Accelerating Molecular Glue Discovery through AI/ML & Computation

Elena Dolgikh | 2nd Molecular Glue Drug Development Summit | January 31st, 2024

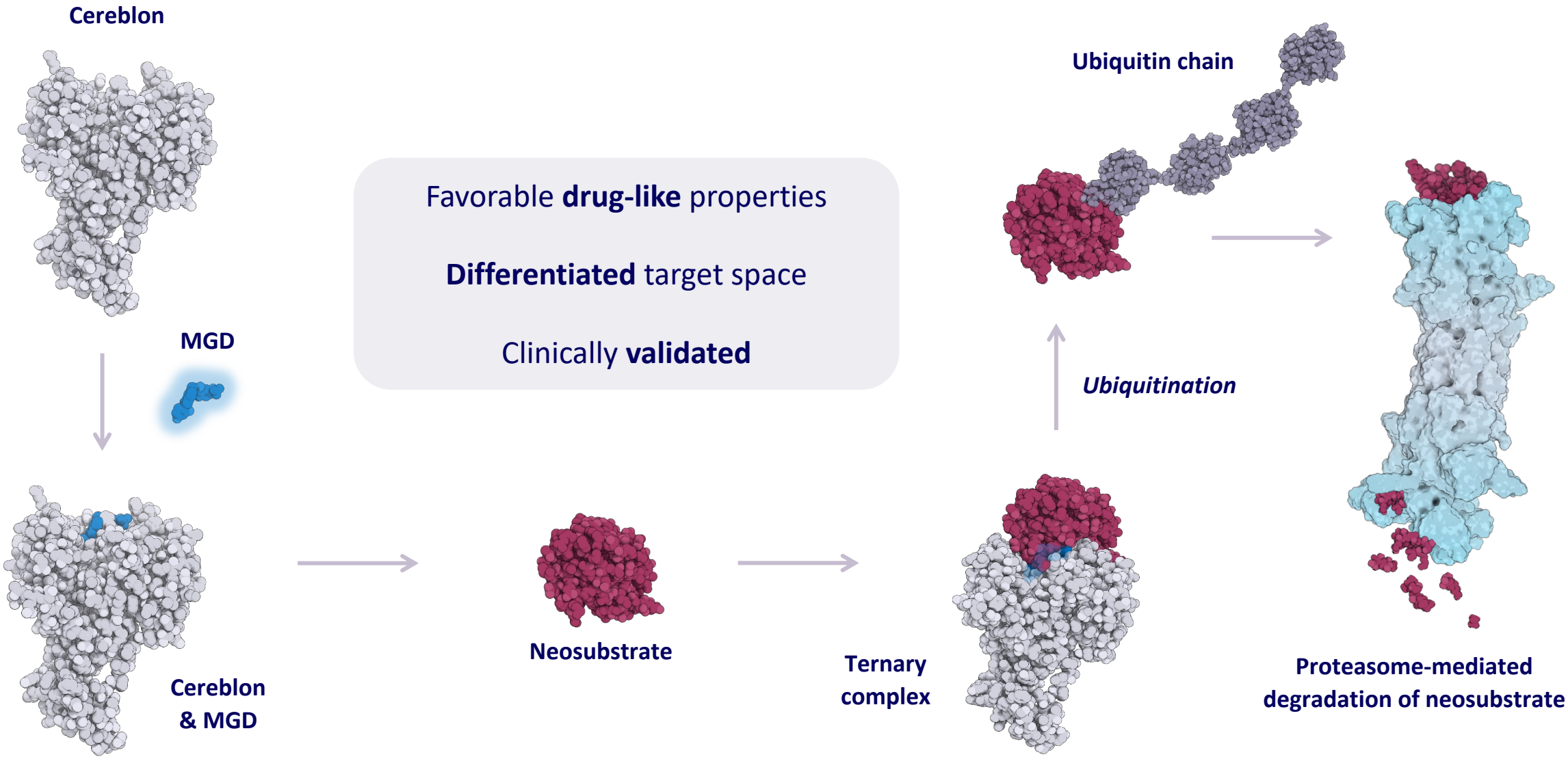
Monte Rosa Therapeutics

Molecular Glue Degraders are a Clinically Validated Modality

Favorable **drug-like** properties

Differentiated target space

Clinically **validated**



Monte Rosa Pipeline

Program/ Target	Indication(s)	Discovery / IND-Enabling / Clinical	Next Anticipated Milestone	Ownership
MRT-2359 (GSPT1)	NSCLC, SCLC and other MYC-driven Malignancies		RP2D in Q2 2024	
MRT-6160 (VAV1)	Autoimmune Disease		IND in 1H 2024	
NEK7	Inflammatory Diseases		Development candidate in Q1 2024	
CDK2	Ovarian Cancer, Breast Cancer		Development candidate in 2024	
Discovery Targets	Multiple		Lead optimization	
Discovery Targets	Oncology and Neurological Diseases		Undisclosed	Roche

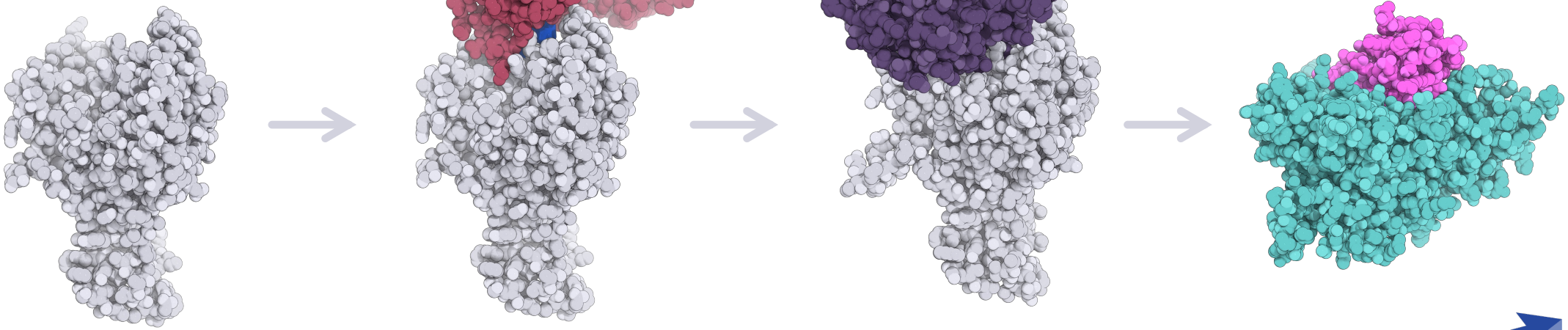
Oncology Immunology Inflammation Various

Our Rational Approach to Unleash the Full Potential of MGDs

Neosubstrate
(e.g. IKZF1)

MGD
(e.g. Lenalidomide)

Cereblon



Unraveling the Canonical Degrome

Beyond the Canonical Degrome

Beyond Cereblon

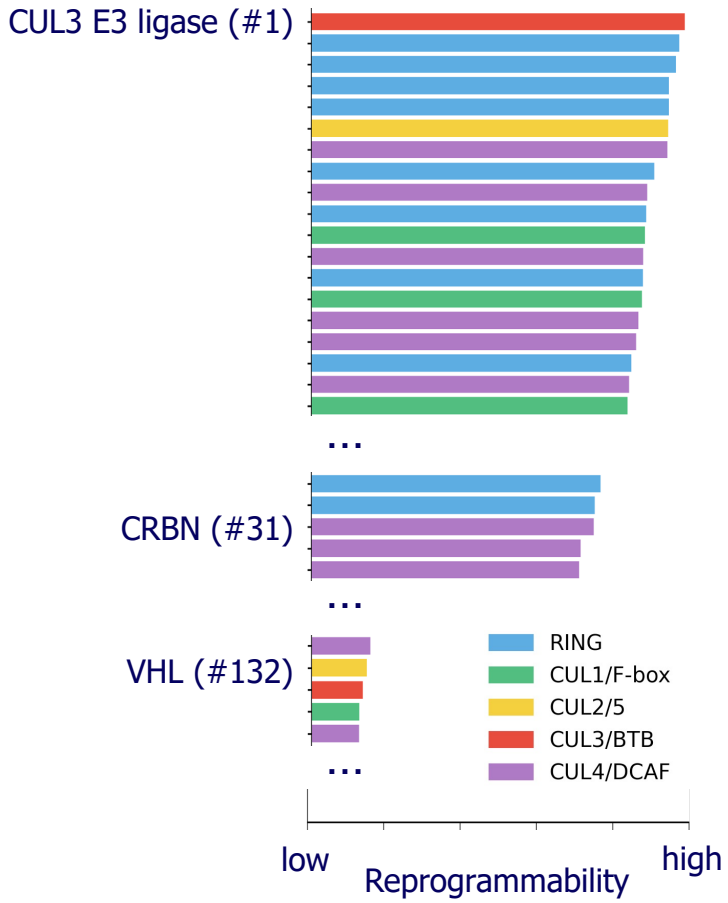
Chemical Space

Target Space

Expanding the Degradable Proteome

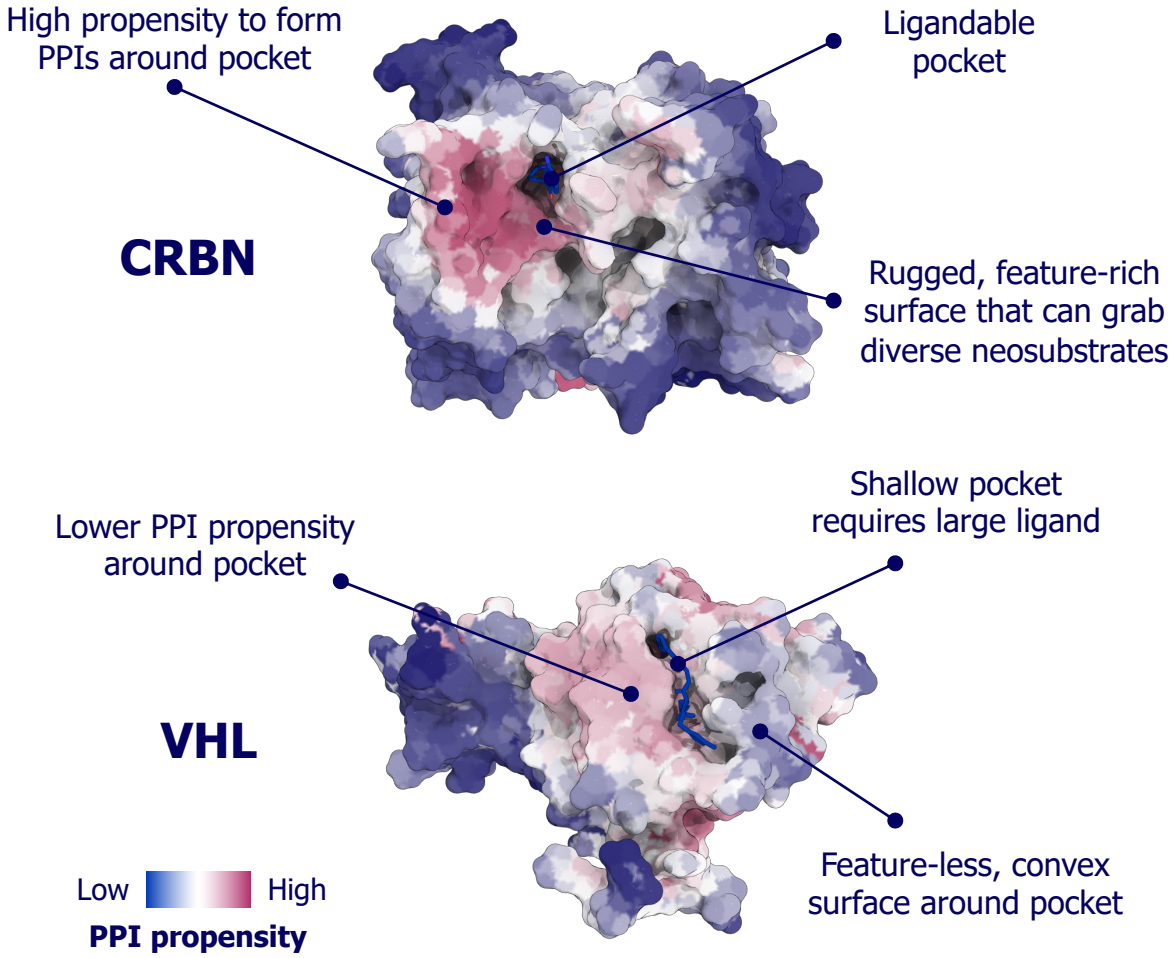
Prioritizing E3 Ligases Based on Reprogrammability Potential

fAIceit AI-powered surface characterization



CRL & RING families shown

CRBN is one of a handful high-scoring E3 ligases
VHL scores much lower

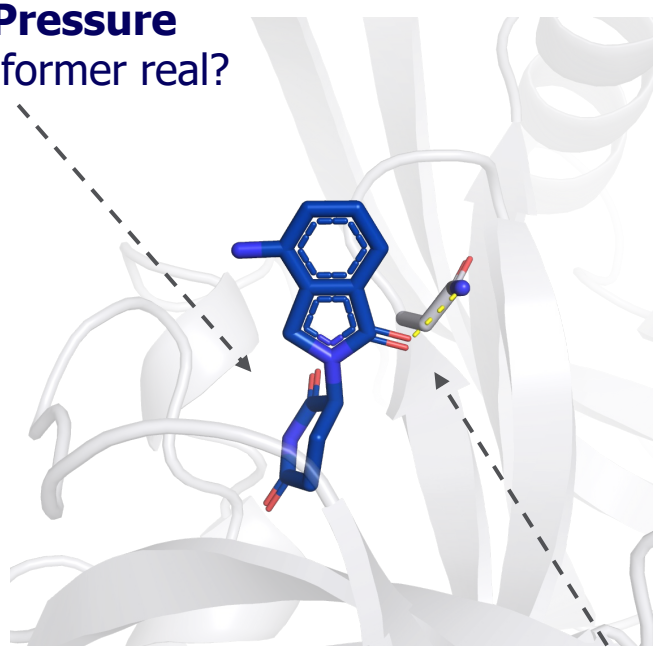


Headlong – In-house Virtual Screening Method

Combines several machine learning methods trained on large datasets

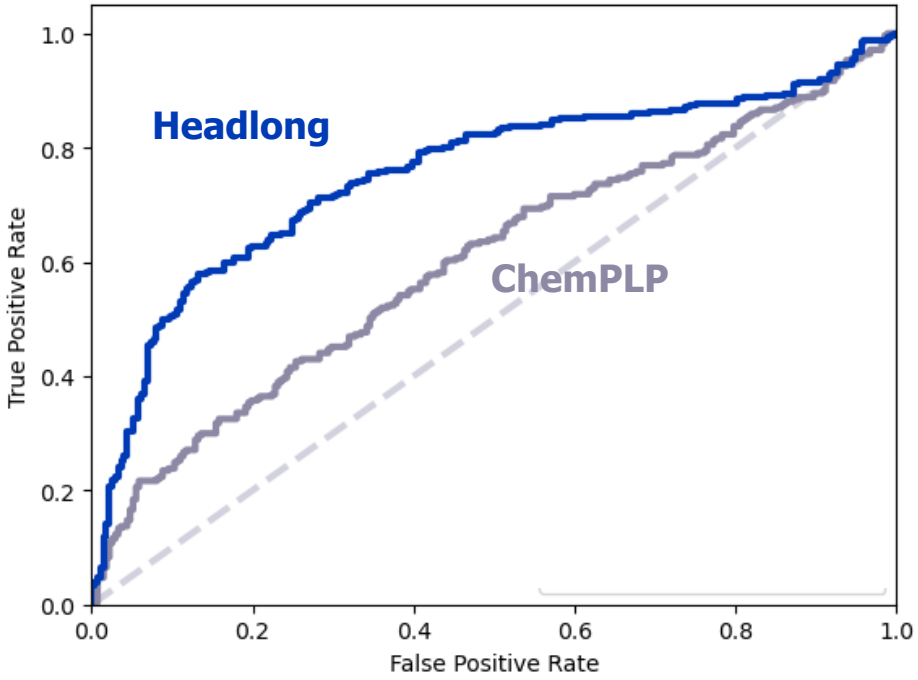
Headlong docks and evaluates docked poses

UnderPressure
Is this conformer real?



DeepScore
Is this interaction real?

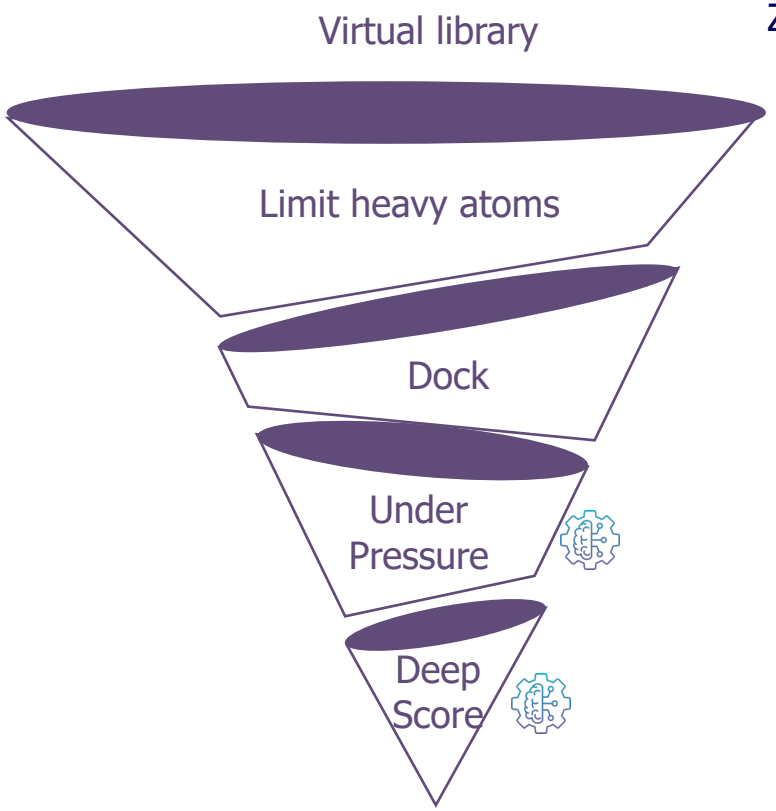
Headlong identifies CRBN active scaffolds



In-house experimentally tested benchmark
using 277 active and 273 inactive CRBN scaffolds

Headlong Identifies Candidate CRBN Scaffolds

Virtual screening of ZINC for CRBN binders



ZINC compounds
651,949,109

1,995,438
(4,404,741 poses)

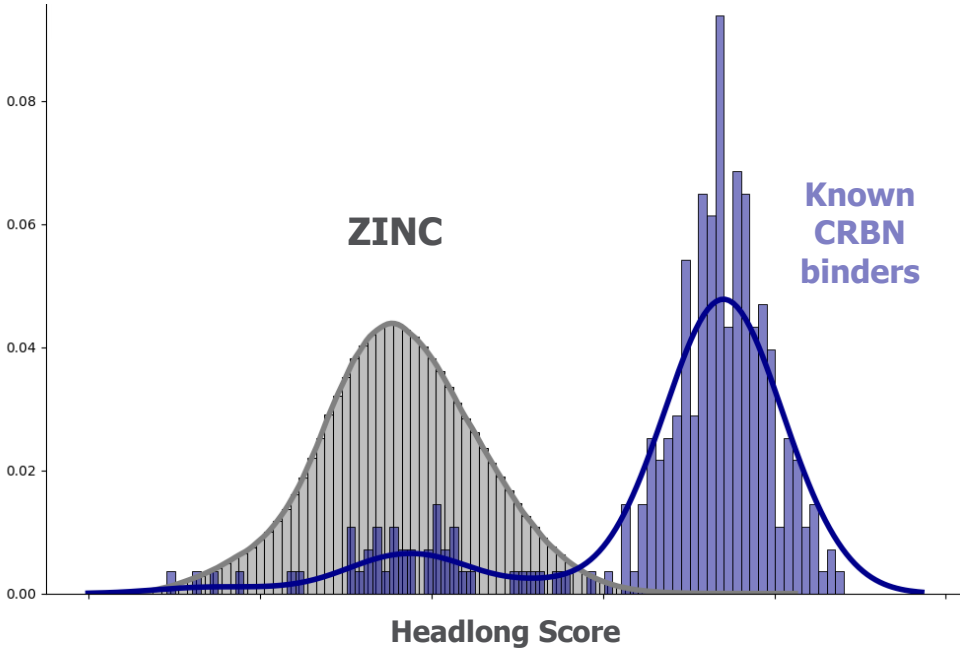
1,457,030
(2,192,037 poses)

1,117,360

1,397

75 selected

Headlong identifies known CRBN binders

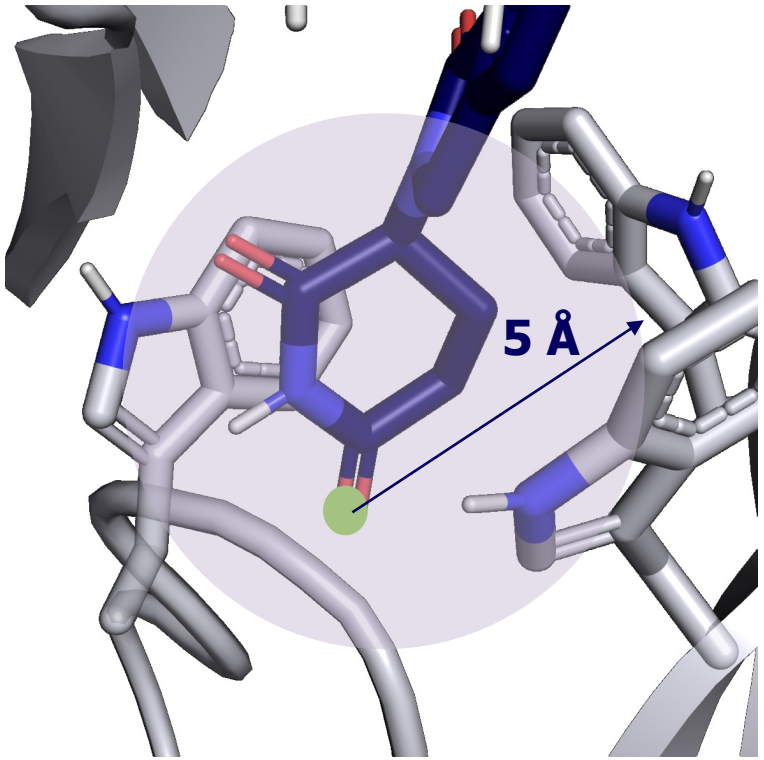


Headlong Virtual Screen
652 million ZINC compounds
200 known CRBN binders

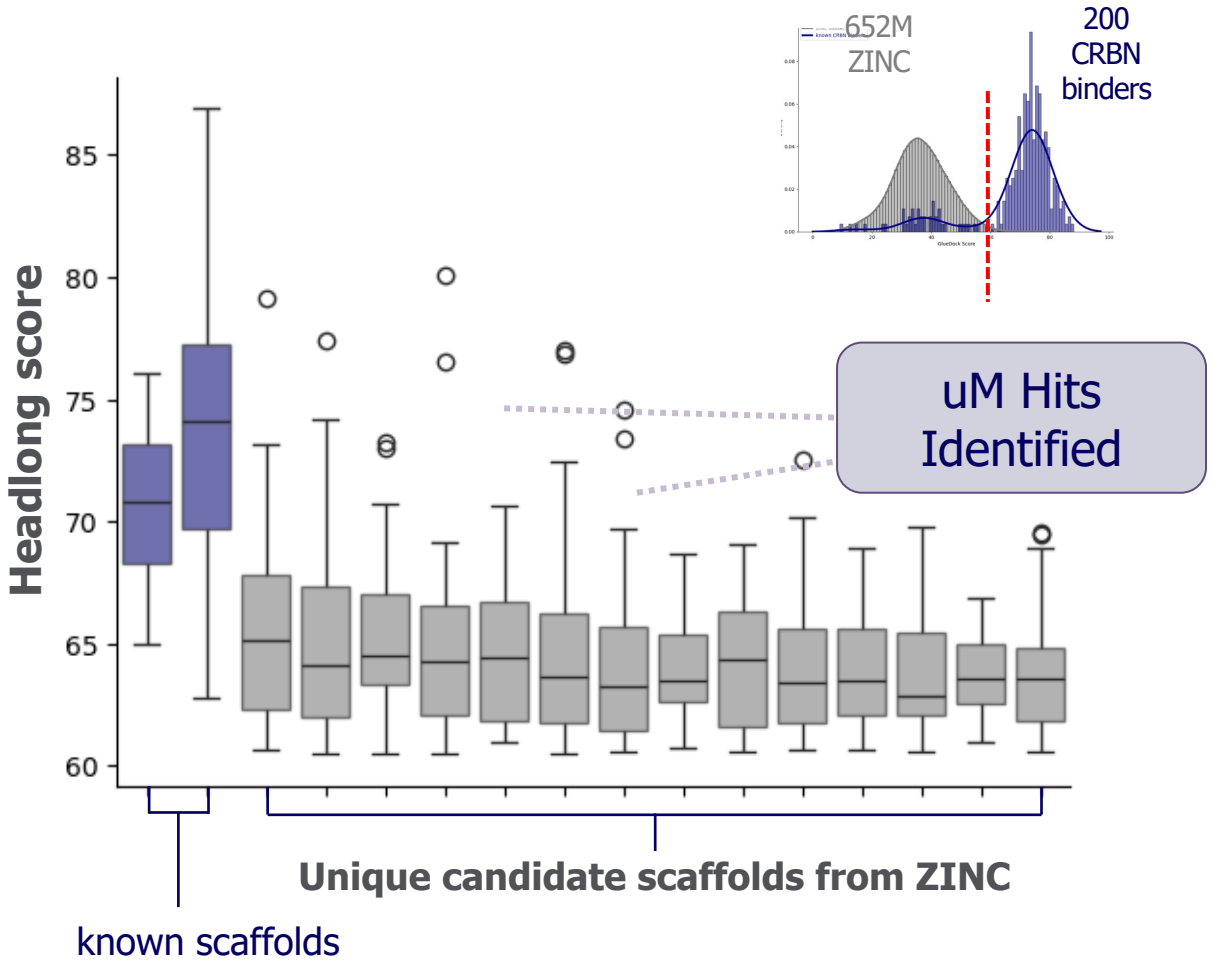
<1.5 days using 200+ AWS cores

Novel Candidate CRBN Scaffolds in Top Compound Poses

'New Scaffold': within 5Å of the CRBN tri-tryptophan cage

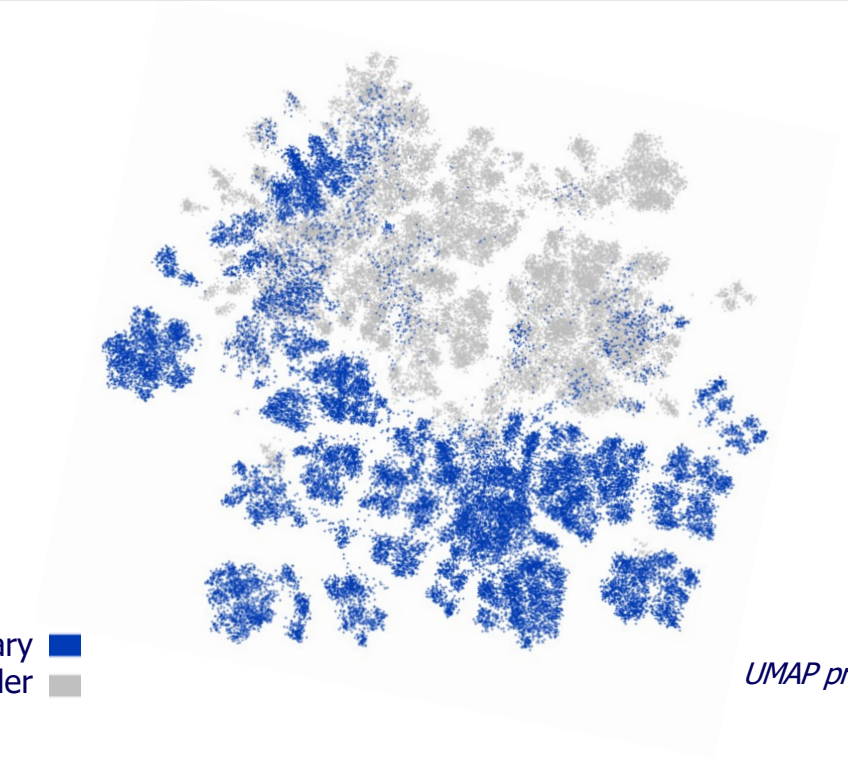


Novel candidate CRBN scaffolds



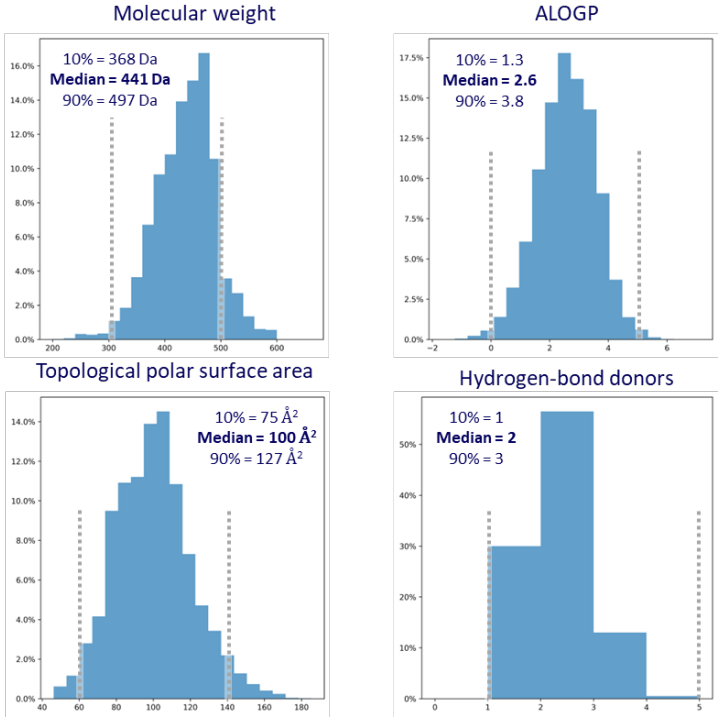
Monte Rosa MGD Library is Diverse and in Good Property Space

MRT library clearly differentiated



UMAP projection

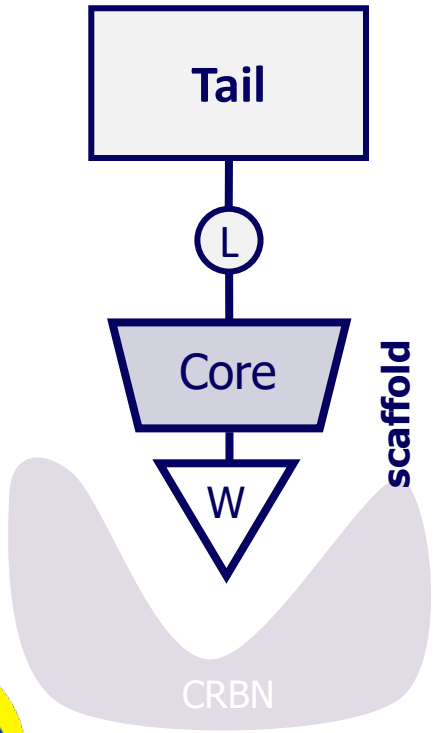
...and has drug-like properties



External space = SciFinder search based on published compounds (40K) containing glutarimide and MW < 500.
The external space may contain PROTAC-like compounds, intermediates, and MolGlue-like compounds, among others.
Internal space comprises Diversity Library, containing glutarimide and MW < 500.

FLASH Virtual Library Creation for Library Expansion, Hit-to-Lead and LO

Virtual CRBN MGD library assembly



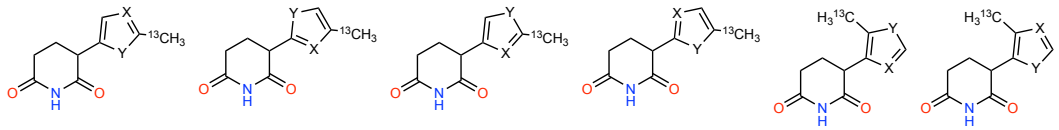
Representative set of commercial building blocks

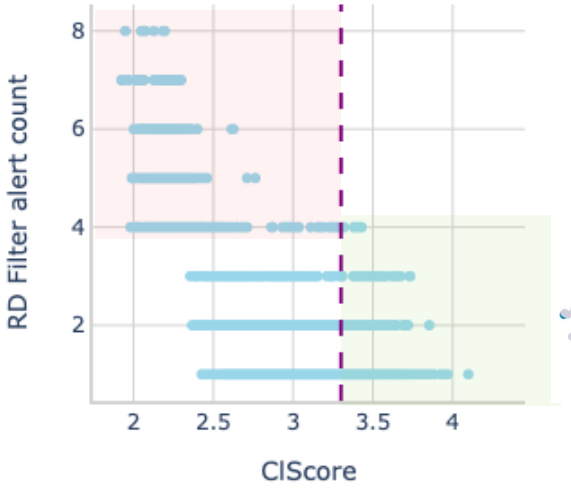
In-house database of reactions

Rule-based scaffold enumeration
- Property filtering
- Synthetic accessibility
- Headlong™ screening
- 3D clustering

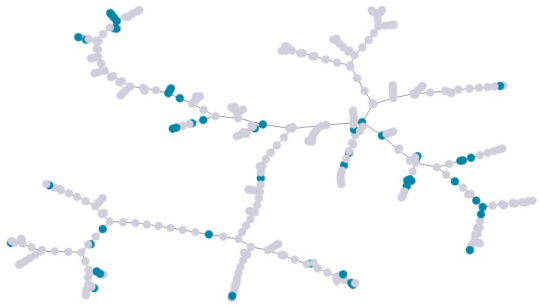
5-atom ring core example

Enumeration example: 1 core → 6 scaffolds





Properties filters



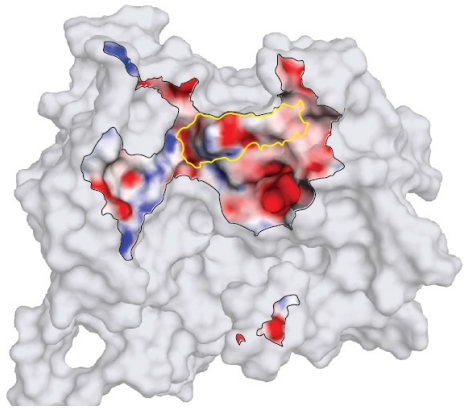
Space coverage

CRBN+MGD Neosurface Clustering to Guide Library Expansion

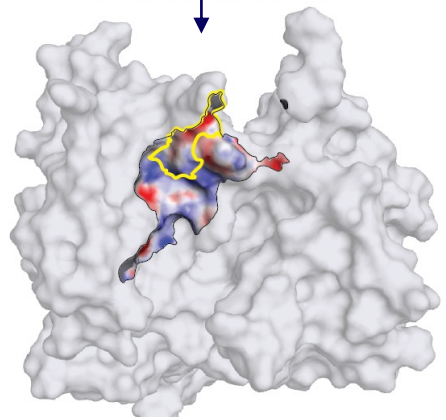
The CRBN neosurface formed by both CRBN and the MGD drives selectivity and potency

Compute electrostatic & geometric similarity between virtual compound pairs using predicted poses

Compound_i neosurface

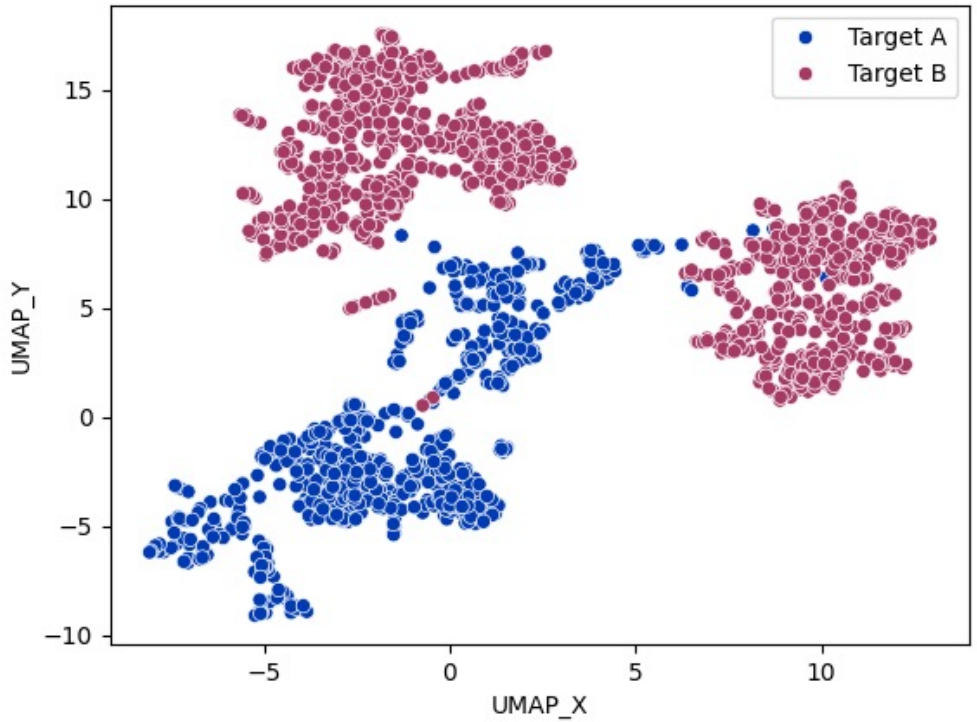


Z_ij: neosurface similarity



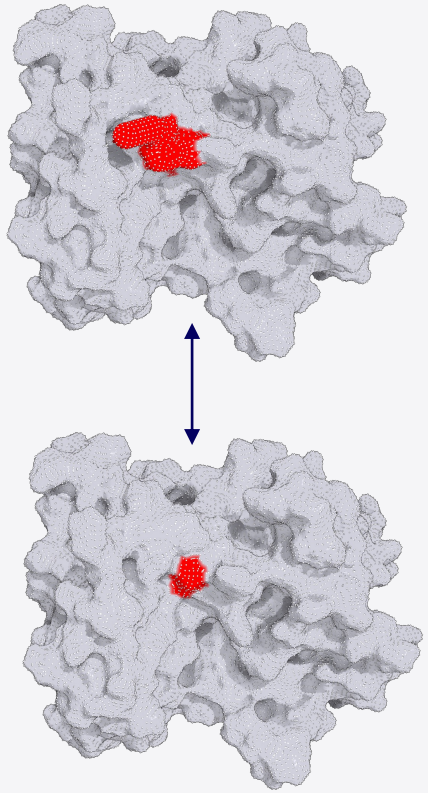
Compound_j neosurface

NEK7 vs. VAV1 active series neosurface-based clustering demonstrates unique CRBN+MGD surfaces

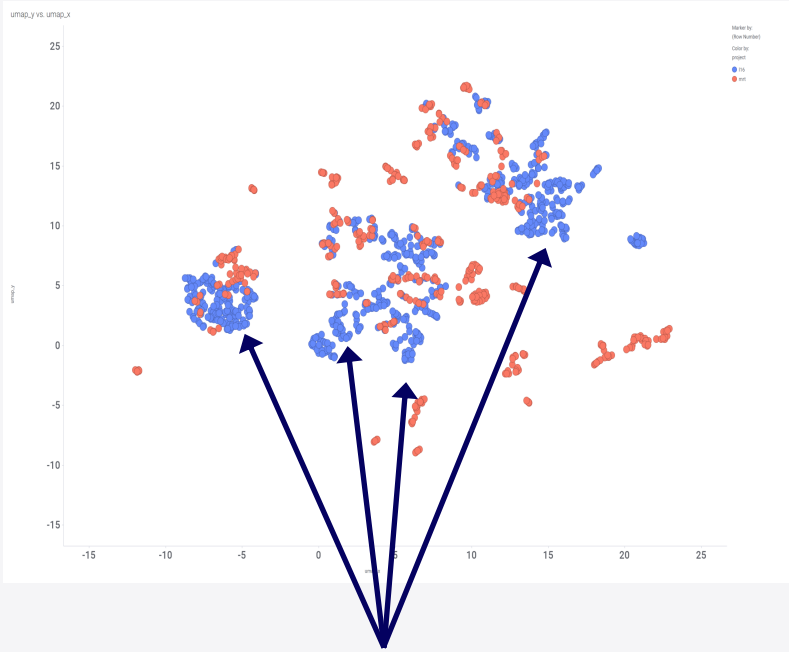


MRT Library Enrichment: Maximum Hybrid Surface Diversity

Metric: compound neosurface similarity

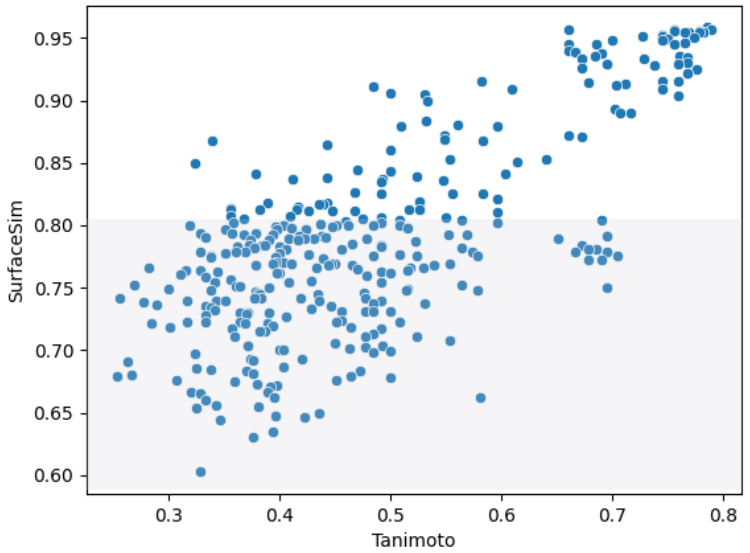


Library **scaffold expansion** beyond current **MRT scaffolds**



Novel territories covered by expansion (UMAP projection)

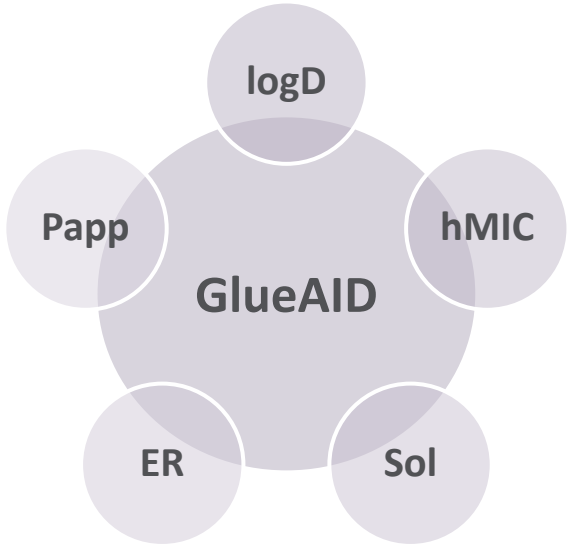
Surface similarity of all proposed monomers vs. those in MRT collection



Final Selection:
monomers with SurfaceSim < 0.8

GlueAID ADME Suite Predicts MGD Properties

MGD-focused ML ADME models



1,000s training set
100s scaffolds

ML algorithms

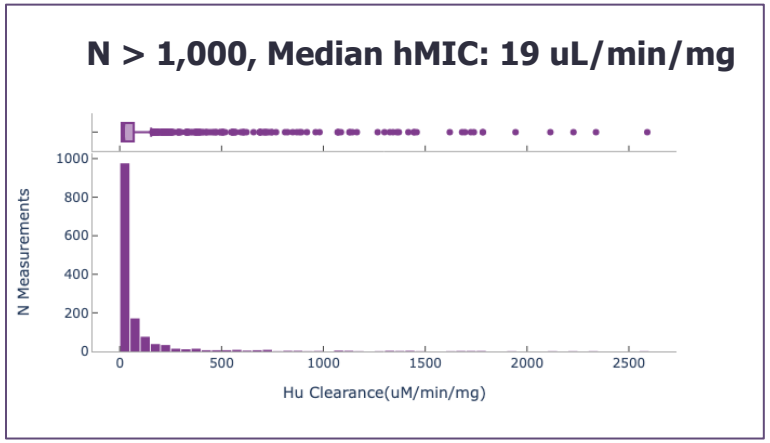
Individual and ensemble models are evaluated at every update

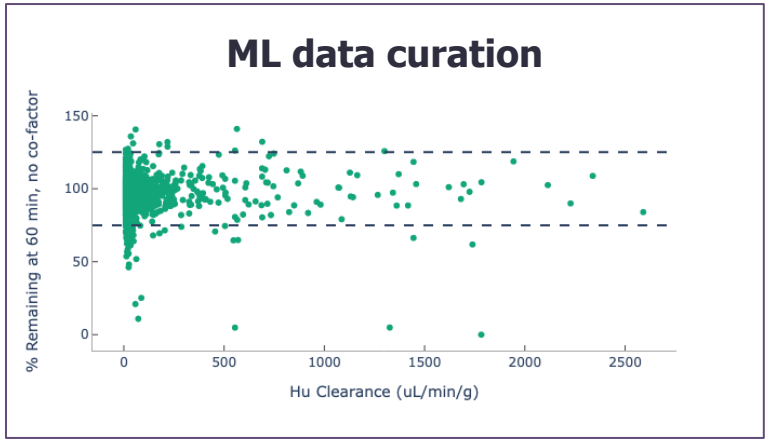
- LightGBM
- XGBoost
- RF
- KNN
- Consensus
- + other

Evaluation Metrics

- R^2
- Rho
- Within 3/5-fold

In-house human clearance data

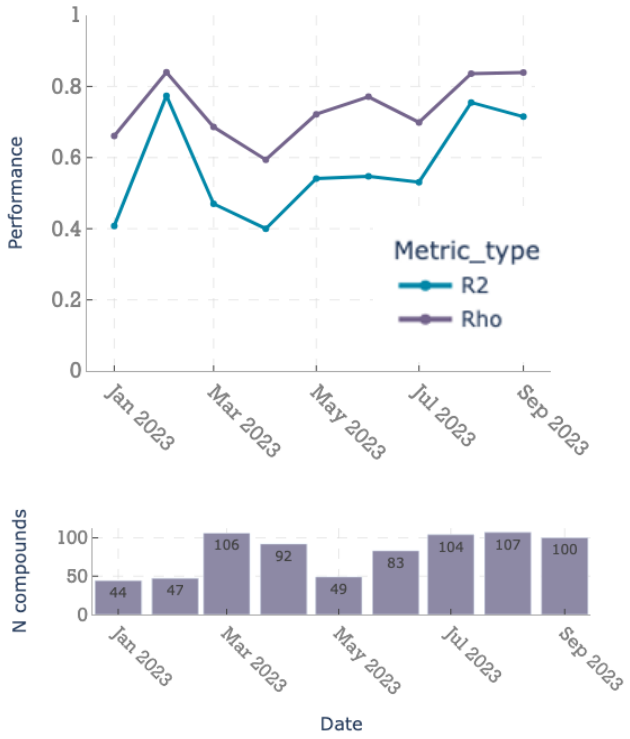




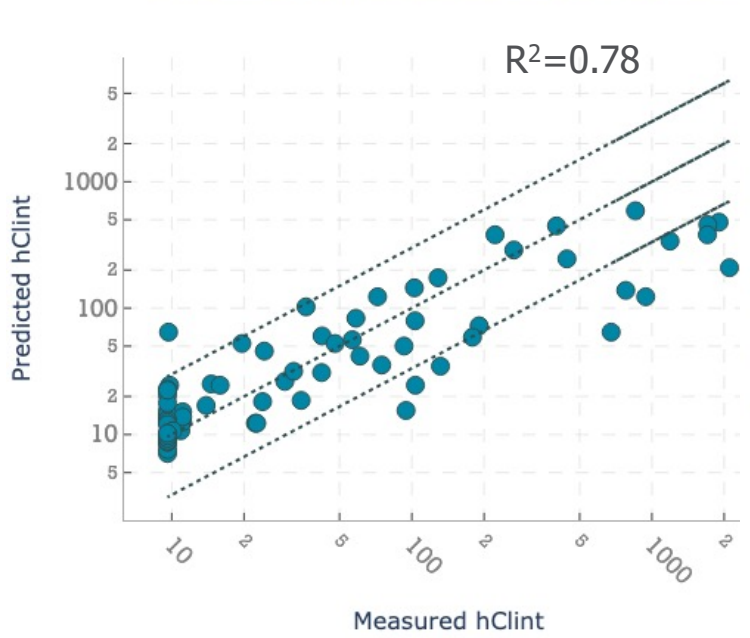
GlueAID Used on Ongoing Projects

Continually improved predictions with incoming data; performance monitored via dashboard

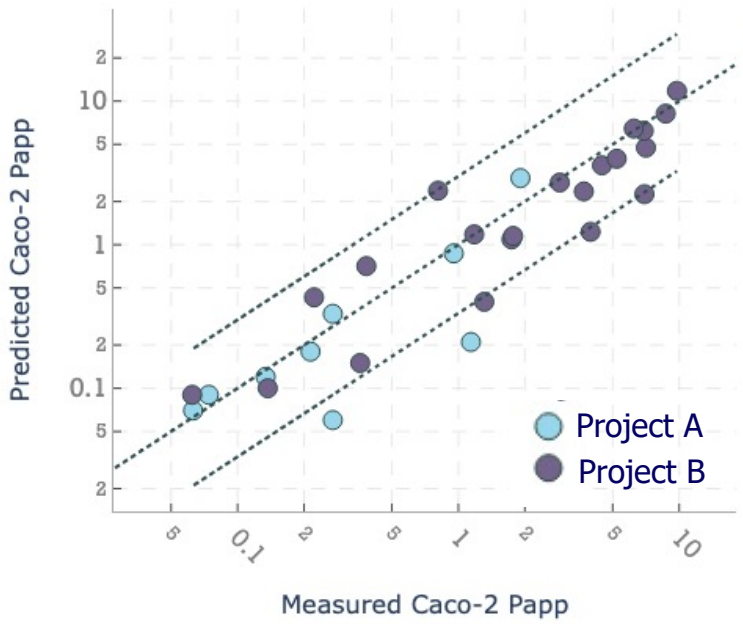
Clearance model performance over time



Clearance model performance last month

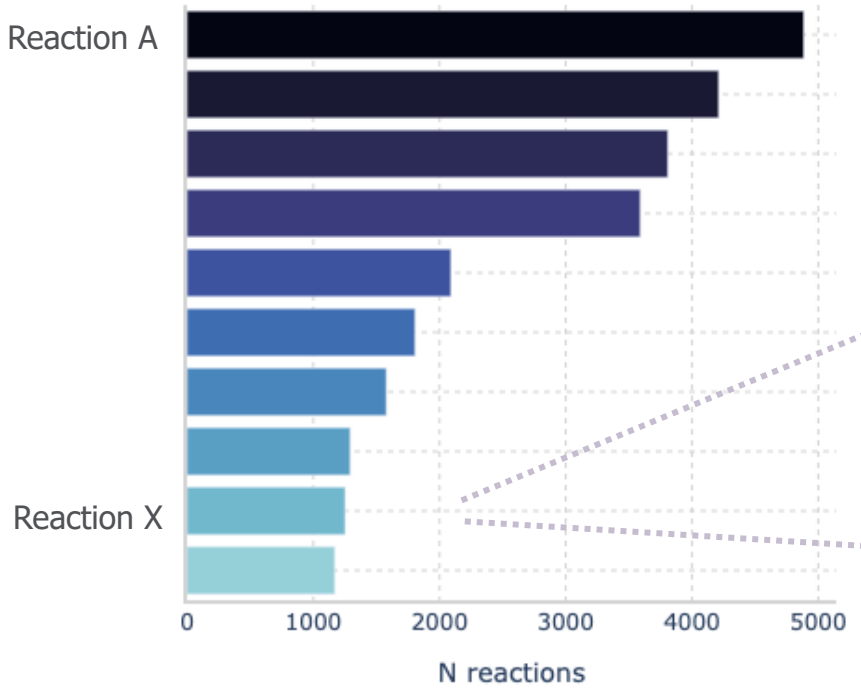


Permeability model performance by project/series

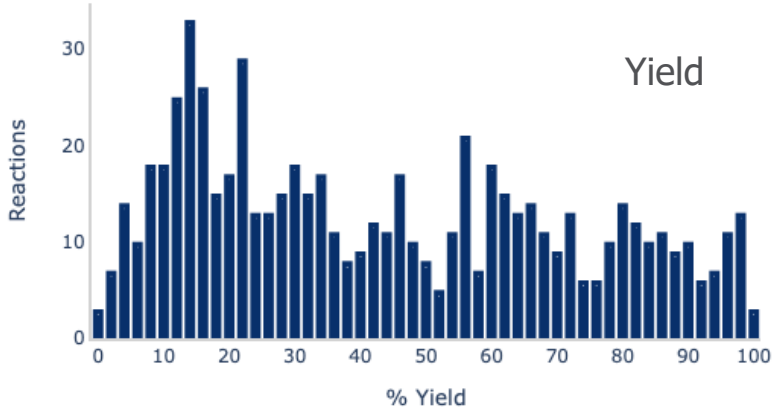


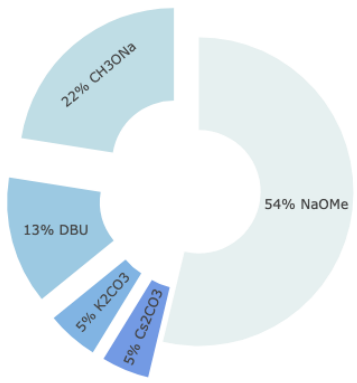
GlueAID Reaction Informatics Accelerates Synthesis

~ 100K In-house reactions and growing

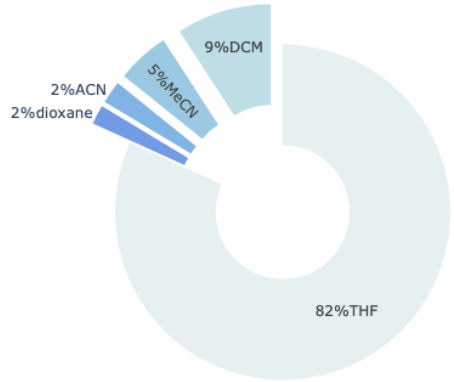


Reaction data analytics





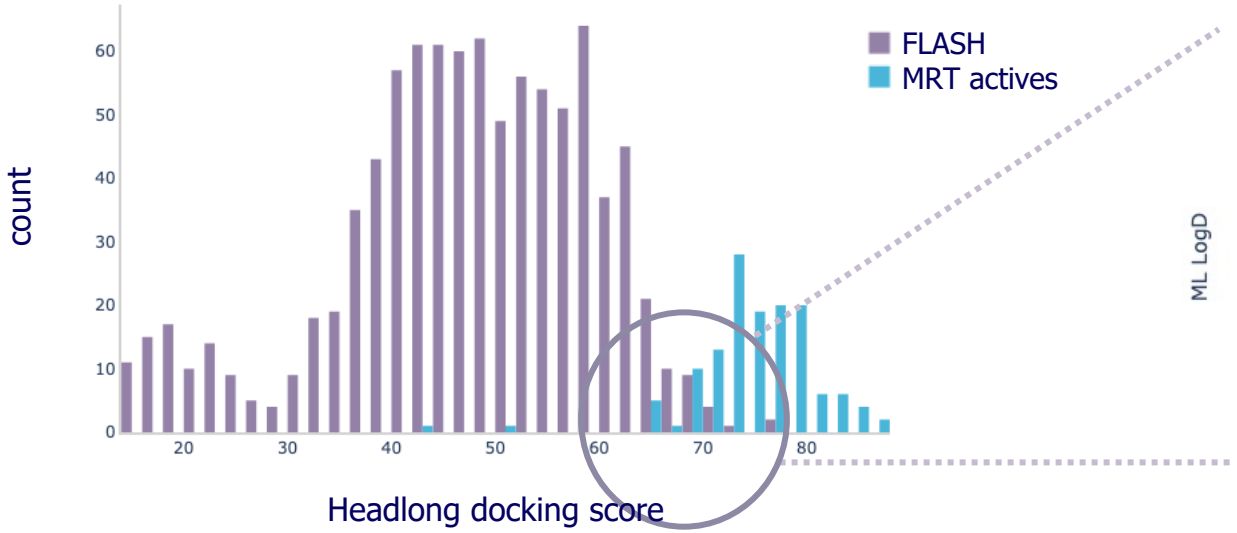
Common Agents



Common Solvents

FLASH Expansion Leverages GlueAID for ADME and Synthesis Optimization

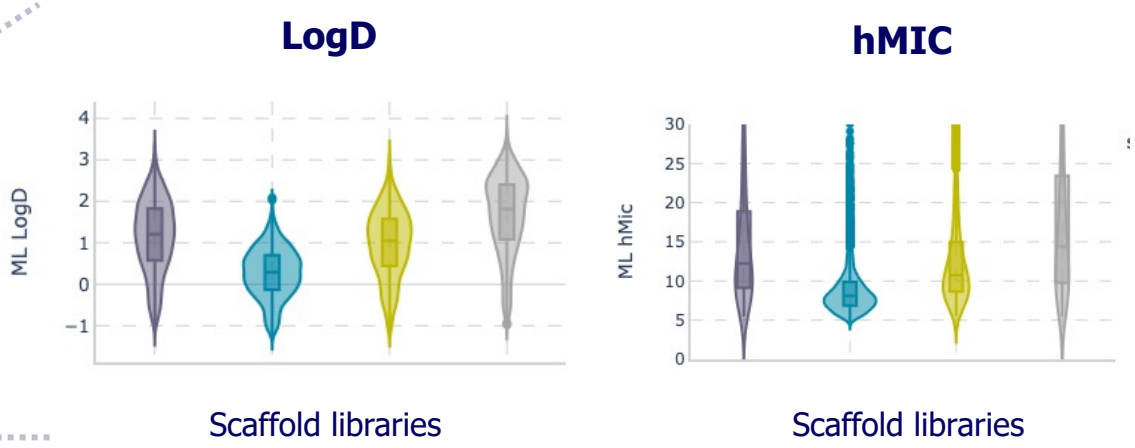
Headlong™ CRBN binding prediction



FLASH: ~1,000 virtual 5-atom ring scaffolds
Actives: ~150 MRT scaffolds

GlueAID™ guides scaffold selection for better ADME profiles

4 scaffolds x 1 reaction x 5,000 building blocks



Scaffold libraries

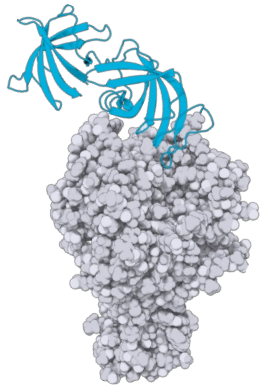
Scaffold libraries

~50 high-scoring (docking & ADME) scaffolds are being explored

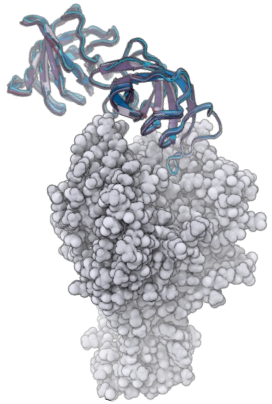
Rhapsody Virtual Screening Predicts Ternary Complex Hits

Hit ID	Hit Expansion/Hit2Lead	Lead Optimization
Find **hit series/active scaffolds** through VS	Identify **active analogs** of a hit w/ good properties	Guide design of **molecules in a program**

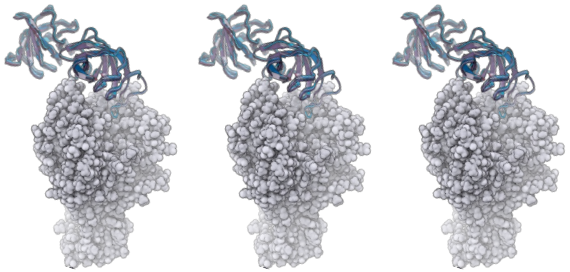
Ternary complex



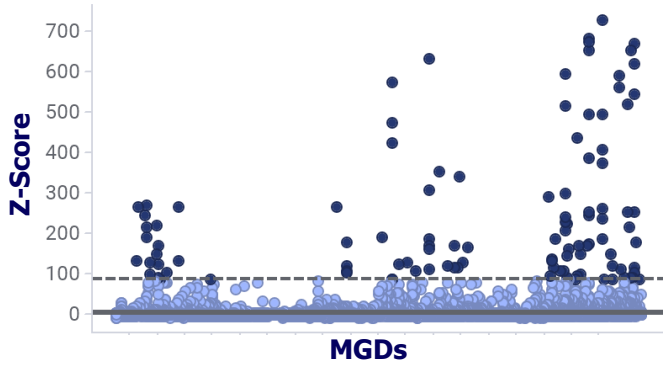
Ensemble

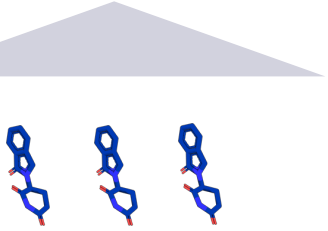


Dock library into ensemble

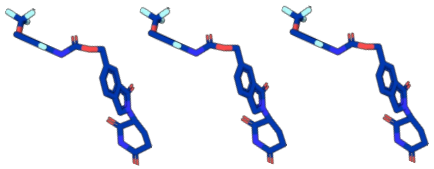


Virtual screening





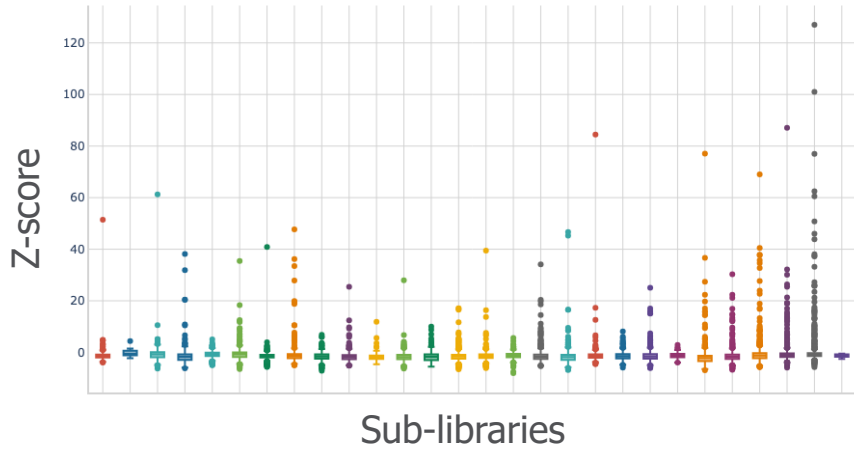
Fragments
E.g., GlueDock scaffolds



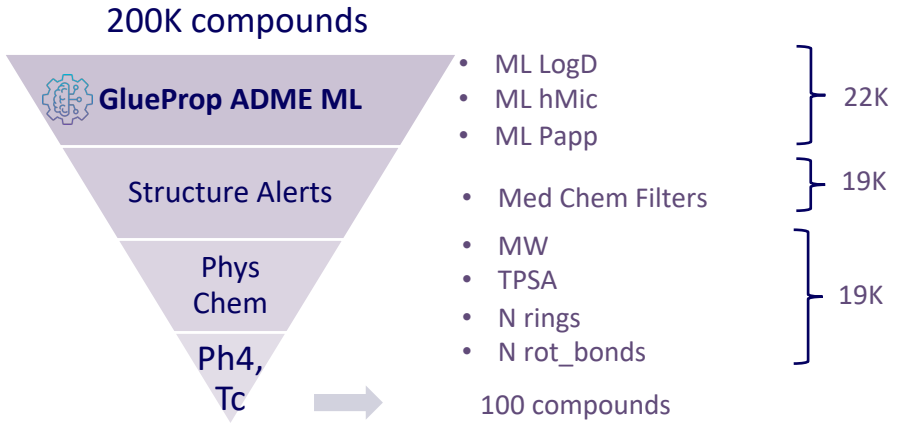
Library
E.g., FLASH virtual library

Putting it Together: Comp Chem Guides Screen to Validated Hit Series in 2 Months for **Target X**

HTS Screen



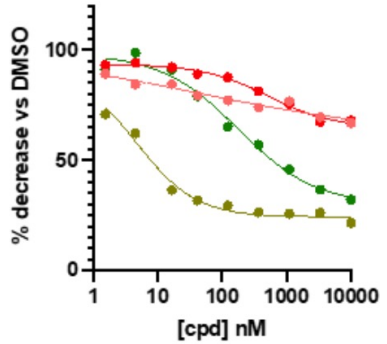
Hit Expansion Workflow



Computational Approaches

- Actives-based Pharmacophore
- ML Model from HTS Data
- Ternary Complex Model VS
- Similarity Selection

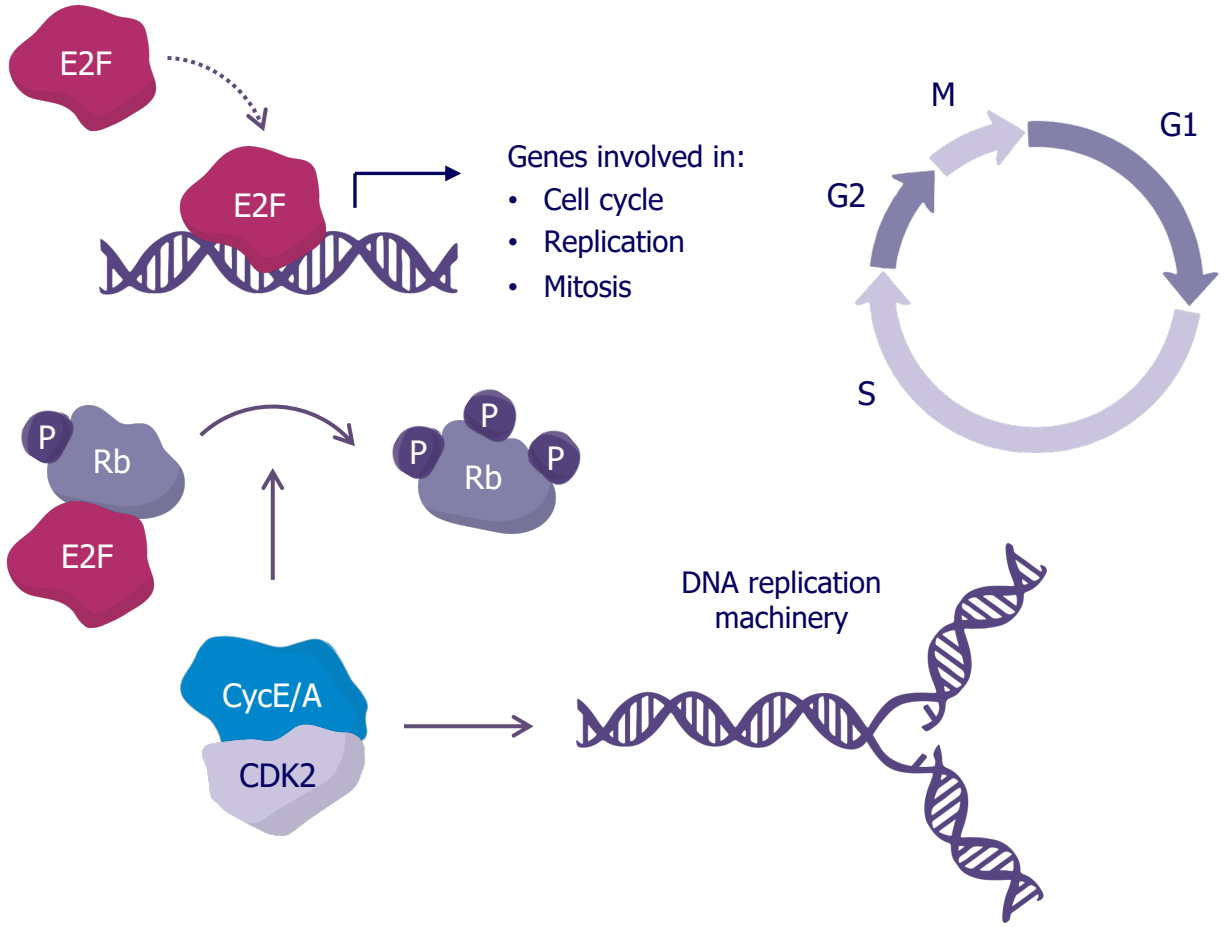
Hit Optimization/SAR exploration



- ✓ Flexible SAR
- ✓ GSPT1 selective
- ✓ Good ADME profile

CDK2 as a Target for Selected Solid Tumors

CDK2 a key cell cycle regulator



Therapeutic hypothesis:

- CDK2 is a key driver of cancers with cyclin dependent kinase pathway alterations
- MGDs will achieve greater selectivity against other CDKs and kinases in general, as well as more sustained pathway inhibition compared to inhibitors

Clinical Opportunity:

- ER positive breast cancer pre and post treatment with CDK4/6 inhibitors (474K patients)
- Ovarian cancer (64K patients), endometrial cancer (124K patients) and other tumors with CCNE1 amplification

Patient diagnosed incidence #s, major markets (US, EU and JP): Decision Resources Group (DRG)

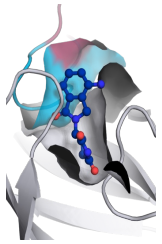
In Silico-driven Scaffold Selection and Compound Design for CDK2

Rich data-set of proprietary Monte Rosa library allows informed selection of designs instead of pure diversity-based compound enumeration

Sidechain

~9,000 tails
- Availability
- Properties

Scaffold



15-20 preferred warheads, cores & linkers
- CRBN affinity
- Stability

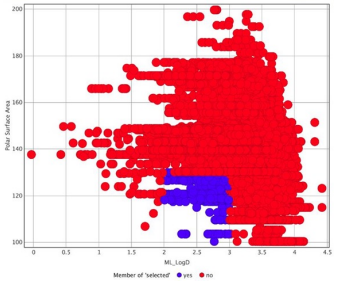
Enumeration

~150K
Compound Designs

Property Selection

Phys-Chem Prop
- MW < 500
- PSA < 150
- clogP 2-3
- Pass filters

~11K
Compound Designs



ADME Selection

Machine Learning Models
- Trained on curated in-house data
- Multi-task learning approach

Rhapsody Ternary Complex Modeling
- 18 x-ray & 3 cryoEM internal structures
- In-house developed docking pipeline validated with project compounds
- Internal x-ray ensemble used to model local SAR

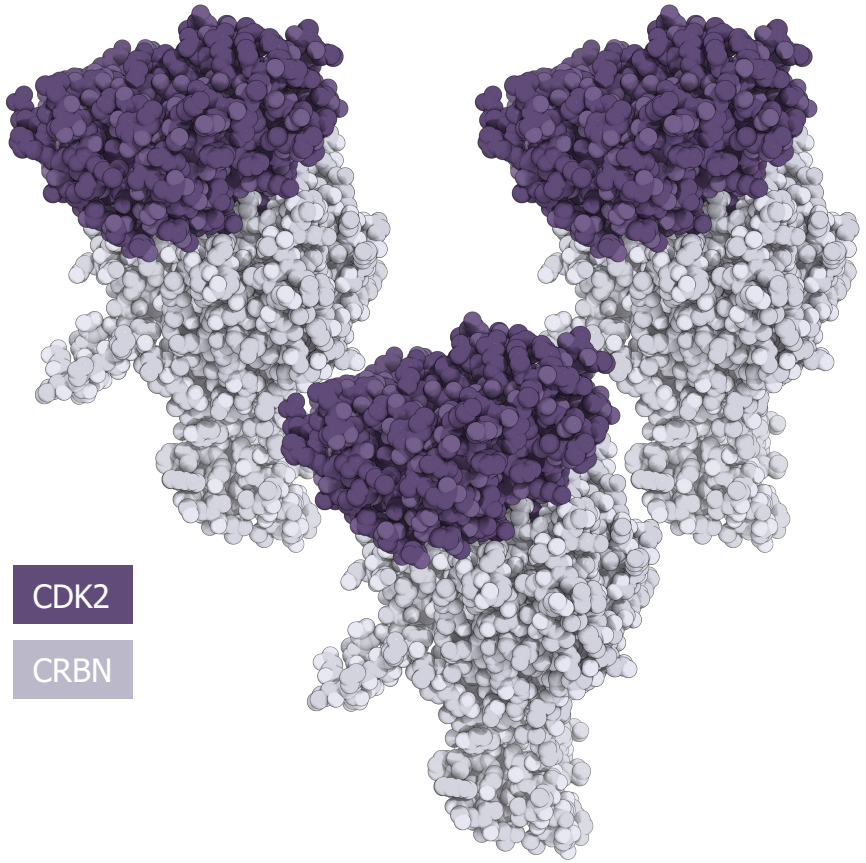
~1K
Compound Designs

Activity Selection

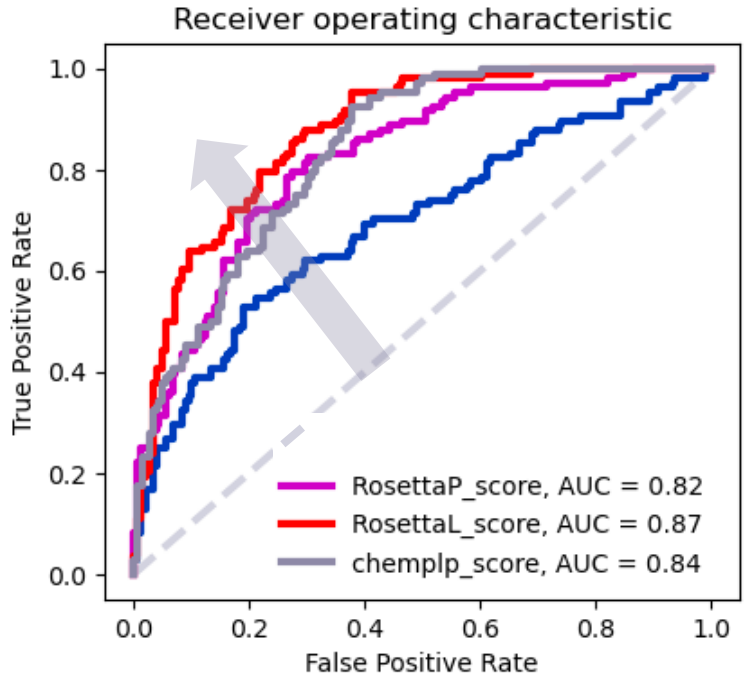
130
Compound Designs

Rhapsody™ Virtual Screening for Lead Optimization

CDK2 ensemble of 12 protein structures



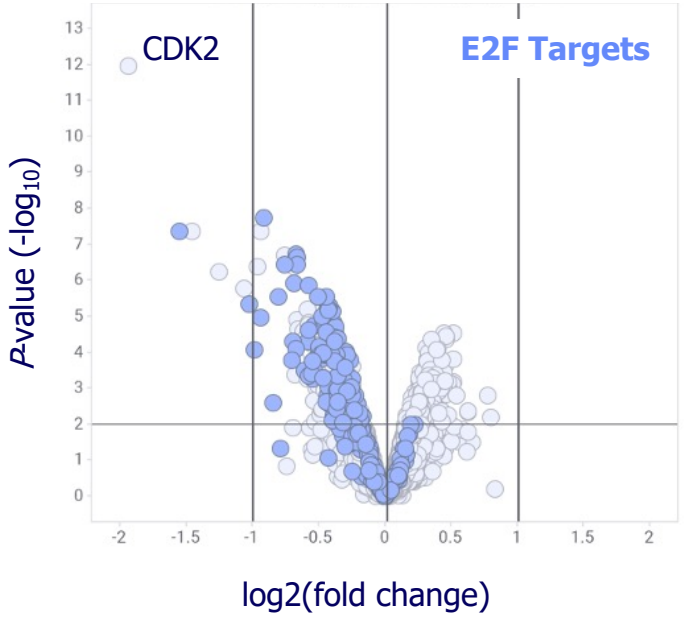
Rhapsody used to select active analogs for lead optimization



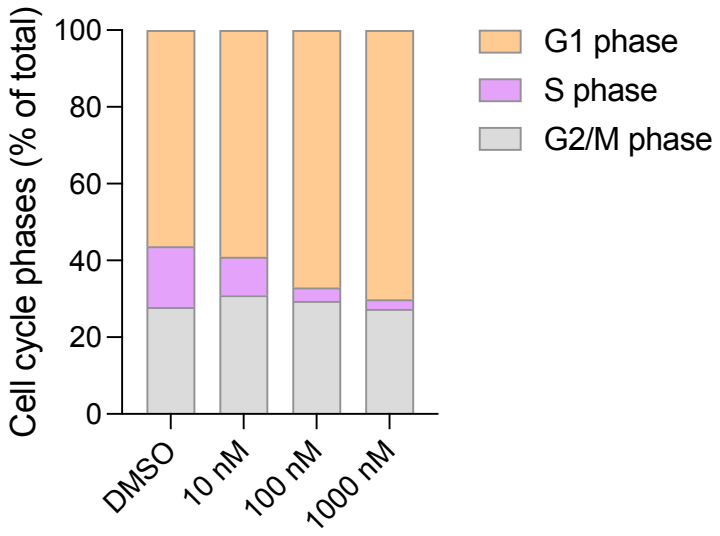
Rhapsody scoring of single-point HTRF screening of library MGDs.

CDK2-directed MGDs are Selective and Inhibit Proliferation of CDK2-dependent Cancer Cells

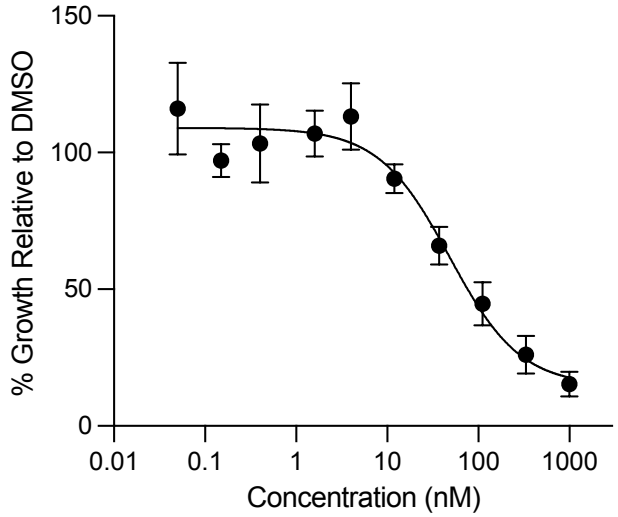
CDK2 degradation results in reduction of E2F pathway proteins



CDK2 degradation arrests CDK2-dependent cells in G1 phase



CDK2 degradation blocks proliferation



CRBN K_i = 129 nM
SPR half-life = 994 s
NanoBiT DC_{50} = 130 nM
CyQuant MDA-MB-157 EC_{50} = 46 nM

TMT Proteomics - MDA-MB-157 (1 uM/24 hr)

Cell cycle analysis (DAPI and EdU) - MDA-MB-157 (24 hr)

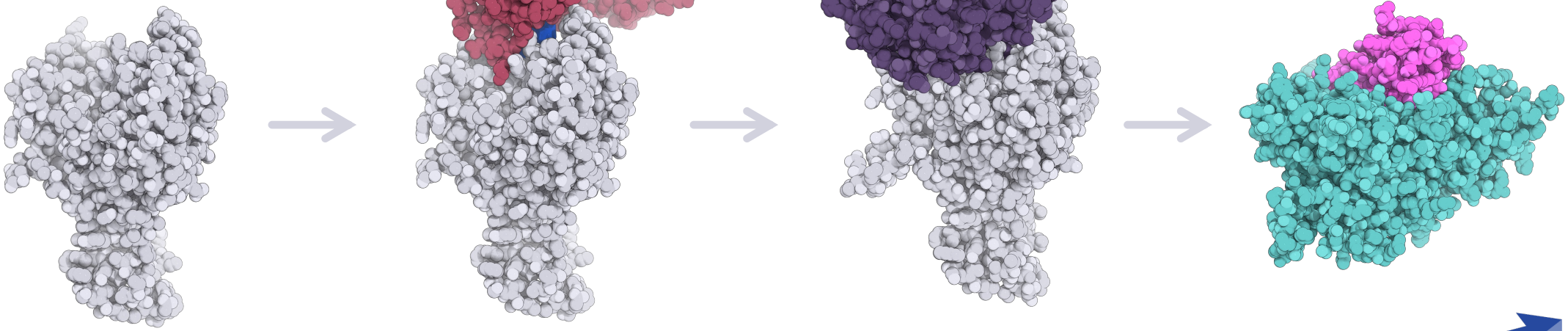
Proliferation (CyQuant) - MDA-MB-157 (7dr)

Our Rational Approach to Unleash the Full Potential of MGDs

Neosubstrate
(e.g. IKZF1)

MGD
(e.g. Lenalidomide)

Cereblon



Unraveling the Canonical Degrome

Beyond the Canonical Degrome

Beyond Cereblon

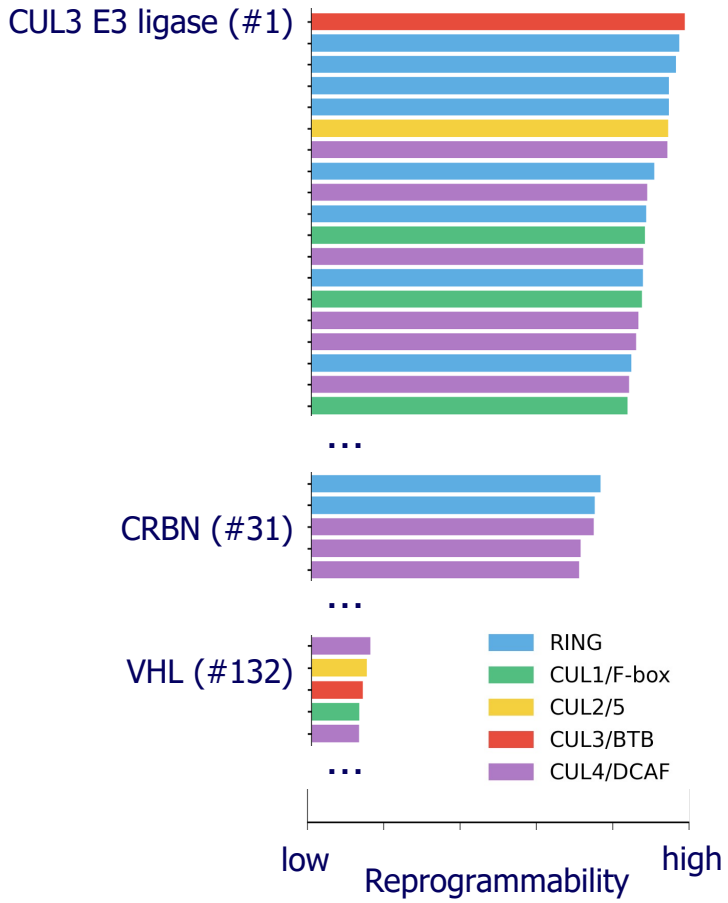
Chemical Space

Target Space

Expanding the Degradable Proteome

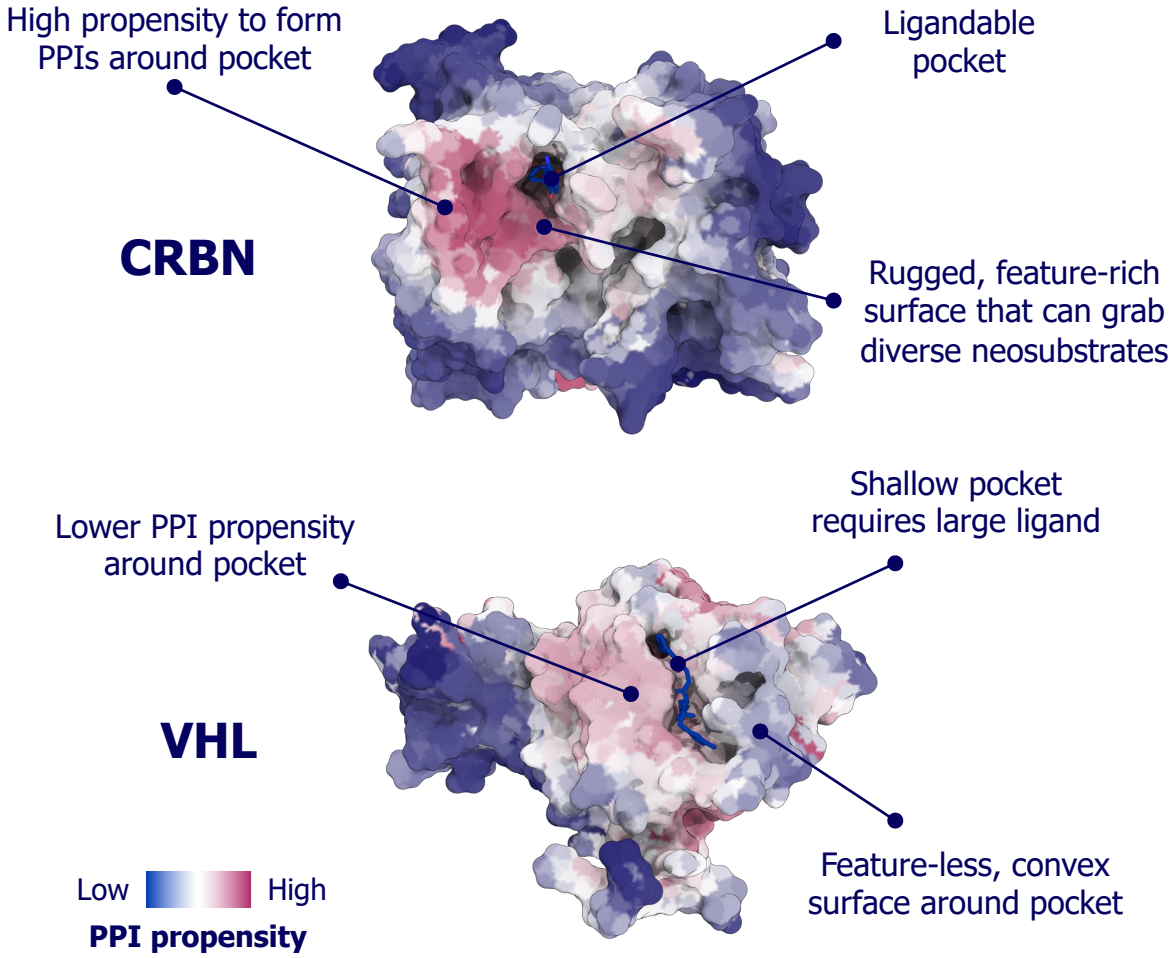
Prioritizing E3 Ligases Based on Reprogrammability Potential

fAIceit AI-powered surface characterization



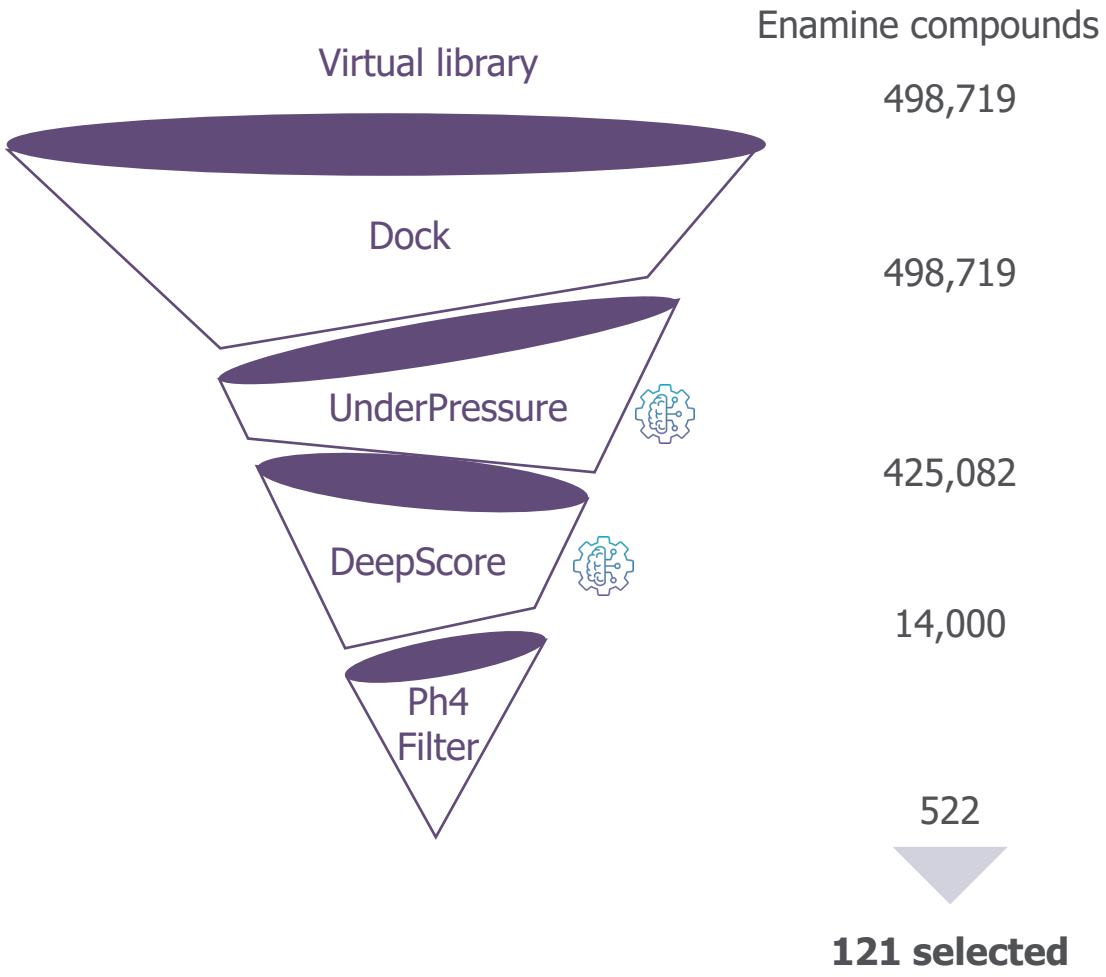
CRL & RING families shown

CRBN is one of a handful high-scoring E3 ligases
VHL scores much lower

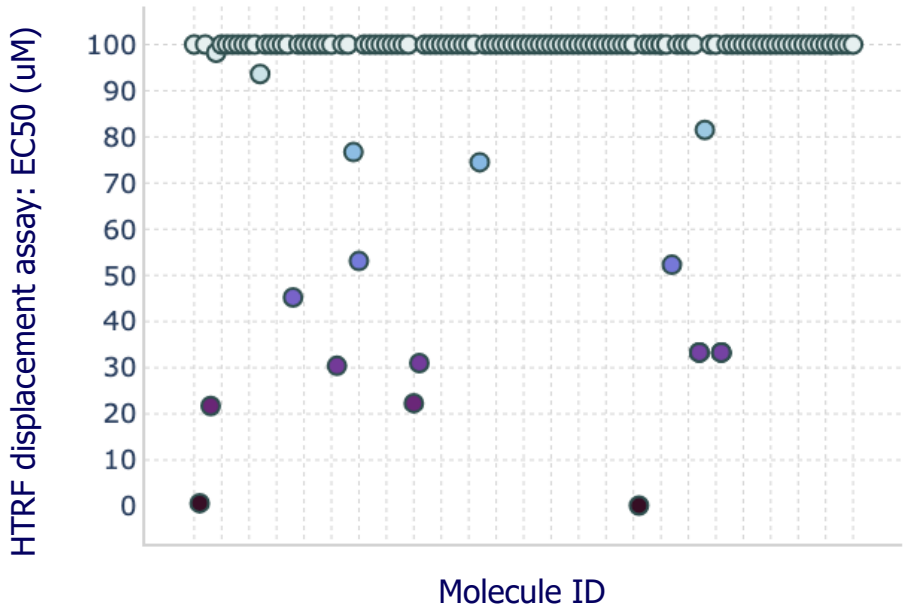


Headlong Virtual Screen Identifies Novel Scaffolds to a New E3 Ligase

E3 ligase Y

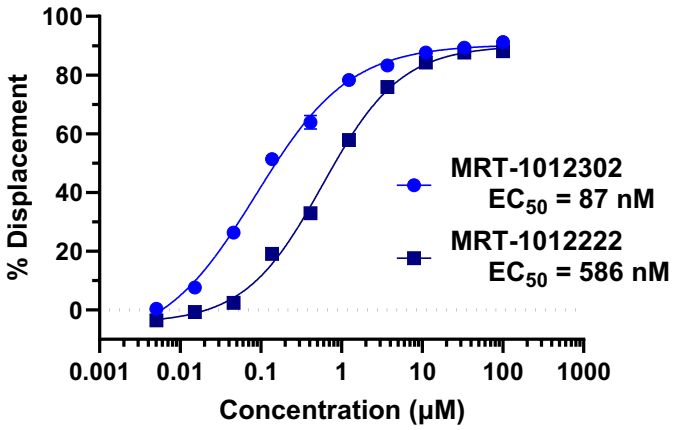


11% hit rate

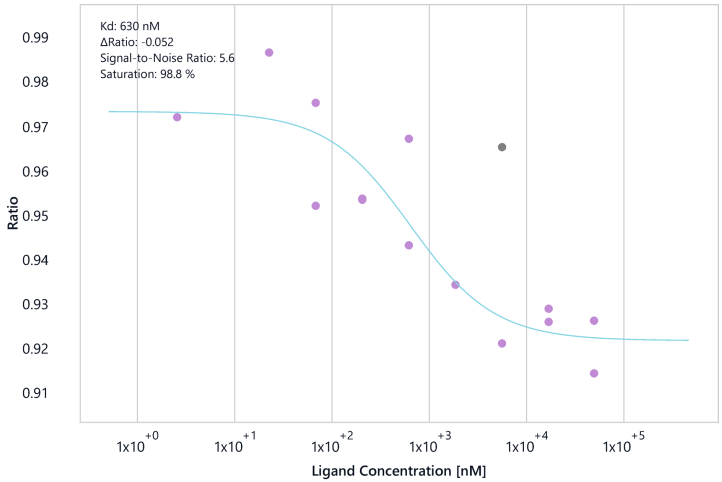


Virtual Screening Hits Confirmed in Dose Response, Including nM Series

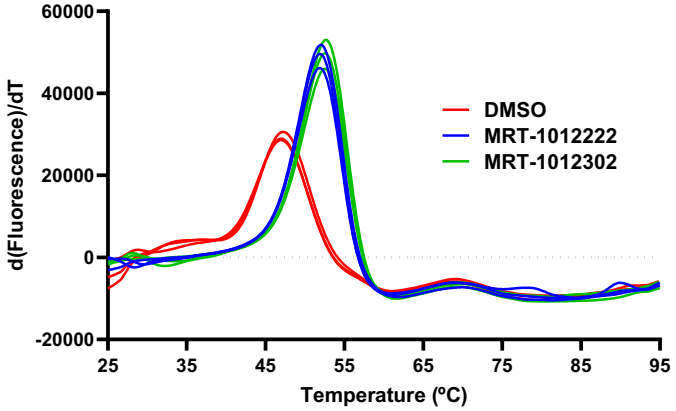
Dose response



Spectral shift



Differential Scanning Fluorimetry



Acknowledgments


Thank You
Monte Rosa
Therapeutics