Titli Sarkar, Ph.D.
ATOM Data Scientist
Biomedical Informatics and Data Science (BIDS), FNLCR
Email: titli.sarkar@nih.gov

ATOM

Accelerating Therapeutics for Opportunities in Medicine

Opportunity for Artificial Intelligence and Machine Learning in Cancer Drug Discovery

Frederick National Laboratory for Cancer Research

ATOM Workflow

AI and computing-driven molecular discovery and optimization

Molecular data library
Open curated and model-ready database

Working Compound Library

1. AMPL

Property Prediction Models
- Efficacy
- Safety
- PK
- Developability

Machine learning parameter models

Human-level system models

2. GMD

Generative Molecular Design
Proposes new molecules with optimized properties

Retrain property prediction models

Active learning
- Analyze prediction uncertainty
- Specify experiments

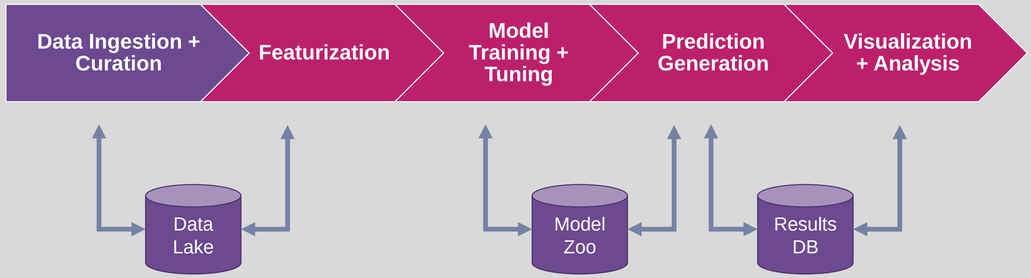
Chemistry Synthesis and Assays

Multi-Parameter Optimization Loop

Design Criteria

Experiment

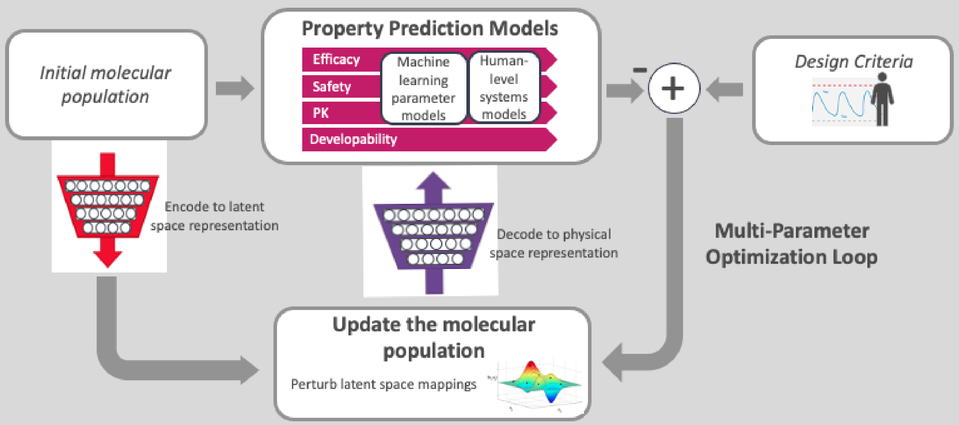
1. AMPL Pipeline



1. AMPL (ATOM Modeling PipeLine) :
Predictive modeling: training, optimization, and property prediction

2. GMD Design Workflow

Cancer Drug Discovery (new molecules Optimized Properties)



2. GMD: A high-performance platform **for parallel optimization** of efficacy, safety, and pharmacokinetics

Publicly Available Data and Models through:

ATOM Science-org/AMPL

GitHub

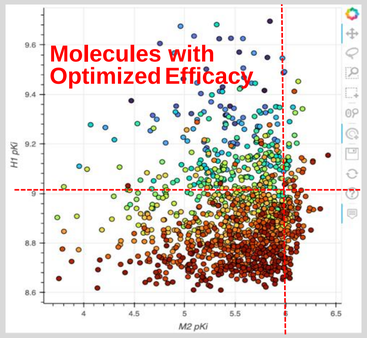
- AMPL Tutorial
- Curated dataset
- AMPL Predictive Models

Model and Data Clearinghouse (MoDaC)

NIH NATIONAL CANCER INSTITUTE Predictive Oncology Model & Data Clearinghouse

- A data-sharing repository
- Resources include datasets and software models
- Searchable against metadata and downloadable

GMD Pilot Project: Neurocrine H1 Design



X-axis: **M2** muscarinic receptor **pKi < 6**
Y-axis:**H1** histamine receptor **pKi > 9**
(Publication will be available soon)