



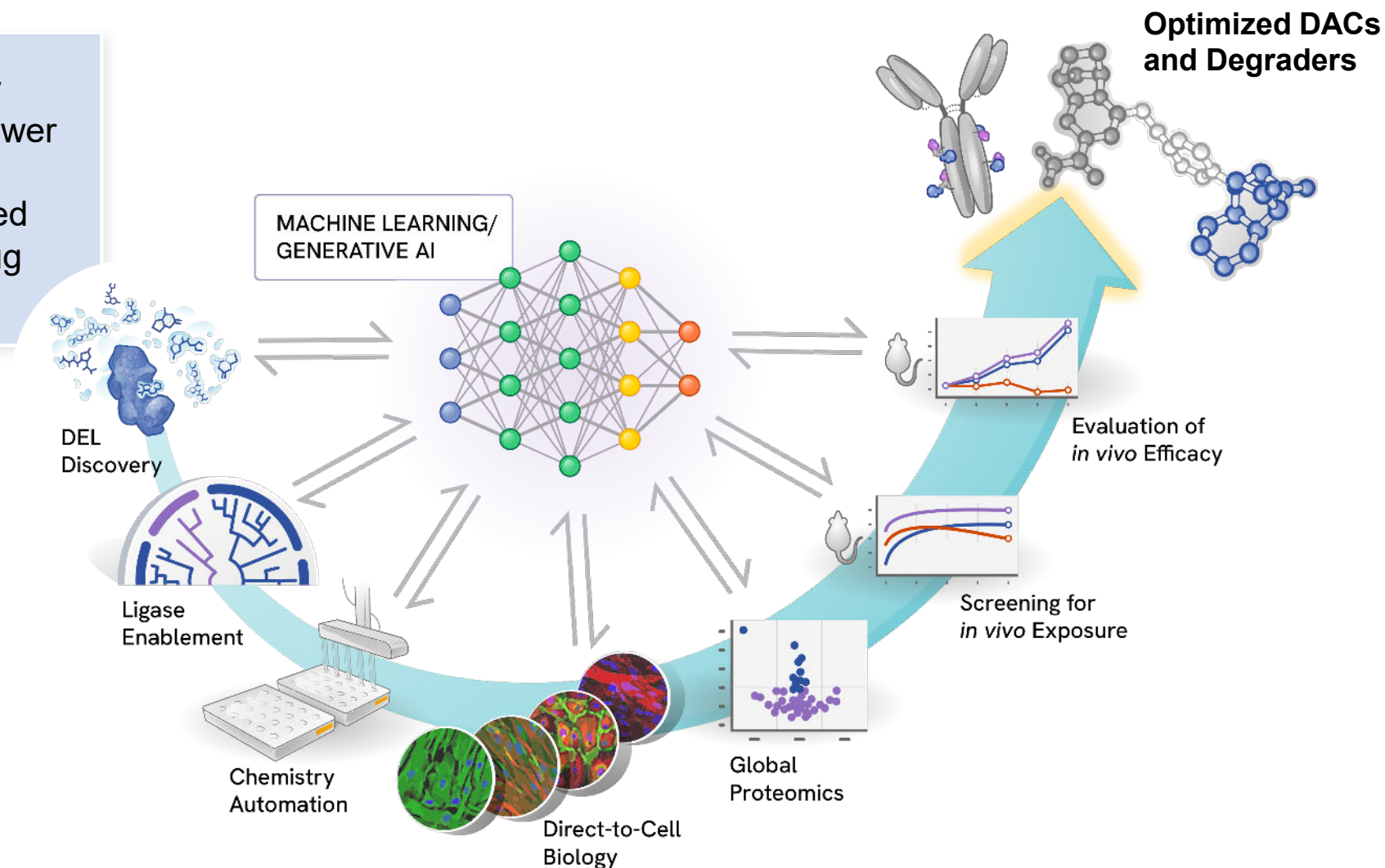
DEL AI: Proteome-Wide Binder Discovery with a DEL- Trained Foundation Model

Paul A Novick, Ph.D.

9th Annual TPD & Induced Proximity Summit
October 29, 2026
Boston, MA

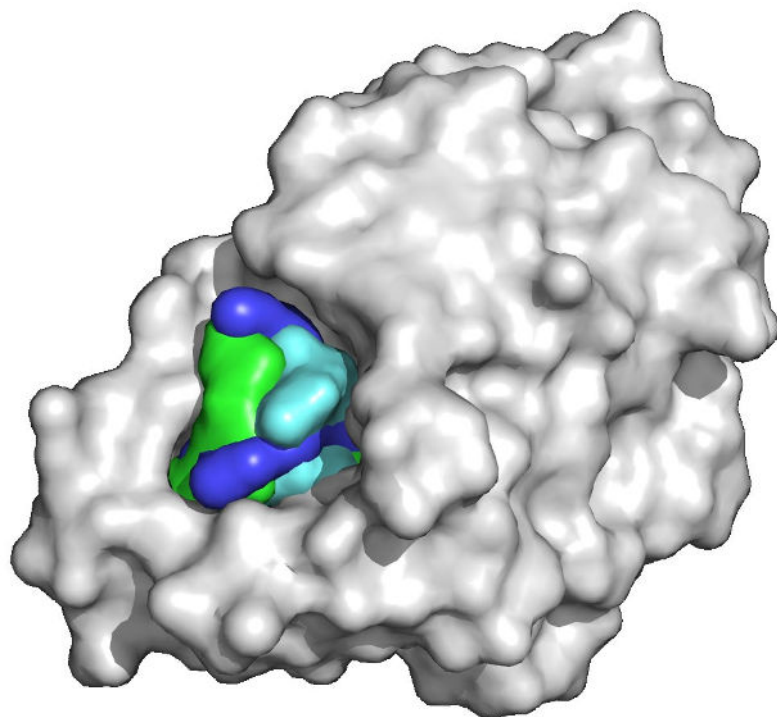
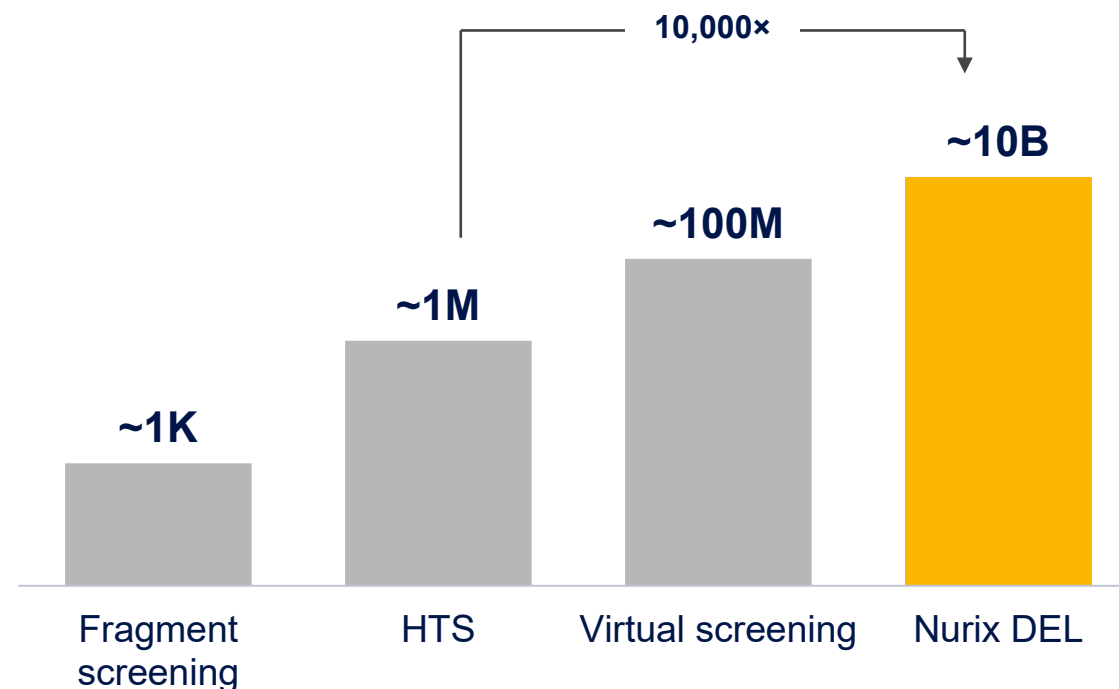
Industry Leading Discovery Engine for Induced Proximity Modalities

Our in-house, integrated discovery engine leverages the combined power of our data-rich DEL capabilities, automated chemistry, and advanced machine learning to accelerate drug discovery



Induced Proximity Drug Discovery Requires Chemical Starting Points

Lead Identification is a central challenge in drug discovery



Library size by screening approach, compounds (log scale)

Scaffold-Based DELs Enable Hit Finding for Challenging Protein Surfaces

>8B

unique compounds, built on complex, custom 3D scaffolds

52%

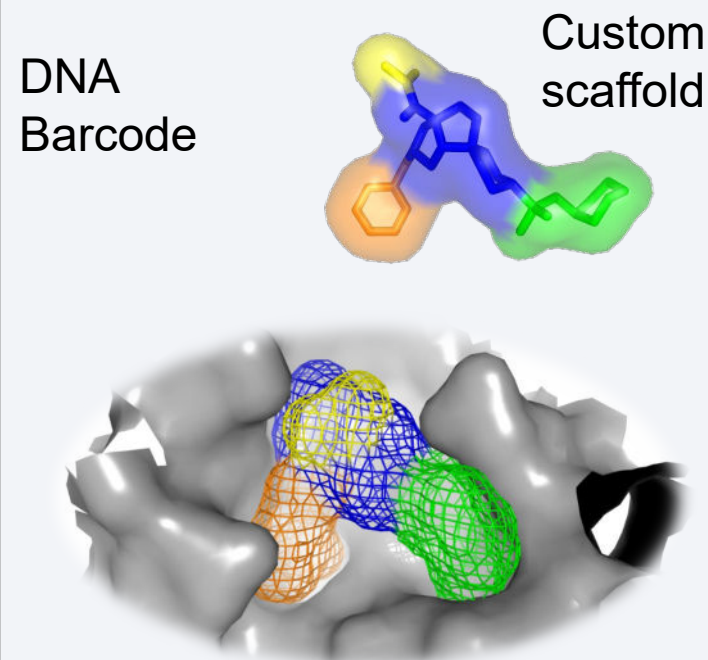
of 'undruggable' targets obtain hits solely from scaffold libraries

DEL collection by library type, % of compounds

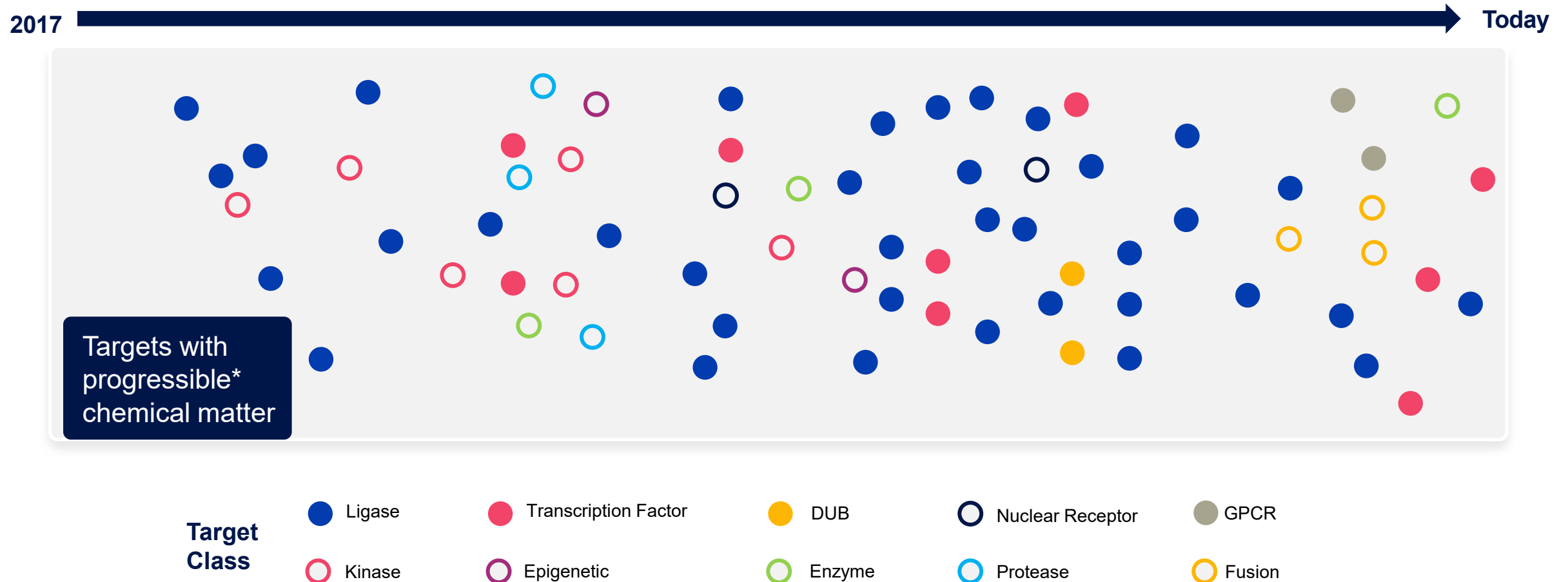


The DEL foundation model (DEL-FM) is trained on these sp³-rich datasets

Custom 3D scaffolds with high sp³ character allow better pocket fit



DEL is a Platform That Delivers Chemical Assets for Traditionally Undruggable Targets

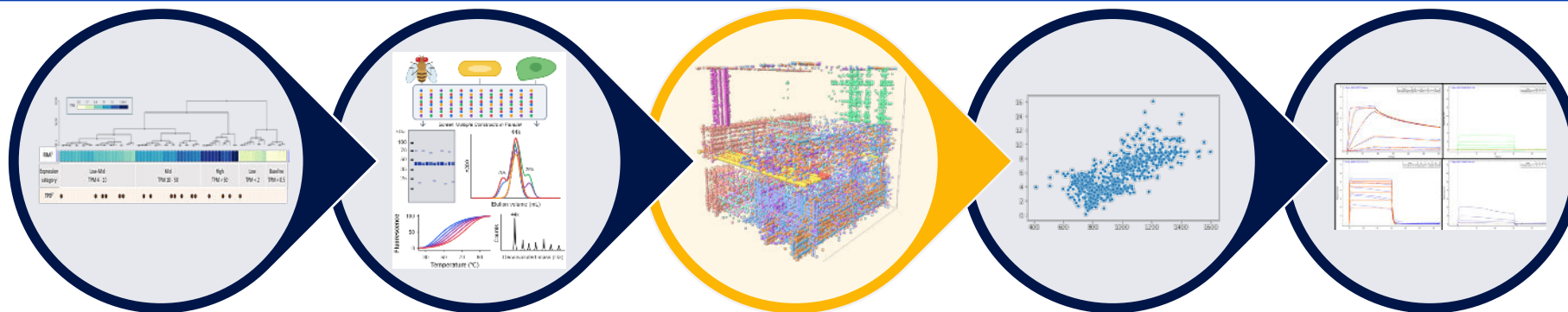


Approximate discovery timeline

*All series validated by more than 2 orthogonal assays binders with IC_{50} or $K_D < 10 \mu M$

DEL, DNA-encoded library

Nurix Generates Massive, High-Quality Datasets Ideal for Machine Learning



Target
Identification

High-Throughput
Protein
Production

DEL Screening

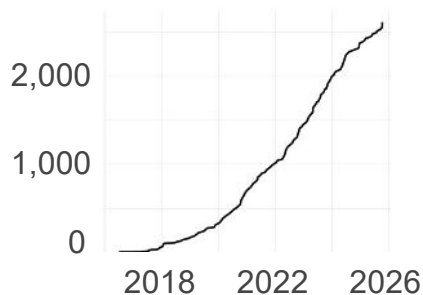
Hit Selection &
Ligand
Acquisition

Hit Confirmation

DEL Screening

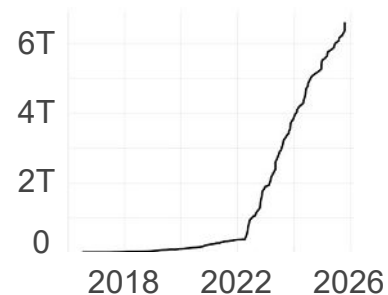
2,700+

DEL screens



7T+

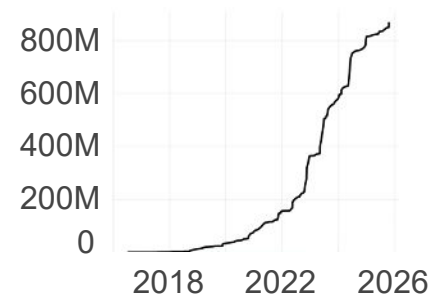
DEL datapoints



Hit Selection

900M+

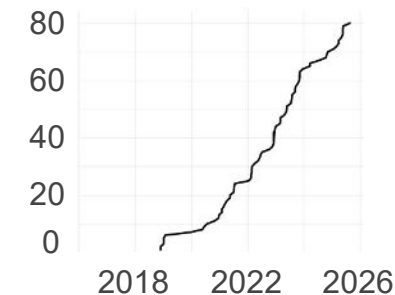
Ligands above threshold



Hit Confirmation

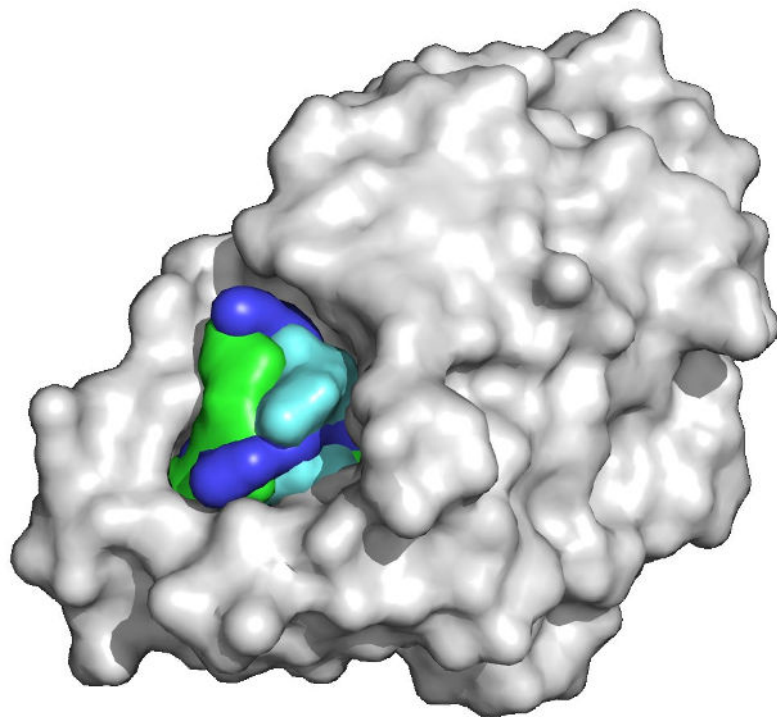
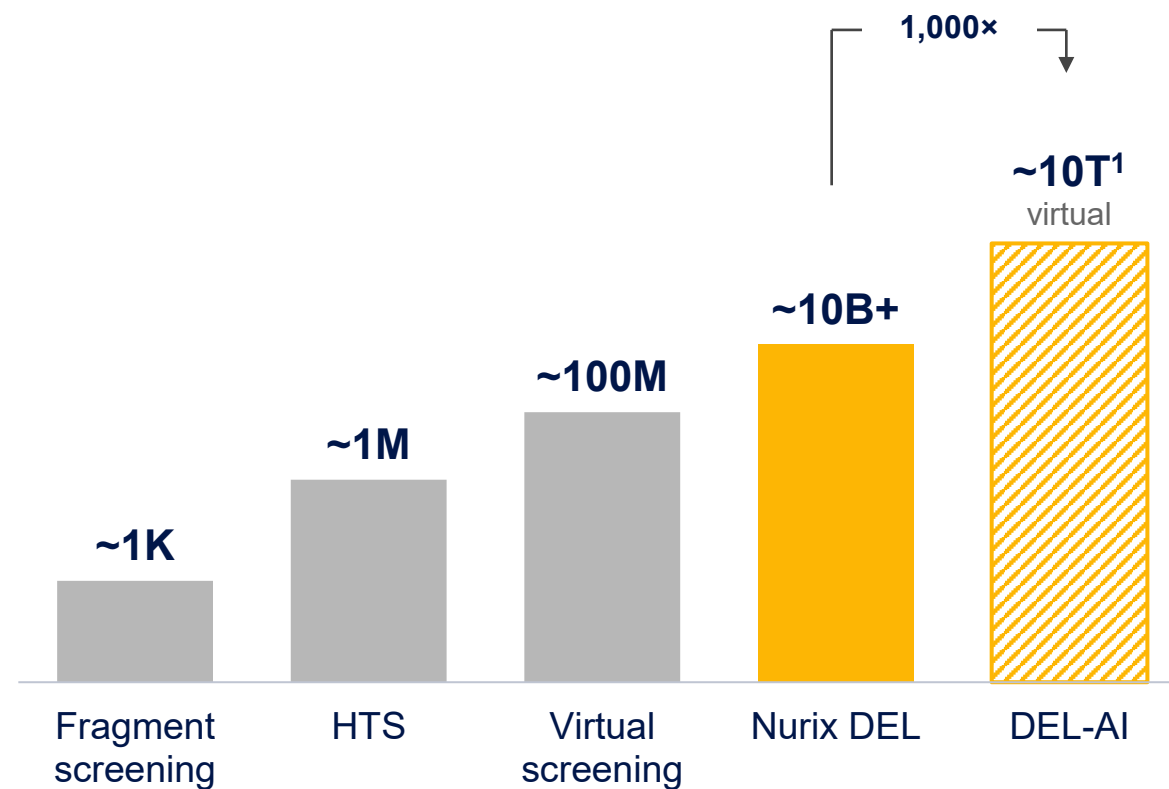
80+

High-affinity series



DEL Foundation Extends Discovery Beyond Physical Screening and is the Key to Unlocking Induced Proximity

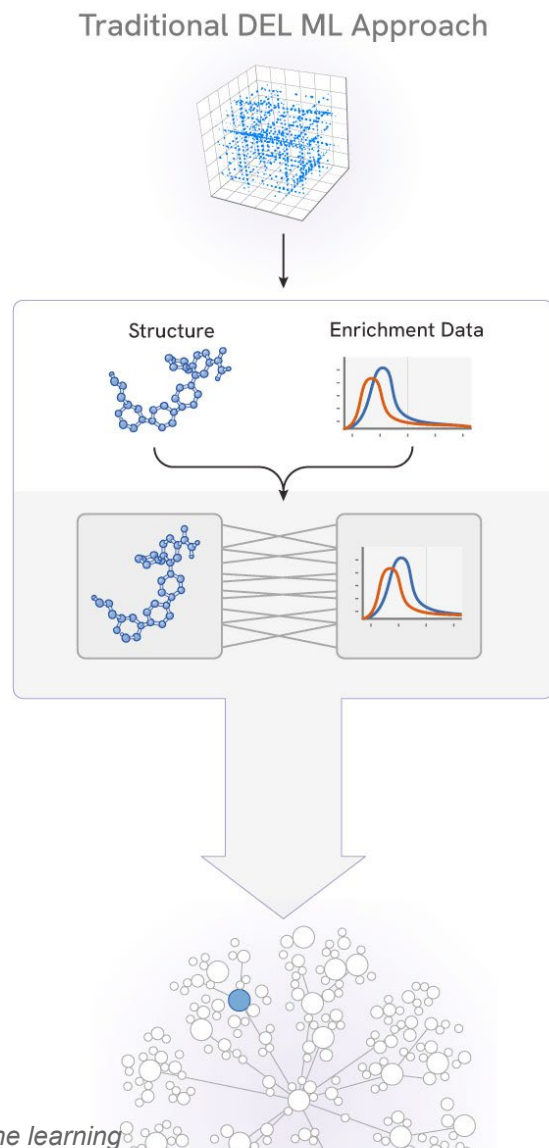
Information-rich DEL screening datasets are ideally suited to advance machine learning



Library size by screening approach, compounds (log scale)

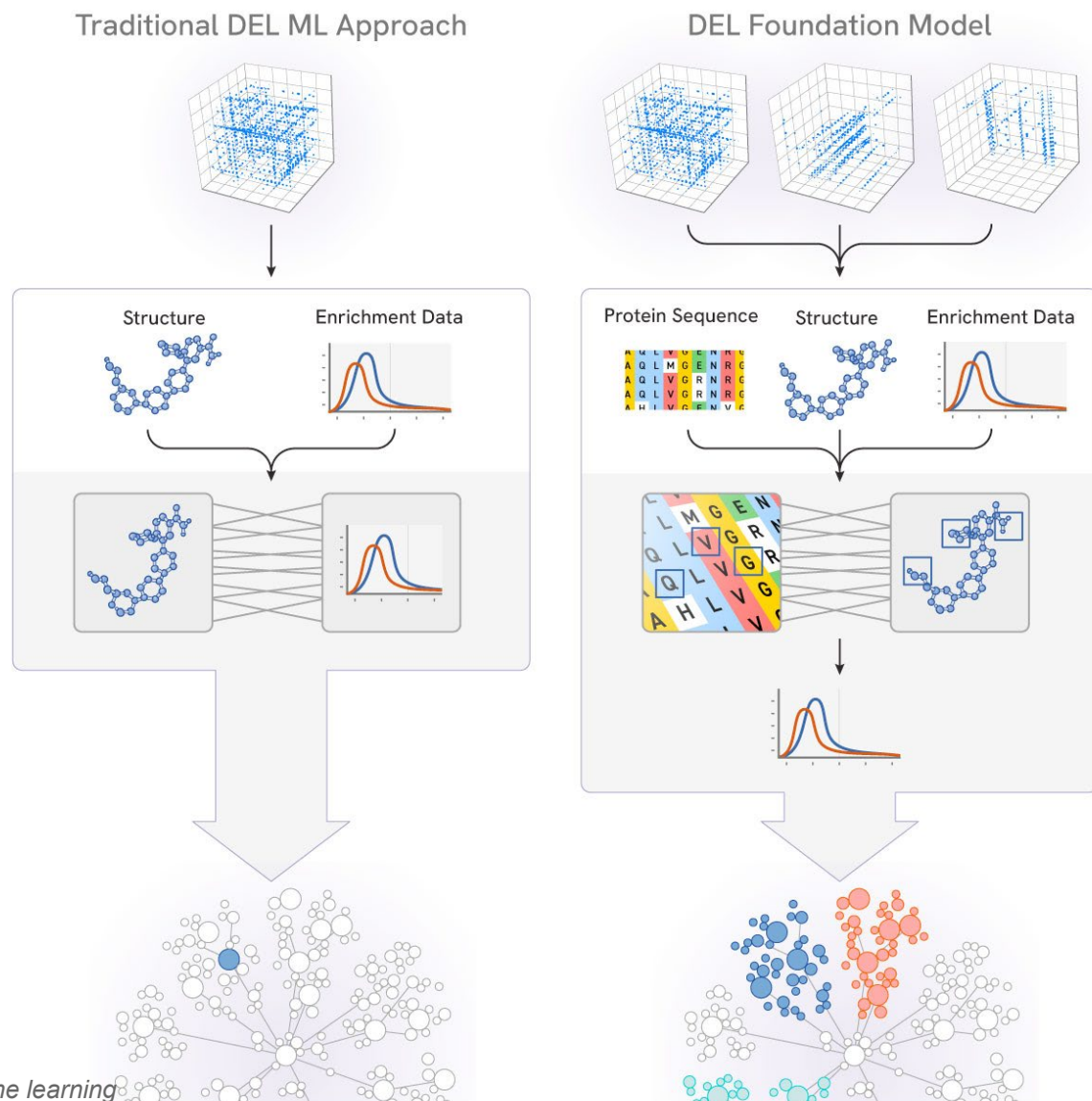
DEL Foundation: A Generalized Model for Proteome-Wide Hit Identification

- Established DEL ML approaches train models on single target datasets
- These models predict binding to a single protein and are inefficient to scale to ligand discovery across diverse targets
- Rely on time and resource intensive quality protein production and DEL screen execution.



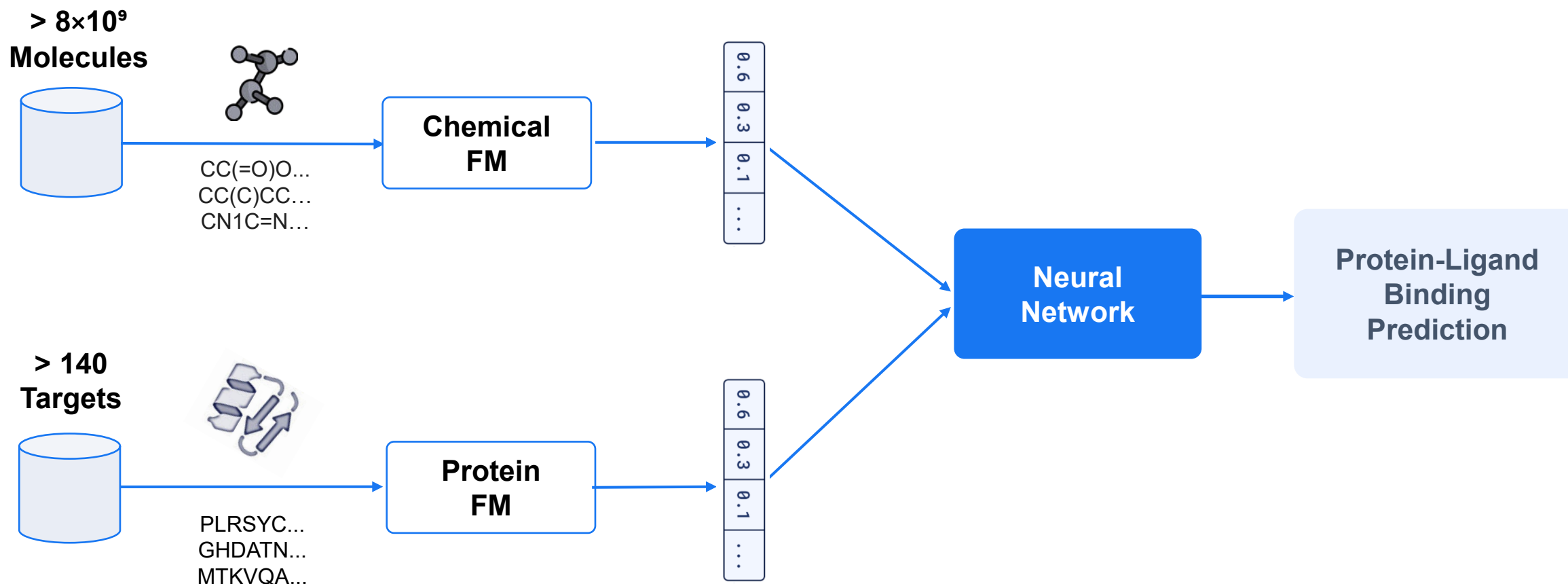
DEL Foundation: A Generalized Model for Proteome-Wide Hit Identification

- Established DEL ML approaches train models on single target datasets
- These models predict binding to a single protein and are inefficient to scale to ligand discovery across diverse targets
- Rely on time and resource intensive quality protein production and DEL screen execution.

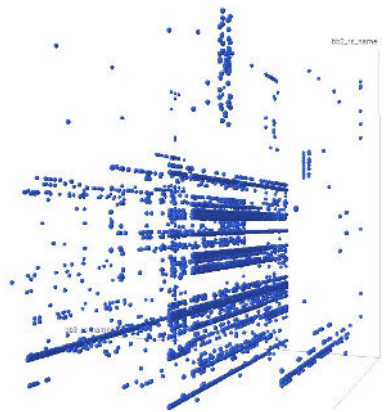


- DEL Foundation Model is trained on DEL datasets from diverse protein targets
- These models learn a generalized relationship between protein sequence and molecular structure
- DEL Foundation is thus able to predict DEL data on proteins without the need for experimental data

DEL Foundation Model Architecture



DEL Foundation Dataset Construction and Definition of Terms

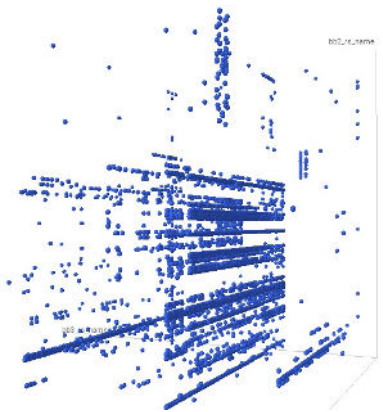


Experimental Actives

Compounds which bound in
the DEL screen

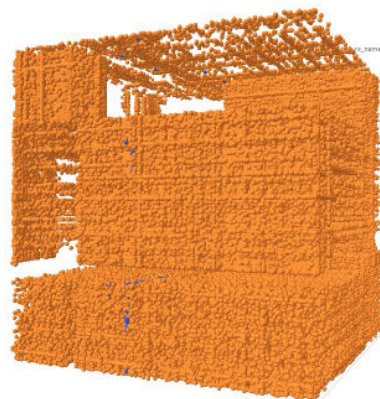
DEL Foundation Dataset Construction and Definition of Terms

Retrospective datasets constructed from experimental data



Experimental Actives

Compounds which bound in the DEL screen

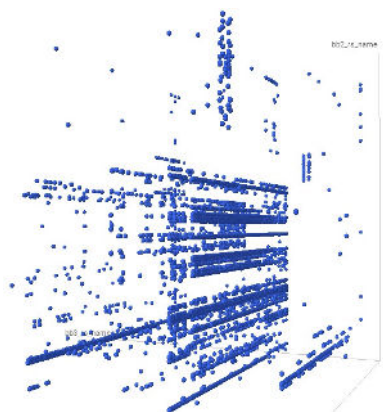


Experimental Inactives

Compounds which did not bind in the DEL screen, sampled from full DEL library

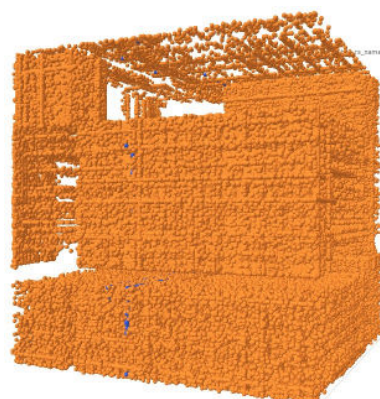
DEL Foundation Dataset Construction and Definition of Terms

Retrospective datasets constructed from experimental data



Experimental Actives

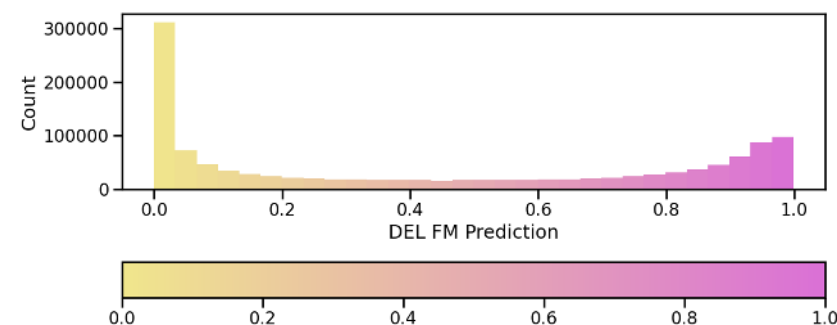
Compounds which bound in the DEL screen



Experimental Inactives

Compounds which did not bind in the DEL screen, sampled from full DEL library

Datasets not in the training dataset are evaluated by DEL Foundation (DEL FM)



DEL FM Inactives

Compounds predicted as non-binders*

**Low DEL FM score*

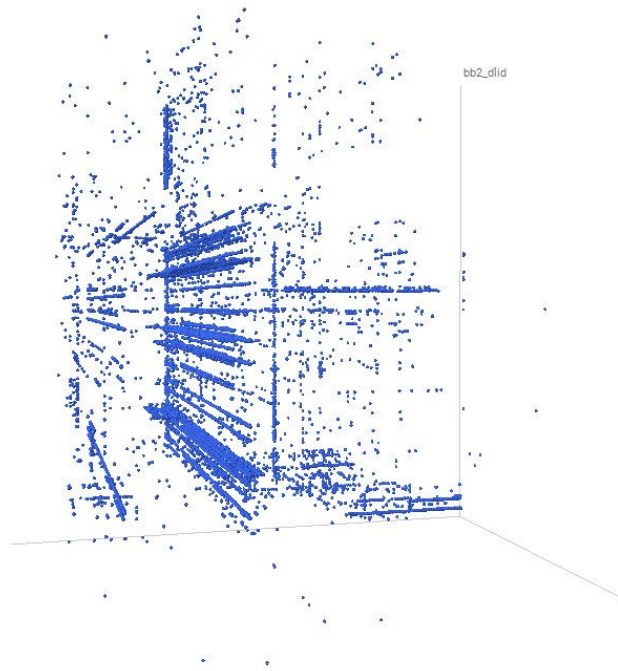
DEL FM Actives

Compounds predicted as binders*

**High DEL FM score*

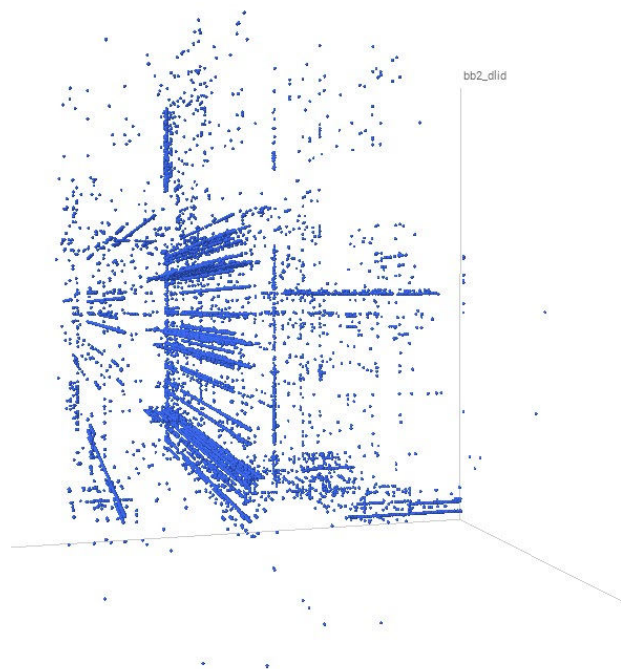
DEL FM Prospectively Predicts DEL Screening Results

Experimental Actives

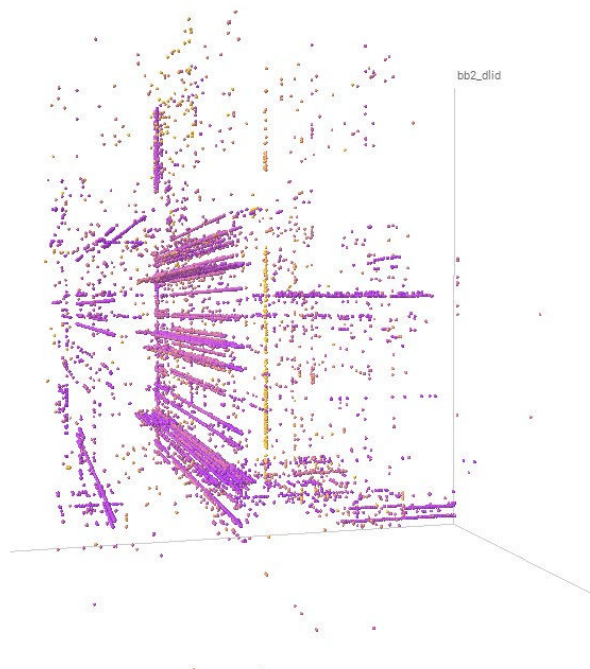


DEL FM Prospectively Predicts DEL Screening Results

Experimental Actives



**DEL FM Scoring of
Experimental Actives**

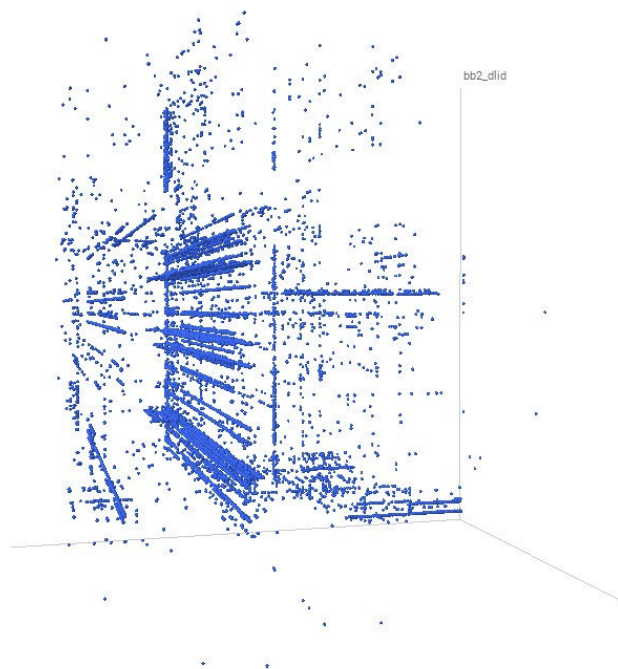


DEL FM Score

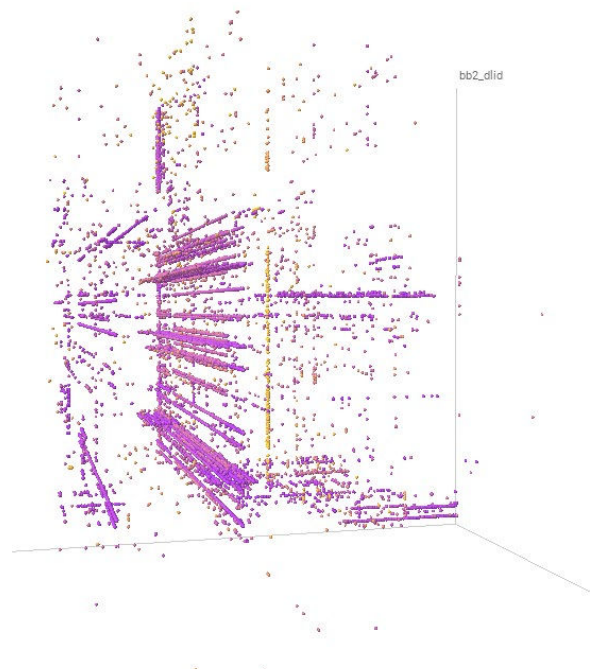
-1  1

DEL FM Prospectively Predicts DEL Screening Results

Experimental Actives



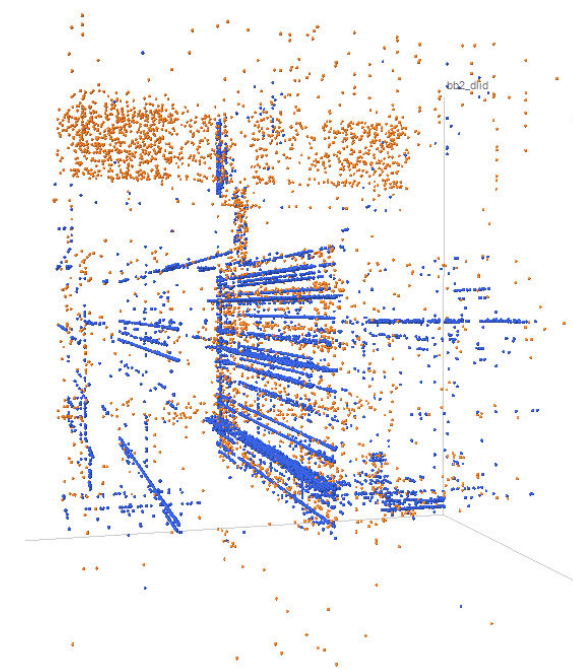
**DEL FM Scoring of
Experimental Actives**



DEL FM Score

-1  1

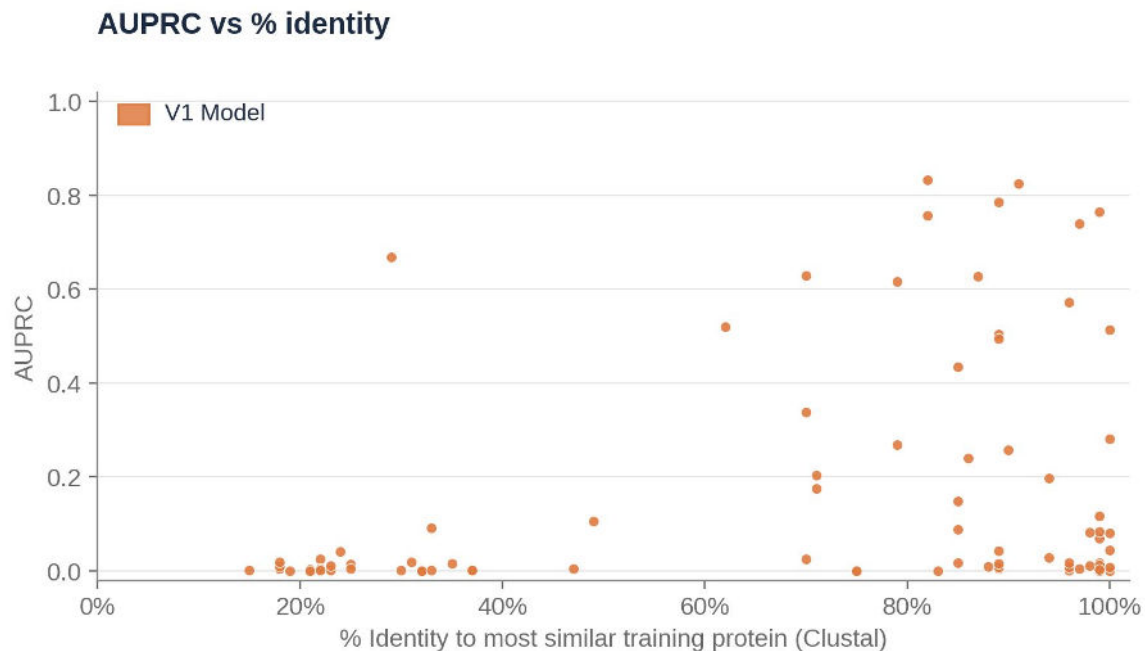
**DEL FM Actives with
Experimental Annotation**



Experimental Actives
Experimental Inactives

Rapid Improvement in Generalization to Novel Protein Space

First Generation DEL Foundation Model Highly Constrained by Sequence Similarity to Training Set

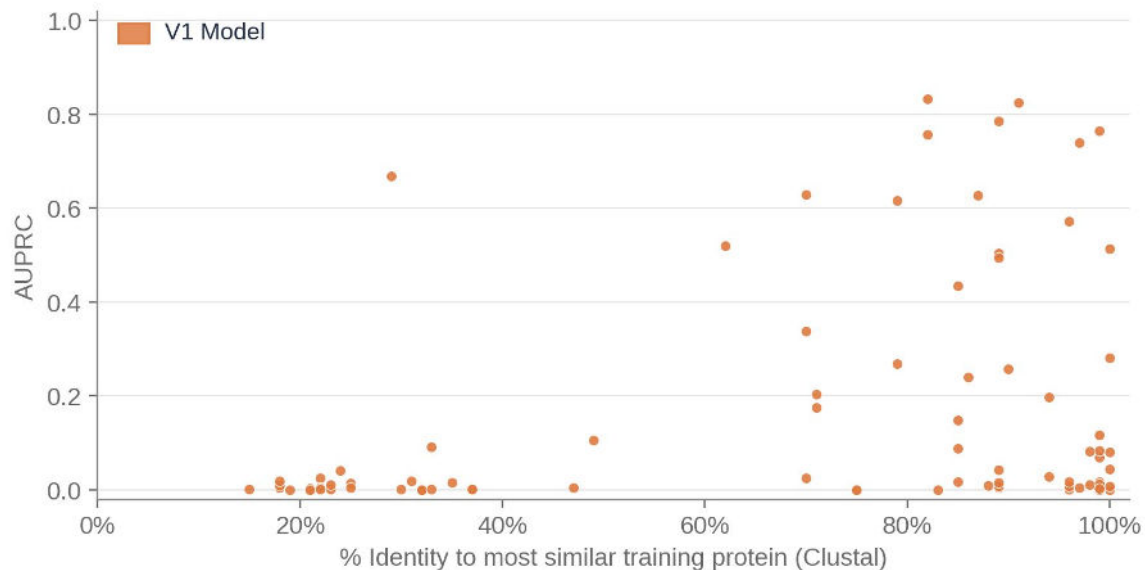


Each point = one held-out protein construct (V1: $n = 90$; V2: $n = 213$).

Rapid Improvement in Generalization to Novel Protein Space

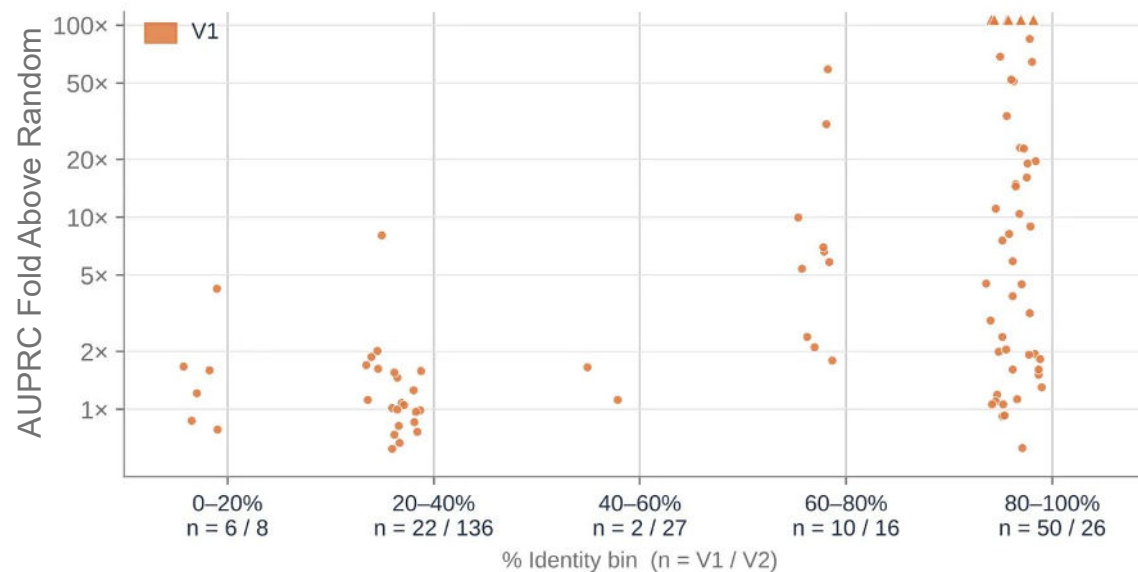
First Generation DEL Foundation Model Highly Constrained by Sequence Similarity to Training Set

AUPRC vs % identity



Each point = one held-out protein construct (V1: $n = 90$; V2: $n = 213$).

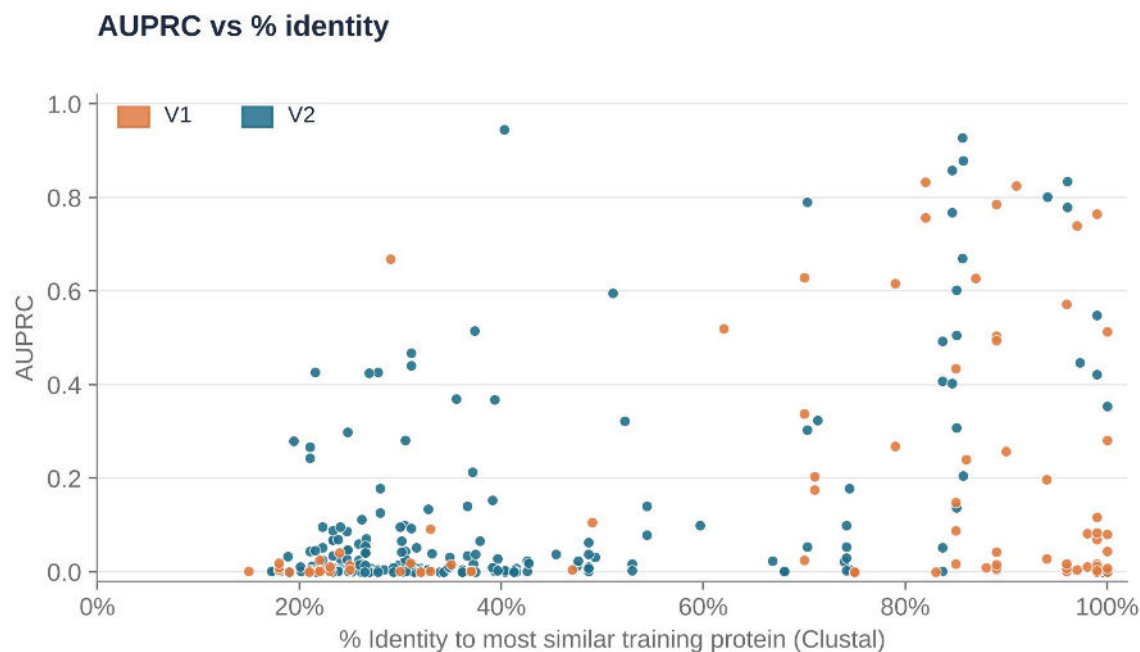
Fold above base rate by % identity (quintiles)



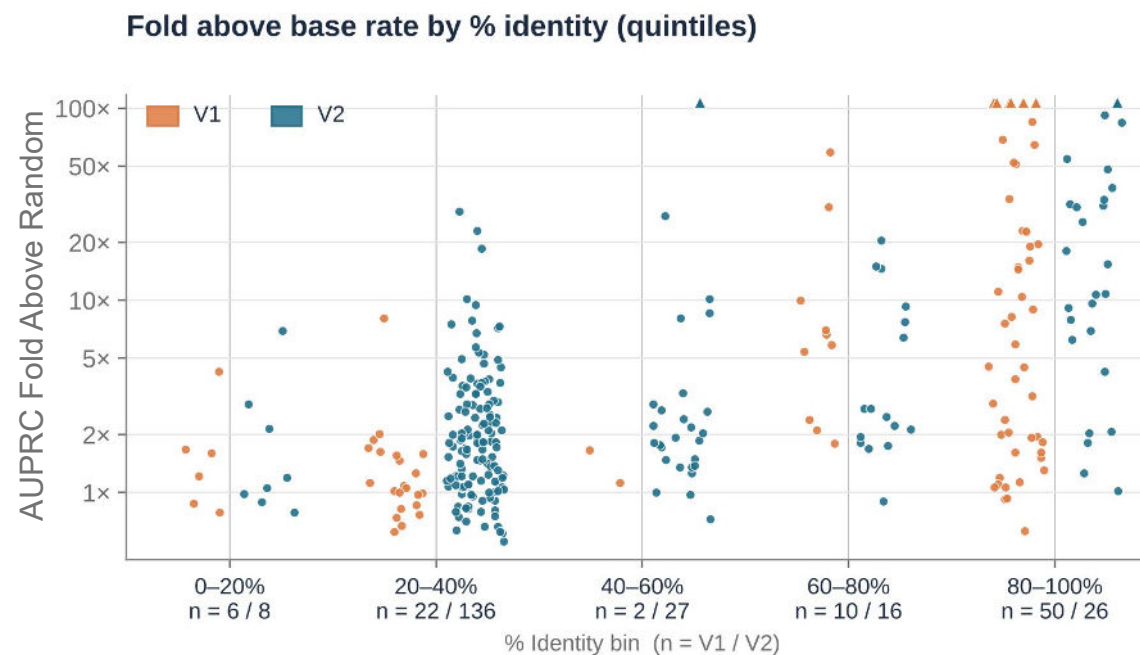
Each point = one held-out protein construct (V1: $n = 90$; V2: $n = 213$); jittered within each bin.
▲ = value above 100x (axis capped).

Rapid Improvement in Generalization to Novel Protein Space

Second Generation Model Outperforms First Generation Model On Low Identity Proteins



Each point = one held-out protein construct (V1: $n = 90$; V2: $n = 213$).

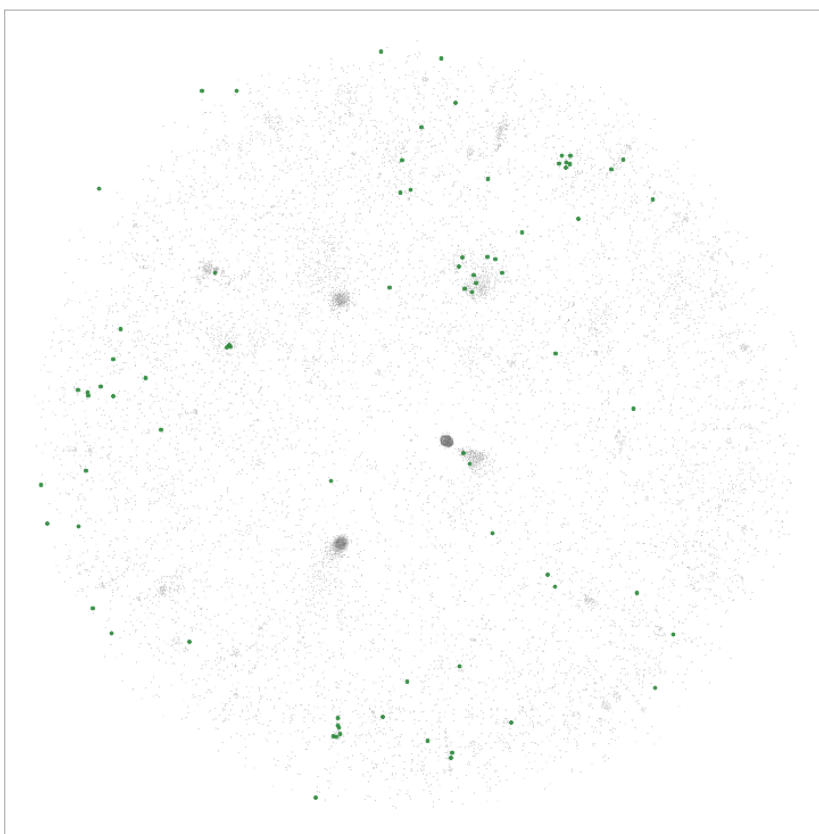


Each point = one held-out protein construct (V1: $n = 90$; V2: $n = 213$); jittered within each bin.
▲ = value above 100x (axis capped).

Broad Regions of the Proteome are Accessible with DEL Foundation Model

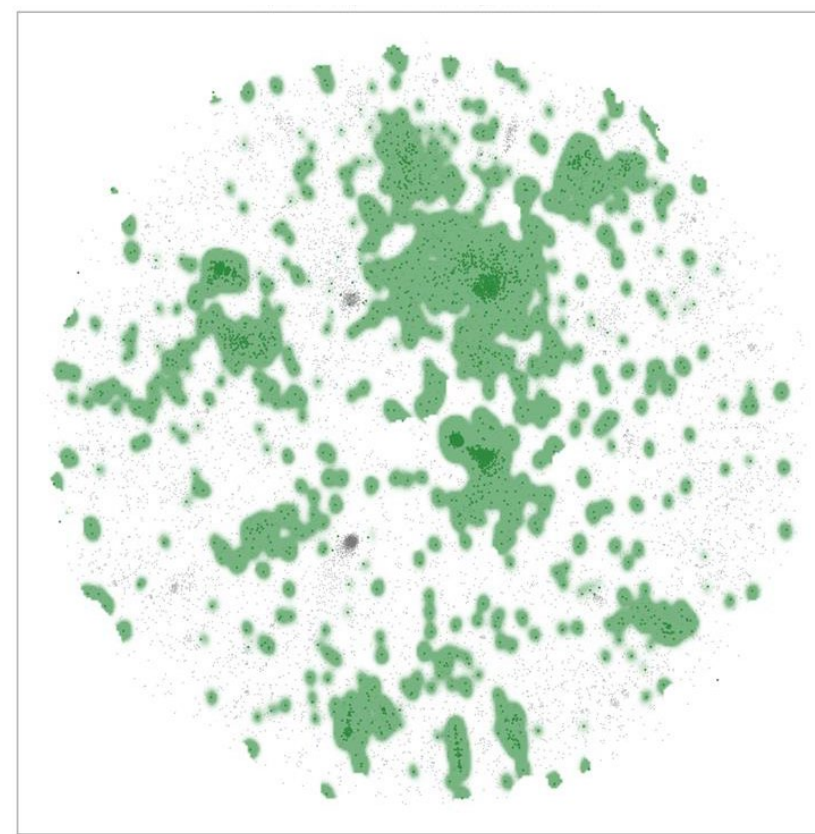
By learning a general relationship between protein sequence and molecular structure, DEL FM transforms our existing wet-lab DEL data into a predictive model for hit and lead ID across a much larger portion of the proteome

What we have accessed with wet-lab DEL



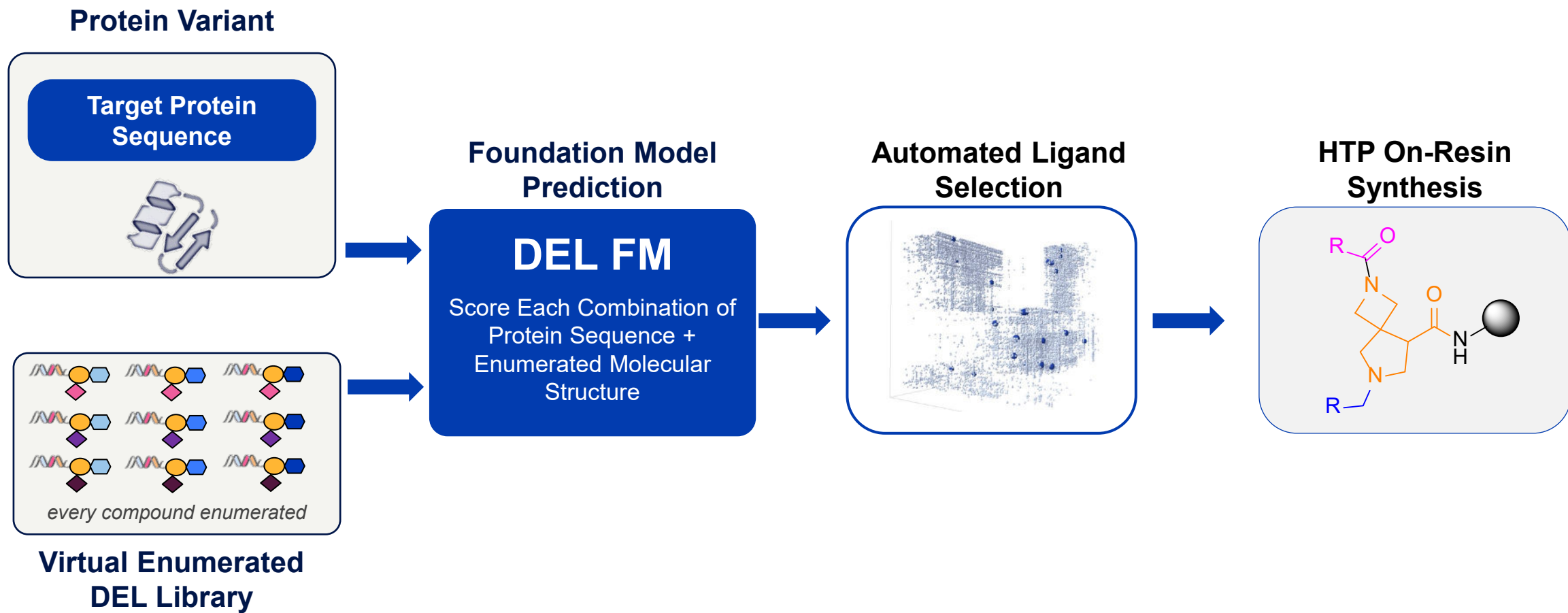
DEL FM Training Set Proteins

What we might access using DEL Foundation

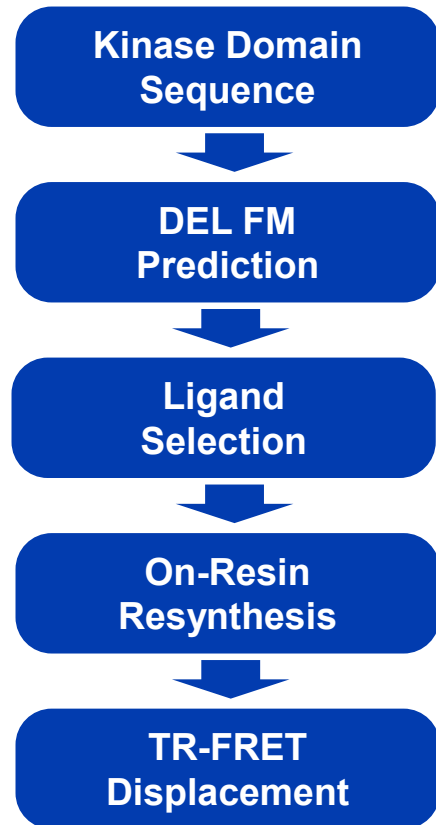
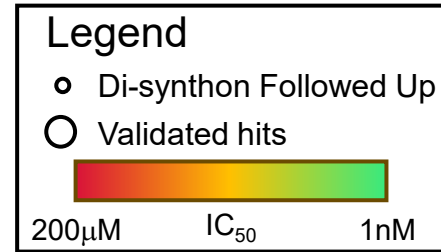
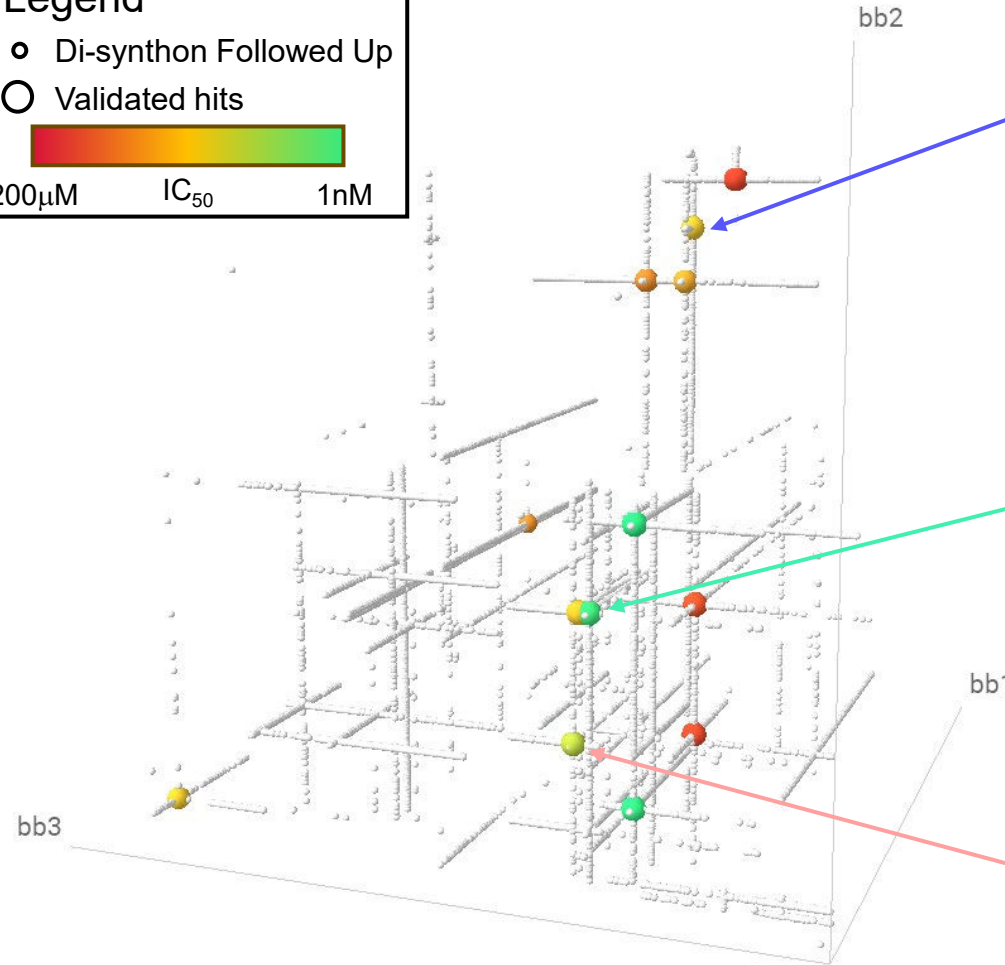
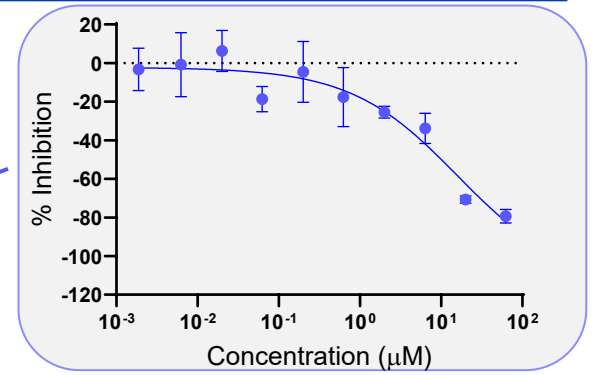
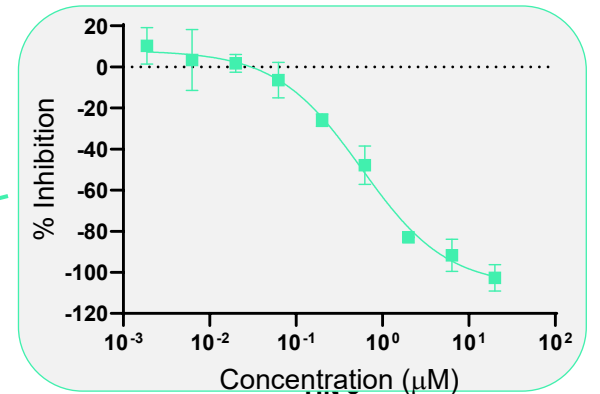
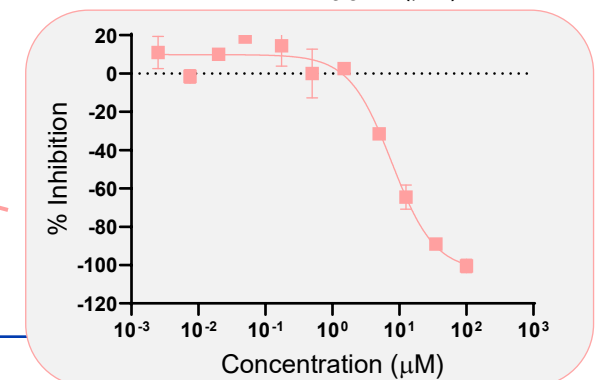


Proteome Coverage of DEL FM

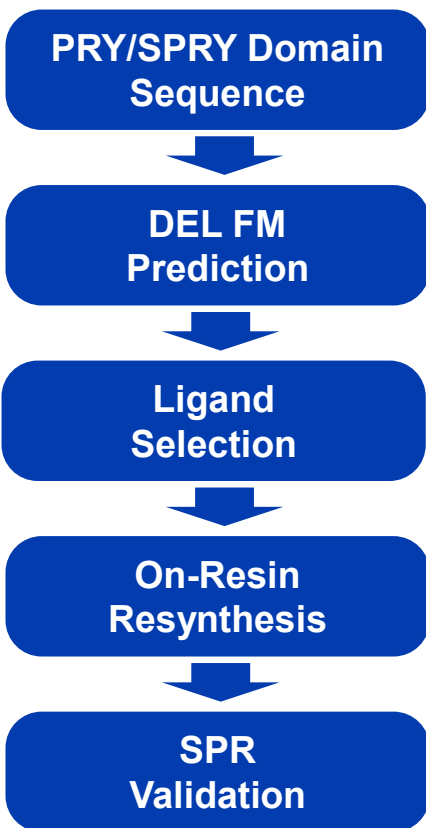
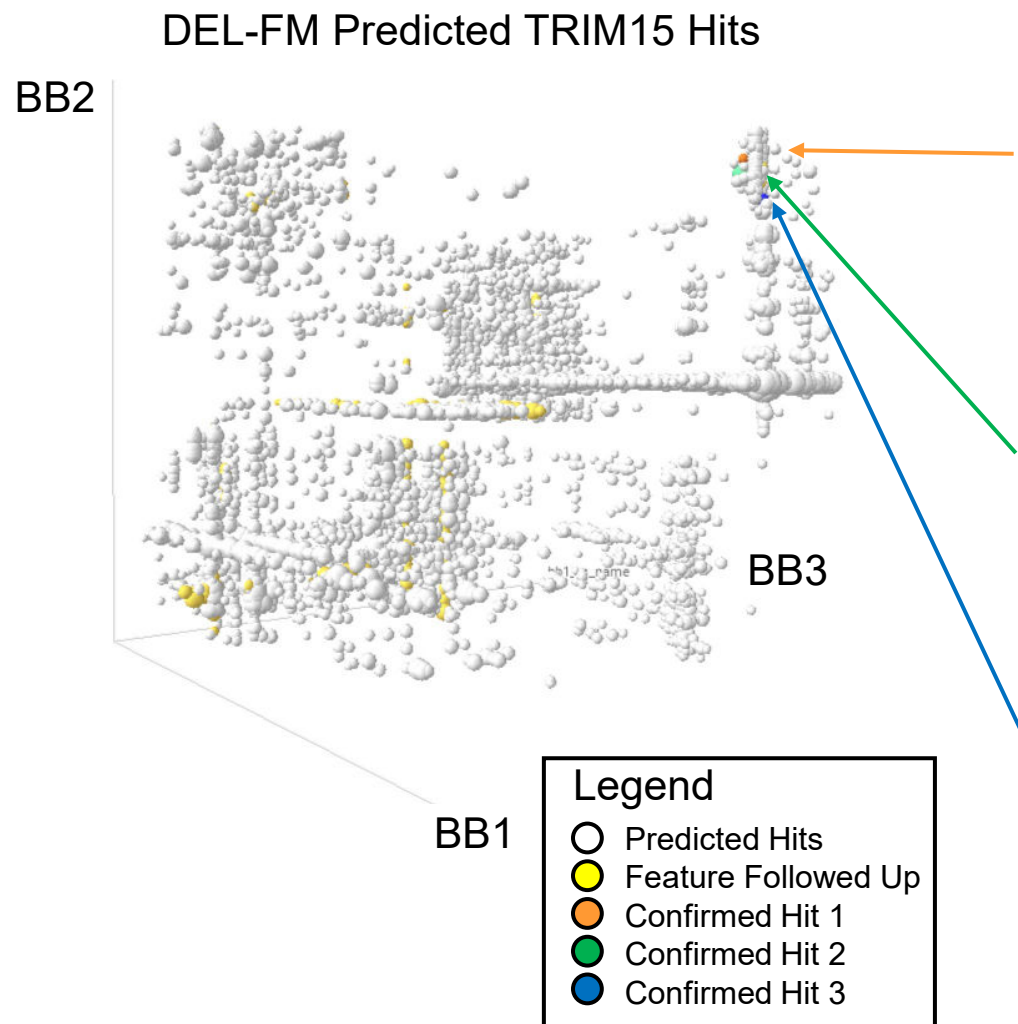
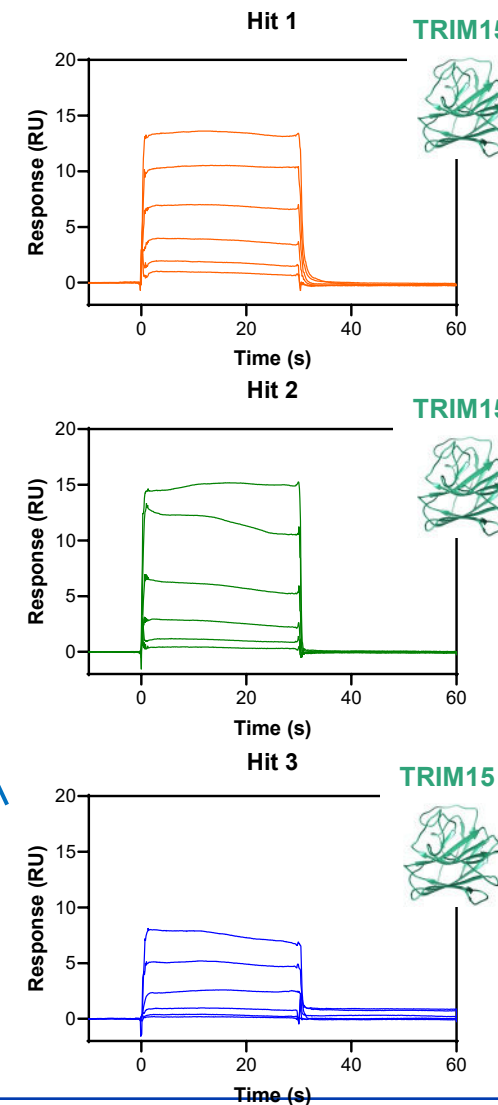
Automated Prediction and Ligand Selection Pipeline Validated for Screening Unseen Targets



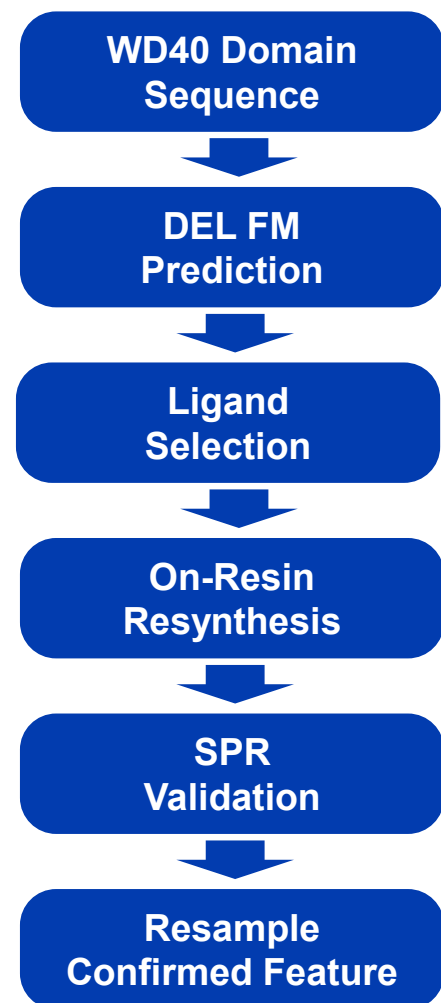
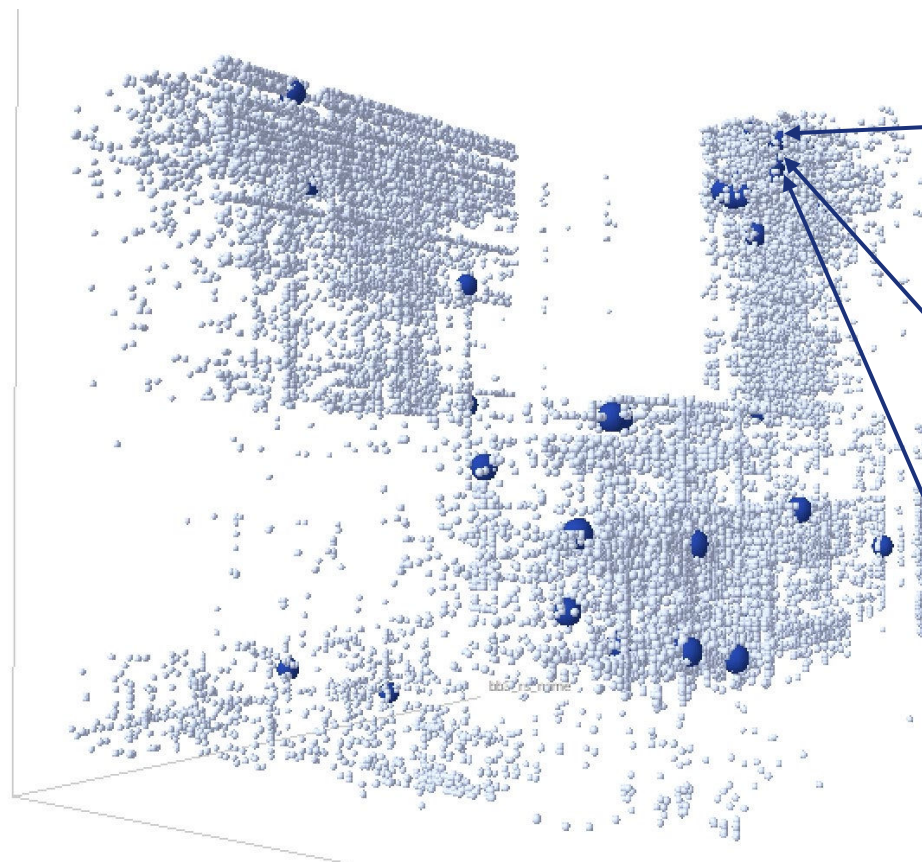
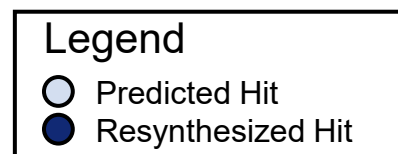
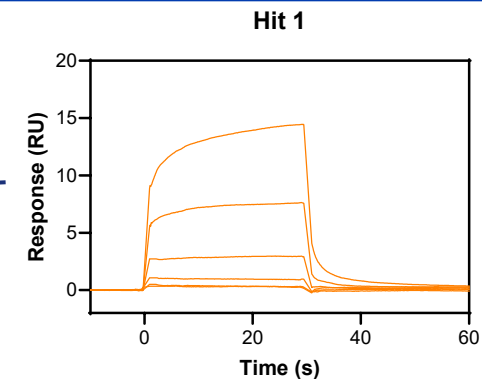
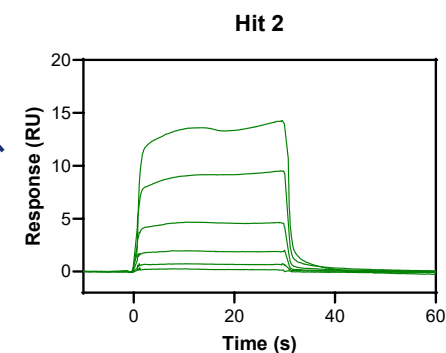
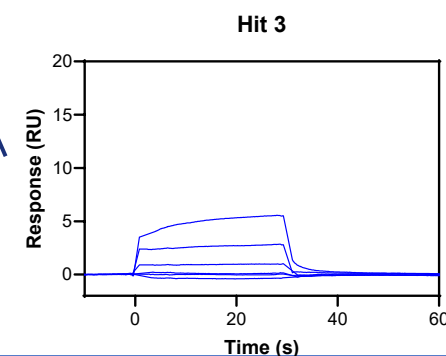
Initial Prospective Test of First Generation Model: DEL Foundation Identified Diverse Set of High Affinity Binders for Kinase X



Second Test: Multiple Distinct Chemical Series Confirmed From Initial Foundation Model Predictions for TRIM15



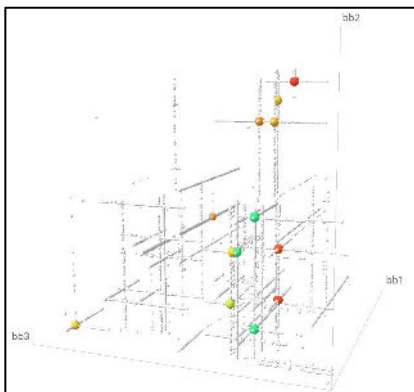
Third Test: Novel Series for An Additional Undrugged Ligase



First Gen DEL Foundation Model Highly Productive in Prospective Screening

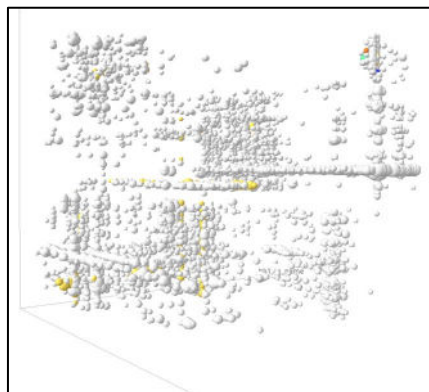
First Generation DEL Foundation Model Has Yielded Binders for All Targets Evaluated to Date

Kinase



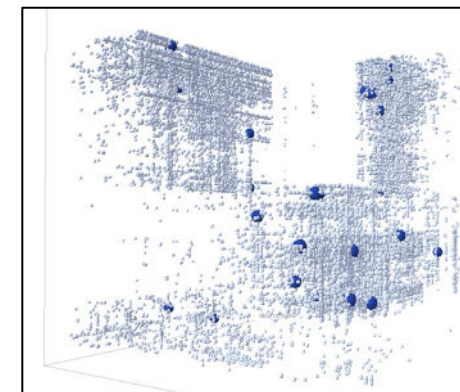
Metric	Value
Compounds Tested	55
Confirmation Rate	20%
Best Affinity	0.17 μ M

TRIM15



Metric	Value
Compounds Tested	55
Confirmation Rate	7%
Best Affinity	11.4 μ M

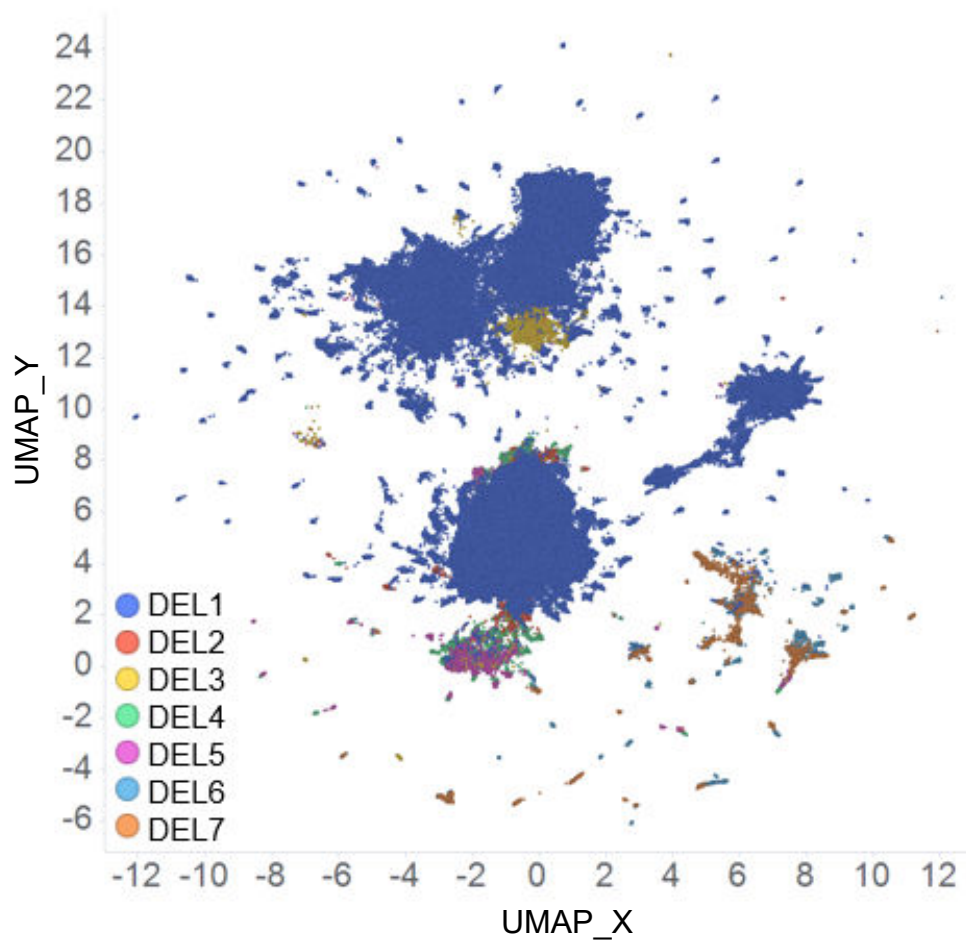
Ligase 1



Metric	Value
Compounds Tested	45
Confirmation Rate	13%
Best Affinity	8.3 μ M

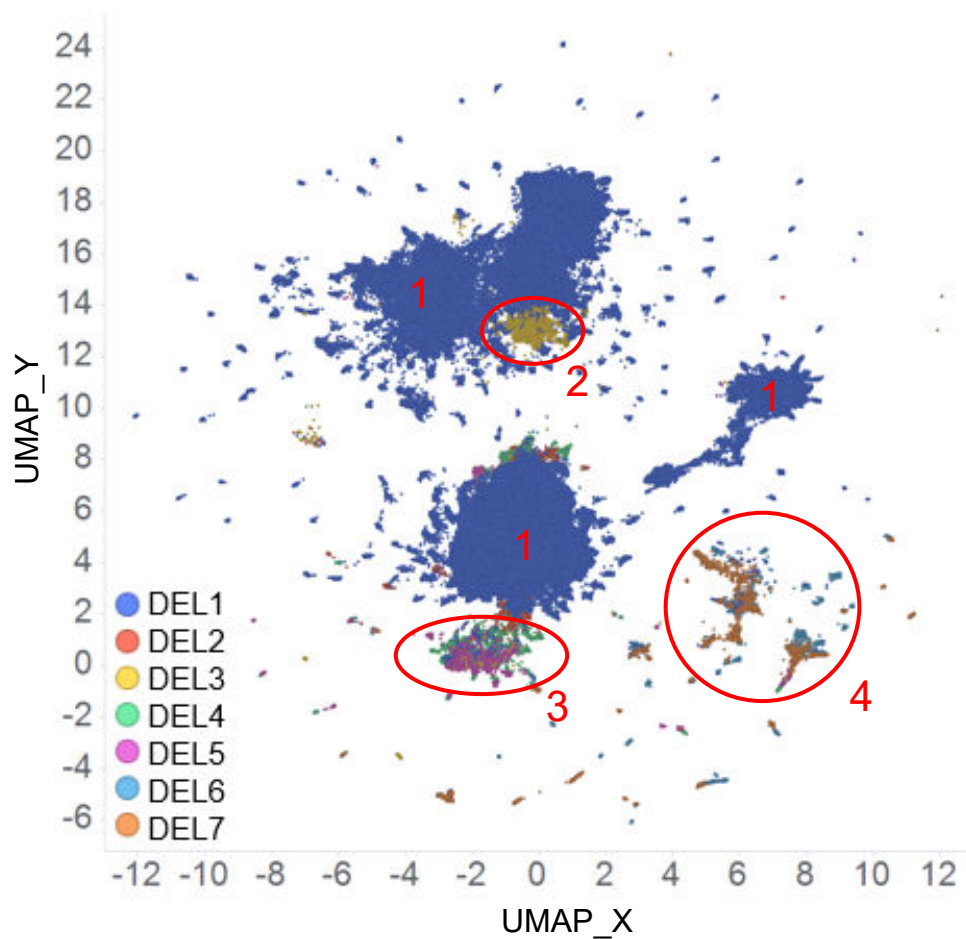
DEL FM Generalizes to New Chemical Space

UMAP of Nurix DELs in DEL FM Training Set

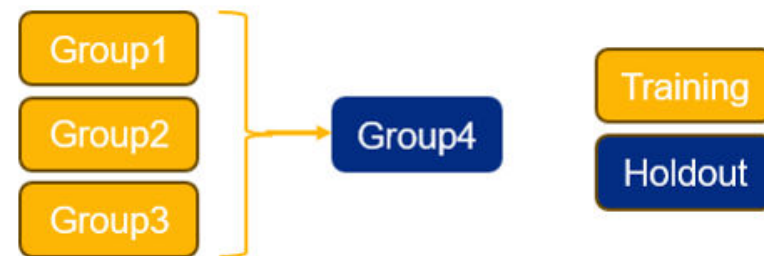


DEL FM Generalizes to New Chemical Space

UMAP of Nurix DELs in DEL FM Training Set

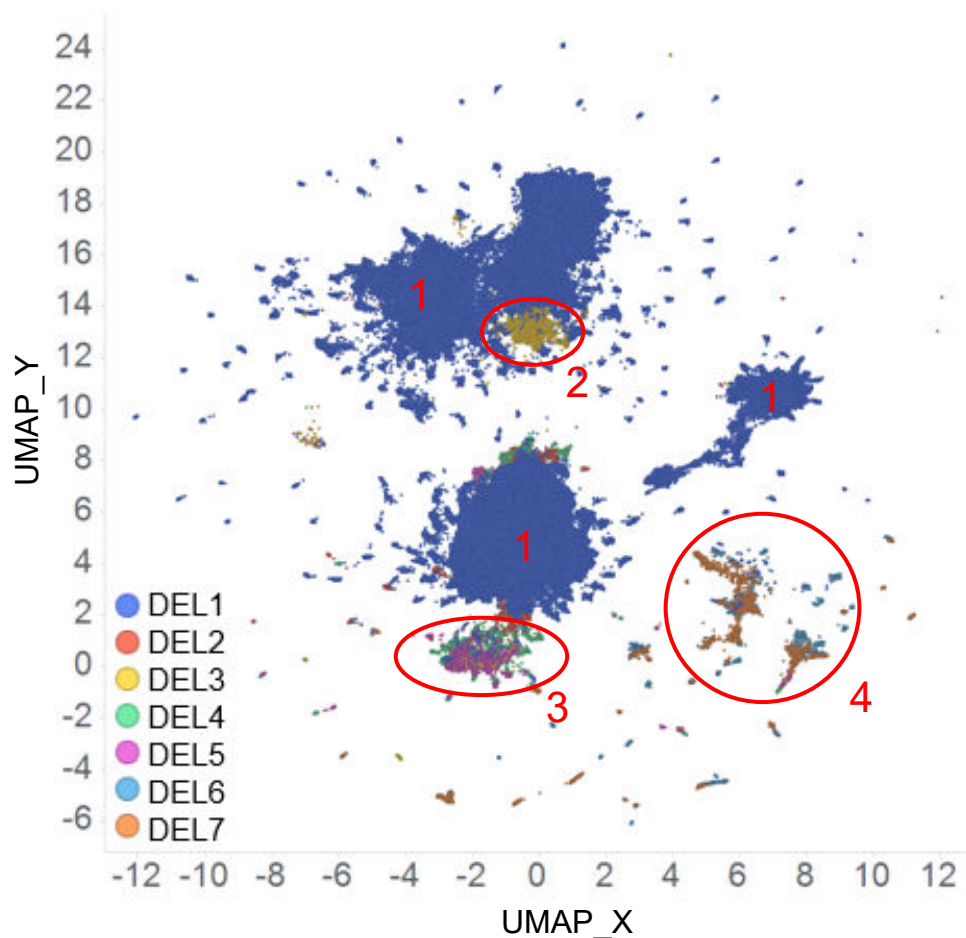


Schematic of Leave-One-Out Experiment

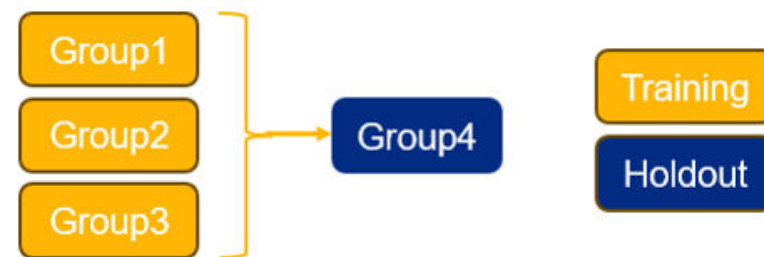


DEL FM Generalizes to New Chemical Space

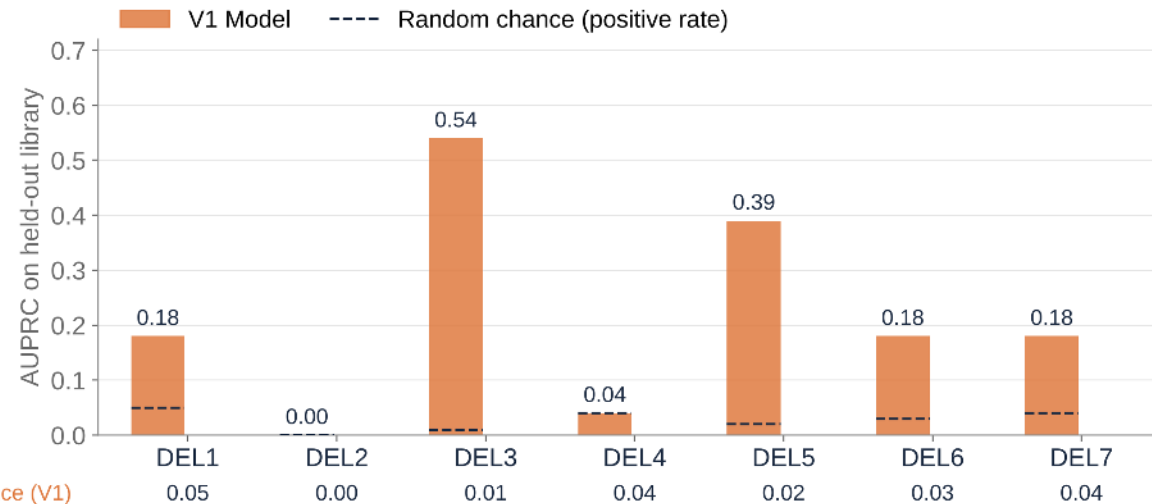
UMAP of Nurix DELs in DEL FM Training Set



Schematic of Leave-One-Out Experiment

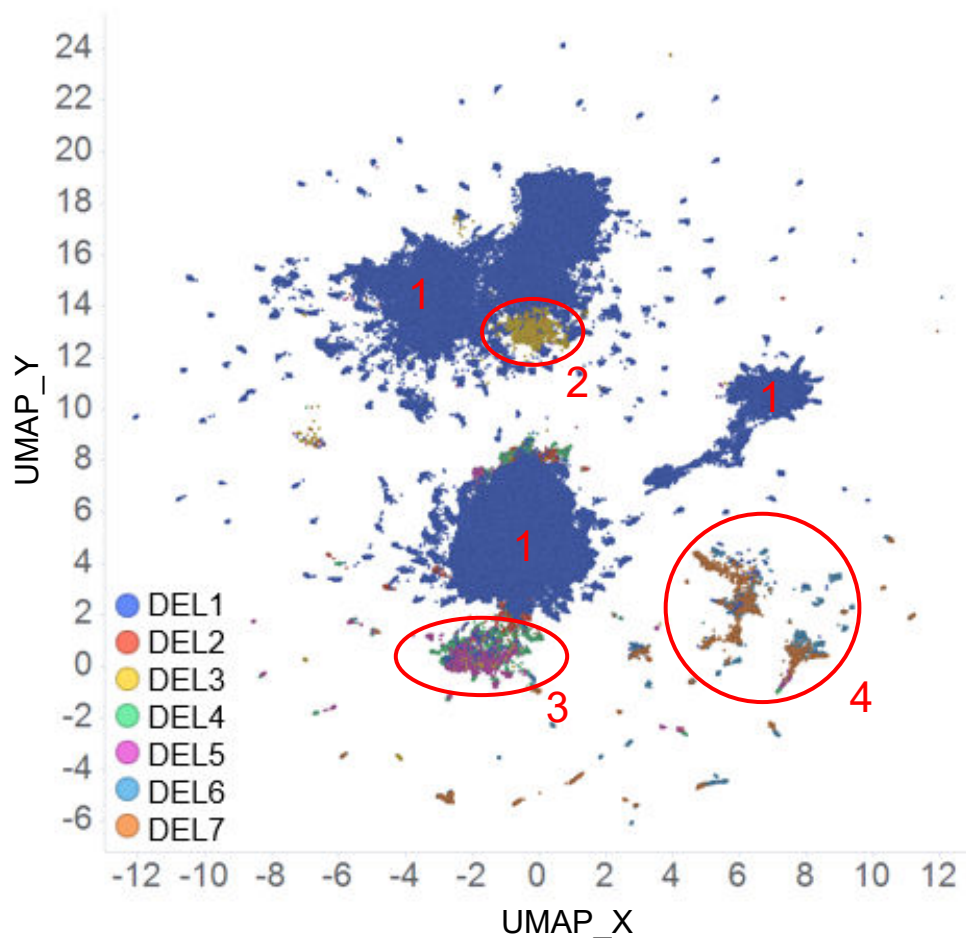


AUPRC on held-out DEL libraries

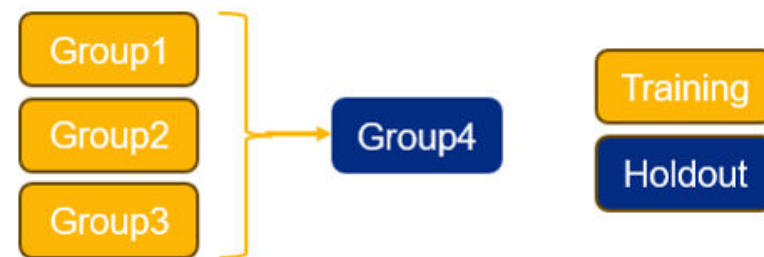


DEL FM Generalizes to New Chemical Space

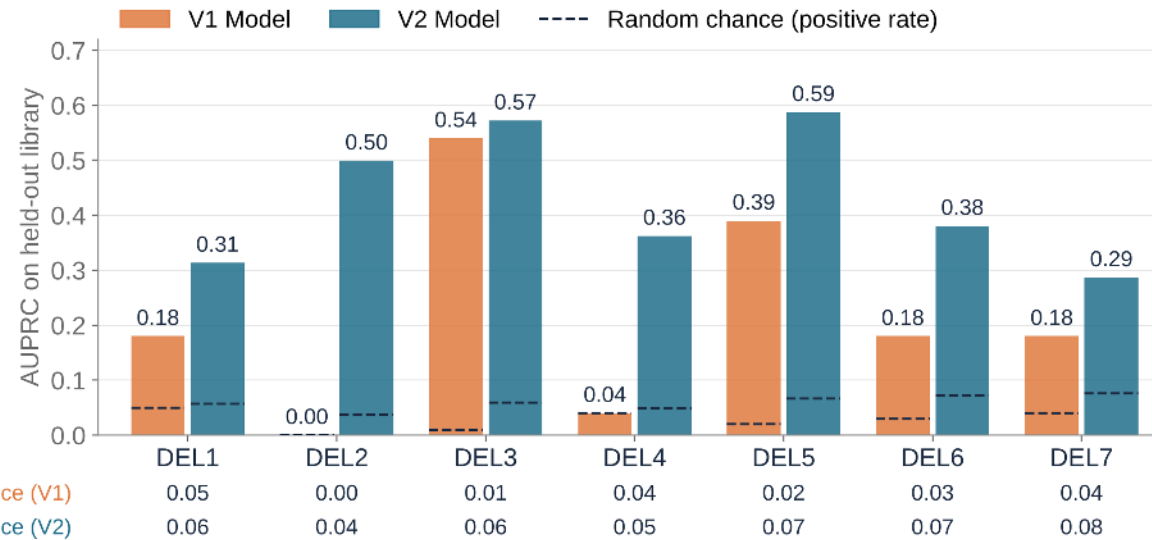
UMAP of Nurix DELs in DEL FM Training Set



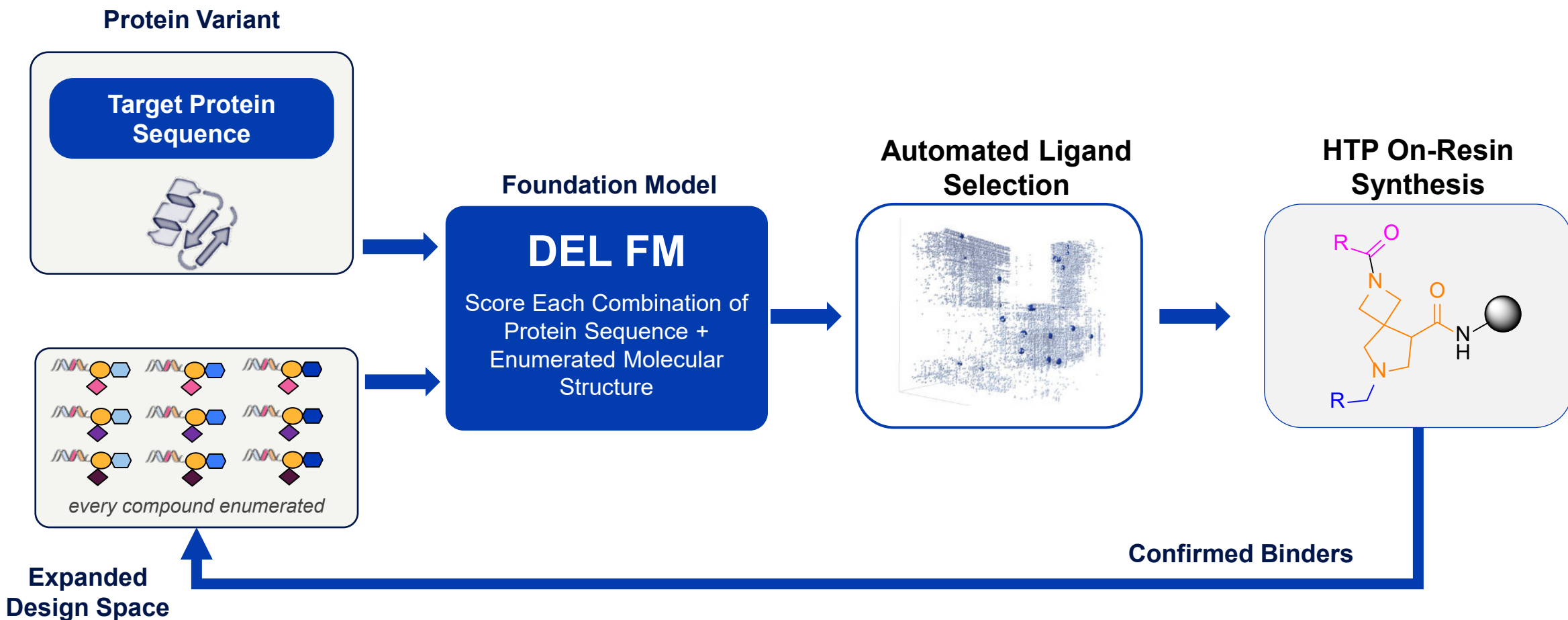
Schematic of Leave-One-Out Experiment



AUPRC on held-out DEL libraries

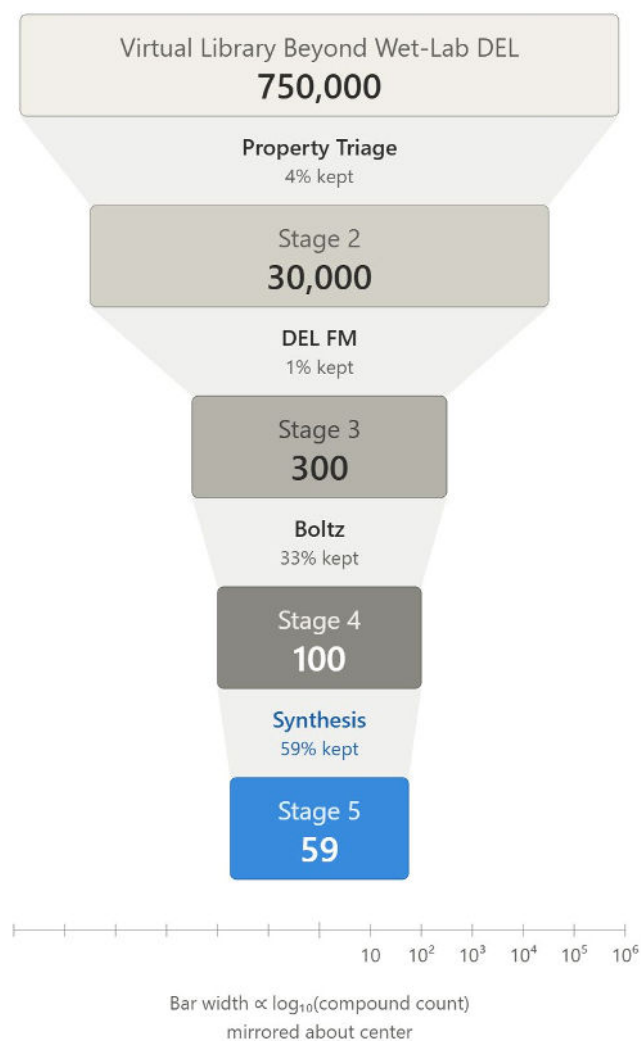
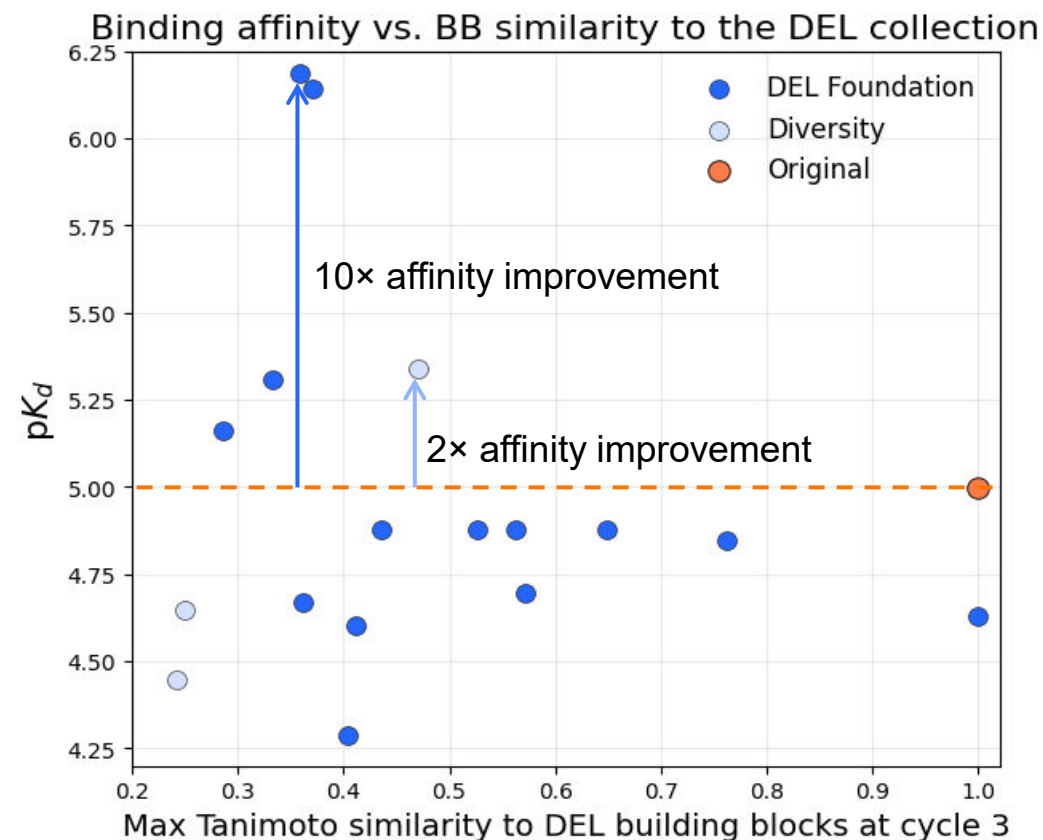


DEL Foundation Model Can Also Be Employed for Hit Optimization



DEL FM Can Also Be Used to Optimize Model-Derived Hits:

Suggests DEL Foundation Generalizes Beyond Training Data

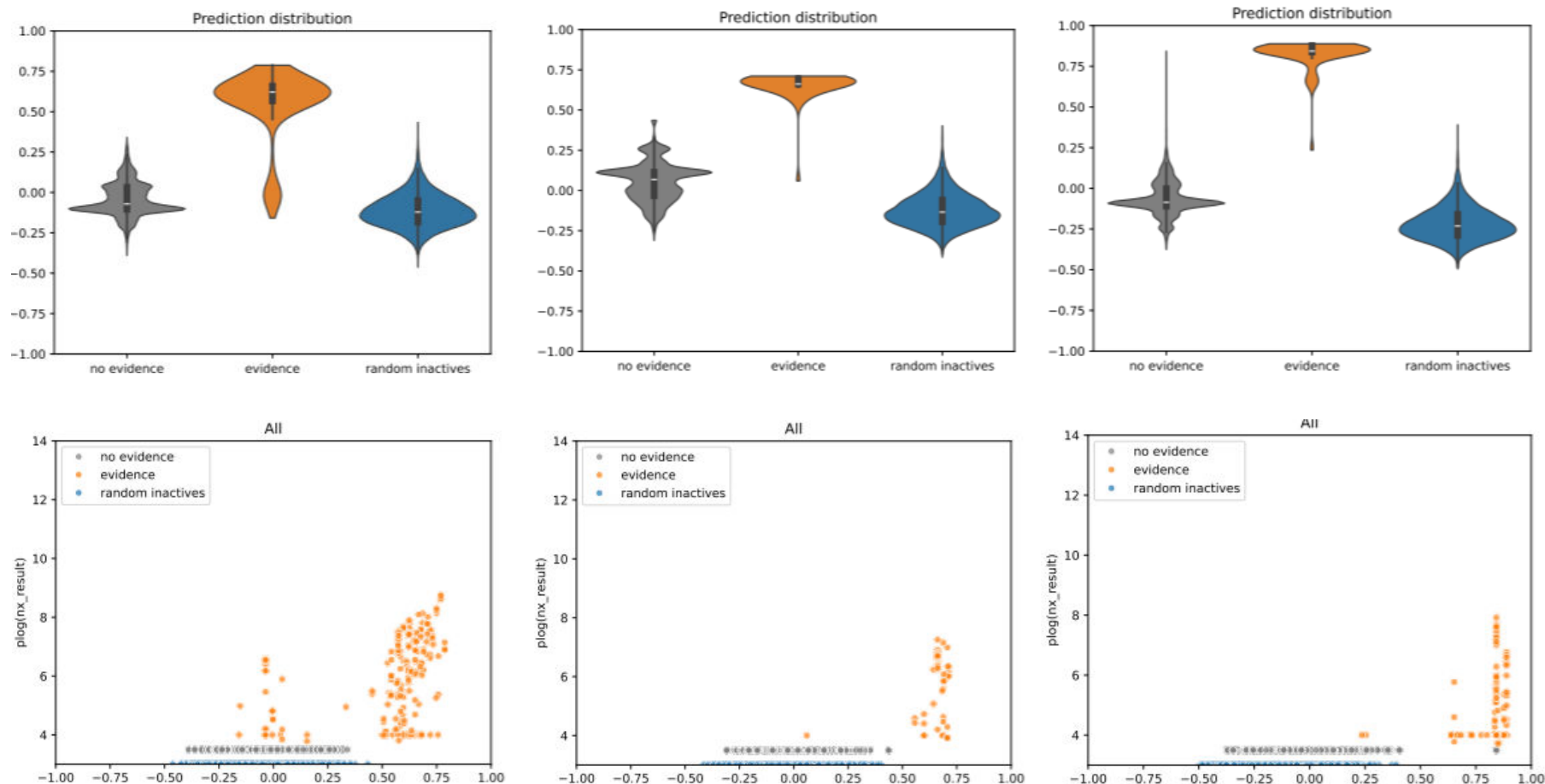


Selection Method	Confirmation Rate
DEL FM + Boltz	23%
Diversity	10%

Identification of Experimental Actives from DEL Adjacent Chemical Space

DEL FM predictions on our aggregated DEL Lead ID and Lead Optimization dataset cleanly separate confirmed binders from non-binding subsets

- Experimentally Tested DEL and DEL Adjacent Ligands
- Resynthesized ligands with no demonstrated target binding
- Random subsampling of the DEL Library



Summary

DEL FM is a powerful tool for drugging the proteome

- Nurix's DEL Foundation Model is trained on our aggregated DEL dataset covering diverse protein classes and proprietary library collection
- DEL FM enables virtual DEL experiments which generalize to both new proteins and new chemical space
- Nurix has validated the application of DEL FM to both Lead Identification and Lead Optimization
- Nurix is leveraging DEL FM to identify ligands for targets across the proteome

Thank You to All Contributors Across Loka, Nurix, and AWS



Machine Learning &
Cheminformatics
Information Technology
Informatics
Protein Sciences
DEL Chemistry
DEL Screening
Biophysics
Drug Discovery Technology

Cristiana Carpinteiro
Julian Fernandez
Zafir Stojanovski
Maria Loureiro
Manuel Ravasqueira
Telmo Felgueira
Jorge Moura Sampaio
Kiril Zelenkovski
Patricia Rocha

Sushmitha Regulapati
Teona Kostova
Emily Kruger
Matheus Deister
Helena Ferreira

Jennifer Gruefe
Jim Davis

