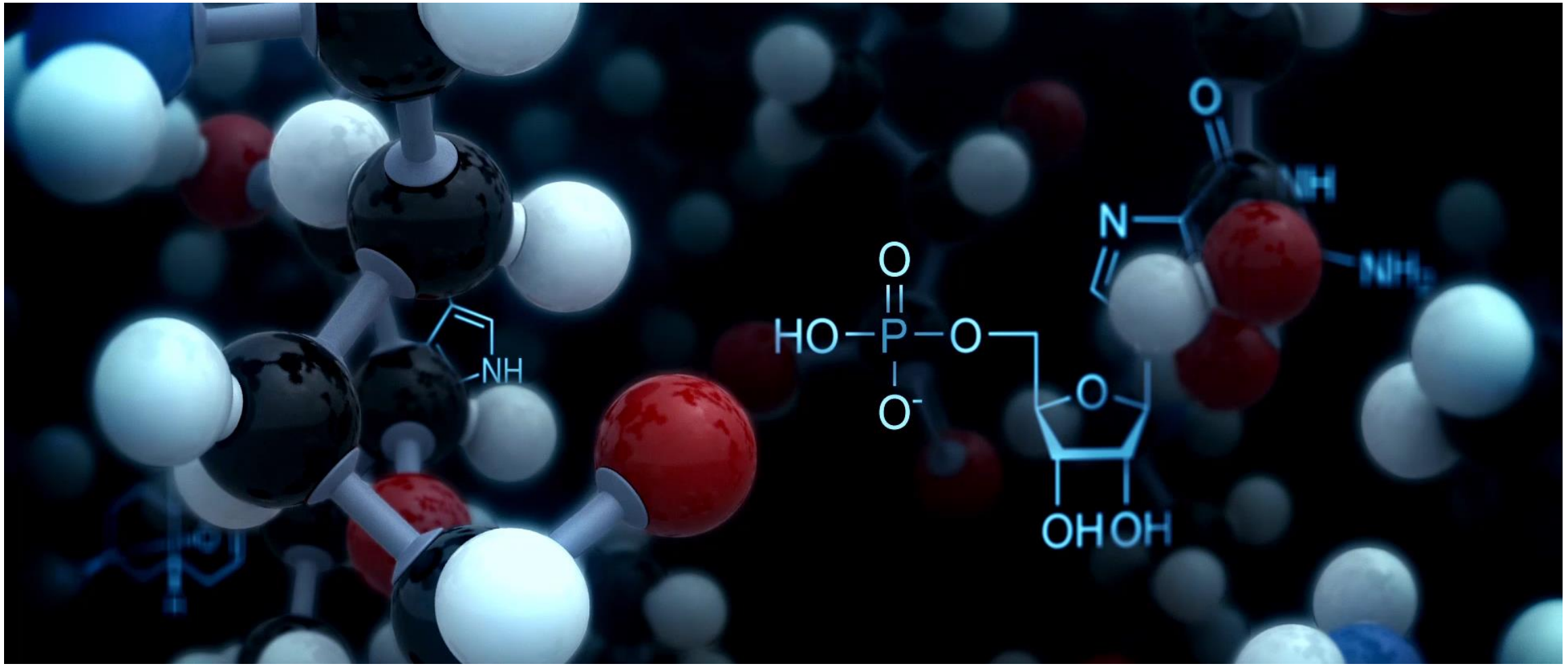
Ai in Drug Discovery



AI in Drug Discovery (D^2)

TEXAS A&M UNIVERSITY

Target identification
- Cavity search
- PPI networks
- Structure-prediction
- Multi-omics

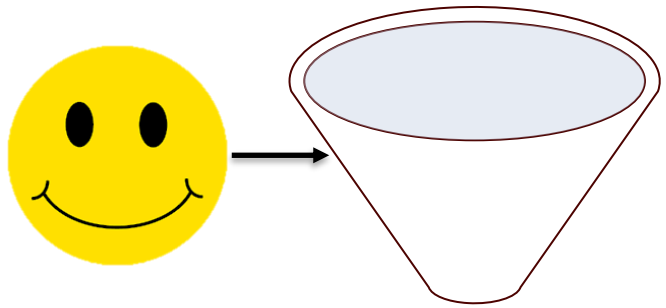
Hit Discovery
- Virtual screening
- Phenotypic screening
- Drug MoA modeling
- De-novo ligand design

Hit to lead
- Quantitative structure-activity relationships (qSAR)
- Computer-aided drug design (CADD)

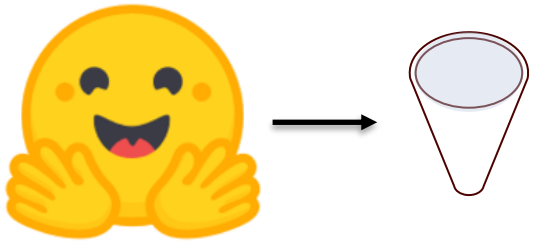
Lead optimization
- Absorption, Distribution, Metabolism, and Excretion (ADME)
- Toxicity

Precision Medicine
- Clinical trial design
- Drug response prediction
- Prediction of Adverse drug reaction (ADR)

Conventional D^2 Pipeline



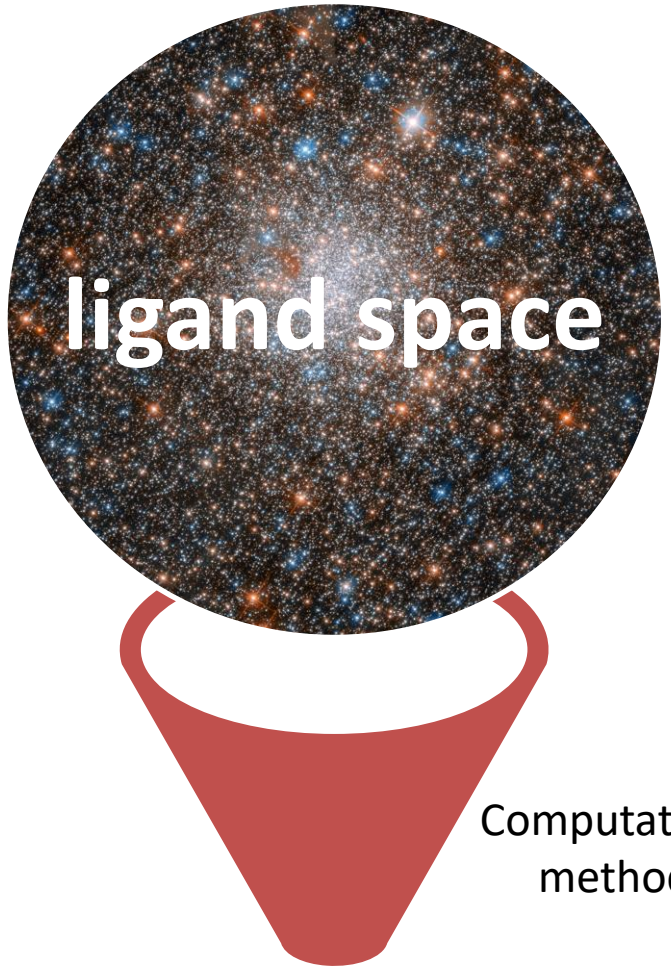
AI enhanced D^2 Pipeline



Hugging Face

Ultra-HT screening using in silico methods

TEXAS A&M UNIVERSITY



^{79}Au Target identification

- Multi-omics analysis
- Bio-informatics/biostatistics
- Structure-based (PDB, Alphafold2, OmegaFold, Openfold, RoseTTAFold, ESMFold)

Location Ligand libraries

- Commercial: ZINC, Chembridge, Maybridge, Enamine
- GenAI: MegaMolBART, molformer, chemGPT, drugGPT

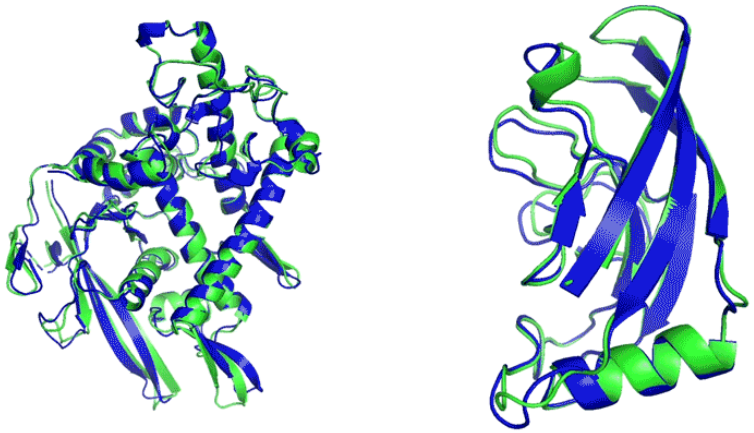
Pan Molecular modeling

- Cavity Search: Fpocket, Mdpocket, PocketMiner
- Docking: Autodock4, Vina, SMINA, GNINA, Glide, GOLD, DiffDock
- qSAR: DeepChem, MPNN, SPMM, ChemBERTa

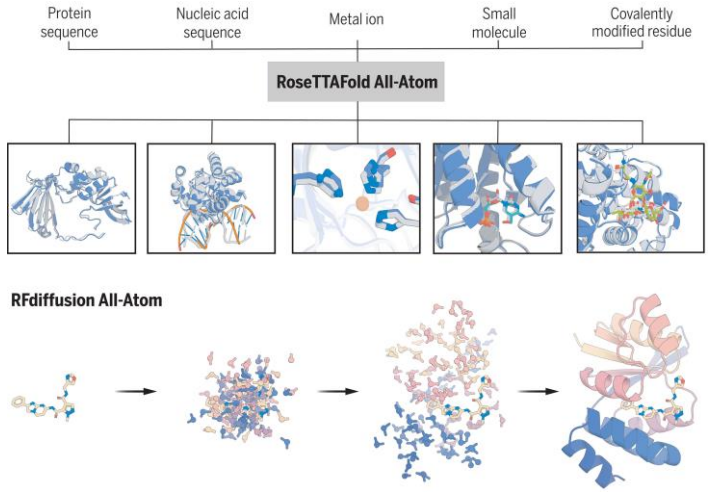
Protein structure & function prediction

TEXAS A&M UNIVERSITY

[alphafold: Open source code for AlphaFold.]

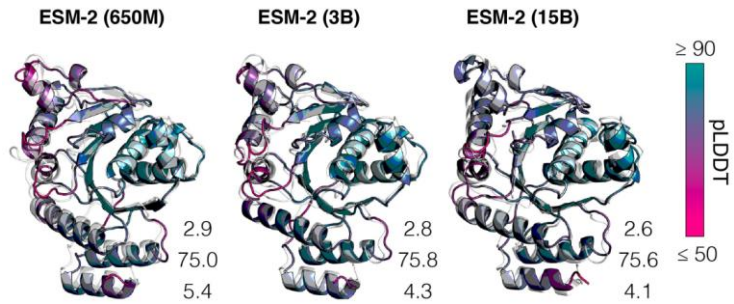


[baker-laboratory/RoseTTAFold-All-Atom]



[openfold: Trainable, memory-efficient, and GPU-friendly PyTorch reproduction of AlphaFold 2]

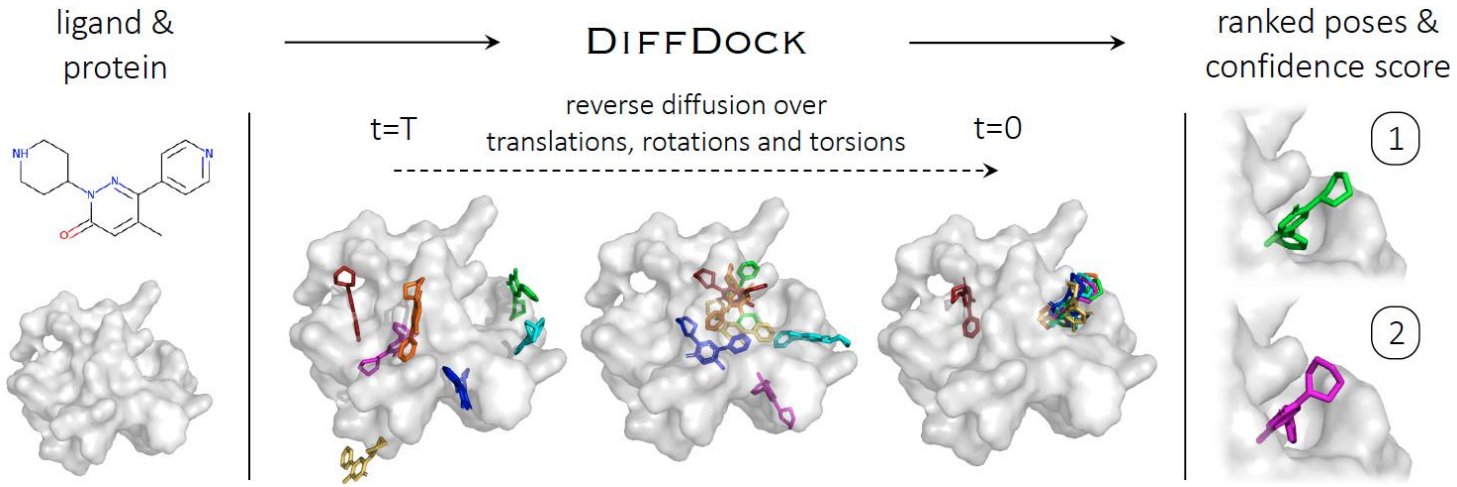
[esm: Evolutionary Scale Modeling (esm): Pretrained language models for proteins]



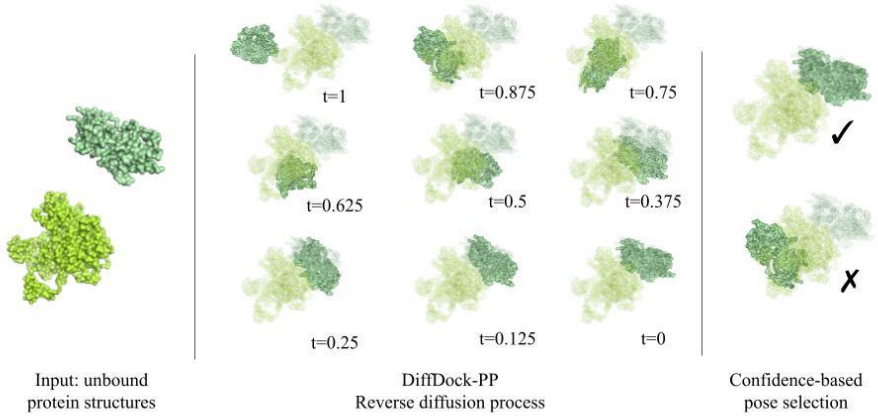
Molecular docking with GenAI

TEXAS A&M UNIVERSITY

[DiffDock: Implementation of DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking]

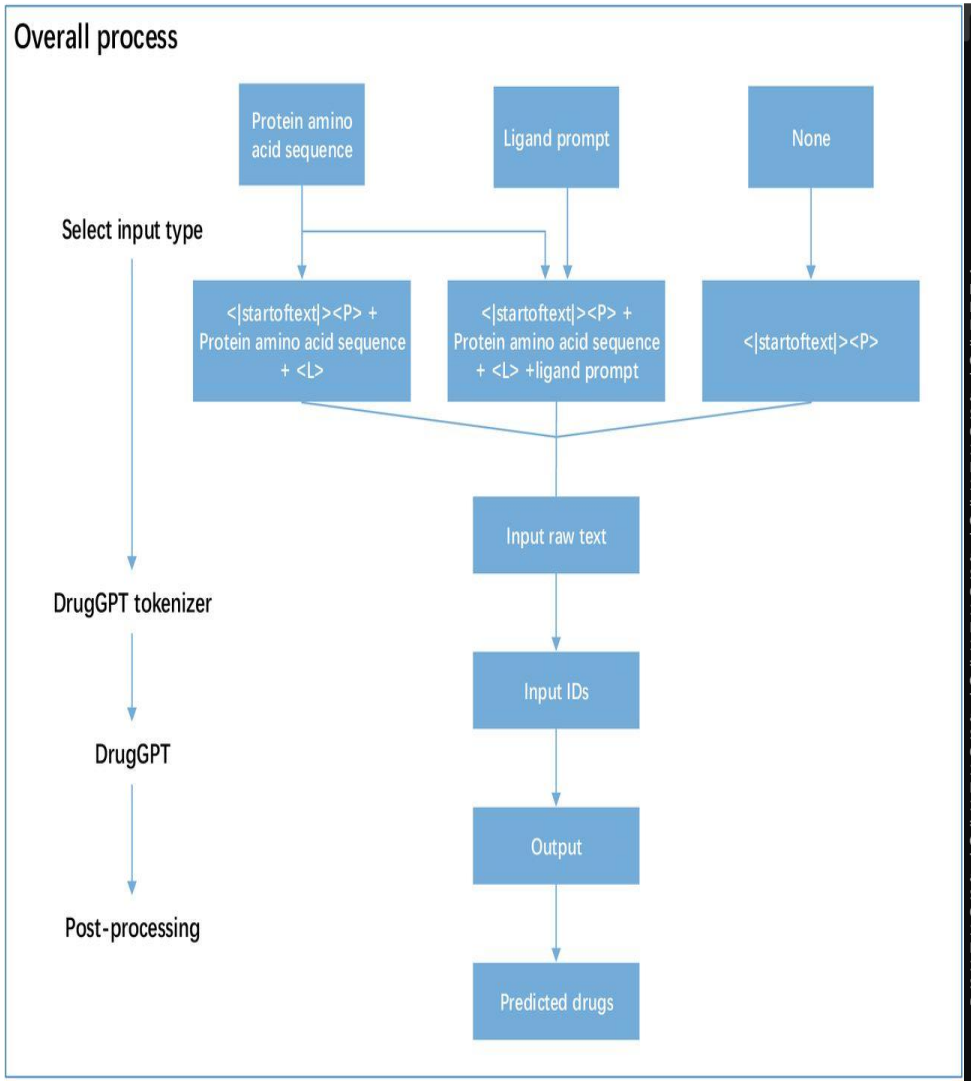


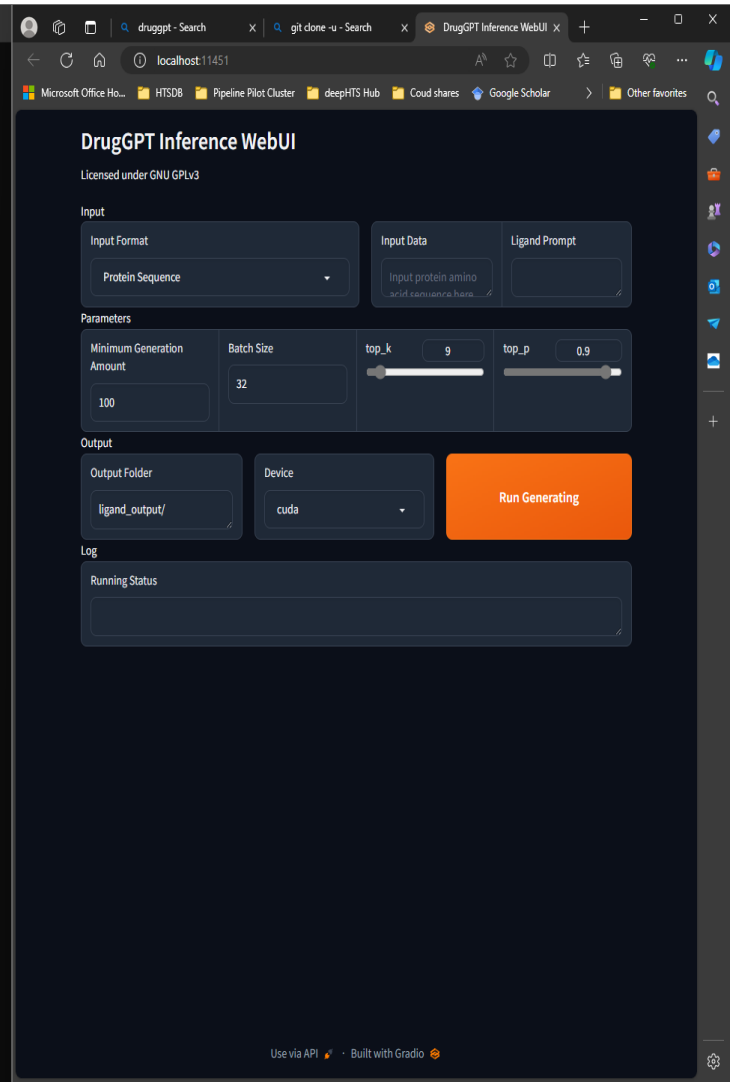
[ketatam/DiffDock-PP: Implementation of DiffDock-PP: Rigid Protein-Protein Docking]



Generative library design

TEXAS A&M UNIVERSITY



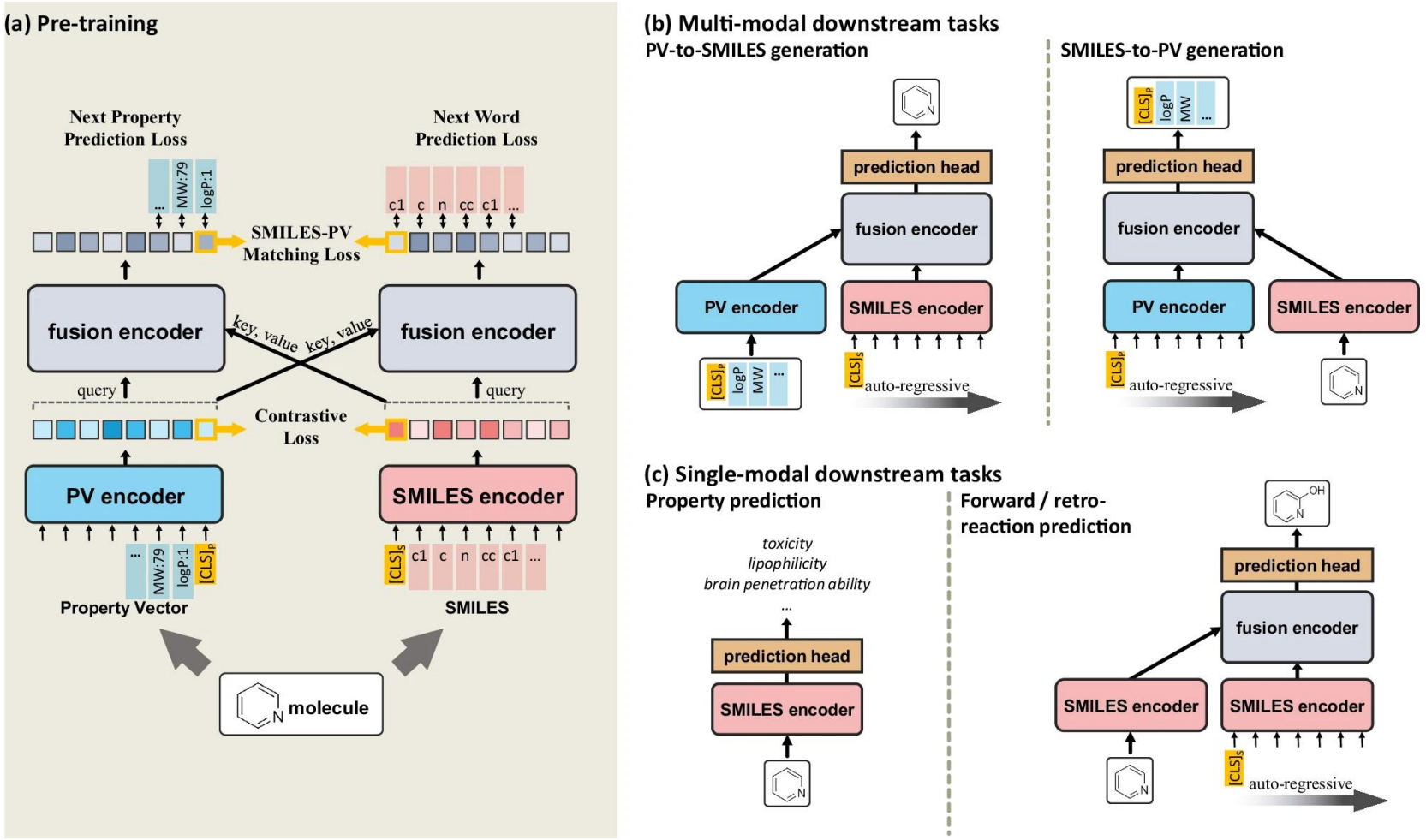


https://github.com/LIYUESEN/druggpt

AI qSAR models (e.g. SPMM)

TEXAS A&M UNIVERSITY

Multimodal learning for chemical domain, with SMILES and properties.

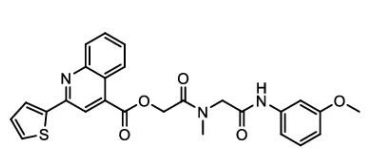
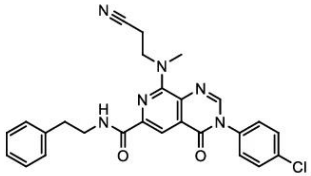
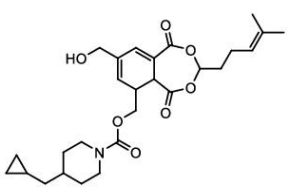
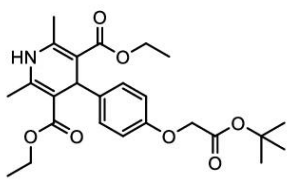
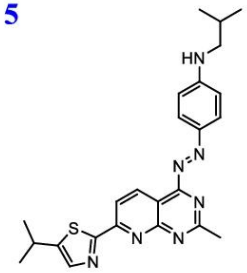
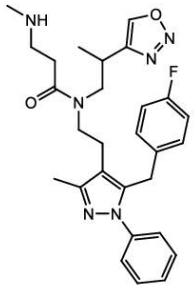


Bidirectional generation of structure and properties through a single molecular foundation model | Nature Communications

AI qSAR models (e.g. SPMM)

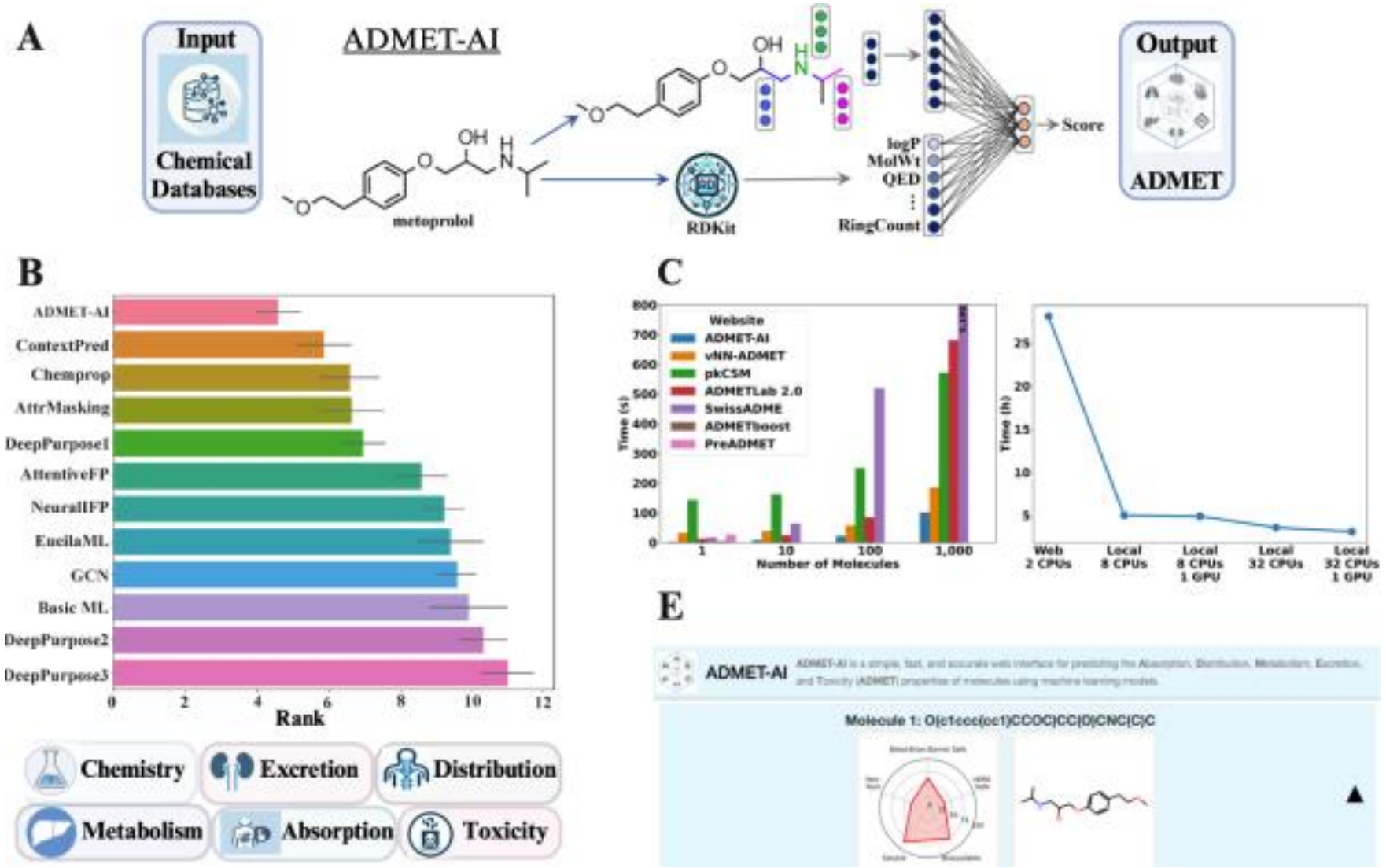
TEXAS A&M UNIVERSITY

Multimodal learning for chemical domain, with SMILES and properties.

	original PV + source molecule	(1) unchanged PV input	(2) #aromatic ring=0	(3) #ring=2 #aromatic ring=1	(4) logP=7	(5) #rotatable bond=12
	1	2	3	4	5	6
						
#HBondDonors	1	1	1	1	1	1
#HBondAcceptors	7	7	7	8	8	7
MolWt	489.1	486.2	487.2	459.2	445.2	504.3
#ring	4	4	4	2	4	4
#aromatic_ring	4	4	0	1	4	4
#rotatable_bond	8	8	8	8	7	12
TPSA	97.83	103.9	102.4	100.2	88.31	89.08
#heavy_atom	35	35	35	33	32	37
logP	4.226	3.757	3.899	3.768	7.063	4.078
MR	134.6	136.1	127.9	122.3	131.2	139.1
#N,O	8	8	8	8	7	8
QED	0.3722	0.4064	0.4092	0.4632	0.3075	0.3128

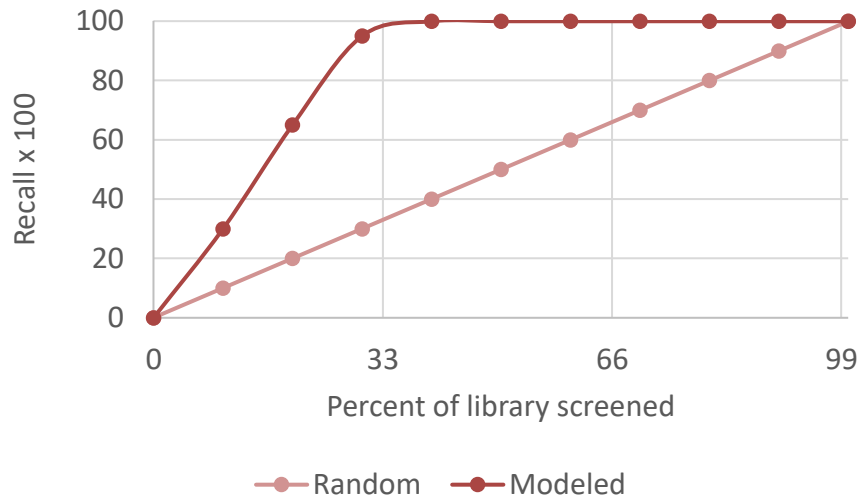
Bidirectional generation of structure and properties through a single molecular foundation model | Nature Communications

Multi-task ADMET prediction



https://doi.org/10.1101/2023.12.28.573531 , https://github.com/swansonk14/admet_ai

In silico workflow



Target

drugGPT

Commercial Library

Negative control

DUD

1° screen

DiffDock

Counter screen

ADMET_AI

Leads

Ad hoc screen

SPMM