# Success and Challenges of Artificial Intelligence in Drug Discovery

Dr Taravat Ghafourian
Nova Southeastern University

Barry and Judy Silverman College of Pharmacy
NOVA SOUTHEASTERN UNIVERSITY
NSU Florida

# A brief biography

## Experience in Academia

- Oct 2022-Present: Nova Southeastern University
- Director of Pharmacy Development and Reader in Pharmacy, University of Bedfordshire, Aug 2020-2022
- Senior Lecturer, School of Life Sciences, University of Sussex, 2015-2020
- Lecturer, Medway School of Pharmacy, University of Kent, 2005-2015
- Research Fellow, Liverpool John Moores University, 2003-2005
- Visiting scientist at University College London 2001
- Assisstant/Associate Professor at Tabriz Faculty of Pharmacy, 1997-2003

## Education:

- Pharmacy Doctorate, Tabriz Faculty of Pharmacy, 1992
- PhD in Pharmaceutical sciences, Liverpool John Moores University, 1996
- Postgraduate Certificate in Higher Education, University of Kent, 2008

# AI and Computational Modeling in Drug Discovery: History

harles Babbage (1791-1871)
First analytical computer

Alan Turing (1912-1954)
father of computer science
and artificial intelligence

Symbolic AI

Machine Learning

Foundation of Generative AI

ChatGPT

1822

1950

2000

2010

2020
Open AI

2022

Data Science
Computer Science
Statistics

Algorithms
Machine Learning
artificial neural networks
learning neural networks

1951 first AI program

1956 term AI coined

Convolutional Neural Networks
Deep Learning
Transformer

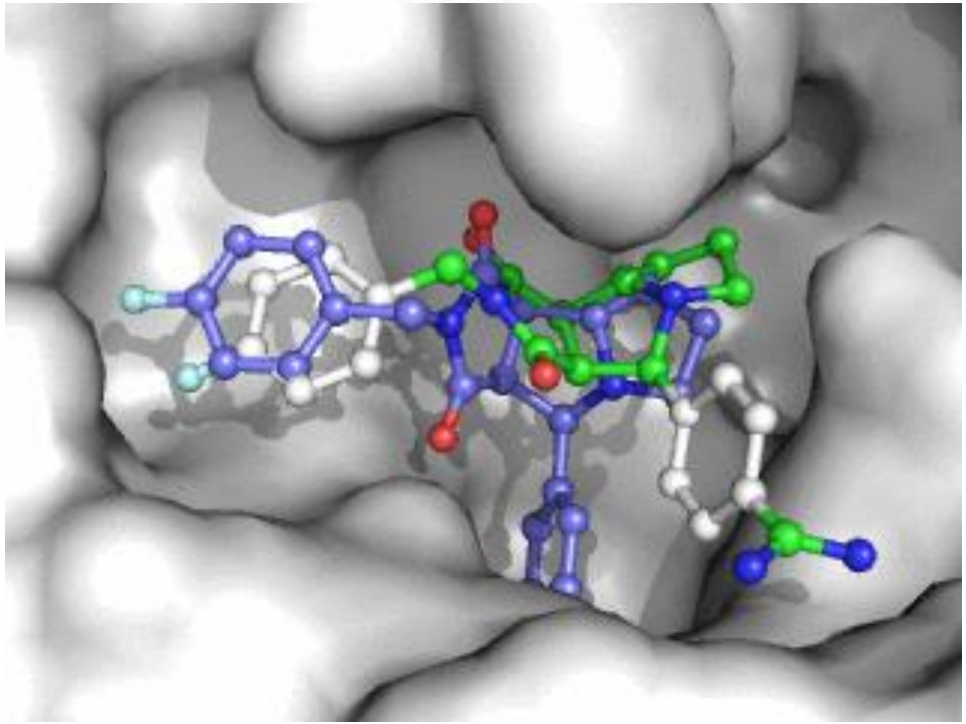
convolutional nets
Artificial neural network tech

Generative adversarial networks

Turing Natural Language Generators

Large Language Models ChatGPT, Perplexity, Claude,

**Corwin Hansch QSAR paper 1953 & 1966**

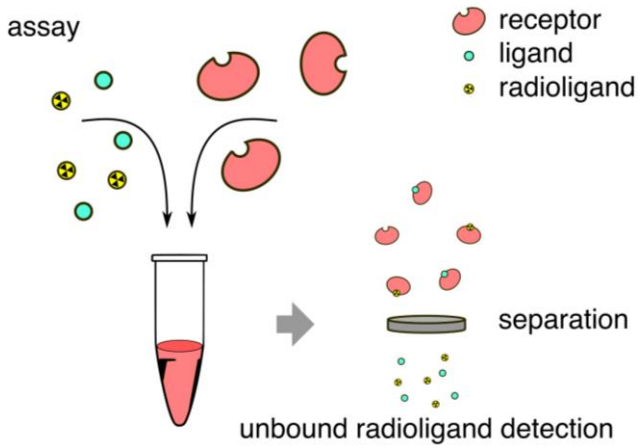


Computer-aided drug discovery has been around for decades

A recent surge in embracing computational technology

# Scientific models in pharmaceutical sciences

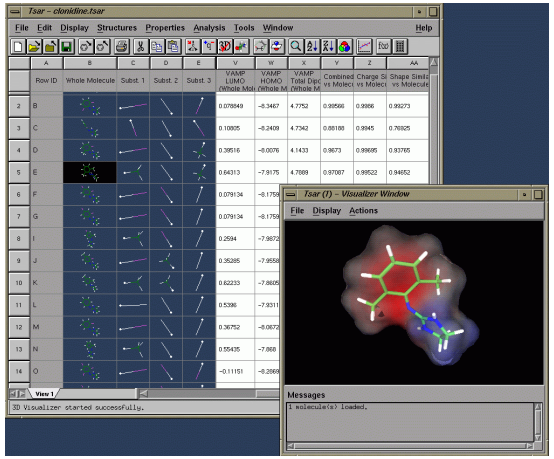
*In vitro models*




*Animal models*

*Clinical trials: clinical models of the real patient population*

*Computational models*



*Hypothetical example: efficacy of an antidepressant in a tablet dosage form*

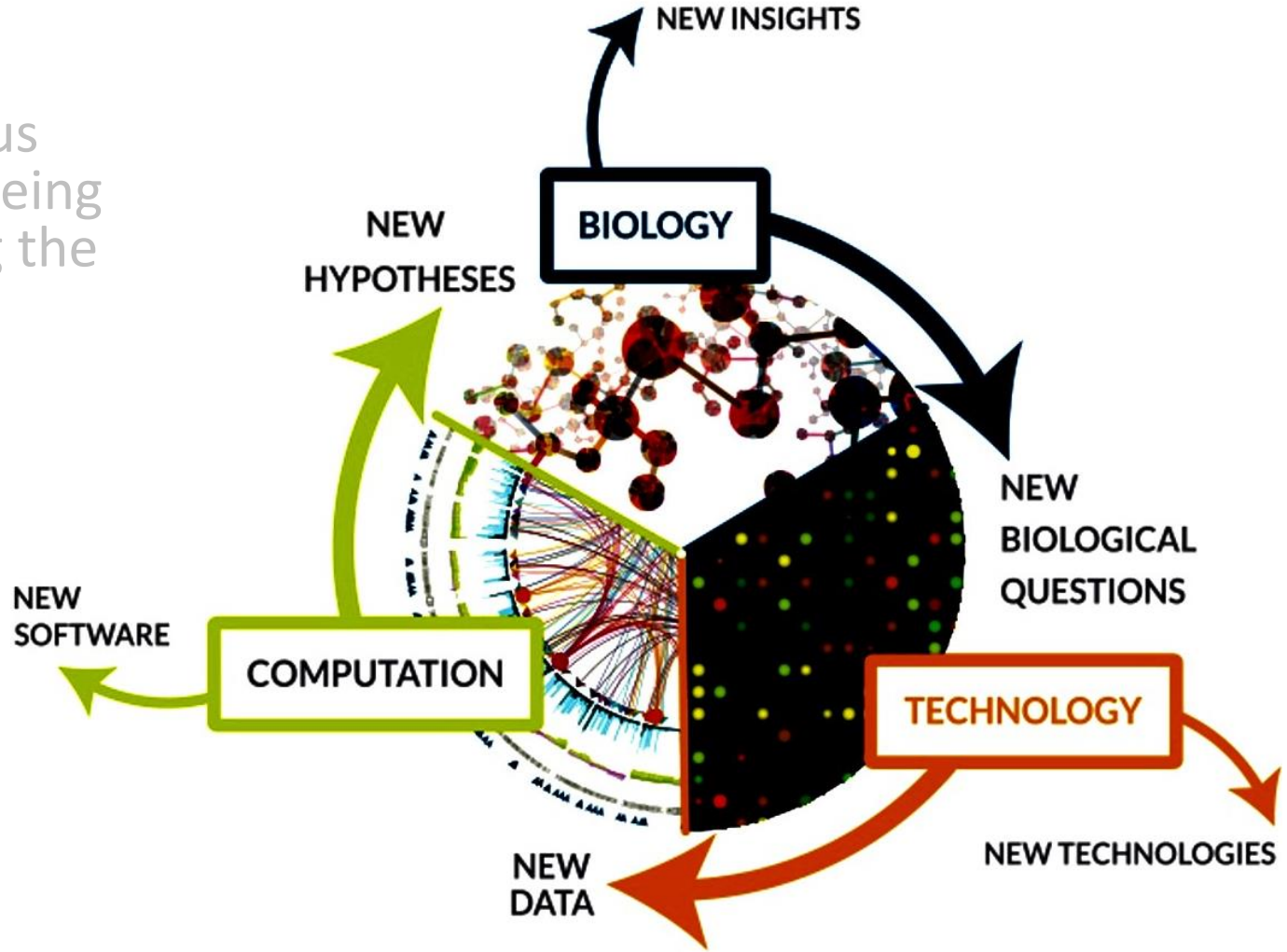
# *All we do in science is modeling*

"... the sciences do not try to explain, they hardly even try to interpret, **they mainly make models**. By a model is meant **a mathematical construct** which, with the addition of certain **verbal interpretations**, describes observed phenomena. **The justification of such a mathematical construct is solely and precisely that it is expected to work—that is, correctly to describe phenomena** from a reasonably wide area."

John von Neumann

# Computational modelling:

The use of computers to simulate and study complex systems using mathematics, physics, and computer science. A computational model contains numerous variables that characterize the system being studied. Simulation is done by adjusting the variables alone or in combination and observing the outcomes. Computer modeling allows scientists to conduct thousands of simulated experiments by computer.



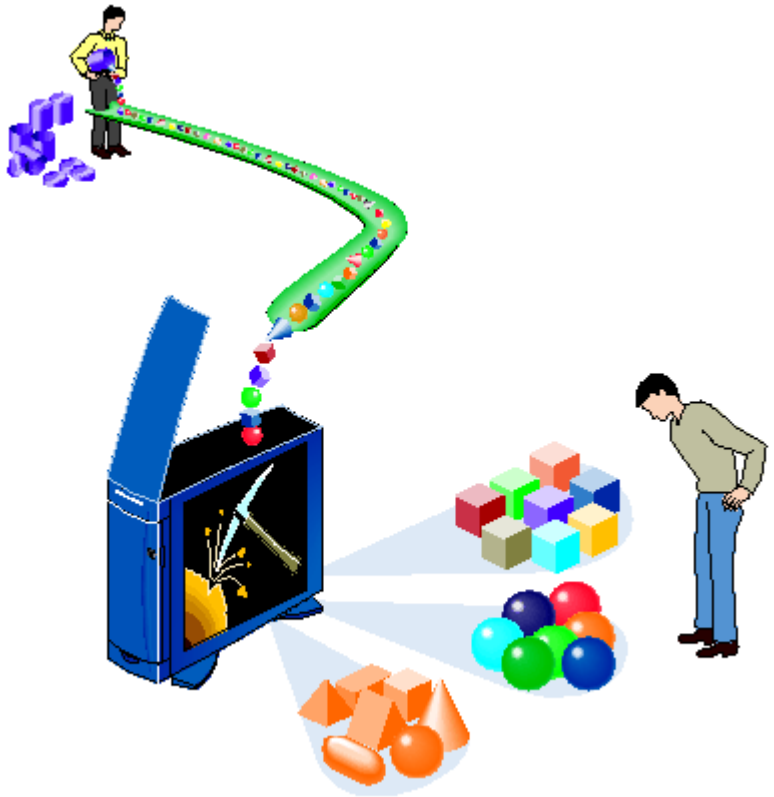

National Institute of Biomedical Imaging and Bioengineering | *Technologies to Shape the Future of Health*

*https://www.nibib.nih.gov/science-education/science-topics/computational-modeling*

Barry and Judy Silverman College of Pharmacy
NOVA SOUTHEASTERN UNIVERSITY | NSU Florida

# AI and ML (Artificial Intelligence and Machine Learning)

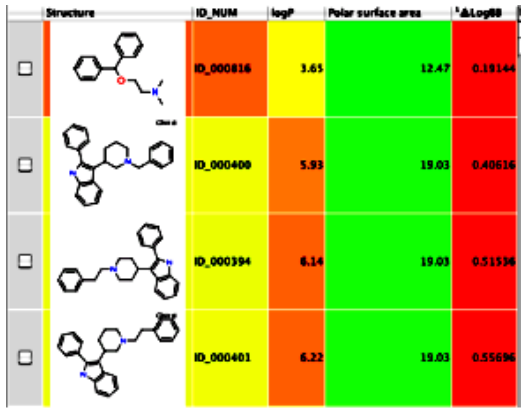
*ML discovers patterns in **data** so that this knowledge can be applied to enable future predictions*

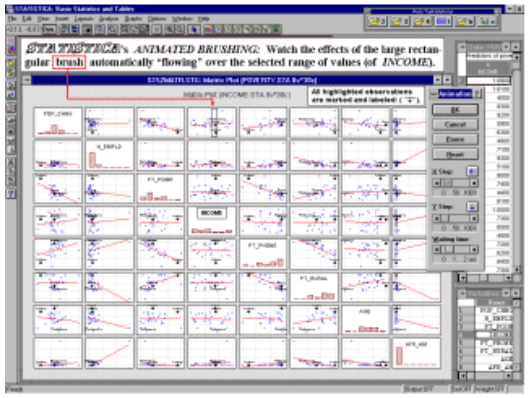


- Specific AI (weak AI) is engineered to execute a specific task, the most prevalent form of AI today
- Universal AI (strong AI) can comprehend, learn, and apply intelligence across a variety of tasks
- Superhuman AI: hypothetical, a topic of debate and research.
- ML is the tool used in AI: creating algorithms that learn from data and make predictions:
  - Supervised learning, Unsupervised learning, Reinforcement learning
  - Neural networks: an MI algorithm to identify patterns in data by simulating the brain's processing methods. Deep learning, employs complex, multilayered neural networks to perform these tasks.

# Data Modelling (Machine Learning)

*Data Source*



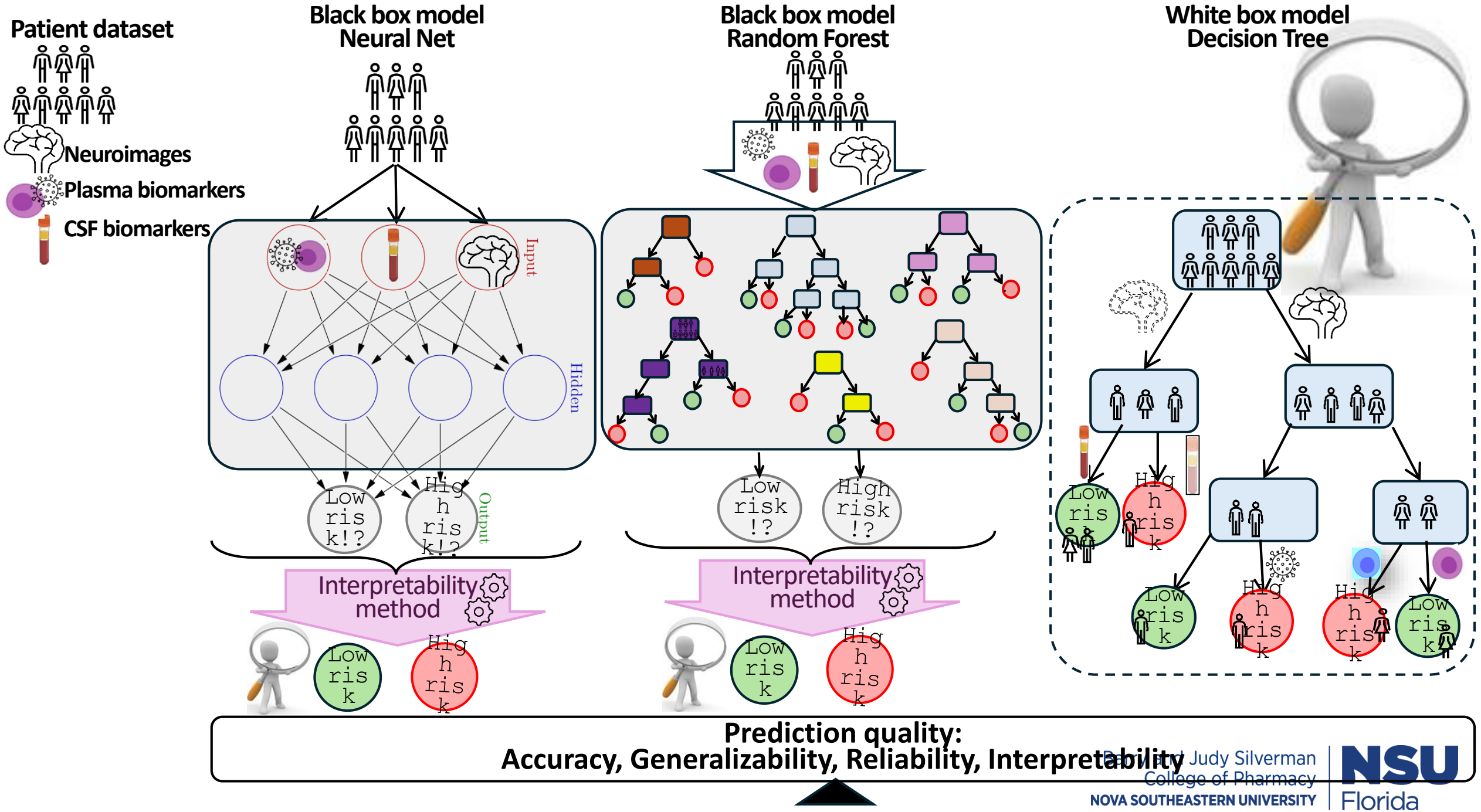



*Machine learning (Data Mining)*

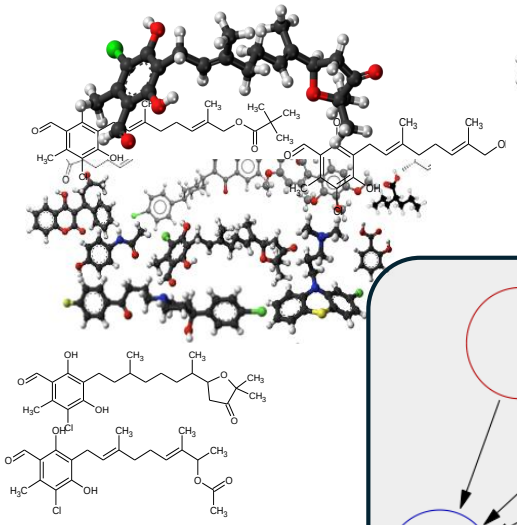
*Learned Patterns or rules (AI system)*

Predictions & Decisions for new cases

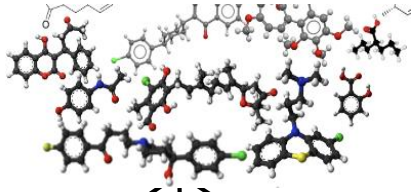

***Knowledge***

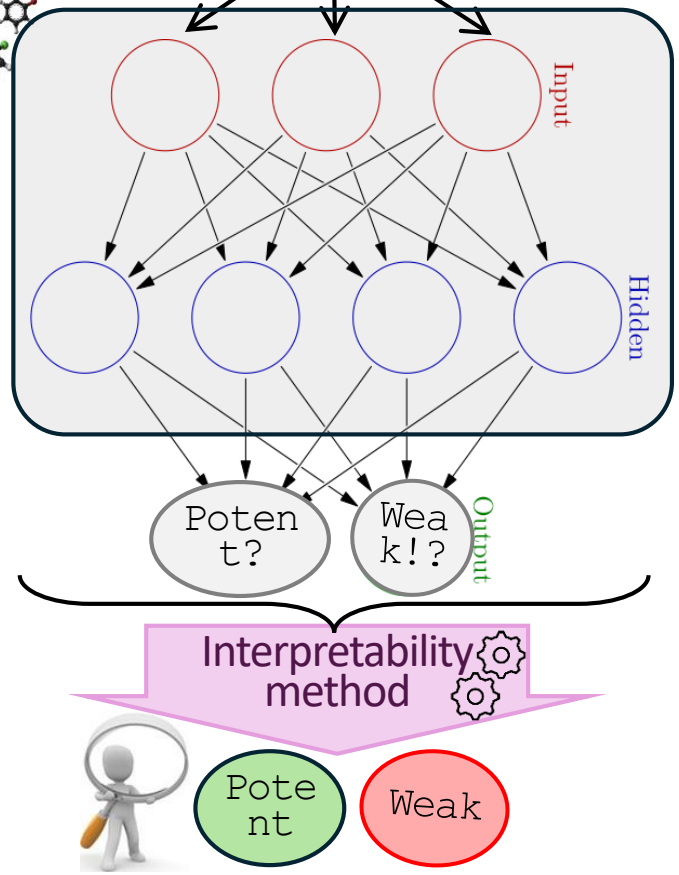

Patient dataset
Neuroimages
Plasma biomarkers
CSF biomarkers
Black box model
Neural Net
Input
Hidden
Output
Low risk!?
High risk!?
Interpretability method
Low risk
High risk
Black box model
Random Forest
Low risk!?
High risk!?
Interpretability method
Low risk
High risk
White box model
Decision Tree
Low risk
High risk
Prediction quality:
Accuracy, Generalizability, Reliability, Interpretability

## Compound dataset

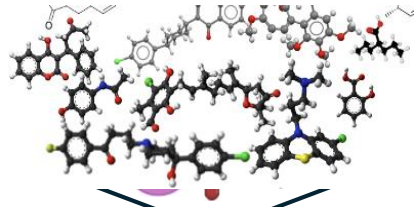


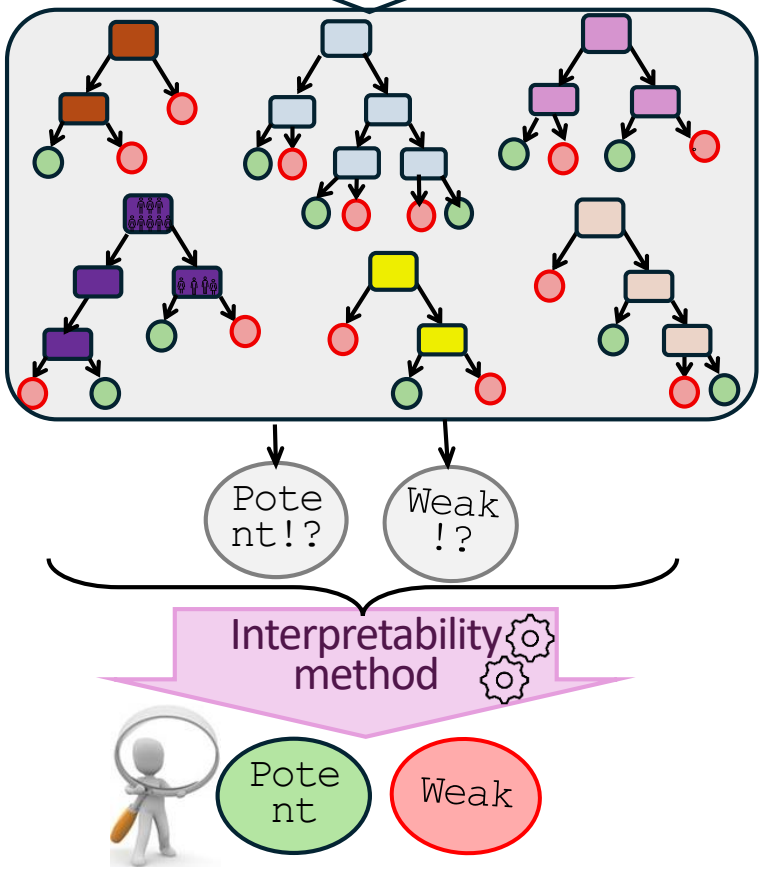
## Black box model Neural Net



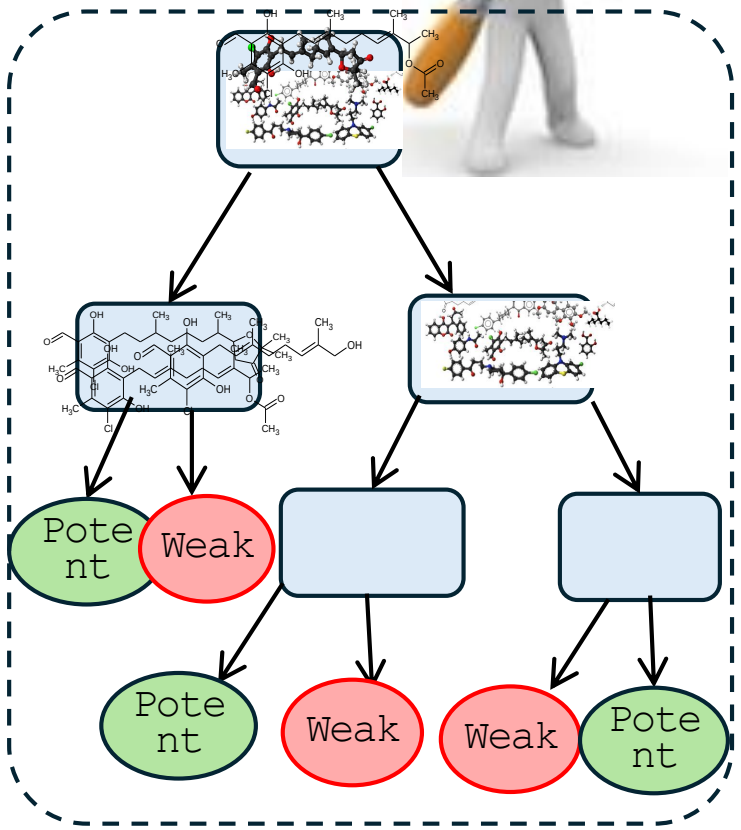



## Black box model Random Forest






## White box model Decision Tree




Barry and Judy Silverman College of Pharmacy
NOVA SOUTHEASTERN UNIVERSITY
NSU Florida

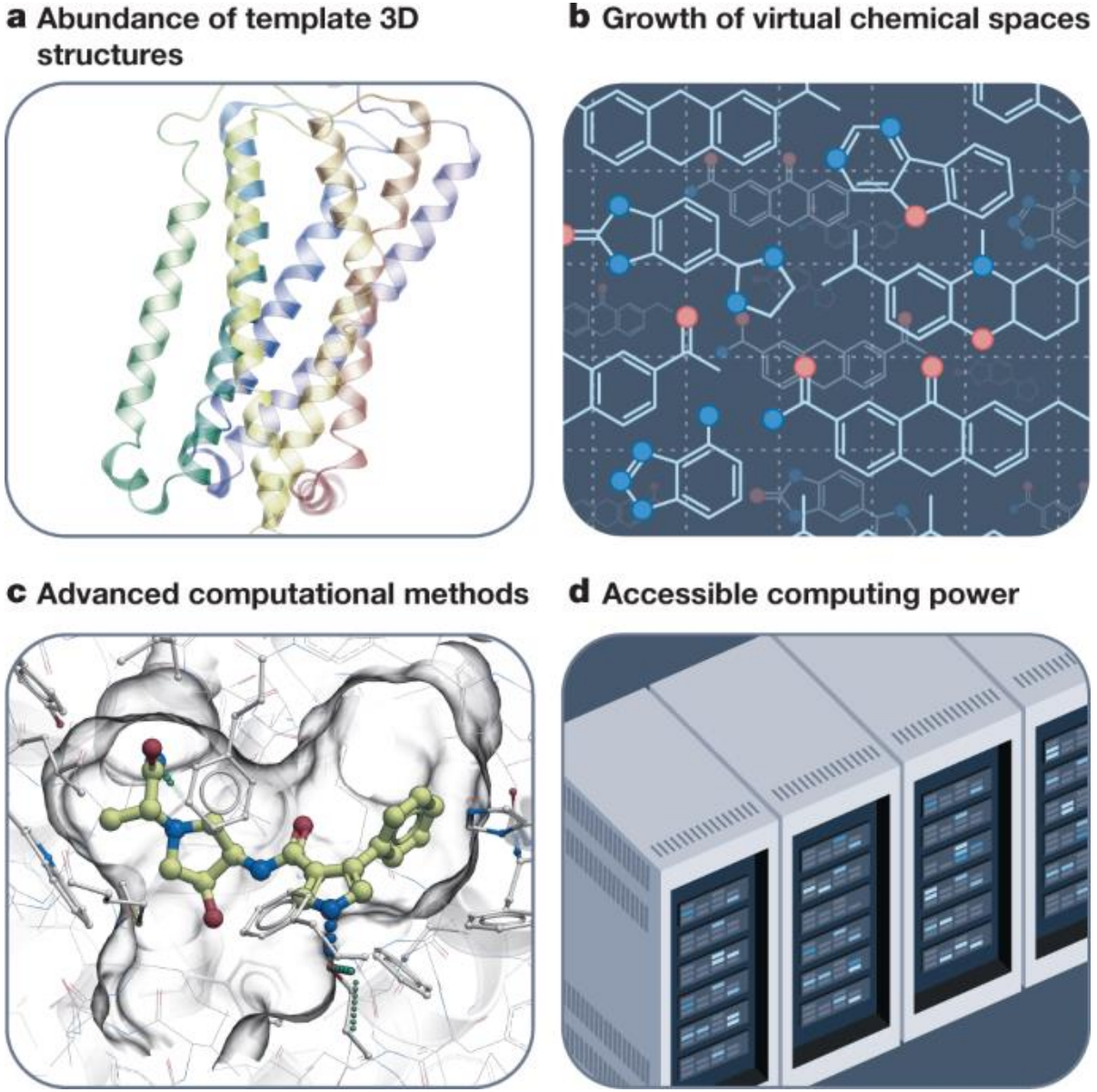
# FDA Modernization Act 2.0

As part of the Omnibus Bill signed in Dec 2022 by President Biden, the act eliminates the requirement to conduct animal testing before human trials

- Opportunities for alternatives to animal testing

# Progress in Computational modelling for Drug Discovery

- The flood of data on ligands and their binding to targets
- the 3D structures of proteins (PDB has 200K proteins, **AlphaFold** can predict the 3D model)
- abundant computing capacities
- on-demand virtual libraries of billions of drug-like molecules ($3 \times 10^{10}$ on-demand compounds in 2022)
- deep learning predictions of ligand properties and target activities

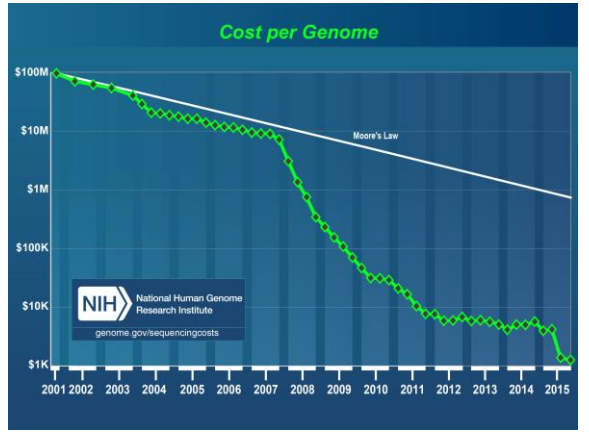



presenting new opportunities for the cost-effective development of drugs; democratizing the drug discovery process

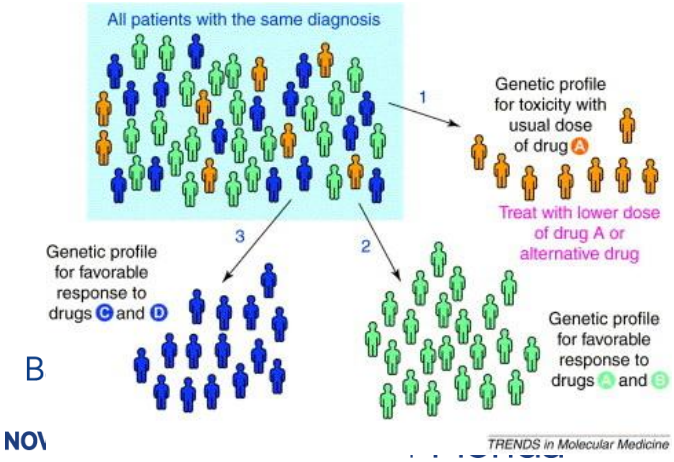
https://www.nature.com/articles/s41586-

# Impact of New ”Revolutionary” Technologies on Drug Discovery and development

- High-throughput automated chemistry: microwave-assisted and flow chemistry, combinatorial chemistry
- Gene chips, genomics and sequencing and now CRISPR/Cas9 for gene editing
- Omics tools
- High-throughput screening & assays: 3D-cell cultures, Organ-on-Chip, in vitro ADME/Tox
- Progress in tissue engineering and organ on a chip
- Bioinformatics & Molecular biology tools
- Biomedical Databases: bioactivity, protein distribution, differential gene expression, network pharmacology
- Longitudinal observational data
- Virtual screening and library design and Docking
- …. and now AI





# Databases: Big Data

- Chemical data and drug data

ChemSpider
Search and share chemistry

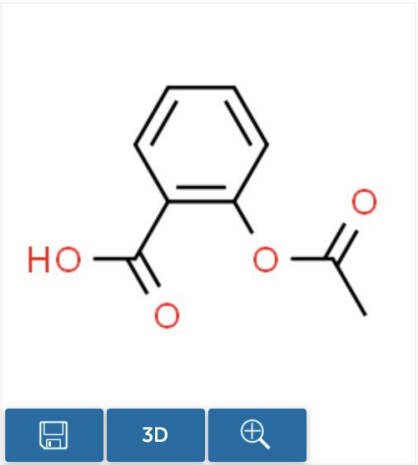
>114 million structures

For medical information relating to Covid-19, please consult the World Health Organisation or local h

Simple Structure Advanced History

Found 1 result

Search term: **aspirin** (Found by approved synonym)




## Aspirin

| | |
|---|---|
| Molecular Formula | $C_9H_8O_4$ |
| Average mass | 180.157 Da |
| Monoisotopic mass | 180.042252 Da |
| ChemSpider ID | 2157 |

## More details:

| | |
|---|---|
| Systematic name | 2-Acetoxybenzoic acid |
| SMILES | CC(OC1=C(C(=O)O)C=CC=C1)=O CC(=O)OC1C=CC=CC=1C(O)=O |

DRUGBANK Online

Browse COVID-19 Search Inter Che

Check out our DrugBank Affiliate Partnerships for biotech and pharma consultants! Learn More

Identification
Pharmacology
Interactions
Products
Categories
Chemical Identifiers
References
Clinical Trials
Pharmacoeconomics
Properties
Spectra
Targets (19)
Enzymes (4)
Transporters (3)

## Aspirin

**Summary** **Aspirin** is a salicylate used to treat pain, fever, inflammation, migraines major adverse cardiovascular events.

**Brand Names** *Aggrenox, Alka-seltzer, Alka-seltzer Fruit Chews, Anacin, Arthriten Inflar cor Aspir-low Bayer Aspirin Bayer Womens Bc Arthritis Bc Original Fc*

**Generic Name** Acetylsalicylic acid
Commonly known or available as Aspirin

**DrugBank Accession Number**

**Background** Also known as *Aspirin*, acetylsalicylic acid (ASA) is a commonly used dr pain and fever due to various causes. Acetylsalicylic acid has both anti- antipyretic effects. This drug also inhibits platelet aggregation and is u blood clots stroke, and myocardial infarction (MI) [Label].

Interestingly, the results of various studies have demonstrated that lor acetylsalicylic acid may decrease the risk of various cancers, including

Barry and Judy Silverman College of Pharmacy
NOVA SOUTHEASTERN UNIVERSITY | NSU Florida

# Databases: bioactivity data

ChEMBL

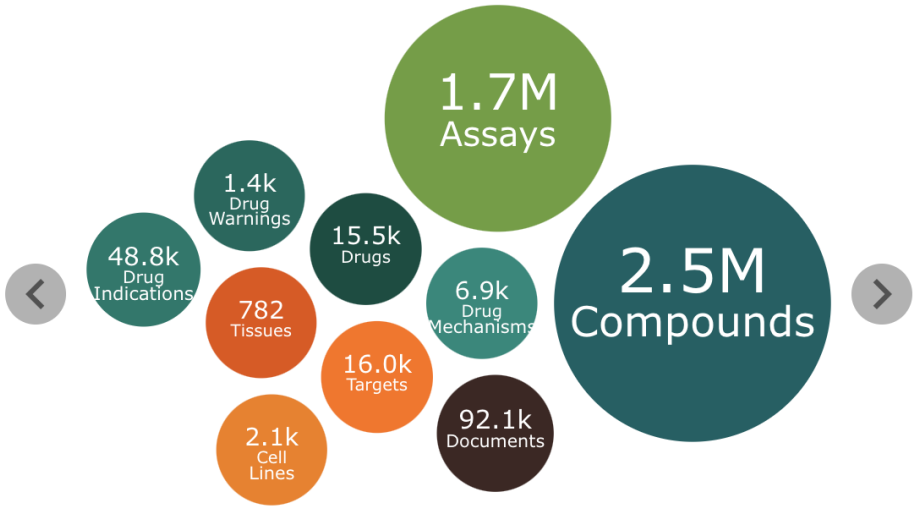
Search in ChEMBL

Examples: Imatinib erbB2 brain MDCK c1ccccc1N

ADVANCED SEARCH

UNICHEM | CHEMBL-NTD | SURECHEMBL | MALARIA INHIBITOR PREDICTION | DOWNLOADS | MORE

ChEMBL is a manually curated database of bioactive molecules with drug-like properties. It brings together chemical, bioactivity and genomic data to aid the translation of genomic information into effective new drugs.




## Explore ChEMBL

**Description:** Shows a summary of the ChEMBL entities and quantities of data for each of them.

**Instructions:** Click on a bubble to explore a specific ChEMBL entity in more detail.

https://www.ebi.ac.uk/chembl/

dy Silverman of Pharmacy
ERN UNIVERSITY
NSU Florida

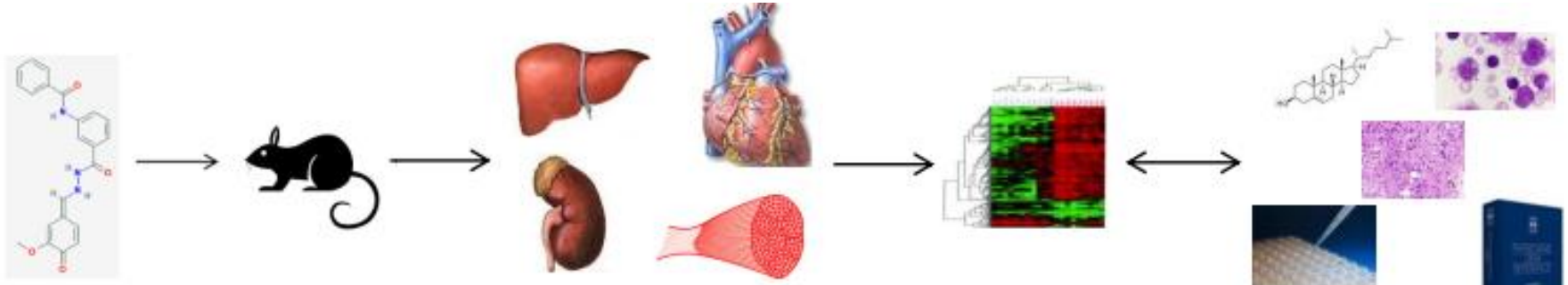
# Toxicogenomic Data

National Toxicology Program
U.S. Department of Health and Human Services

What We Study | How We Work | Data & Resources | Publication

Home » Data & Resources » **DrugMatrix/ToxFX**

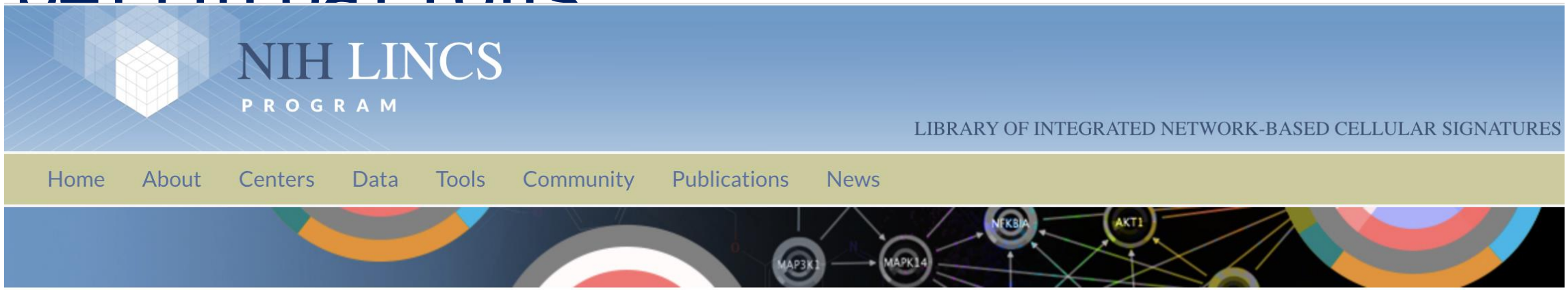
## DrugMatrix/ToxFX



https://tools.niehs.nih.gov/cebs3/views/index.cfm?action=main.dataReview&bin_id=15670

Barry and Judy Silverman College of Pharmacy
IOVA SOUTHEASTERN UNIVERSITY | NSU Florida

# Differential Gene Expression due to many exogenous and endogenous perturbations




## LINCS Data and Signature Generation Centers (DSGCs)

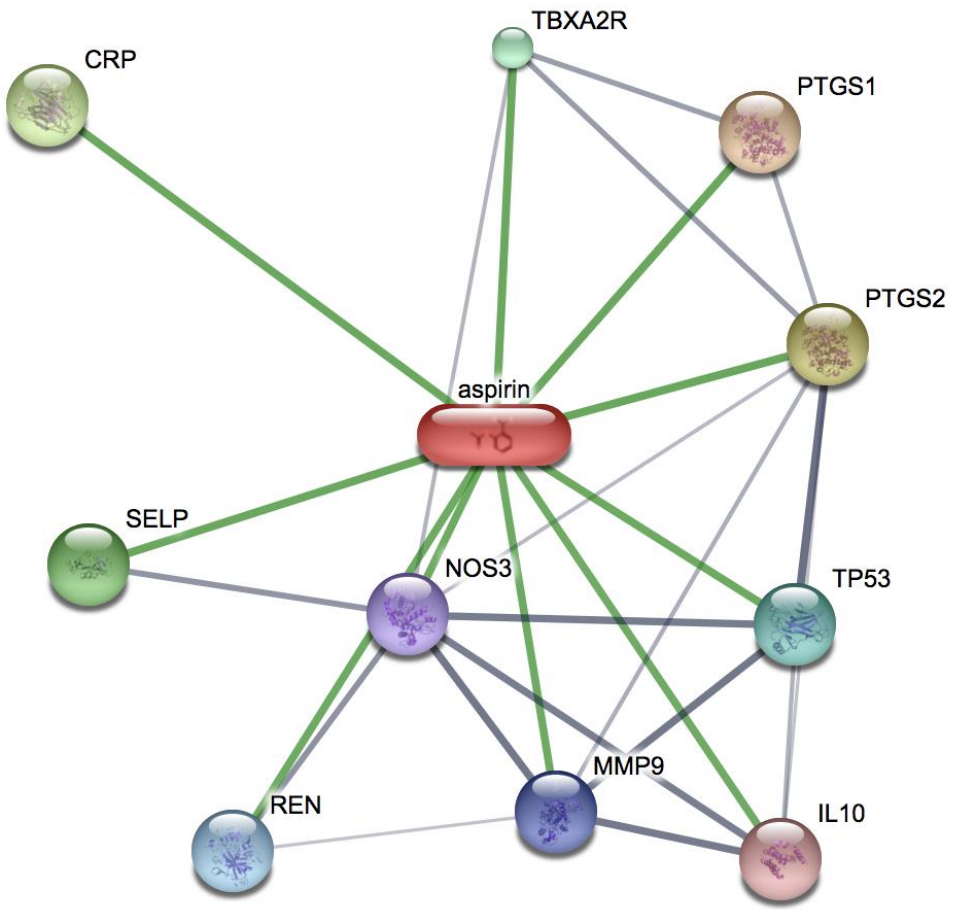
1. Drug Toxicity Signature Generation Center:  cell signatures to predict adverse drug reactions
2. The HMS LINCS Center: normal and diseased human cells response to perturbation by drugs, mutations, environment
3. The LINCS Center for Transcriptomics is studying up to 50 cell types perturbed by a large number of compounds
4. The LINCS Proteomic Characterization Center for Signaling and Epigenetics studies phosphorylation-mediated signaling
5. Microenvironment Perturbagen (MEP) LINCS Center: impact on cellular phenotype, proteins, genes
6. The NeuroLINCS Center concentrates on human brain cells

*https://lincsproject.org/LINCS/centers/data-and-signature-generating-centers*

# **STITCH:** Chemical-Protein Interaction Networks

STITCH

Search Download




https://www.embl.org/

Barry and Judy Silverman College of Pharmacy
NOVA SOUTHEASTERN UNIVERSITY
NSU Florida

# Longitudinal Medical Data

biobank uk
Enabling scientific discoveries that improve human health

Enable your research | Explore your participation | Learn more about UK Biobank


UK Biobank is a large-scale biomedical database and research resource, containing in-depth genetic and health information from half a million UK participants. The database is regularly augmented with additional data and is globally accessible to approved researchers undertaking vital research into the most common and life-threatening diseases. It is a major contributor to the advancement of modern medicine and treatment and has enabled many scientific discoveries that improve human health.

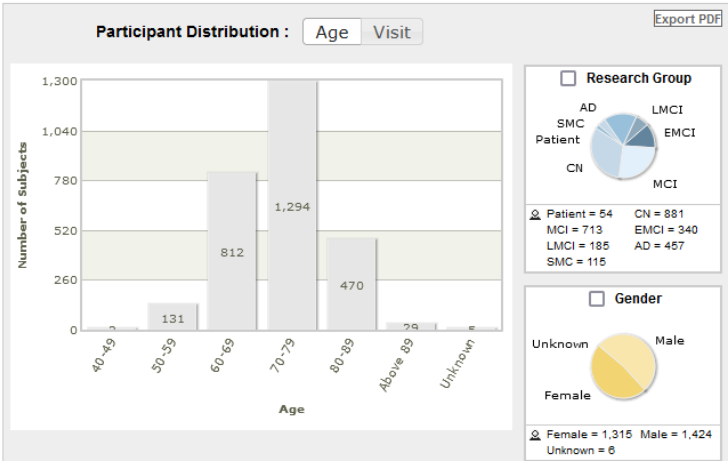
IDA | Select Study: ADNI | ADNI@LONI | Download | Search

## Alzheimer's Disease Neuroimaging Initiative

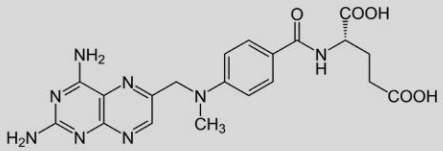
The Alzheimer's Disease Neuroimaging Initiative (ADNI) seeks to develop biomarkers of the disease and advance the understanding of AD pathophysiology, improve diagnostic methods for early detection of AD and improve clinical trial design. Additional goals are examining the rate of progress for both mild cognitive impairment and Alzheimer's disease, as well as building a large repository of clinical and imaging data.

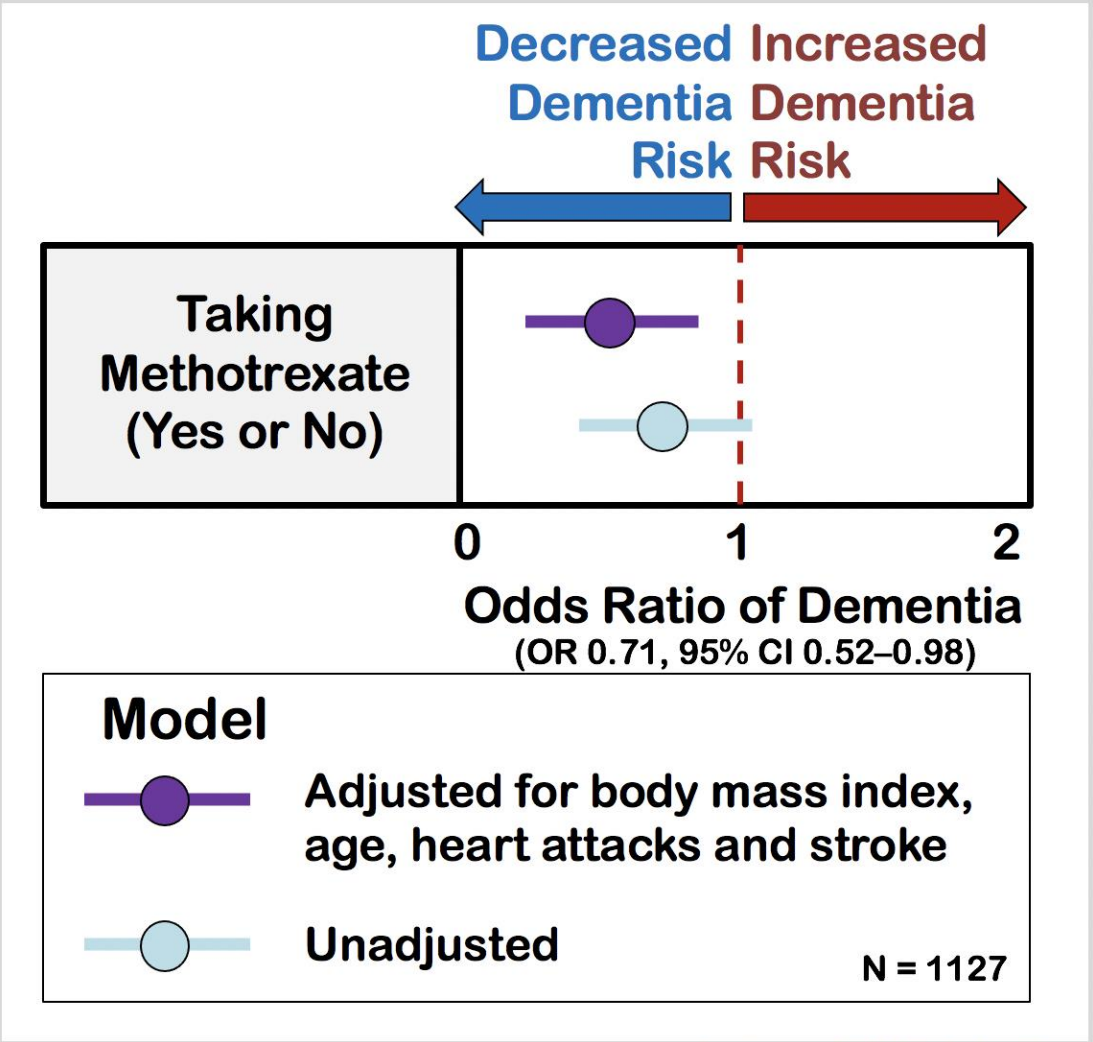
Related Links

- News and Announcements
- MRI Recall Notices
- Searchable Data Dictionary
- ADNI Data Use Agreement



# Methotrexate is associated with decreased dementia risk in people with rheumatoid arthritis

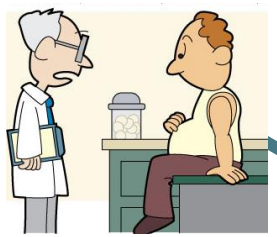




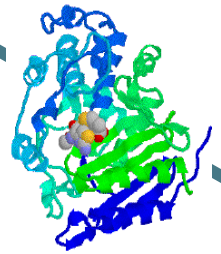

Newby et al (2020) AlzResTher

# Drug Discovery & Development

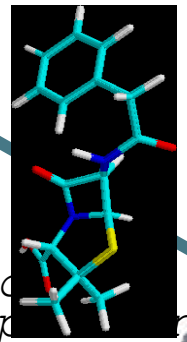
Unmet medical need



Isolate protein involved in disease



Find a drug effec
against disease p r

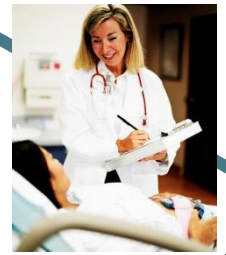


Preclinical testing

Formulation &
Scale-up

File IND

Human clinical trials



File NDA

Regulatory Approval

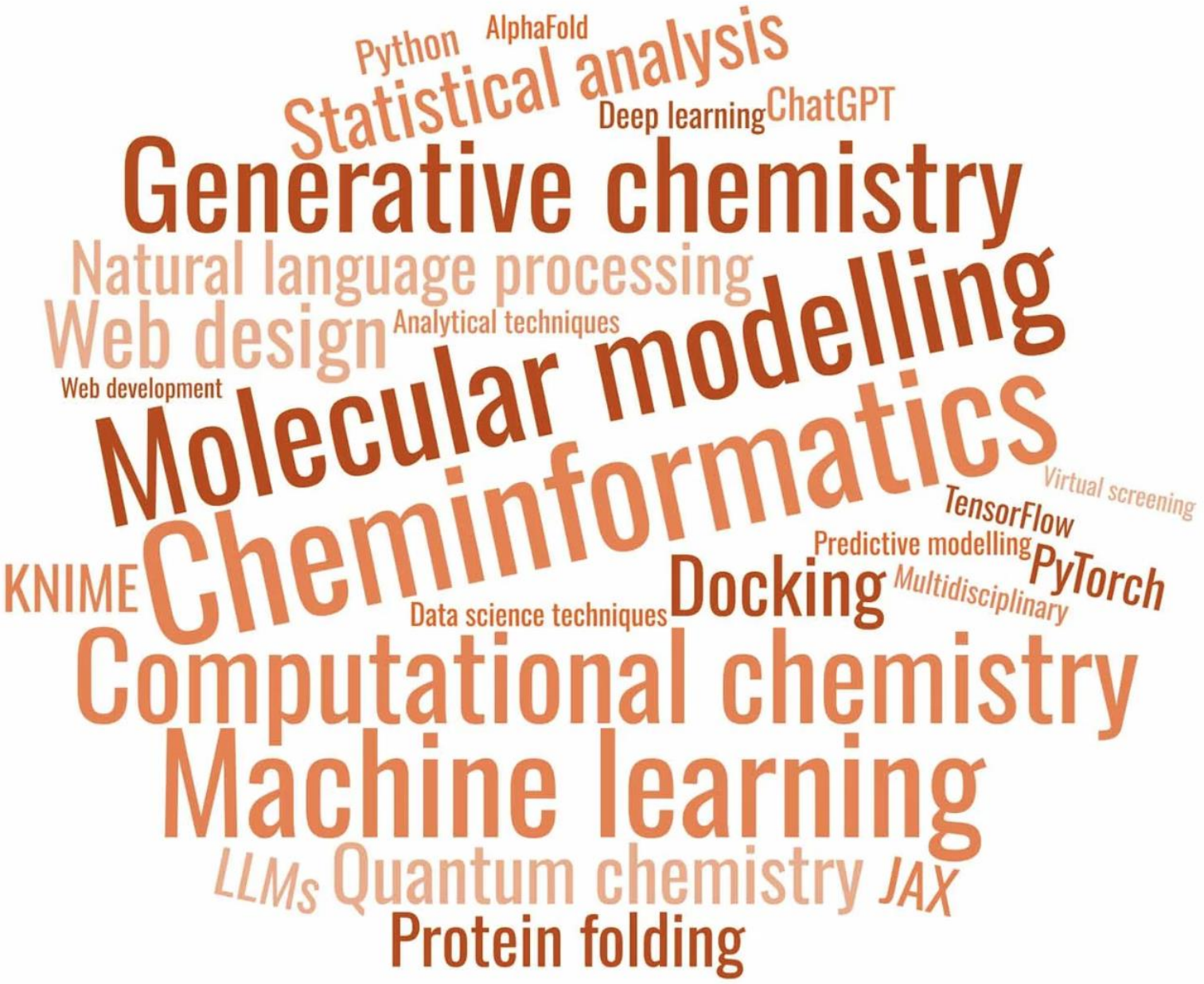
Profit decline/ withdrawal from market

Patent expiration and competitor's products

Post-marketing surveillance for safety controls

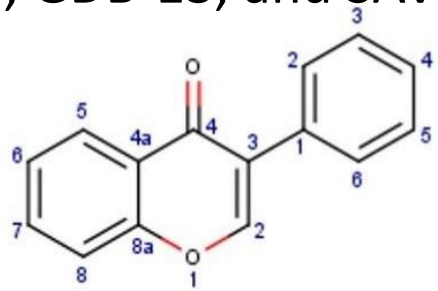
**Computer-based modelling**
**In vitro high-throughput testing**

# Skills for the AI-enabled chemist of the future

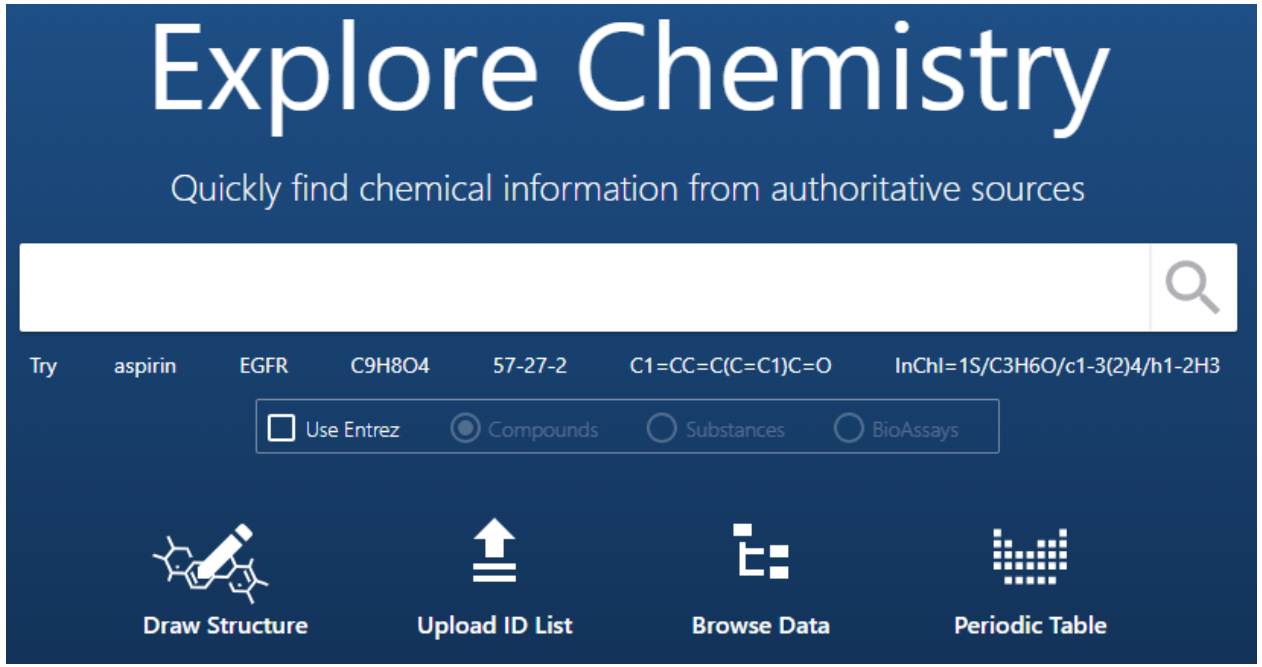


# Molecular Representations in Chemical Librar: Used for virtual screening

- Large virtual chemical information generated by pharm, food, and agriculture industry with synthetic and stability feasibility
- Contain structural data, stereochemical information, physicochemical and spectroscopic properties
- PUBCHEM has 119m compounds
- ZINC, Enamine Real, GDB-13, and SAVI ($10^9$ compounds)



| | |
|---|---|
| **IUPAC Name** | 3-phenyl-4*H*-1-benzopyran-4-one |
| **Smiles ACDChemSketch$_1$** | O=C1c2ccccc2OC=C1c1ccccc1 |
| **Smiles ChemAxon$_2$** | O=C1C(=COC2=C1C=CC=C2)C1=CC=CC=C1 |
| **Smiles JME editor$_3$** | O=c2c(c1ccccc1)coc3ccccc23 |
| **Smiles JME editor$_3$** | C=13(C(=CC=CC=1)C(C(C=2(C=CC=CC=2))=CO3)=O) |
| **Smiles PubChem editor$_4$** | C1=CC=CC2=C1C(C(=CO2)C3=CC=CC=C3)=O |
| **SMART PubChem editor$_4$** | c1cccc-2c1-[#6](-[#6](=[#6]-[#8]-2)-c3ccccc3)=[#8] |




InChI is the International Chemical Identifier developed by the IUPAC, a unique label for each compound

InChI=1S/C15H10O2/c16-13-10-15(11-6-2-1-3-7-11)17-14-9-5-4-8-12(13)14/h1-10H

Barry and Judy Silverman College of Pharmacy
NOVA SOUTHEASTERN UNIVERSITY | NSU Florida

# Molecular modelling software capabilities

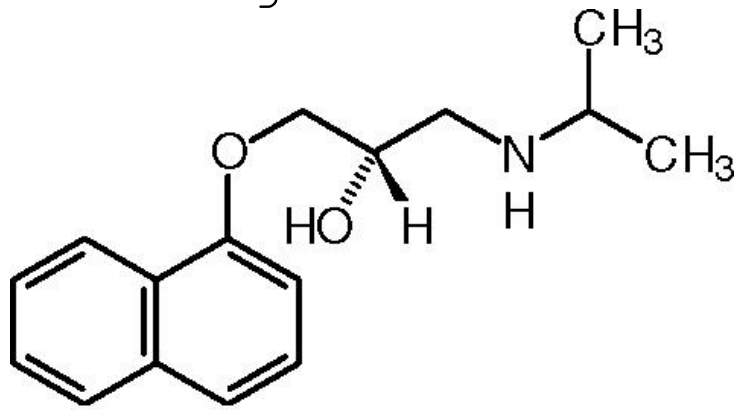
Model or mimic the behavior of molecules using computational chemistry for drug design

Molecular graphics

Perform theoretical chemistry calculations (forcefield or quantum mechanics)

interactive molecular drawing and

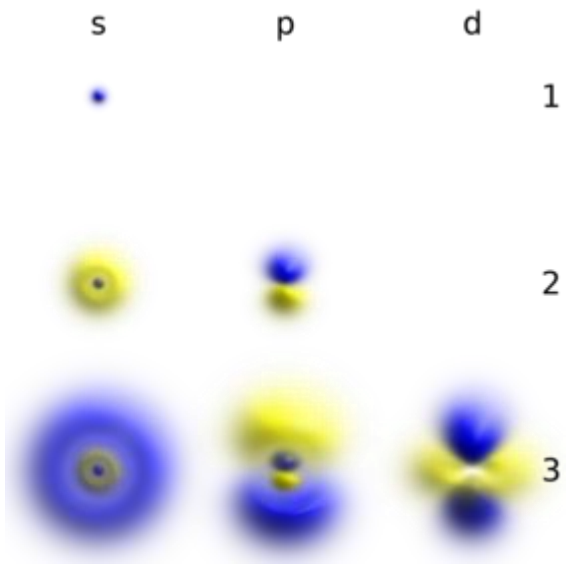
The software recognizes the molecular structures in a chemically meaningful way, *i.e.*: Not just a graph, but a chemical entity



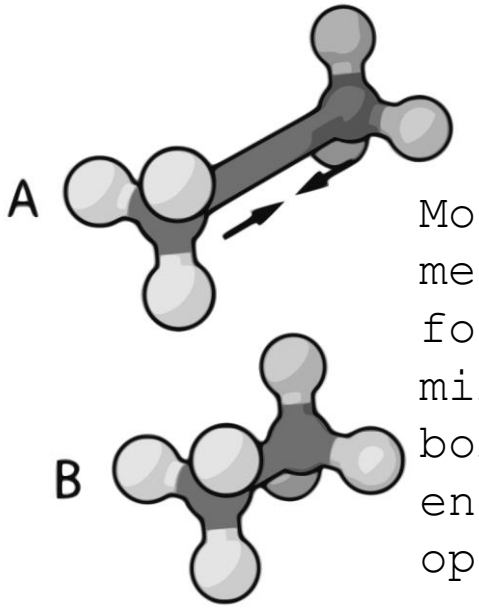

(*S*)-form

# Molecular modelling

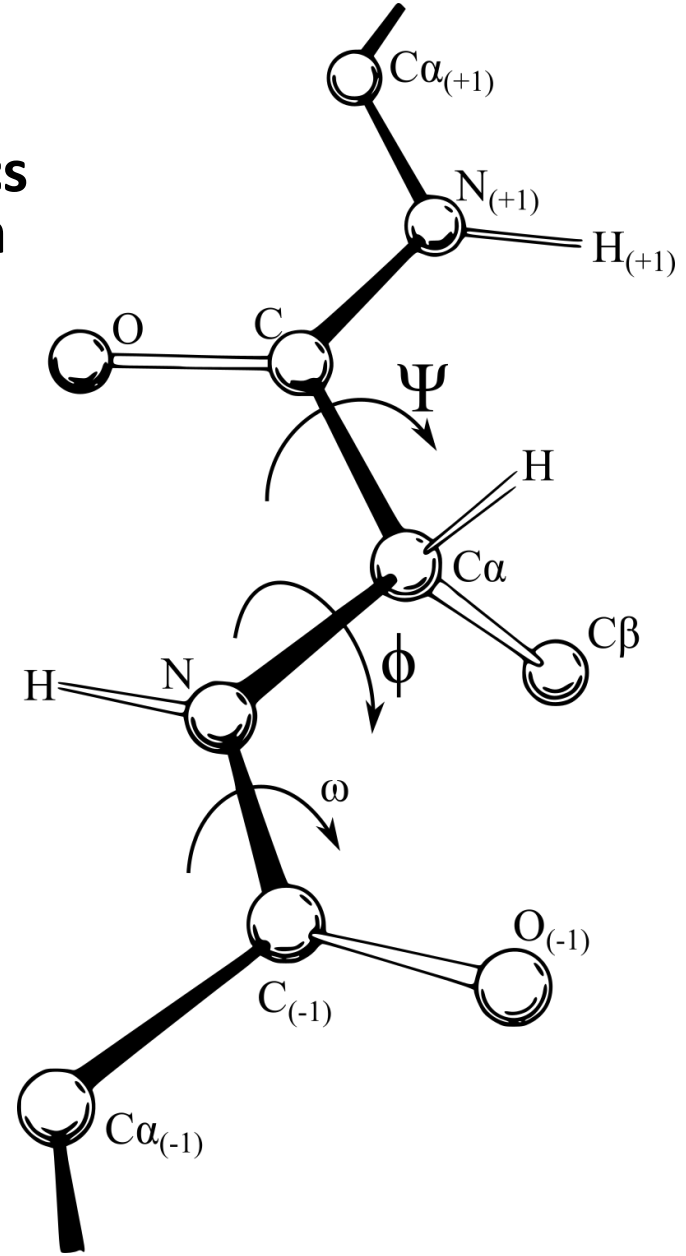
- They may treat atoms as the smallest unit (a **molecular mechanics** approach), or model protons, neutrons and electrons (a **quantum chemistry** approach)
- Methods that take advantage of data are winning as we get more data and more compute! Hence Machine Learning methods for molecular simulations (Amin, Raja, Krishnapriyan, 2025)




Probability densities of the wave functions of an electron in hydrogen atom




Molecular mechanics uses a force field to minimize the bond stretching energy and optimize angles.

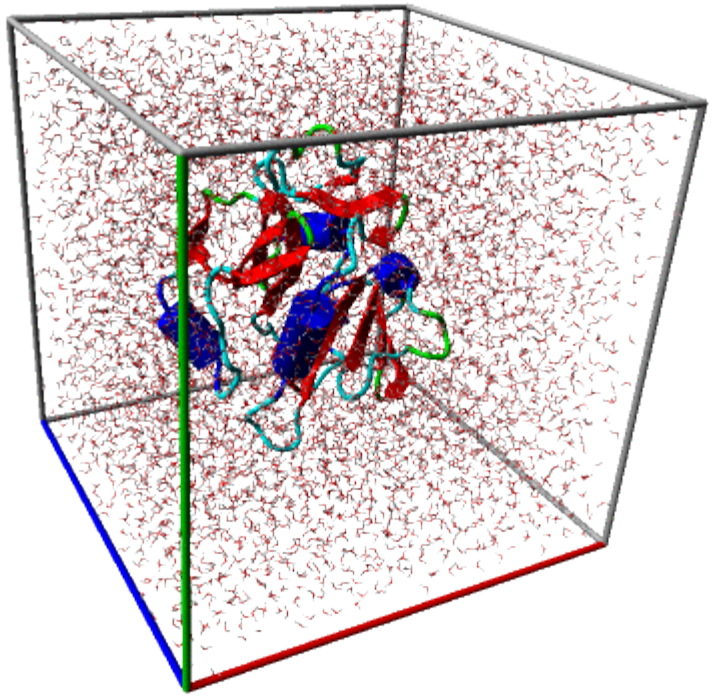



Barry and Judy Silverman College of Pharmacy
NOVA SOUTHEASTERN UNIVERSITY | NSU Florida

# Molecular modelling empowered by AI

- More accurate behavior is captured with:
  - Large scale simulations
  - Longer time scales

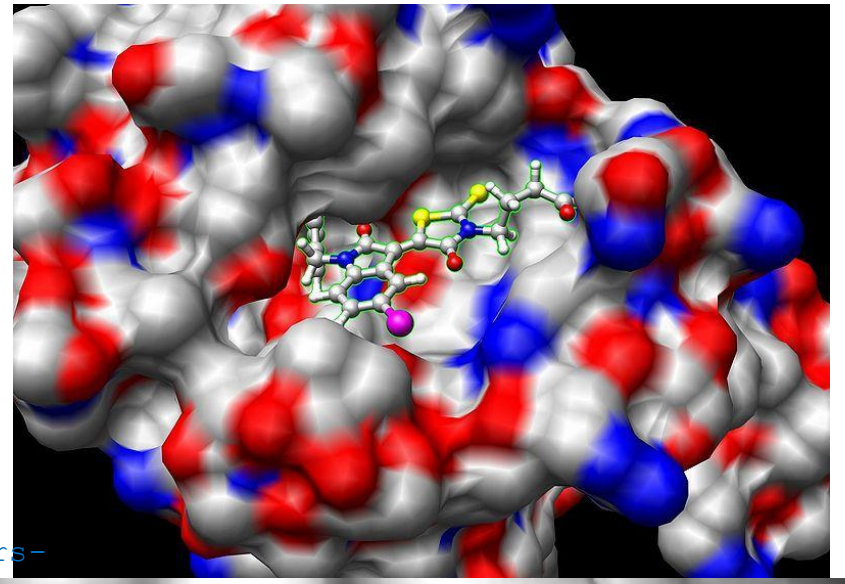
- Use available data to train AI models, one trained prediction cheap



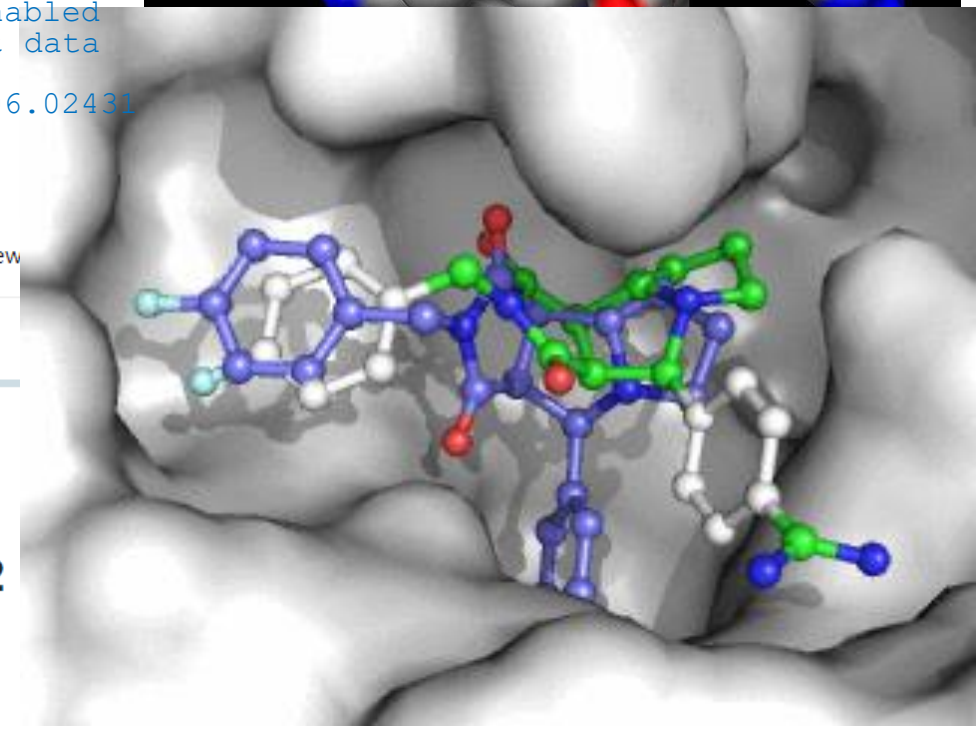
Towards Fast, Specialized Machine Learning Force Fields: Distilling Foundation Models via Energy Hessians - Atomsbysics Data System

# Protein–ligand docking for virtual screening to identify hits

- Predict binding geometry of drugs to target proteins and interaction energy.
- Available software AutoDock and AutoDock Vina, rDock, FlexAID, Molecular Operating Environment, and Glide.
- Available compound libraries ($10^9$ compounds with fingerprints, molecular descriptors, etc)
- Accelerated docking protocols are necessary: the protein-ligand docking results between the prepared receptors and compounds are used to build ML data-frames for each binding site.



Babuji, Y. et al. Targeting sars-cov-2 with ai-and hpc-enabled lead generation: A first data release. http://arxiv.org/abs/2006.02431 2020.



scientific reports

View

Explore content ▾ About the journal ▾ Publish with us ▾

nature › scientific reports › articles › article

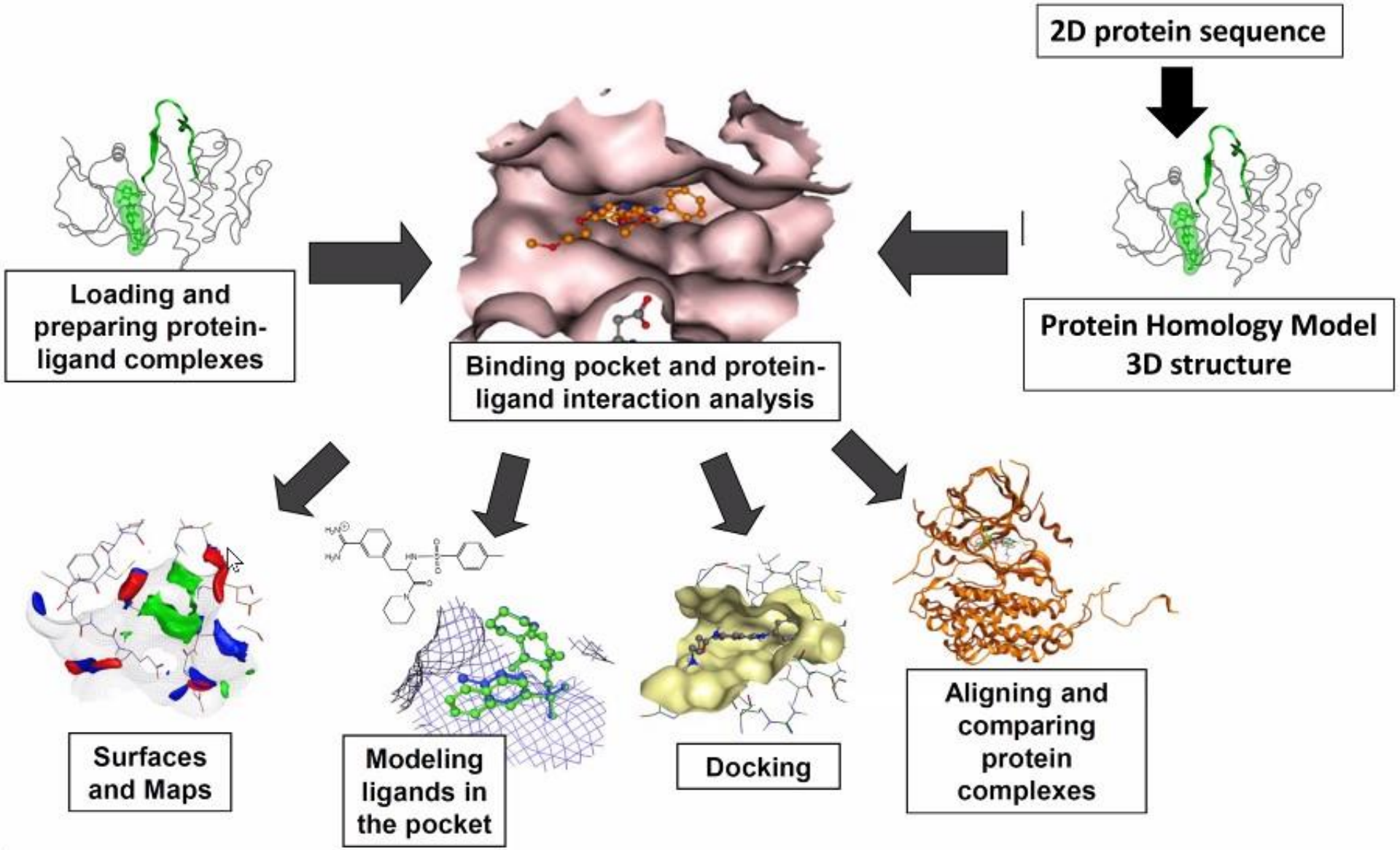
Article | Open access | Published: 06 February 2023

## AI-accelerated protein-ligand docking for SARS-CoV-2 is 100-fold faster with no significant change in detection

Austin Clyde, Xuefeng Liu, Thomas Brettin, Hyunseung Yoo, Alexander Partin, Yadu Babuji, Ben Blaiszik, Jamaludin Mohd-Yusof, Andre Merzky, Matteo Turilli, Shantenu Jha, Arvind Ramanathan & Rick Stevens

Barry and Judy Silverman College of Pharmacy
NOVA SOUTHEASTERN UNIVERSITY
NSU Florida

# Typical Structure-Based Drug Design (SBDD) Workflows



# Protein data bank

1. Protein Databank has the 3D structure of proteins **https://www.rcsb.org**
2. Various resolutions and accuracy

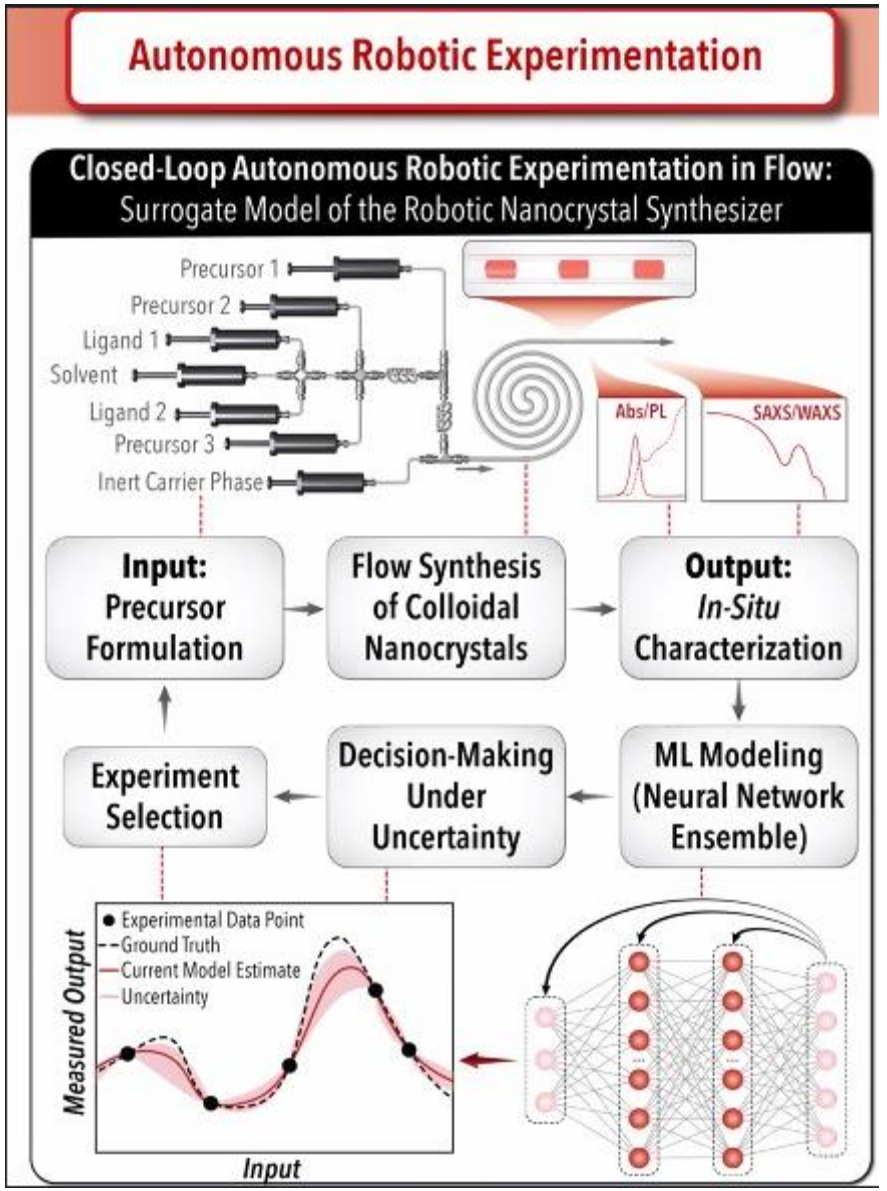
# AI in Protein Folding: AlphaFold

- The problem of predicting protein folding from amino acid sequences a grand challenge (Nobel Laureate Christian Anfinsen in 1972)
- A biennial competition called Critical Assessment of Protein Structure Prediction (CASP) was launched in 1994 to encourage efforts around the protein folding problem
- DeepMind, a UK-based AI company now part of Alphabet Inc developed a neural network model named AlphaFold which was trained on known protein structures in the protein data bank (PDB) and won the 14th CASP competition with remarkably accurate predictions within error of experiments.

https://youtu.be/cx7l9ZGFZkw?si=JvFNEcINqW-uuJjs

# Do we have enough data?

- No!
- Compare the amount of available human text that powers ChatGPT
- Complexity of systems requires complex data collection strategies
- High-throughput data collection is focused on specific research questions: (confirmation bias?)
- Phenotypic data is lagging behind




*AlphaFlow is a self-driven fluidic lab capable of autonomously discovering complex multi-step chemistries.*

*Volk AA, Abolhasani, et al, Nature Comm. AlphaFlow, 2023, https://www.nature.com/articles/s41467-023*

# A combined in vitro and in silico approach in the study of drug-induced mitochondrial dysfunction

Barry and Judy Silverman
College of Pharmacy
NOVA SOUTHEASTERN UNIVERSITY
NSU Florida

# Drug-Induced Mitochondrial Dysfunction related to organ toxicities

*"Failures to predict adverse drug reactions have immense financial implications, result in adverse human suffering and erode in trust of regulatory and pharmaceutical processes"*

Dykens, J.A. and Y. Will, *The significance of mitochondrial toxicity testing in drug development.* Drug Discovery Today, **12**(17): p. 777-785.

## Examples of drugs with Black Box warnings

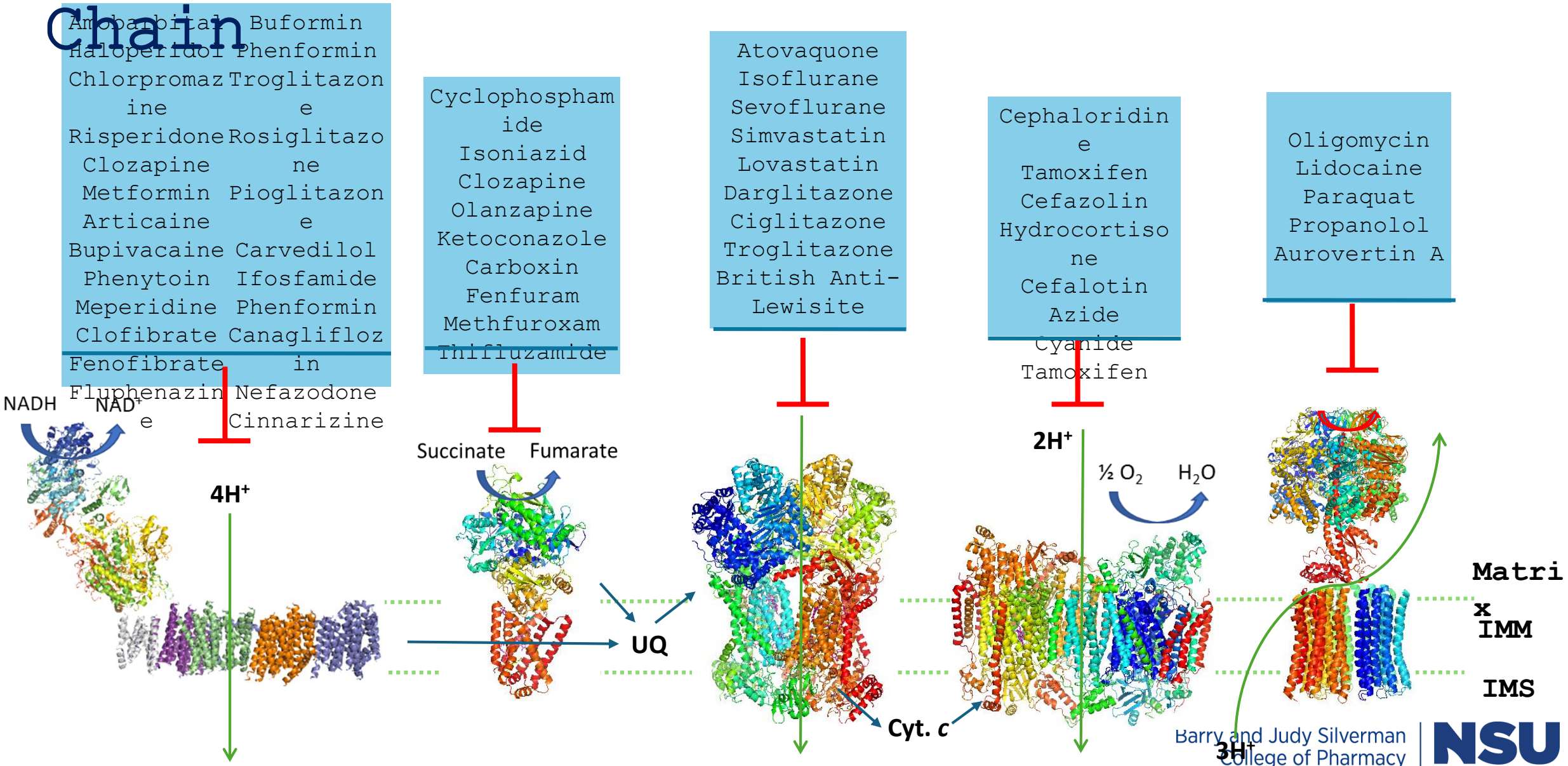
Drugs with mitochondrial liabilities in red

| Hepatic toxicity | | Cardiovascular toxicity | | | Renal toxicity |
|---|---|---|---|---|---|
| **Antivirals**<br>Abacavir<br>Didanosine<br>Emtricitabine<br>Entecavir<br>Lamivudine<br>Nevirapine<br>Telbivudine<br>Tenofovir<br>Tipranavir<br>Stavudine<br>Zalcitabine<br>Zidovudine | **CNS Drugs**<br>Dandrolene<br>Divalproex Sodium<br>Felbamate<br>Naltrexone<br>Nefazodone<br>Valproic acid | **NSAIDs**<br>Celecoxib<br>Diclofenac<br>Diflunisal<br>Etodolac<br>Fenoprofen<br>Ibuprofen<br>Indomethacin<br>Ketoprofen<br>Mefenamic acid<br>Meloxicam<br>Naproxen<br>Nabumetone | Oxaprozin<br>Piroxicam<br>Salsalate<br>Sulindac<br>Thioridazine<br>Tolmetin | **Anti-cancer**<br>Arsenic trioxide<br>Cetuximab<br>Denileukin diftitox<br>Mitoxantrone<br>Tamoxifen<br><br>**Antiarhythmic**<br>Amiodarone<br>Disopyramide<br>Dofetilide<br>Ibutilide | **Antibiotics**<br>Gentamicin<br><br>**Immunosuppressants**<br>Cyclosporin A<br><br>**Antivirals**<br>Tenofovir<br><br>**Anti-cancer**<br>Doxorubicin<br>Cisplatin<br>Ifosfamide |
| **Anti-cancer**<br>Flutamide<br>Dacarbazine<br>Gemtuzumab<br>Methotrexate<br>Pentostatin<br>Tamoxifen | **Antibiotics**<br>Isoniazid<br>Ketoconazole<br>Streptozocin<br>Trovafloxacin | **Anthracyclines**<br>Daunorubicin<br>Doxorubicin<br>Epirubicin<br>Idarubicin<br><br>**Anaesthetics**<br>Bupivacaine | **Beta-blocker**<br>Atenolol<br><br>**Anti-diabetic**<br>Pioglitazone<br>Rosiglitazone | **CNS Depressants and Stimulants**<br>Amphetamines<br>Atomoxetin<br>Droperidol<br>Methamphetamine<br>Pergolide | |

Of the 38 drugs withdrawn from the market between 1994 and 2006, the majority were hepatotoxic or cardiotoxic

Barry and Judy Silverman College of Pharmacy
NOVA SOUTHEASTERN UNIVERSITY
NSU Florida

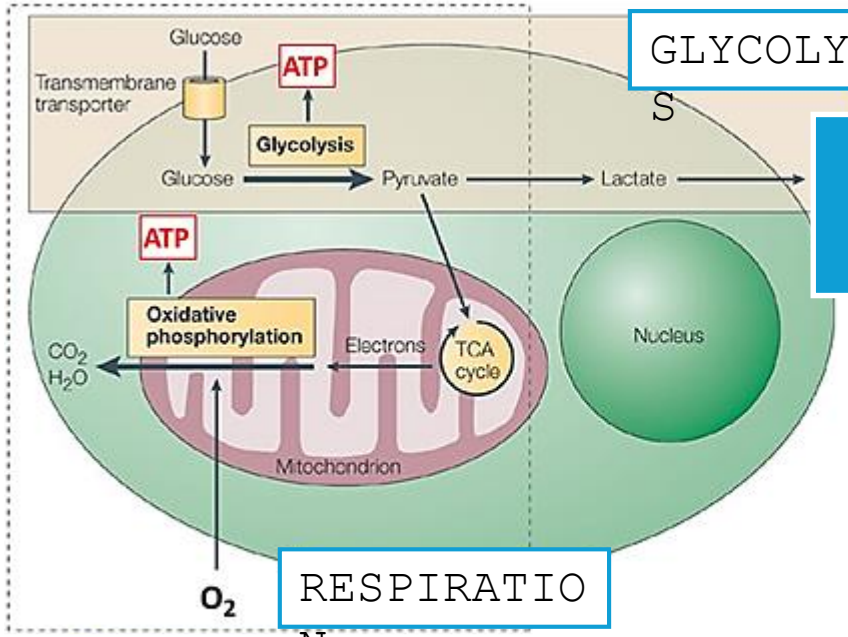
# Electron Transport Chain




Figured adapted from Ashton, T.M., et al., *Oxidative Phosphorylation as an Emerging Target in Cancer Therapy.* Clin Cancer Res, 2018. **24**(11): p. 2482-2490.

Barry and Judy Silverman College of Pharmacy
NOVA SOUTHEASTERN UNIVERSITY
NSU Florida

Seahorse Bioscience XF96

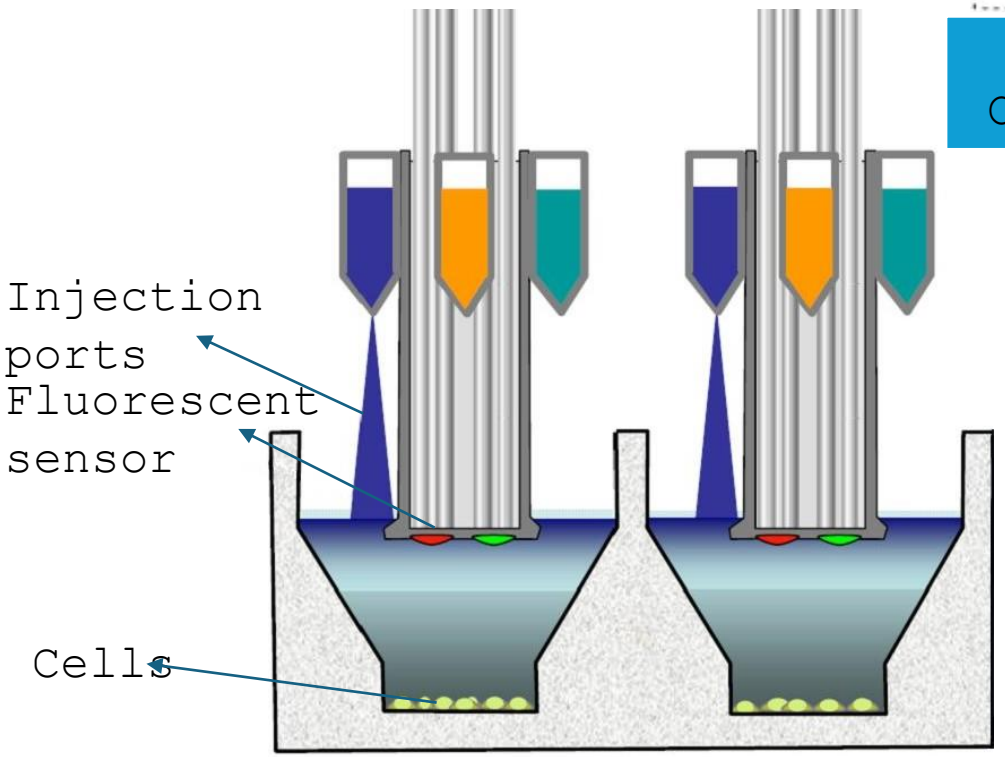



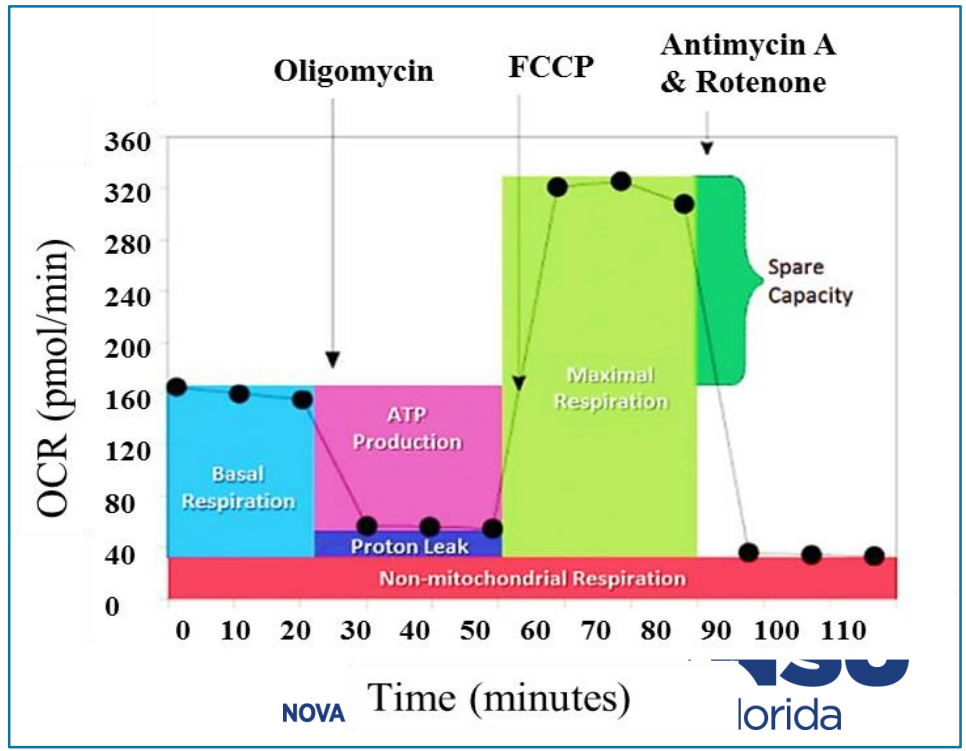

GLYCOLYSIS

**ECAR**
Extracellular acidification rate

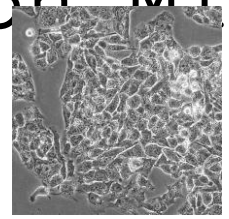
RESPIRATION

**OCR**
Oxygen consumption rate



# The Acute Extracellular Flux Assay to Investigate Real-Time Effects of Fungicides on Mitochondrial Function

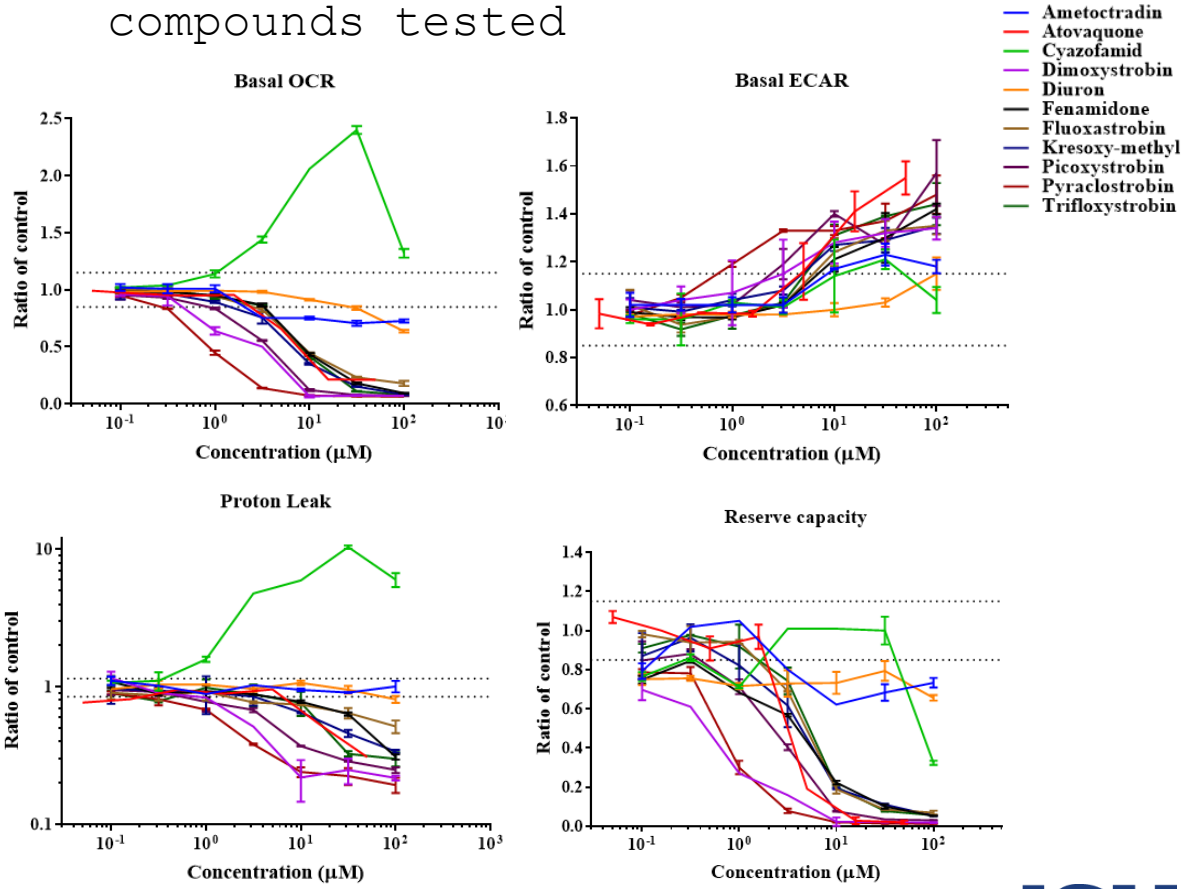


Intact HepG2 cells

Seahorse Bioscience XF96

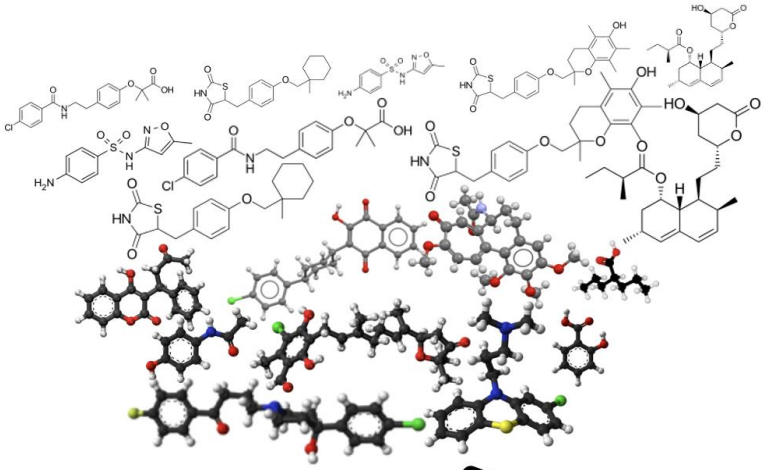
| Compound | Concentration range (µM) | Direction of change | | | | | Summary mechanism |
|---|---|---|---|---|---|---|---|
| | | OCR | Reserve capacity | ECAR | ATP | Proton leak | |
| Ametoctradin | 0.1 – 100 | ↓ | ↓ | ↑ | ↓ | NR | ETC inhibitor |
| Atovaquone | 0.05 – 50 | ↓ | ↓ | ↑ | ↓ | ↓ | ETC inhibitor |
| Cyazofamid | 0.1 – 100 | ↑ | ↓ | ↑ | ↓ | ↑ | Uncoupler |
| Dimoxystrobin | 0.1 – 100 | ↓ | ↓ | ↑ | ↓ | ↓ | ETC inhibitor |
| Diuron | 0.1 – 100 | ↓ | ↓ | ↑ | ↓ | ↓ | ETC inhibitor |
| Fenamidone | 0.1 – 100 | ↓ | ↓ | ↑ | ↓ | ↓ | ETC inhibitor |
| Fluoxastrobin | 0.1 – 100 | ↓ | ↓ | ↑ | ↓ | ↓ | ETC inhibitor |
| Kresoxim-methyl | 0.1 – 100 | ↓ | ↓ | ↑ | ↓ | ↓ | ETC inhibitor |
| Picoxystrobin | 0.1 – 100 | ↓ | ↓ | ↑ | ↓ | ↓ | ETC inhibitor |
| Pyraclostrobin | 0.1 – 100 | ↓ | ↓ | ↑ | ↓ | ↓ | ETC inhibitor |
| Trifloxystrobin | 0.1 – 100 | ↓ | ↓ | ↑ | ↓ | ↓ | ETC inhibitor |
| Rotenone | 0.003 – 1 | ↓ | ↓ | ↑ | ↓ | ↓ | ETC inhibitor |

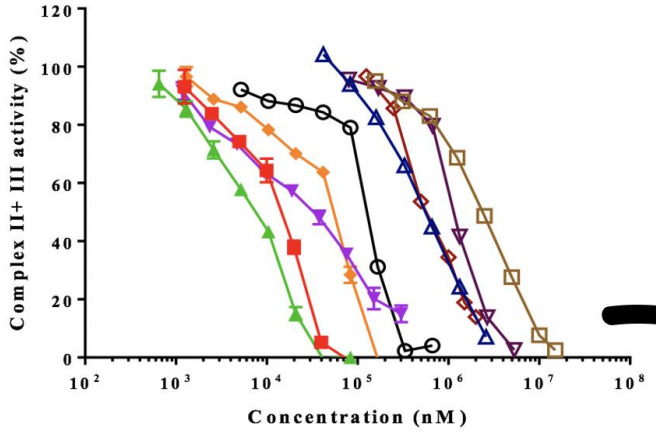
Bioenergetic profile of compounds tested



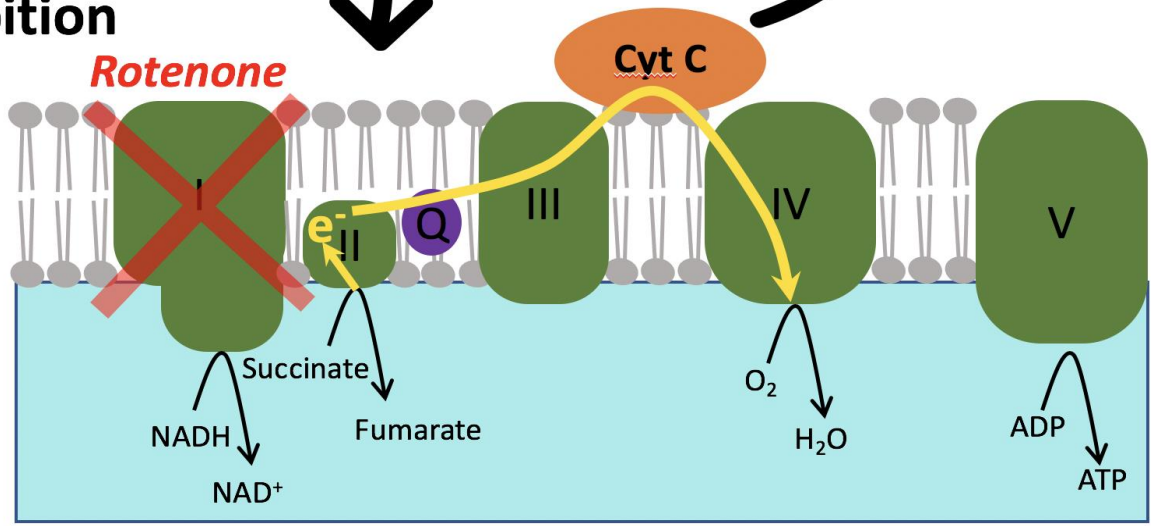

Barry and Judy Silverman College of Pharmacy
NOVA SOUTHEASTERN UNIVERSITY | NSU Florida

# Succinate-cytochrome *c* reductase activity followed by Computer-based modelling

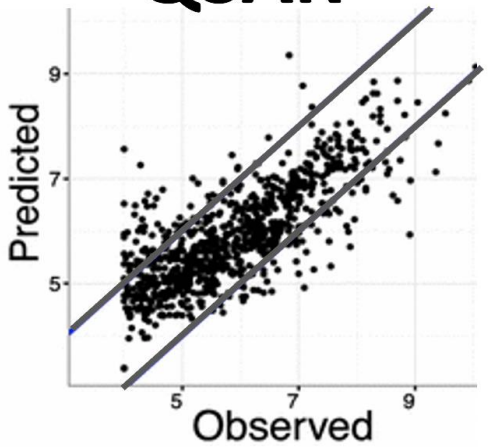





**Mitochondrial Succinate-Cytochrome-C Inhibition**




**QSAR**



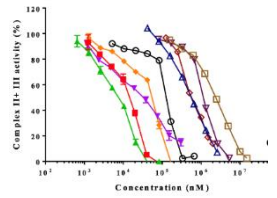

verman armacy
NOVA SOUTHEASTERN UNIVERSITY
NSU Florida

# A combined in vitro and in silico approach in the study of drug-induced mitochondrial dysfunction

**Molecular docking + PLIF analysis**

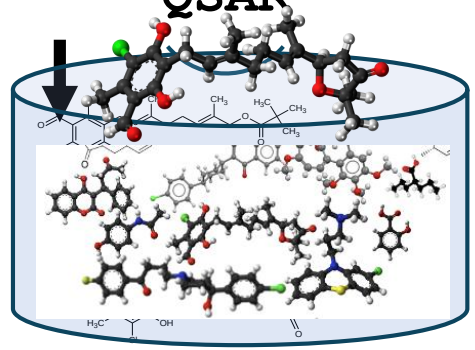
***In silico* methodologies**

***To predict future inhibitors***



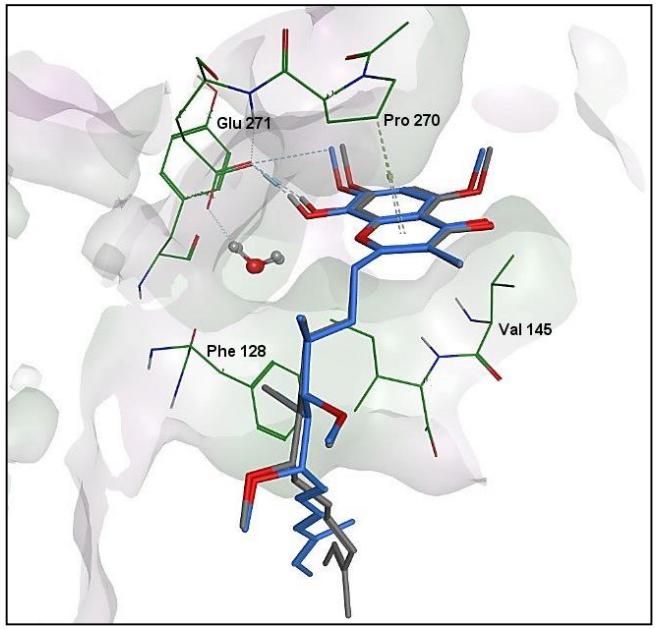

**SCR $IC_{50}$ values**

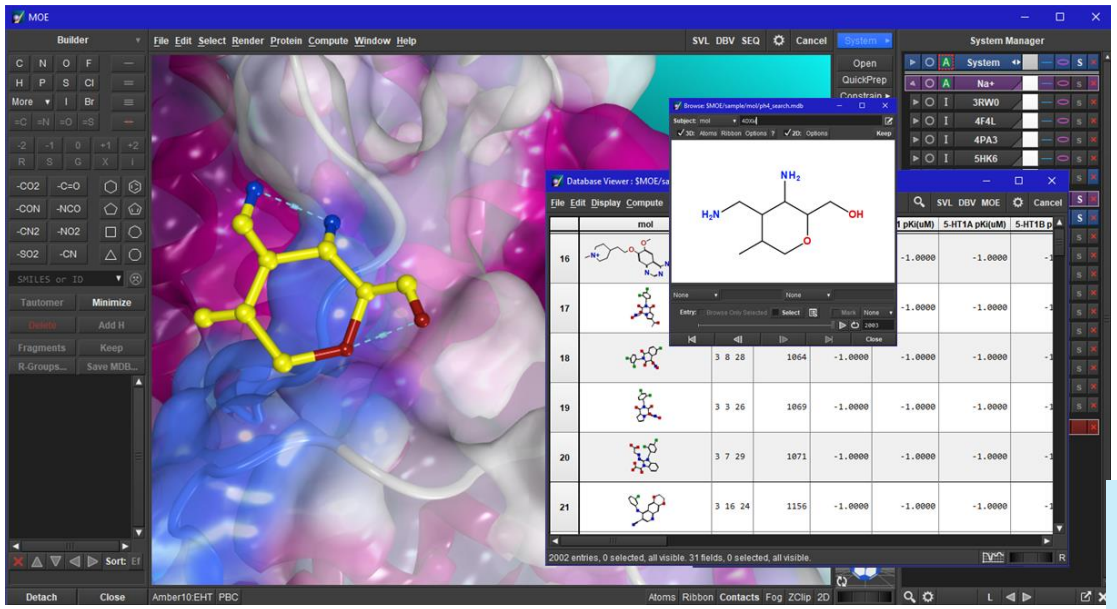
**QSAR**



**Molecular Modelling**

**Molecular Descriptors**




**Machine Learning Model**

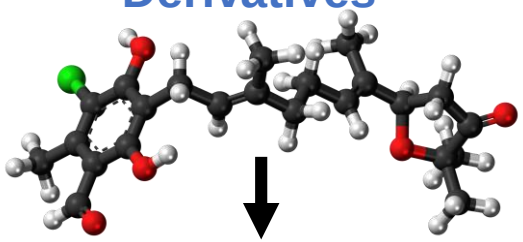
- Investigate the interaction mechanisms between the mitochondria and its inhibitors
- Calculate binding energies to mt Enzymes

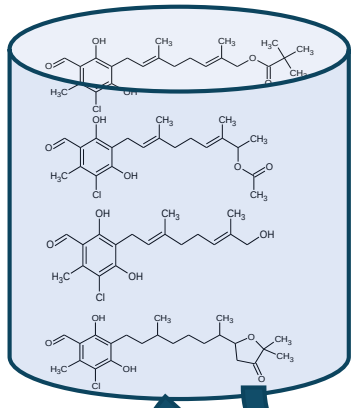
- Identify the molecular features responsible for the inhibitory activity of compounds
- Calculate IC $_{50}$s

Barry and Judy Silverman College of Pharmacy
NOVA SOUTHEASTERN UNIVERSITY
NSU Florida

# Prediction of AOX inhibitors as antifungals

Derivatives

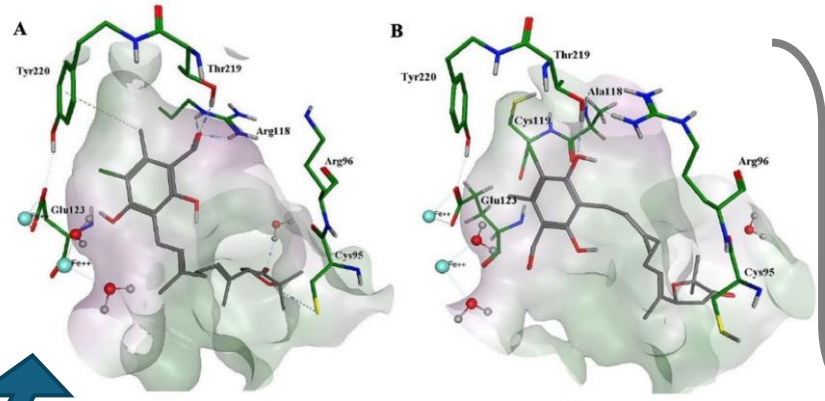




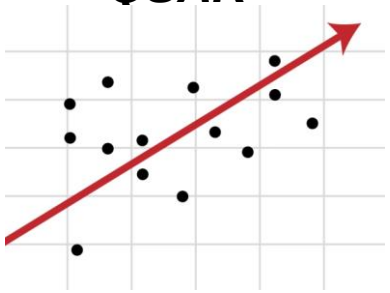
Molecular Descriptors

High-resolution respirometry to experimentally determine the $IC_{50}$ values of AOX inhibitors (recording Oxygen Consumption Rates) using an Oroboros® Oxygraph-2K and Seahorse XF

***Molecular Docking***




***QSAR***



**Predicted AOX Inhibit**

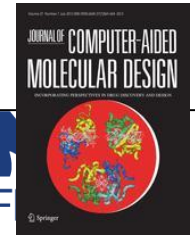
Journal of Computer-Aided Molecular Design
https://doi.org/10.1007/s10822-020-00360-8

Check for updates

**QSAR and molecular docking for the search of AOX inhibitors: a rational drug discovery approach**

Alicia Rosell-Hidalgo[1] · Luke Young[1] · Anthony L. Moore[1] · Taravat Ghafourian[1,2]

Received: 21 January 2020 / Accepted: 12 November 2020

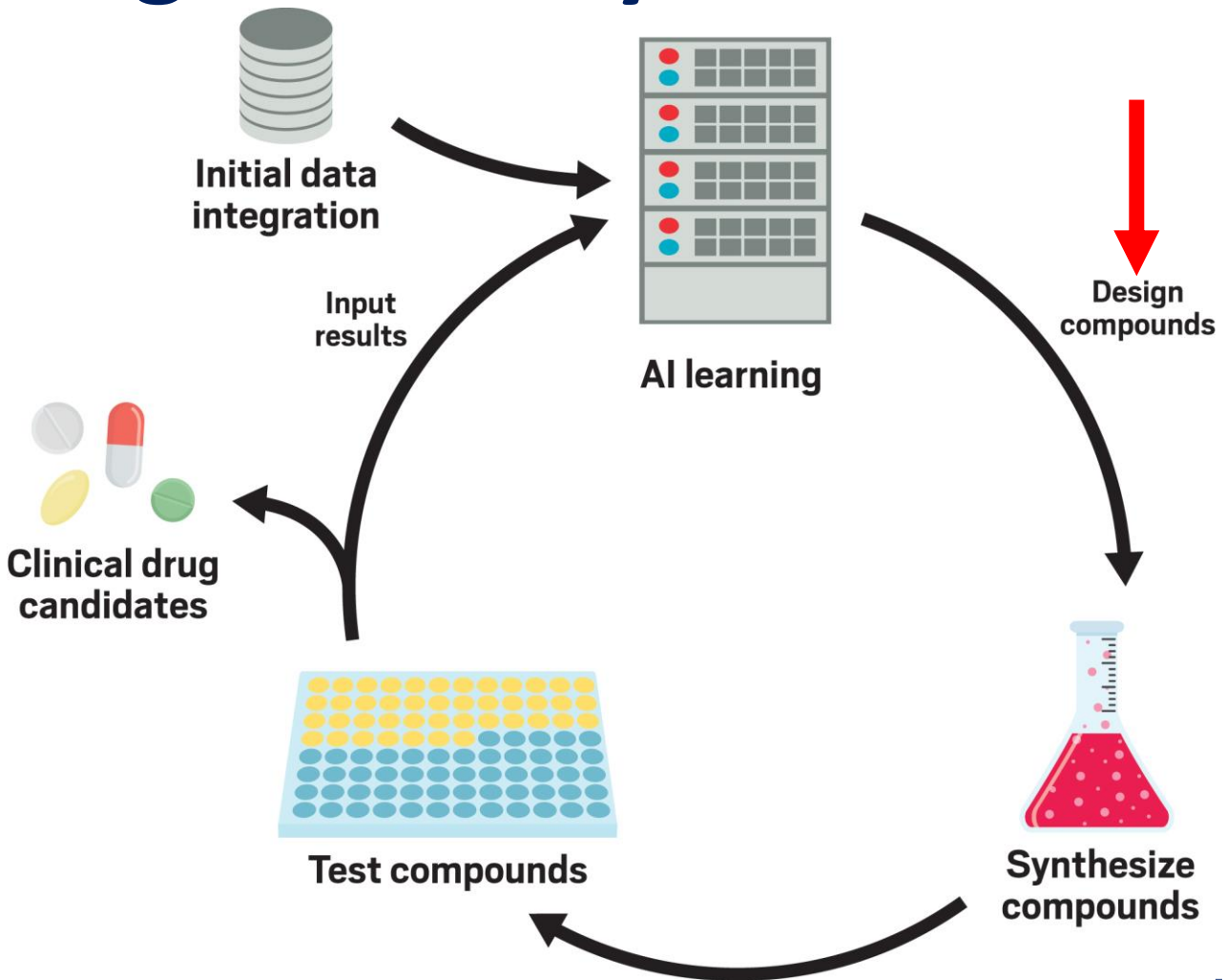




Rosell-Hidalgo A, Young L, Moore AL, Ghafourian T, QSAR and molecular docking for the search of AOX inhibitors: a rational drug discovery approach *J Comput Aided Mol Des* (2020)

Barry and Judy Silverman College of Pharmacy
NOVA SOUTHEASTERN UNIVERSITY

# AI in the Pharmaceutical Industry: artificial intelligence (AI)-based drug-discovery

- Design-Build-Test model
- The entire decision-making and design process is driven by algorithms
- **Exscientia** (Andrew Hopkins): collaborations with other drug companies—including Evotec, GlaxoSmithKline, Roche, and Sanofi
- **Insilico Medicine** (Zhavoronkov, 2013)
- **BenevolentAI** (Bryn Williams-Jones): Partnered with AstraZeneca and Merk



| Treatment | Organization | Description | Phase | Lead indication |
|---|---|---|---|---|
| REC-2282 | Recursion | Small molecule pan-HDAC inhibitor | 2/3 | Neurofibromatosis type 2 |
| REC-994 | Recursion | Small molecule superoxide scavenger | 2 | Cerebral cavernous malformation |
| REC-4881 | Recursion | Small molecule inhibitor of MEK1 and MEK2 | 2 | Familial adenomatous polyposis |
| INS018_055 | InSilico Medicine | Small molecule inhibitor | 2 | Idiopathic pulmonary fibrosis |
| BEN-2293 | BenevolentAI | Topical pan-tyrosine kinase inhibitor | 2a | Atopic dermatitis |
| EXS-21546 | Exscientia and Evotec | $A_{2A}$ receptor antagonist | 1b/2 | Solid tumors carrying high adenosine signatures. |
| RLY-4008 | Relay Therapeutics | Inhibitor of FGFR2 | 1/2 | FGFR2-altered cholangiocarcinoma |
| EXS-4318 | Exscientia | PKC-θ inhibitor | 1/2 | Inflammatory and autoimmune conditions |
| BEN-8744 | BenevolentAI | Small molecule PDE10 inhibitor | 1 | Ulcerative colitis |
| Undisclosed | Recursion | Small molecular inhibitor of RBM39, a CDK12-associated protein | Pre-clinical | HRD-negative ovarian cancer |

*C. Arnold, Nature Medicine (2023) pp.1292-1295 https://www.nature.com/articles/s41591-023-02361-0/tables/1*

Aug 2024: Merger of Exscientia with Recursion:

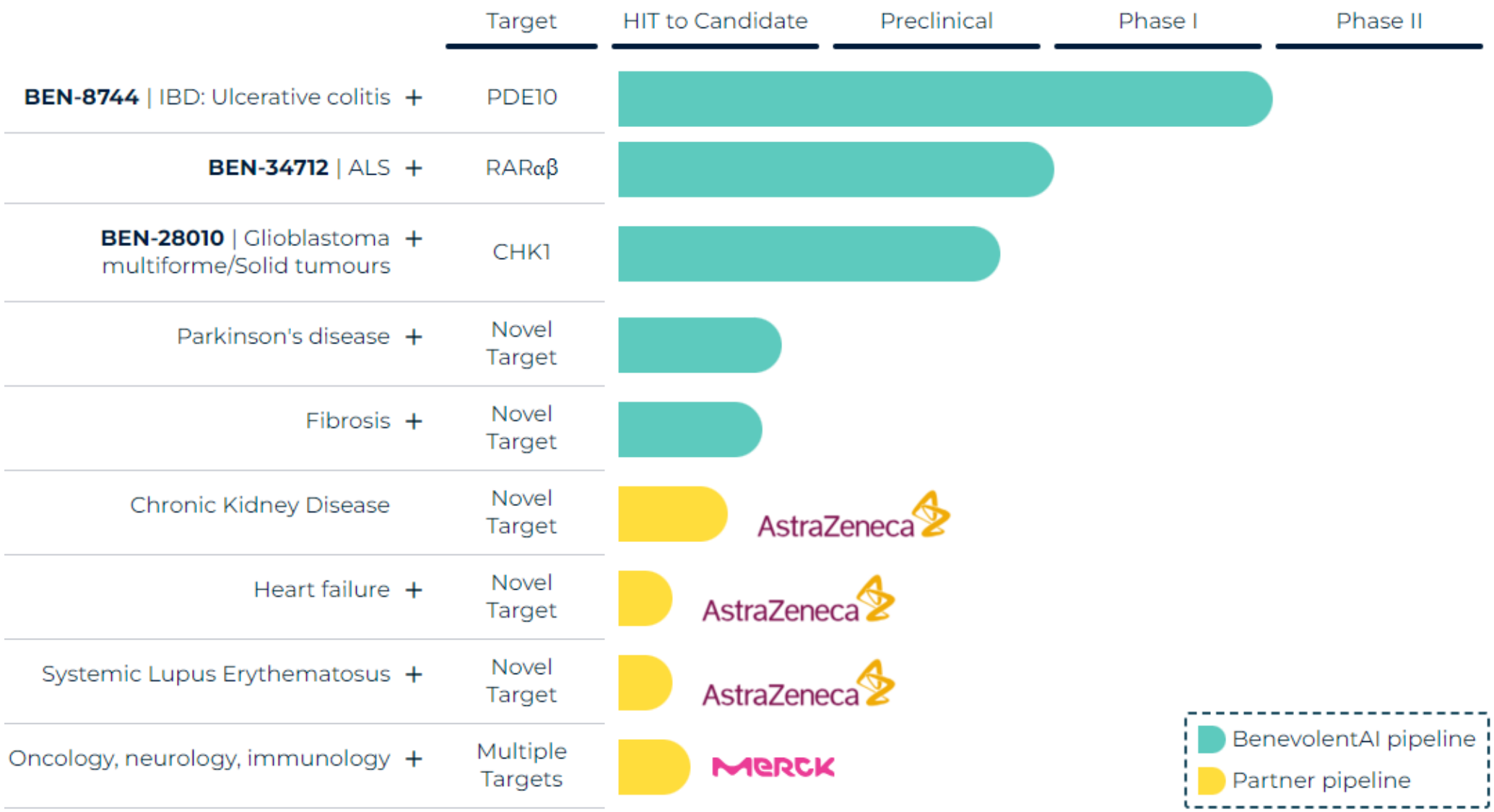
Dec 2024: Recursion Reports Interim Phase 1 Clinical Data for REC-617 Monotherapy, a CDK7 Inhibitor

REC-3565 is a potential best-in-class MALT1 inhibitor for multiple hematology indications, designed to reduce the risk of hyperbilirubinemia, a common side effect of other MALT1 inhibitors

REC-4539, an LSD1 inhibitor, is the first designed to be reversible and CNS penetrant for small-cell lung cancer

*https://ir.recursion.com/news-releases/news-release-details/recursion-announces-two-key-investigational-*

# BenevolentAI drug pipeline




*Retrieved: Feb, 2025*

# Example AI system in the Pharmaceutical Industry

- R2E (Retrieve to Explain) system combines AI-driven information retrieval (scientific literature, genomics, and clinical information) with explainable predictions to prioritize drug targets

# Conclusion

- AI can enable wider accessibility of drug discovery to smaller, minimally funded entities (democratization of drug discovery)
- Researchers and start-ups can benefit greatly within the competitive market of the pharmaceutical industry.
- A scientist and a health professional of the future will need to be able to operate the AI systems (prompt engineering etc), if not build them!
- Complex machine-learning algorithms may grow in the future, for example, by incorporating biological/clinical data into AI models for personalized dosing and precision medicine.
- The prediction quality of AI models, including their accuracy, generalizability, reliability, and interpretability, is a major consideration, especially when used for regulatory purposes (e.g., FDA submissions) and clinical decision-making.
- The black-box nature of AI systems hinders their