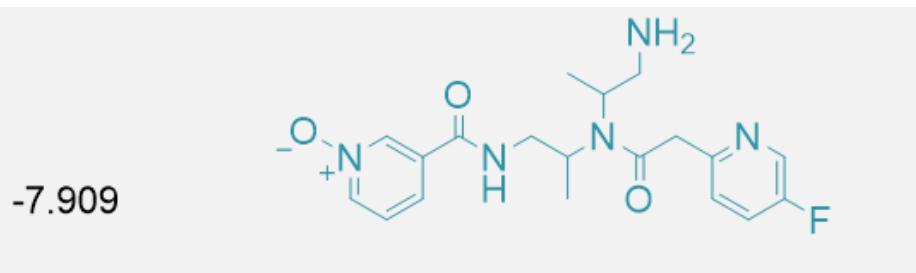
CAS Reaction Number: 31-614-CAS-24837821  
 Angewandte Chemie, International Edition (2021), 60(52), 27070-27077.



# De-Novo Generative Compounds: Top 10 Scoring (Glide)



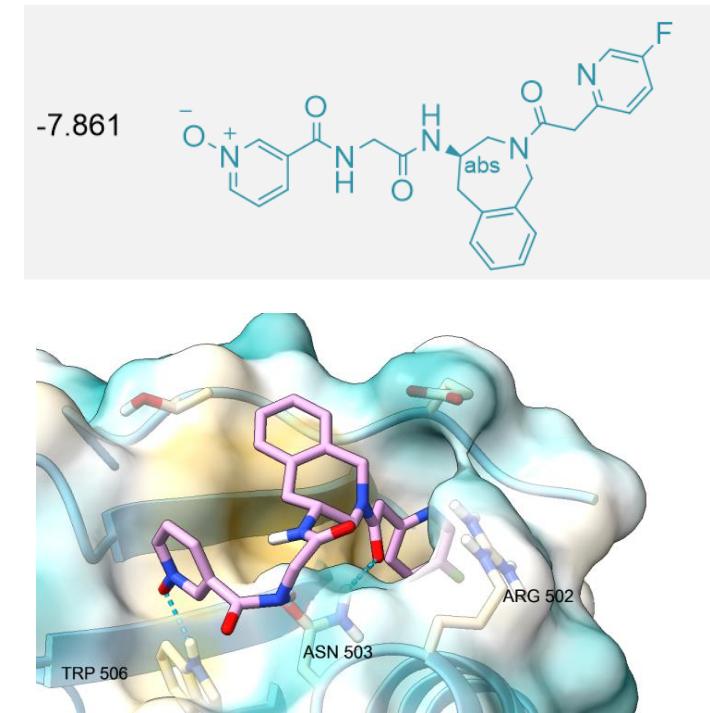
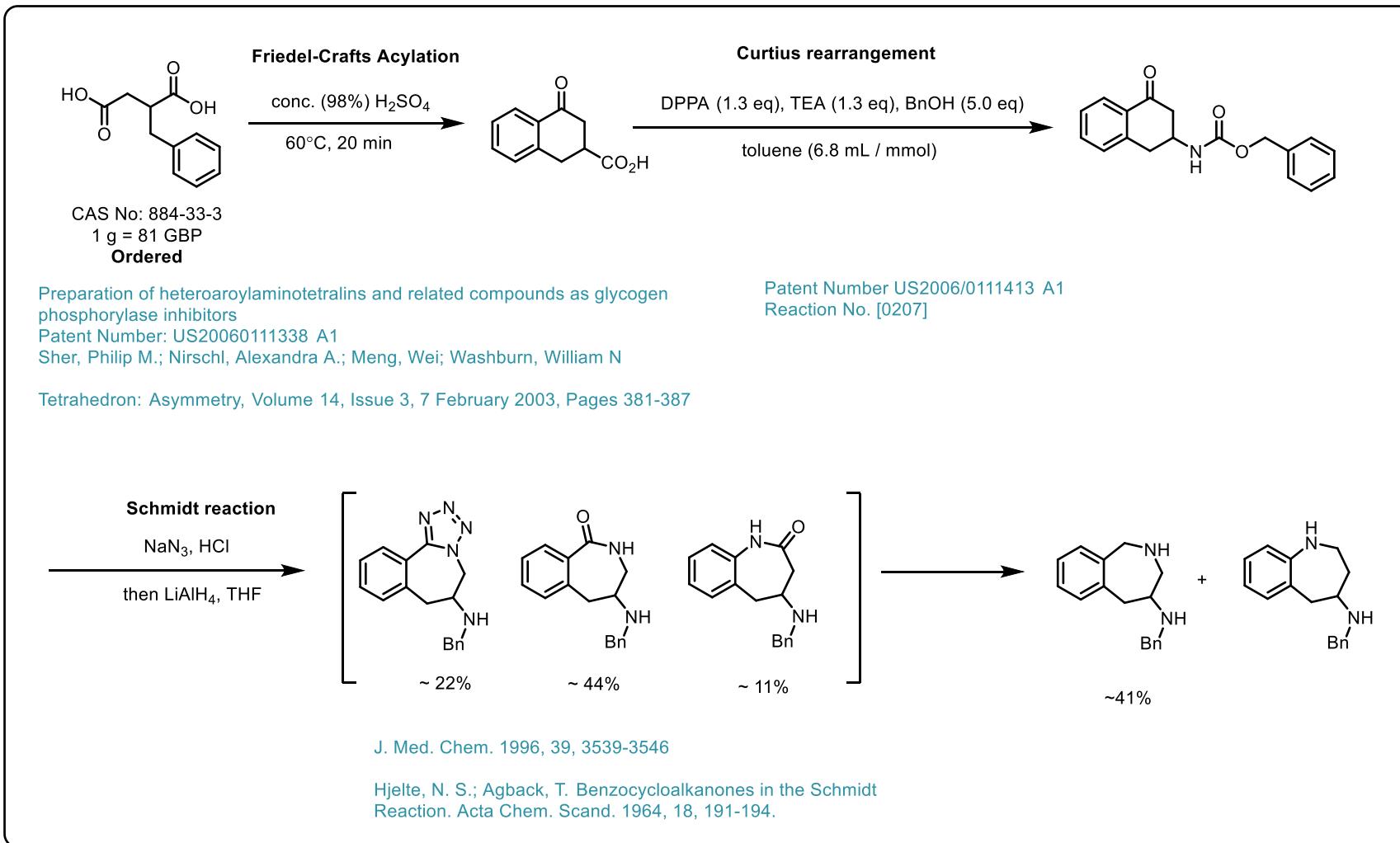
CAS No. 95656-86-3



# De-Novo Generative Compounds: Top 10 Scoring (Glide)



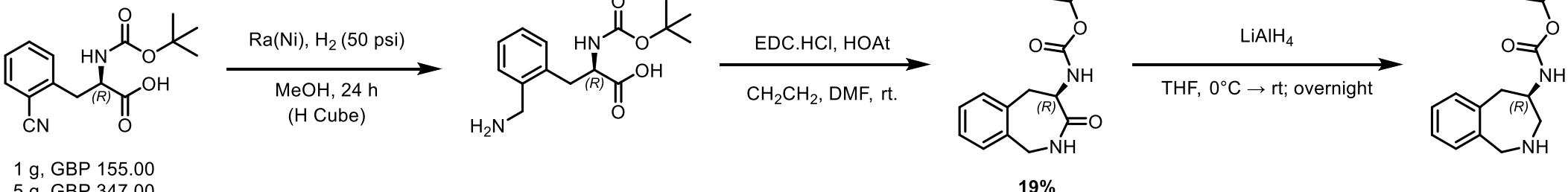
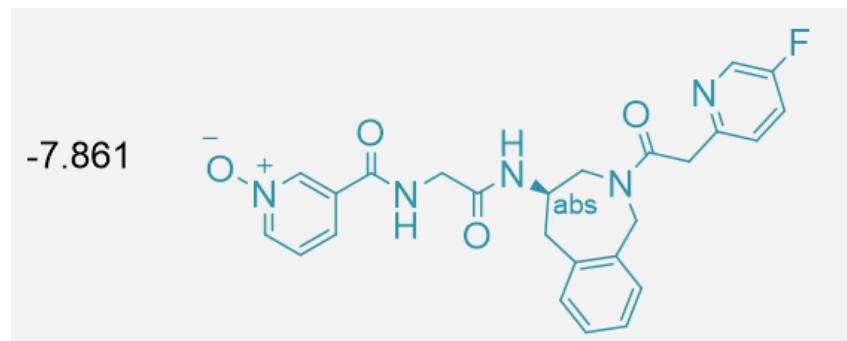
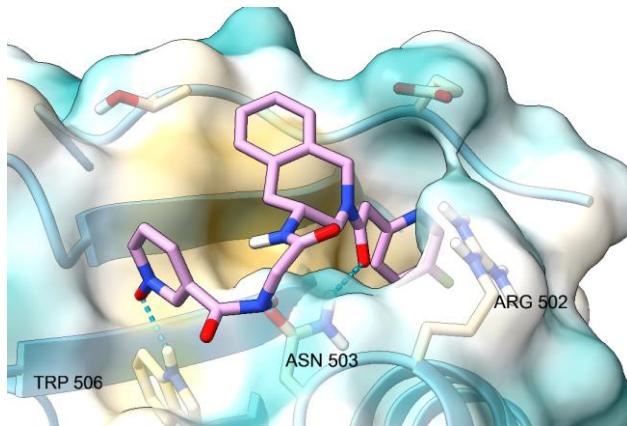
- Suggestion from James...



# De-Novo Generative Compounds: Top 10 Scoring (Glide)



- Alternative route...

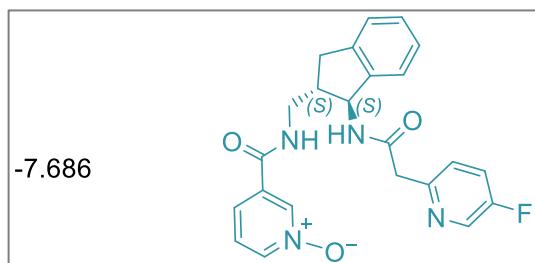
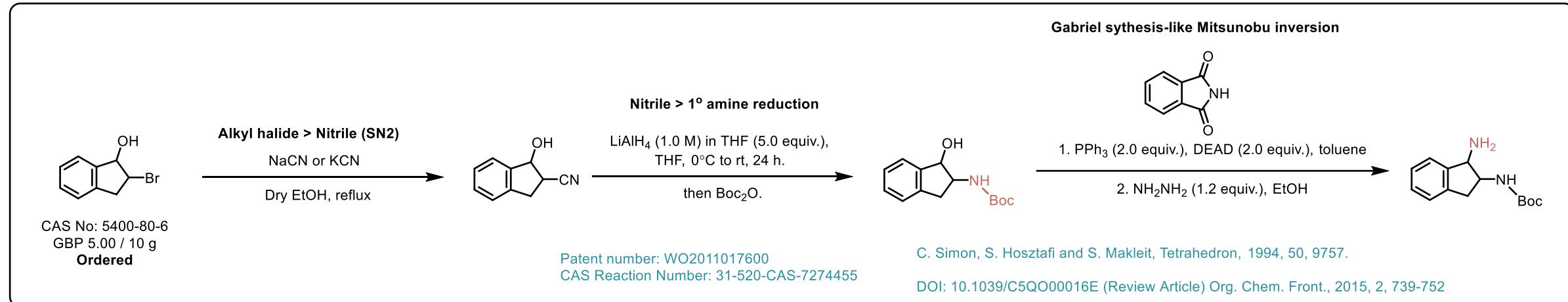


Simonin, Frederic; et al. World Intellectual Property Organization, WO2019170919 A1 2019-09-12

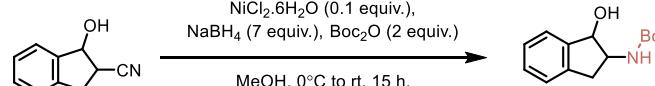
Zawodny, Wojciech; et al. Journal of the American Chemical Society (2018), 140(51), 17872-17877

Shonberg, Jeremy; et al. Journal of Medicinal Chemistry (2015), 58(13), 5287-5307.

# De-Novo Generative Compounds: Top 10 Scoring (Glide)

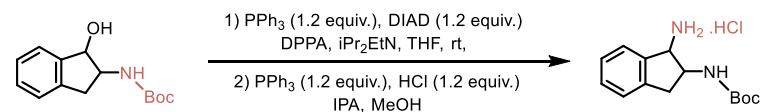


**Alternate route: Ni-catalysed reduction ( $\text{CN} > \text{NH}_2$ )**

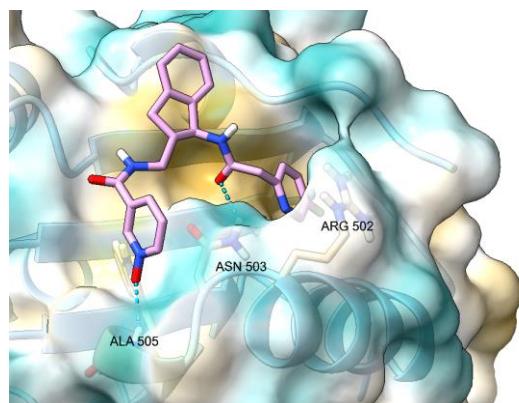


A generic approach for the catalytic reduction of nitriles  
 S. Caddick et al. / Tetrahedron 59 (2003) 5417–5423.

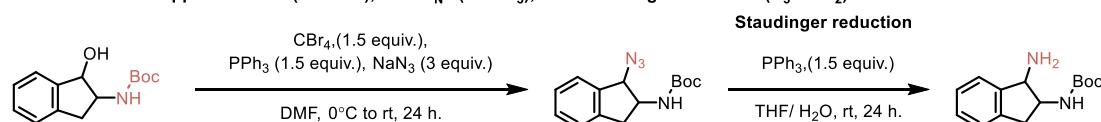
**Alternate route: One-Pot Mitsunobu inversion then Staudinger reduction**



Mitsunobu Inversion of a Secondary Alcohol with Diphenylphosphoryl azide.  
 Application to the Enantioselective Multikilogram Synthesis of a HCV Polymerase Inhibitor  
[dx.doi.org/10.1021/op20002u](https://dx.doi.org/10.1021/op20002u) , Org. Process Res. Dev. 2011, 15, 1116–1123



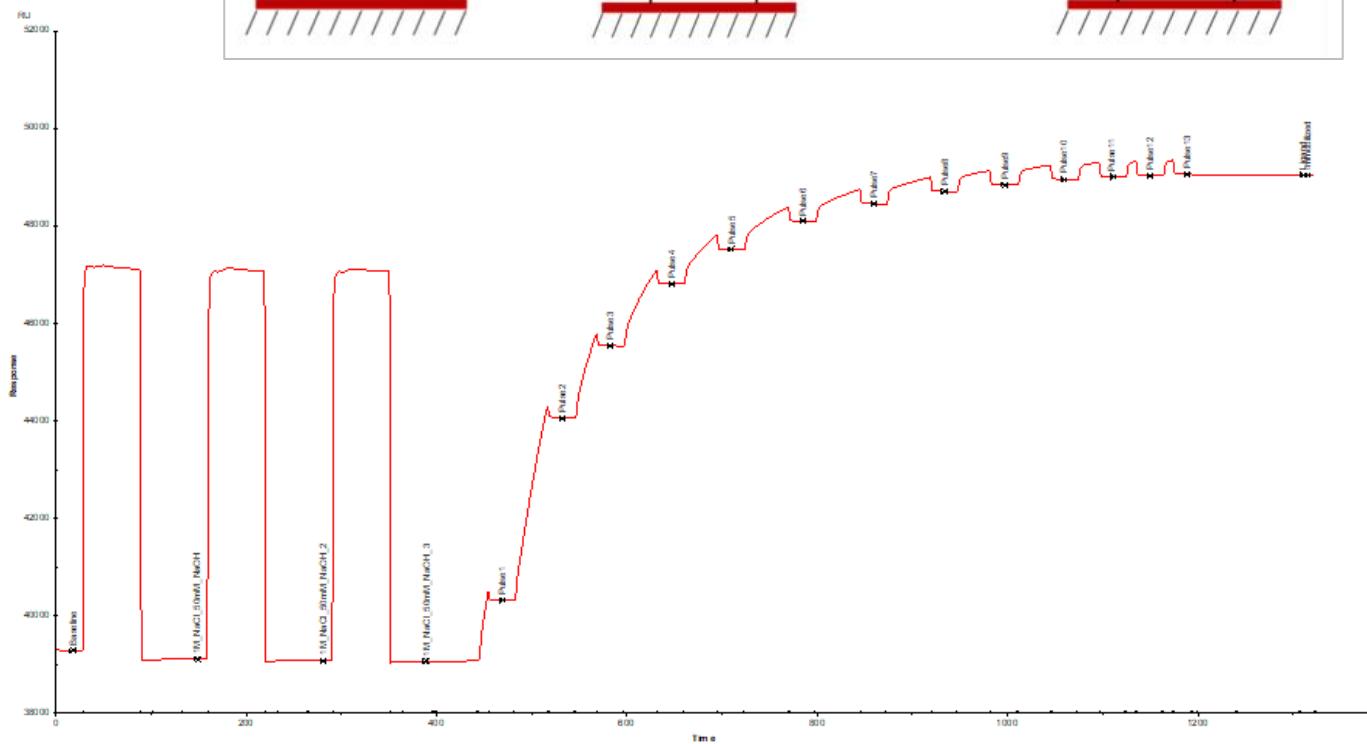
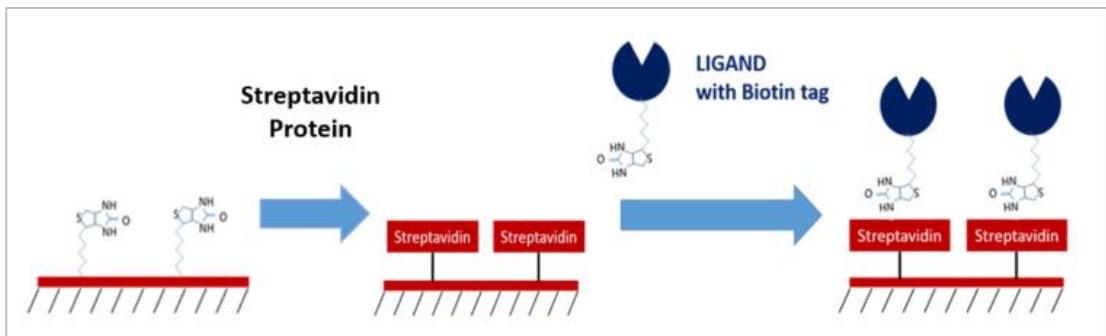
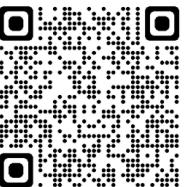
**Alternate route: Appel Reaction ( $\text{OH} > \text{Br}$ ), then  $\text{S}_\text{N}2$  ( $\text{Br} > \text{N}_3$ ), then Staudinger reduction ( $\text{N}_3 > \text{NH}_2$ )**



N-(3-Acyloxy-2-benzylpropyl)-N'-(4-(methylsulfonylamino)benzyl)thiourea Analogues: Novel Potent and High Affinity Antagonists and Partial Agonists of the Vanilloid Receptor  
 J. Med. Chem. 2003, 46, 3116-3126

# SPR Assay Progress: Biotin-Streptavidin Immobilisation protocol

SPR Assay - Getting up and running · Issue #25 · StructuralGenomicsConsortium/CNP4-Nsp13-C-terminus-B (github.com)



## Summary:

- Final response was 9756.1 RU using 5 mM protein in HBS-EP+ buffer.
- Biotin capture: Flow channel 1: HBS-EP+ running buffer (reference.) Flow channel 2: Nsp13 (5 mM) in HBS-EP+ buffer.

## Biotinylated-Nsp13 / SA Chip (Cytiva) Immobilisation protocol:

*The Biacore SA capture wizard protocol was used, briefly:*

### 1. Immobilisation buffer used:

1x HBS-EP+ : (Cytiva 20x HBS-EP) 50 mL diluted to normal concentration with 950 mL de-ionised water, then 450 µL of P20 surfactant was added).

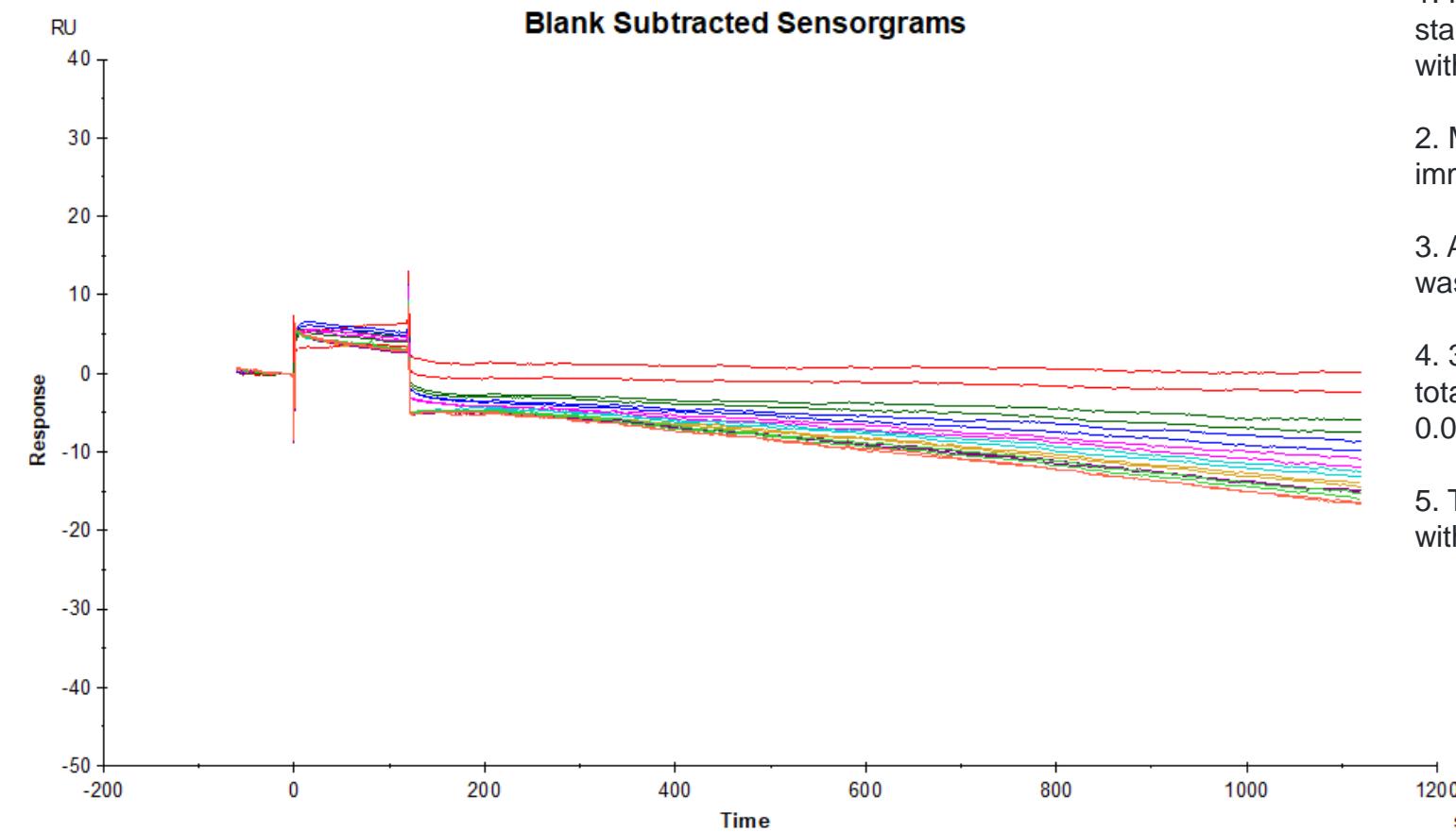
### 2. To condition the SA chip surface for capture:

Washed flow channels with 3x NaOH (50 mM) / NaCl (1.0 M).

### 3. To immobilise the protein:

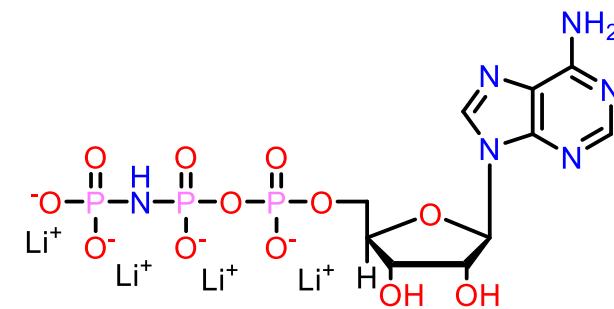
Set target response level to 10'000 RU, then flow biotinylated-nsp13 (0.05 mg/ml) in HBS-EP+ buffer through flow channel 2 until saturation is reached or until the ligand runs out (9756.1 RU reached).

# SPR Assay Progress: Kinetics with positive control compounds (AMP-PNP)



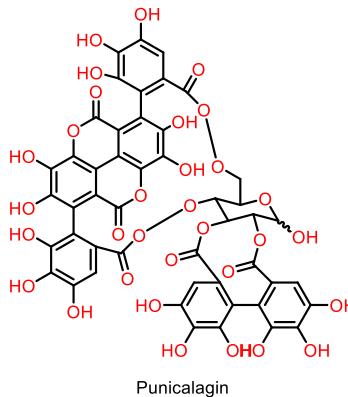
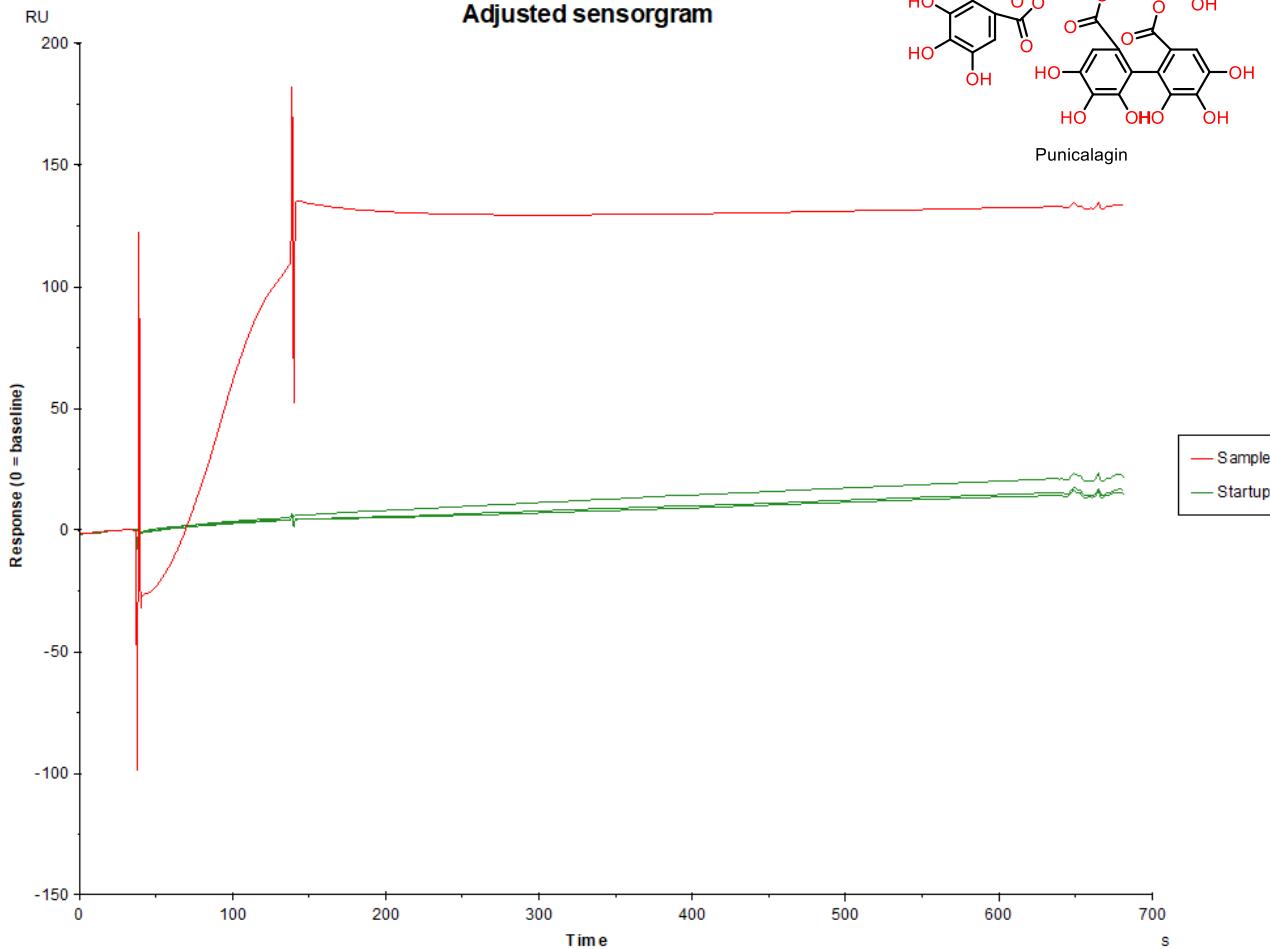
## Kinetic screen protocol (AMP-PNP):

1. PBS-P+ buffer, 10x (Cytiva) was diluted ten times to form 1 litre of standard concentration solution, 20 mL was then removed and replaced with an equal volume of DMSO to make PBS-P+ with 2% DMSO.
2. MgCl<sub>2</sub> (476 mg) was dissolved in the stock solution to make the final immobilisation buffer; PBS-P+ (2% DMSO, 5 mM MgCl<sub>2</sub>).
3. AMP-PNP (CAS No: 25612-73-1, Sigma Aldrich,>93% HPLC, 2.4 mg) was dissolved in 600 µL PBS-P+ buffer with 2% DMSO.
4. 300 µL of the stock solution was taken for 2x serial dilution (10 aliquots total: 8.0 mM, 4.0 mM, 2.0 mM, 1.0 mM, 0.50 mM, 0.25 mM, 0.125 mM, 0.0625 mM, 0.03125 mM, 0.0156 mM).
5. The kinetic screen was run overnight (8 h) and the rack was covered with a film top to prevent solvent evaporation.



AMP-PNP

# SPR Assay Progress: Kinetics with positive control compounds (Punicalagin, PUG)



## Kinetic screen protocol (PUG):

- The positive control compound Punicalagin (PUG, KD = 21.6 nm) was flowed over the captured nsp13 protein to confirm it as a nanomolar binder to Nsp13; <https://www.sciencedirect.com/science/article/pii/S0166354222001589?via%3Dihub>
  - Briefly, PUG [CAS-No: 65995-63-3], (1.5 mg) was diluted in DMSO (1 mL) to give a 1.3828 mM stock solution, which was then diluted to 216 nM (10x KD) by dissolving 1.5 uL up to 10 mL with phosphate-buffered saline (PBS, [https://info.cytivalfesciences.com/Biacore-help-files\\_Buffers-for-small-molecule-screening.html](https://info.cytivalfesciences.com/Biacore-help-files_Buffers-for-small-molecule-screening.html)) with 2% DMSO.
  - PUG does show nM binding affinity, but binding is irreversible. Spikes due to solvent compatibility (DMSO), but sensorgram is consistent with that previously reported.
  - Protein regeneration conditions are required if PUG is to be used as a positive control for SPR.
  - Attempted regeneration conditions (Glycine.HCl @ pH 2.5) provided by Professor Zhonghui Lin @ Fuzhou University, but currently trying to interpret the assay.
  - Sumera has suggested that the inclusion of MgCl<sub>2</sub> in the running buffer avoids the need for regeneration conditions as PUG is released (60 s capture, 100 seconds desorb); currently leasing to get exact assay conditions.