New inhibitors of oncology targets discovered using the affinity-mediated selection of DNA-encoded chemical libraries followed by machine-learning

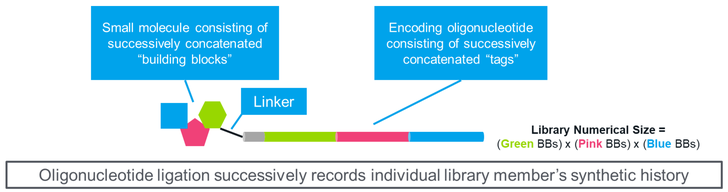
Anthony D. Keefe, *Mohamed AbouZleikha, AJ Baghaie, Matthew A. Clark, Marie-Aude Guié, John P. Guilinger, Ryan Walsh, Ying Zhang* X-Chem, Inc., Waltham MA, USA

X-CHEM
TAKE THE LEAD. DRIVE DISCOVERY.

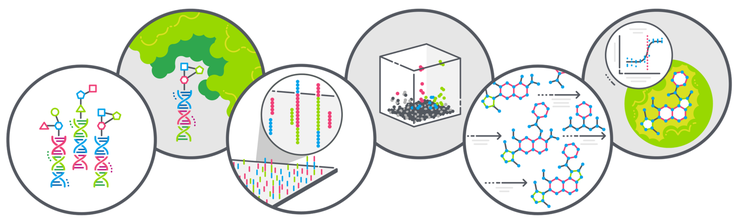
Poster Abstract # 1826

DNA-Encoded Chemistry

DNA-Encoded Chemical Library (DEL) technology enables the rapid discovery of small molecules that bind to distinct protein targets of therapeutic value. Individual compounds are synthesized and screened as mixtures enabling the screening of very large numbers of compounds. Affinity-mediated selection enriches target-engaging compounds which are subsequently identified by deep sequencing to recover encoded chemical information.

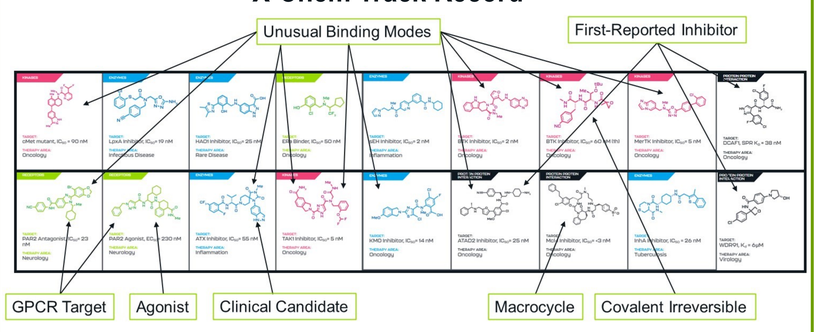


X-Chem Lead Discovery Platform



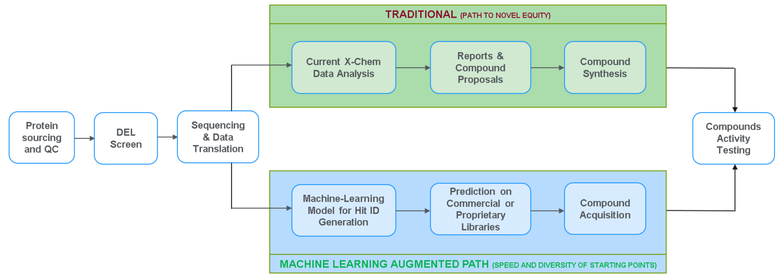
Design and synthesize DNA-encoded libraries. >100 billion compounds → Enrich for selective binders via affinity-mediated selection → Characterize selection output using deep sequencing → Identify enriched library compounds through statistical analysis → Synthesize representative compounds → Assay compounds and identify target engagers and inhibitors

X-Chem Track Record



Published compounds discovered using the X-Chem DEL Platform

DEL>ML>Catalog: An Alternative Approach to Hit ID

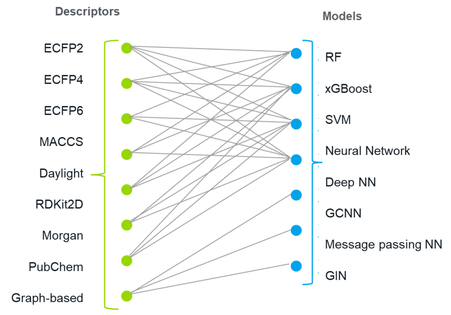


Searching for the Right Combination of Descriptor and Model

We utilize over 12 different representations and over 10 different ML methods.

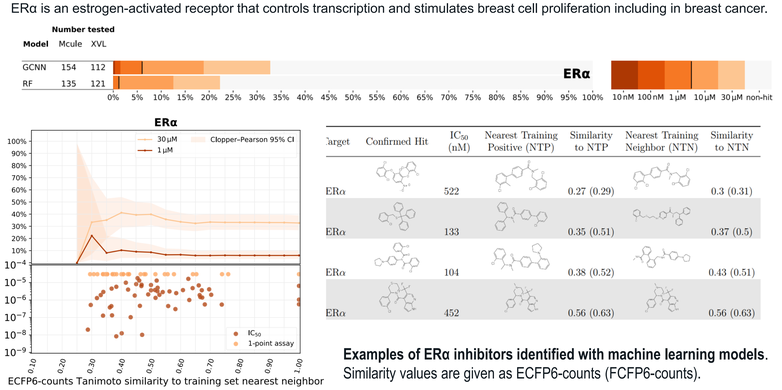
Each training process outputs a pool of up to 40 ML models, each model is the result of hyperparameter tuning of around **350** experimental variations.

High-performance computing and scalable infrastructure is required to enable training completion in days.



Case Study 1: ERα

ERα is an estrogen-activated receptor that controls transcription and stimulates breast cell proliferation including in breast cancer.



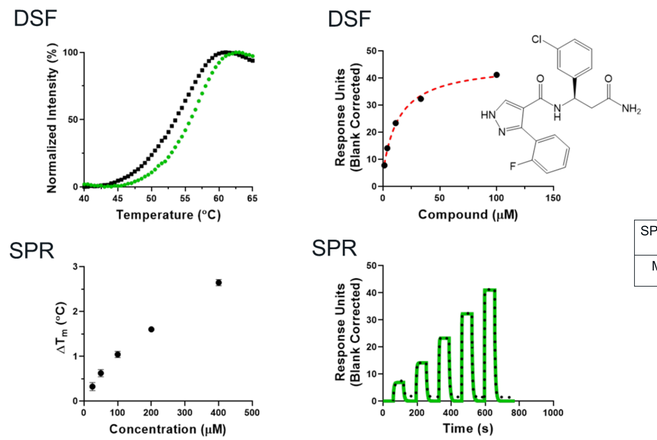
Target	Confirmed Hit	IC_{50} (nM)	Nearest Training Positive (NTP)	Similarity to NTP	Nearest Training Neighbor (NTN)	Similarity to NTN
ERα		522		0.27 (0.29)		0.3 (0.31)
ERα		133		0.35 (0.51)		0.37 (0.5)
ERα		104		0.38 (0.52)		0.43 (0.51)
ERα		452		0.56 (0.63)		0.56 (0.63)

Examples of ERα inhibitors identified with machine learning models. Similarity values are given as ECFP6-counts (FCFP6-counts).

McCloskey et al (2020) Journal of Medicinal Chemistry 63 (16), 8857-8866

Case Study 2: DCAF1

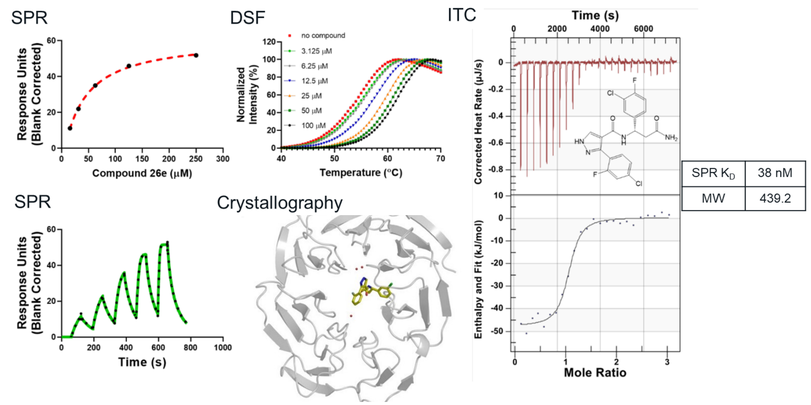
DCAF1 is the substrate-recognition component of an E3 ligase complex that is associated with the proliferation of colon cancer cells.



SPR K_D	11 µM
MW	386.8

Li et al (2023) Journal of Medicinal Chemistry 66 (7), 5041-5060

DCAF1 Compound Advancement



SPR K_D	38 nM
MW	439.2

Li et al (2023) Journal of Medicinal Chemistry 66 (7), 5041-5060

Summary

We have established a DEL>ML>Catalog>Hit ID platform that utilizes our deck of DNA-Encoded chemical compounds, affinity-mediated selection and DNA sequencing to identify populations of enriched compounds. The structures of these enriched compounds are then featurized using a range of representations and the resultant fingerprints are subjected to machine-learning using a range of architectures. The most successful models are combined and used to score catalog compounds, the highest scoring of which are then sourced and tested. This strategy has rapidly provided attractive inhibitors for a range of oncology targets including those for which no inhibitors were previously known. Two case studies, DCAF1 and ERα, are presented as examples.