OpenDEL™

Starting a Journey to Access the Vast DEL Space

HitGen Inc.

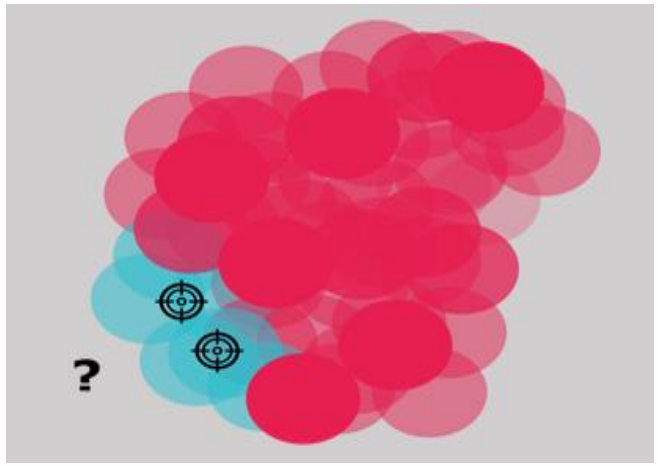
OpenDEL™: Empowering Your Drug Discovery Journey

➢ **AI/ML**

- Post-selection DEL data for the prediction of new chemical space outside DELs

➢ **Assessment of Target Ligandability**

- To perform screening of novel targets for assessment of target ligandabilities



➢ **Hit Discovery**

- To directly discover novel compounds through screening for the purpose of drug development

OpenDEL™: A Self-service DEL Product

OpenDEL™ - Small molecules


DEL selection sample


PCR forward primer (Primer F)


PCR reverse primer (Primer R)


Standard selection protocol


qPCR standard curve sample


Salmon sperm DNA (ssDNA)

OpenDEL™ - Macrocycle

✓ **To Access**

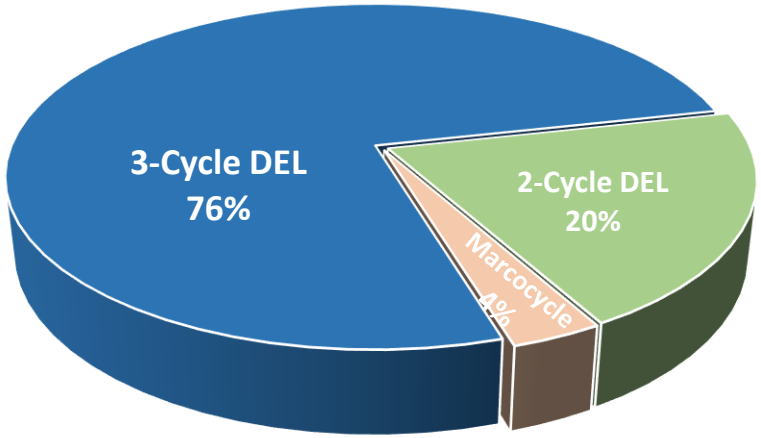
- 10 DEL samples for experiments (support 2-3 targets screening per kit)
- Fully enumerated molecules
- Building Block Structures
- DNA Codon Sequences
- Scaffolds Information

✓ **No Structure Disclosure Fee**

✓ **No Compound IP License Fee**

Users only need to prepare biological targets!

OpenDEL™: Library Content



>4,000,000,000

Diverse and Druglike Compounds

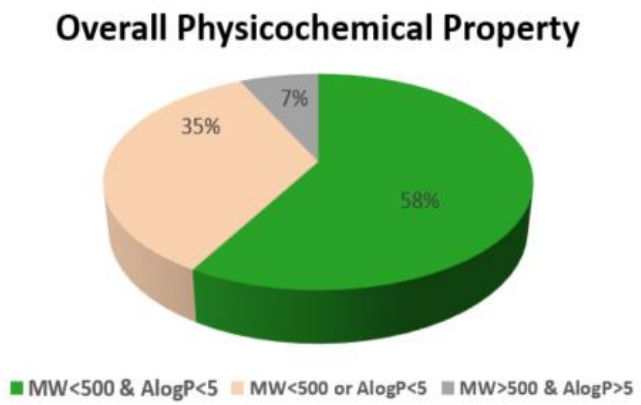
OpenDEL™ - Small molecules

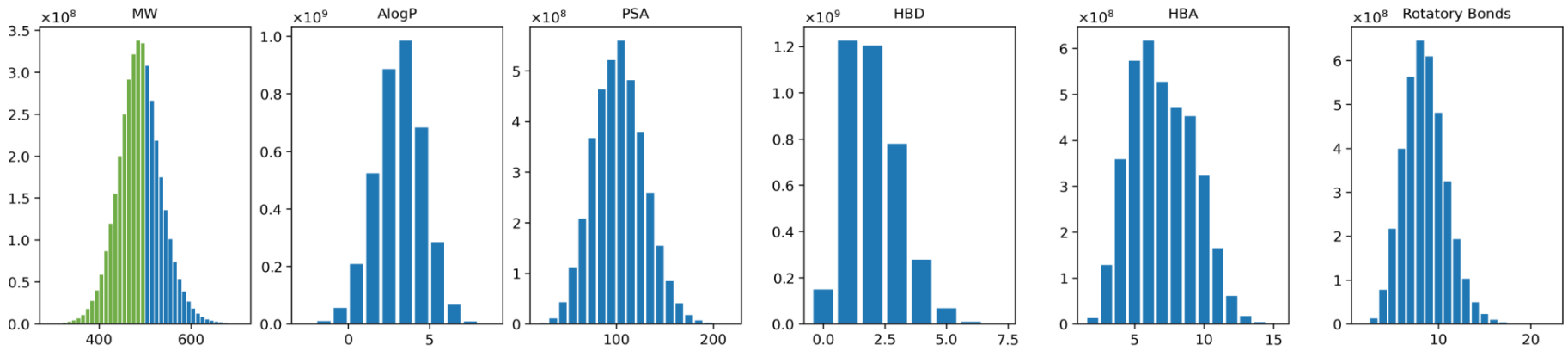
- **57 Small Molecules Encoded Libraries**
- **>10 2-Cycle Libraries, ~20M Compounds**
- **>40 3-Cycle Libraries, ~3.8 Bn Compounds**

OpenDEL™ - Macrocycle

- **1 Macrocycle+1 Linear control, ~200M Compounds**

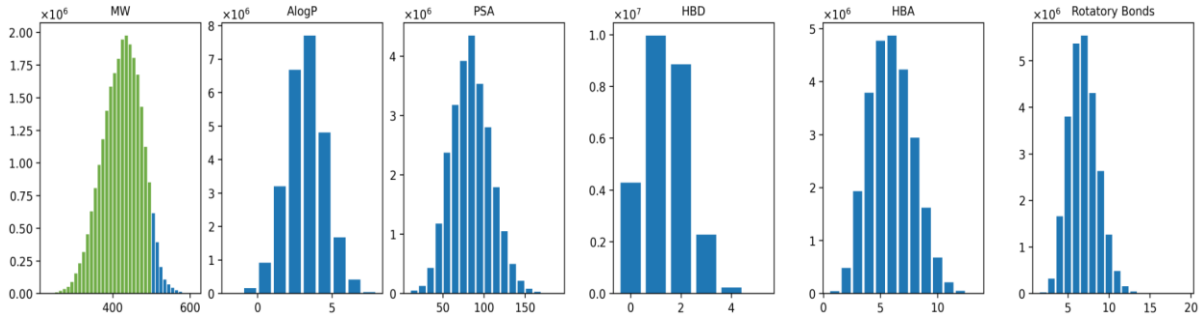
Physicochemical Property of OpenDEL™ Small Molecules



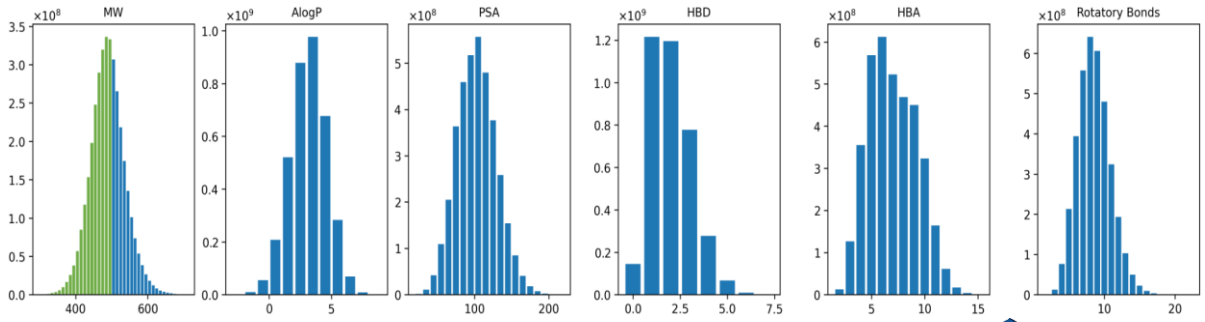


	Average MW (Da)	MW<500 & AlogP<5	MW<550	MW<500	MW<450
2-Cycle DELs	419	86.8%	99.7%	94%	67%
3-Cycle DELs	484	58%	90.6%	60.7%	19%

Physicochemical Property of 2-cycle Library Molecules

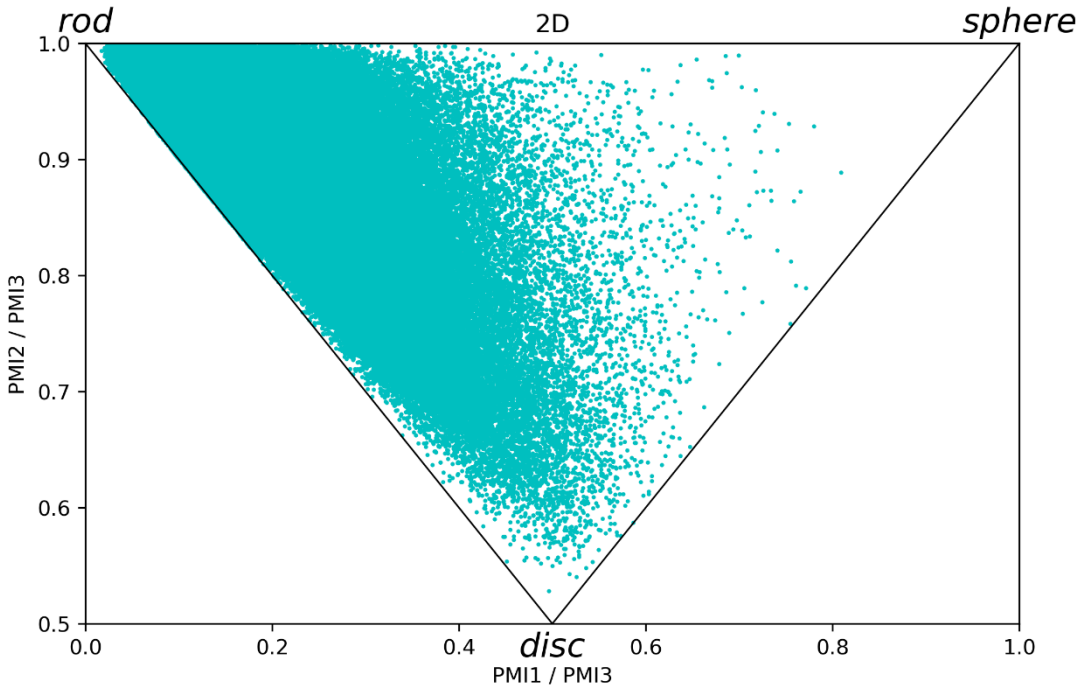


Physicochemical Property of 3-cycle Library Molecules

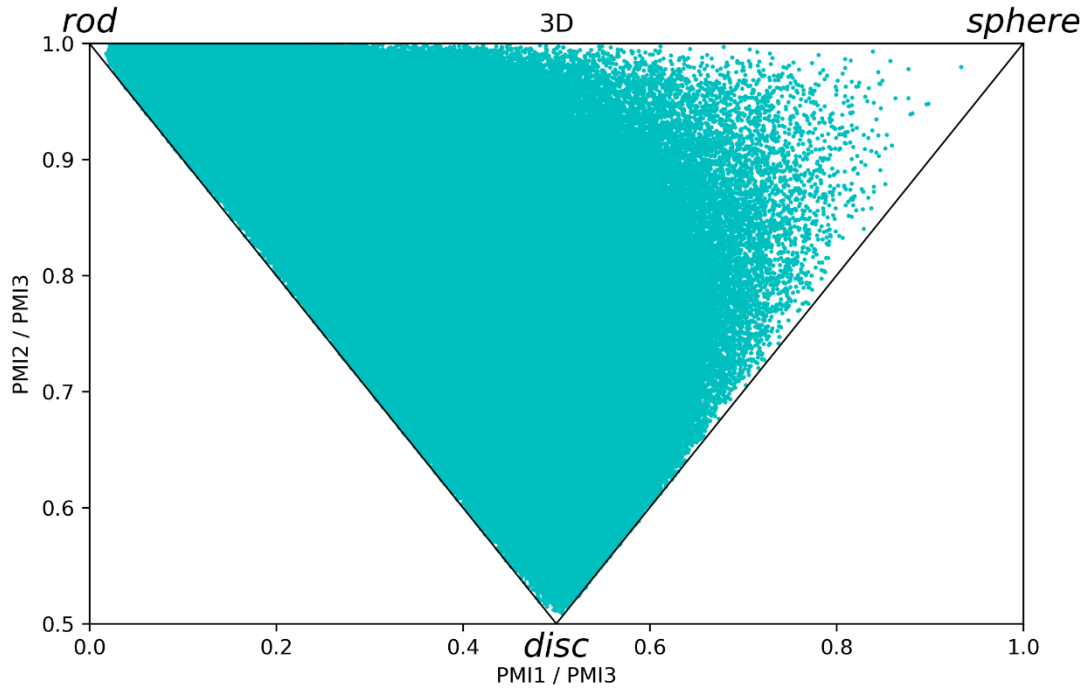


Topology Diversity of OpenDEL™ Small Molecules

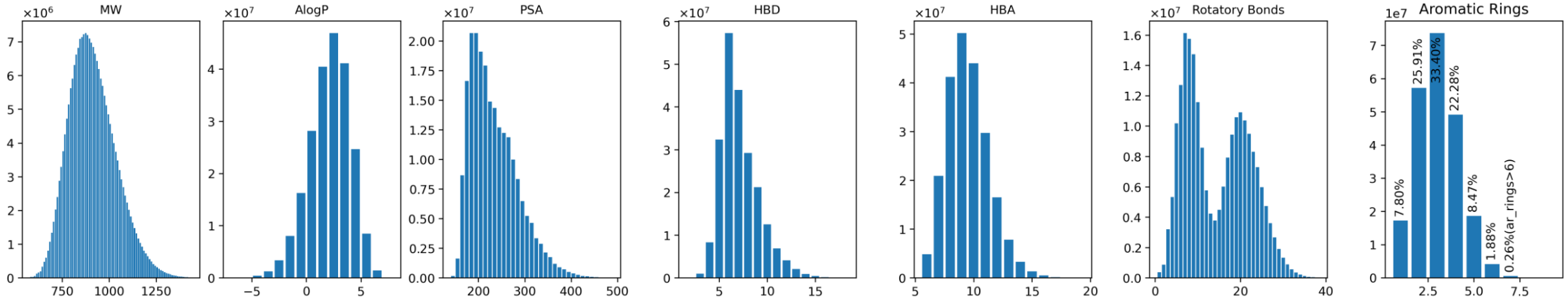
- OpenDEL™ 2-Cycle DEL Molecules



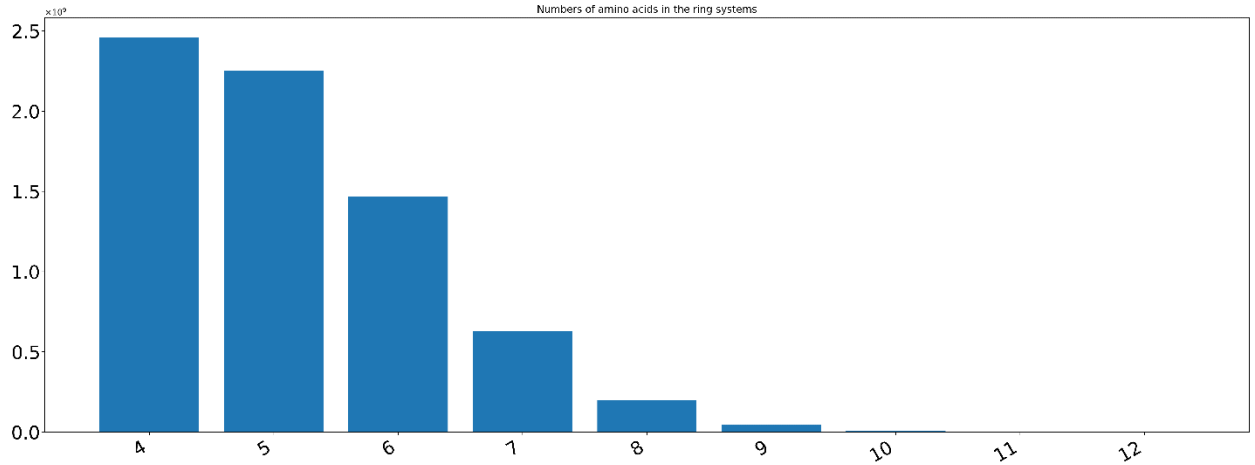
- OpenDEL™ 3-Cycle DEL Molecules



Macrocycle Physicochemical Property Distribution



Number of amino acids in the ring systems



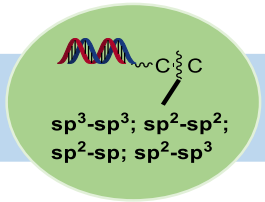
OpenDEL™ — Meeting Needs Through Choice

Option	Kit	Content	Size	Kit content
1	OpenDEL™5.0 Standard Kit	57 small molecule DELs	~3.8Bn	10 tubes, 10^6 copy
2	OpenDEL™5.0 Standard Kit +OpenDEL™-Macrocycle	59 DELs (57 small molecule DELs+ 1 macrocyclic DEL and 1 linear control)	~4Bn	1.Two separate kits will be delivered 2. 10 tubes per kit,10^6 copy 3. one for small molecules, one for macrocycle
3	OpenDEL™- Macrocycle	1 macrocycle+1 linear control	~230M	10 tubes, 10^6 copy

HITGEN

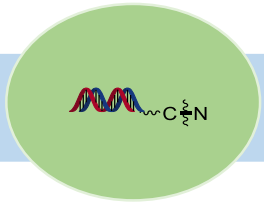
Diversity of On-DNA Chemistry and Building Blocks

Chemistry Diversity



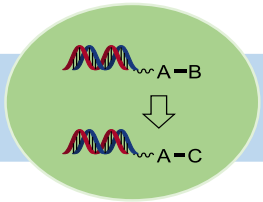
C-C Bond Formation

- Indole C3 alkylation
- Suzuki coupling
- Photoredox
- Sonogashira coupling



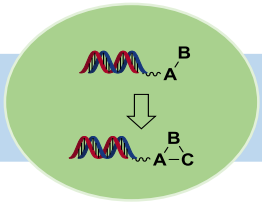
C-N Bond Formation

- Acylation;
- Reductive Amination with Aldehyde
- Reductive Amination with Ketone
- Buchwald-Hartwig Cross Coupling
- SN_2/SN_{Ar}
- Sulfonylation



Functional Group Transformation

- NO_2 to Amine
- Azide to amine
- Amine to azide
- Halide to azide
- Halide to amine
- Halide to alkyne
- Halide to acid
- Halide to cyanide
- Aldehyde to alkyne
- Alkene to aldehyde
- Thioether oxidation



In-situ Heterocycle Formation

- Triazoles
- Imidazolindinone
- Pyridones
- 1,2,4-oxadiazole
- isoindolinone
- Benzimidazole
- Benzotriazole
- Indazolone
- 2-pyridinone

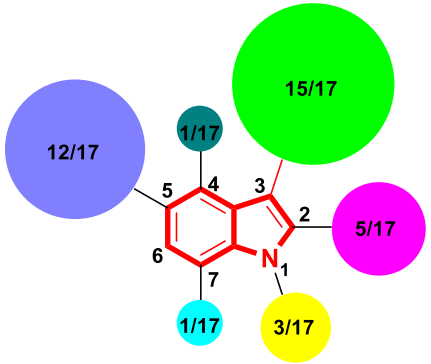
Building Block/Scaffold Diversity

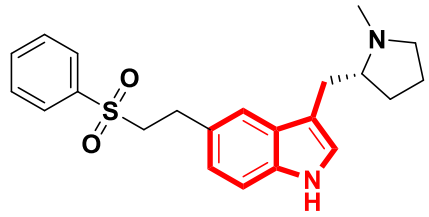
- ☐ Mono-functional group BBs: >20,000
- ☐ Bi-functional group BBs: >3,000
- ☐ Novel scaffolds: >550

BBs: amines, acids, aldehydes, boronates, protected amino acids, free amino acids, amino esters, diamines, acid-aldehydes, acid-aryl-halides, etc.

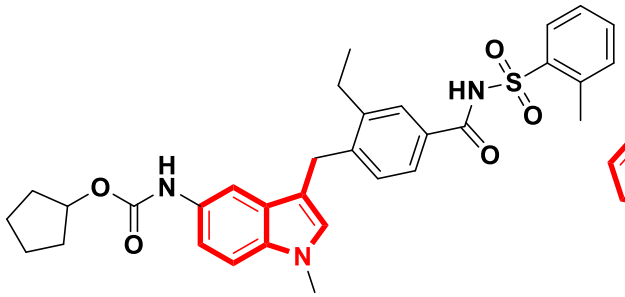
Example of Novel Chemistry: C3-Alkylations of Indoles

Preferred substitution patterns with a vast majority of indole-cored drugs containing a substituent at C3 (88%, green)

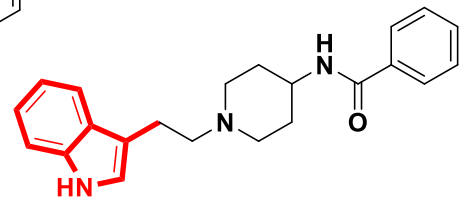




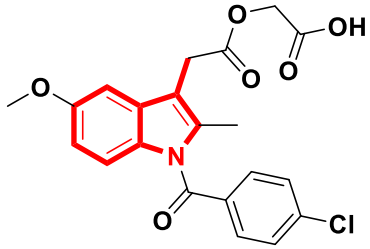
Eletriptan



Zafirlukast



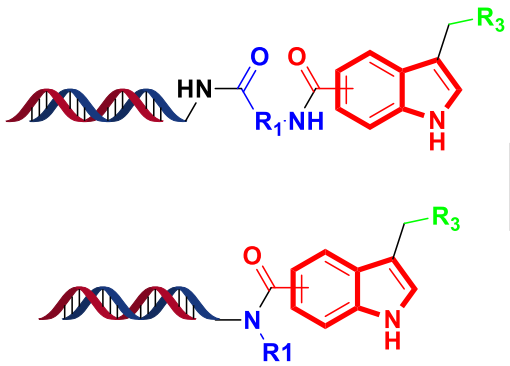
Indoramin



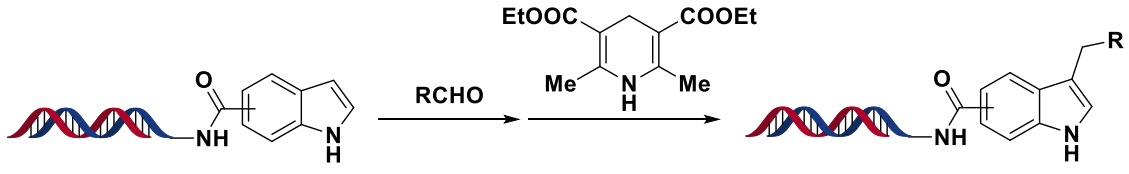
Acemetacin

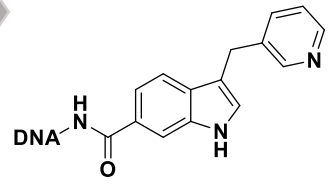
J. Med. Chem. **2014**, *57*, 10257−10274

Indole-based focused libraries

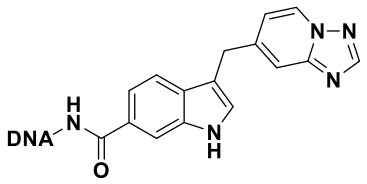


On-DNA C-3-alkylation approach from aldehydes

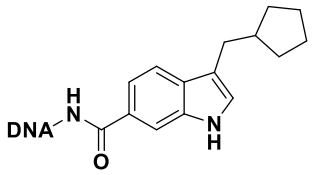




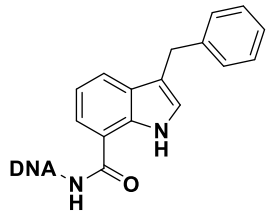
Conversion of step 1: 99%
Conversion of step 2: 90%
2-step conversion: 89%



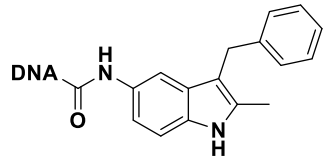
Conversion of step 1: 99%
Conversion of step 2: 95%
2-step conversion: 94%



Conversion of Step 1: 40%
Conversion of Step 2: 90%
2-step conversion: 36%



Conversion of step 1: 99%
Conversion of step 2: 85%
2-step conversion: 84%



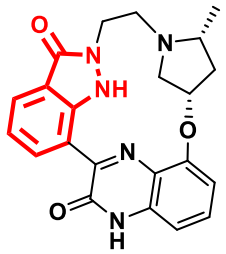
Conversion of step 1: 80%
Conversion of step 2: 95%
2-step conversion: 76%

Org. Lett. **2019**, *21*, 6633–6637

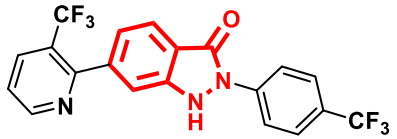
HITGEN

Example of Novel Chemistry: Indazolone Formation

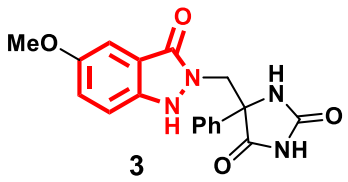
Active compounds containing indazolone cores



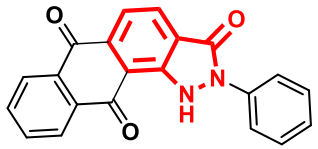
1



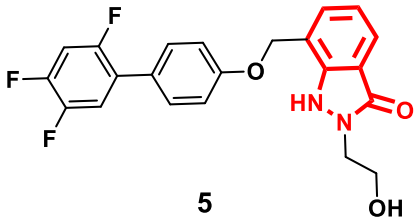
2



3

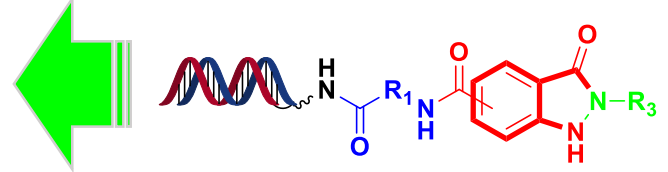


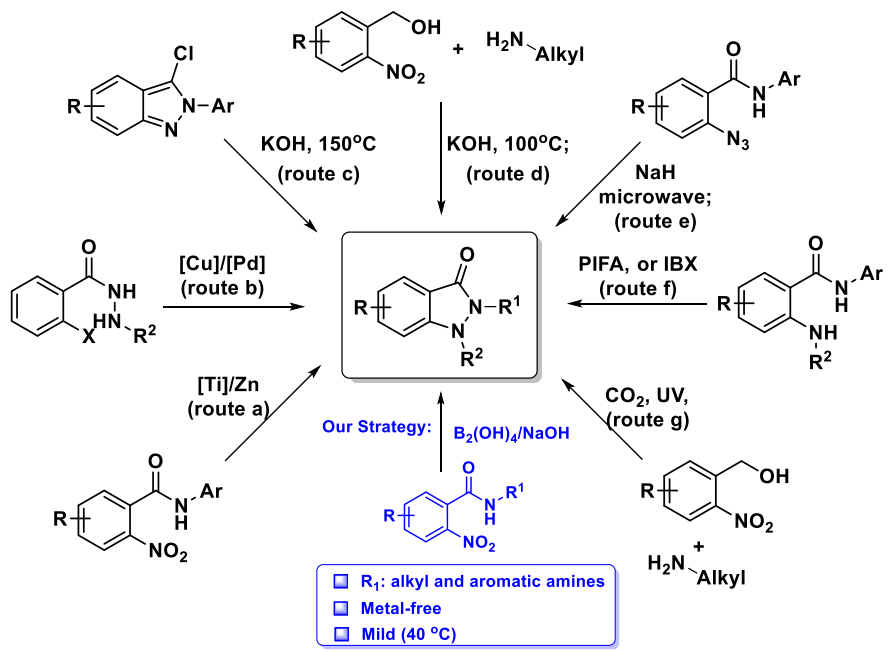
4



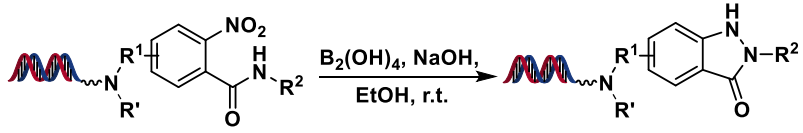
5

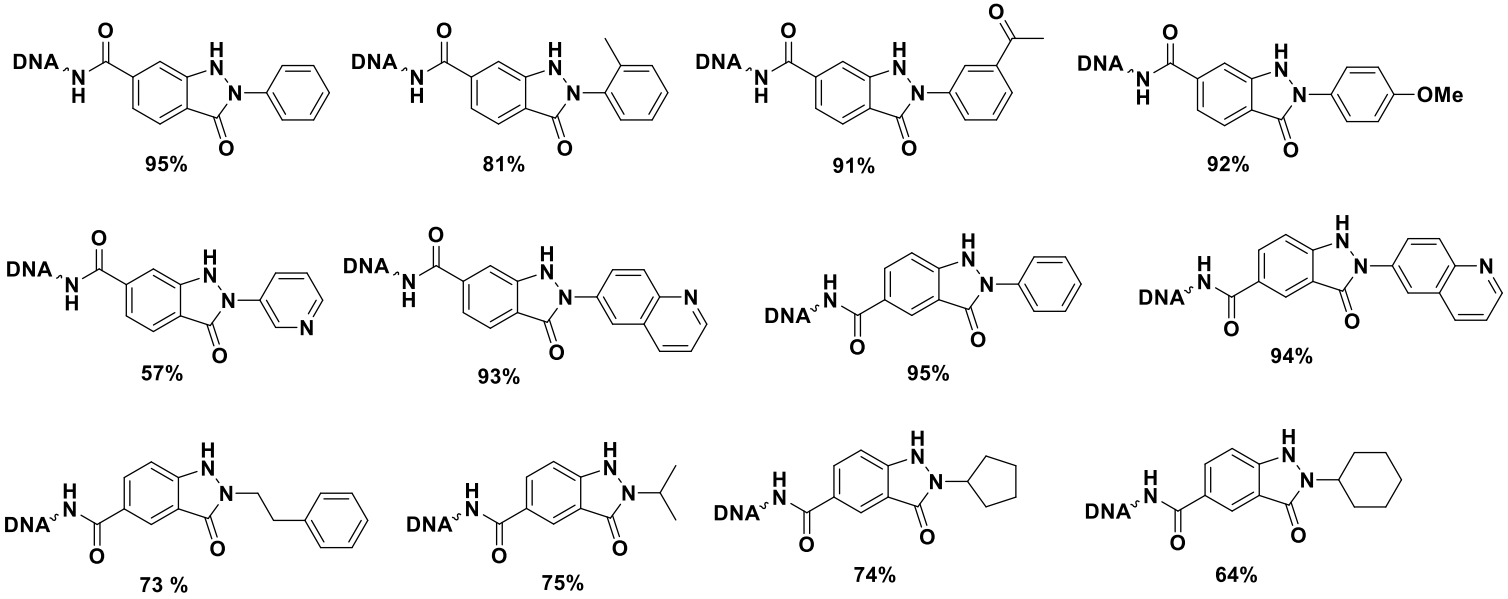
Library design





Most reported conditions are NOT compatible with DNA

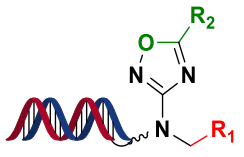
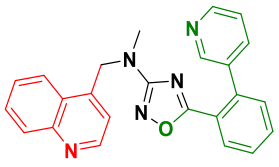
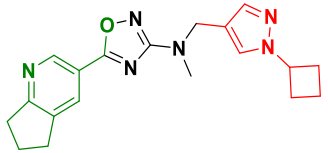
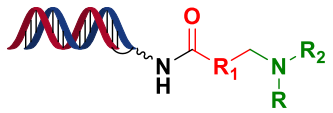
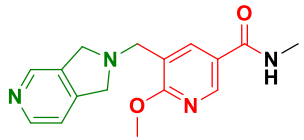
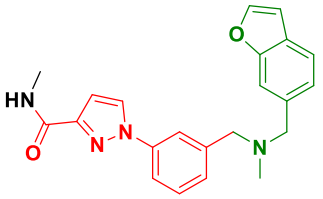
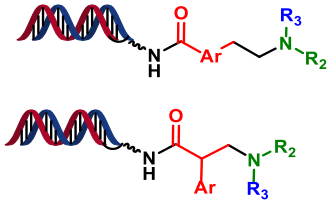
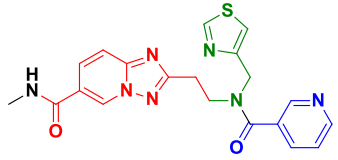
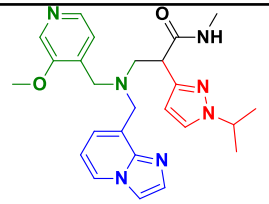
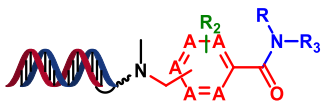
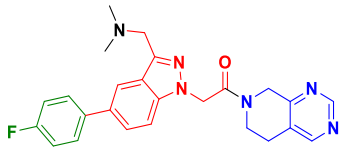
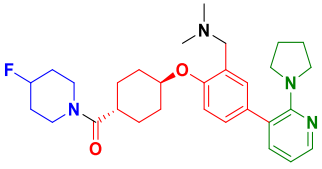
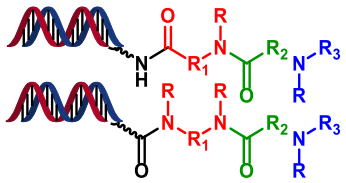
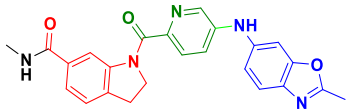
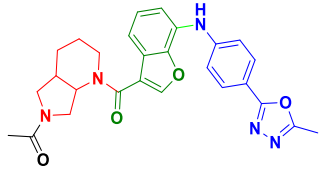
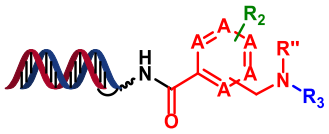
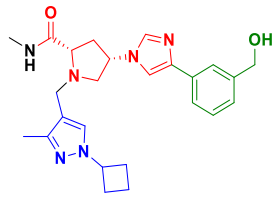
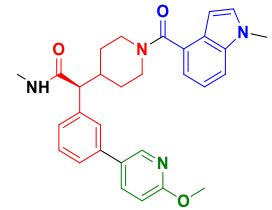
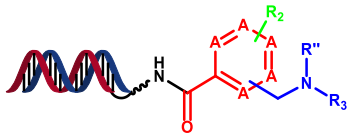
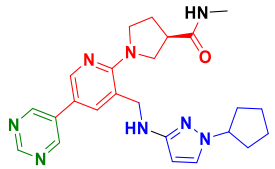
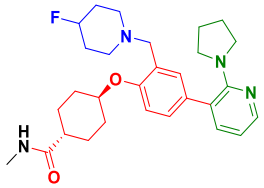
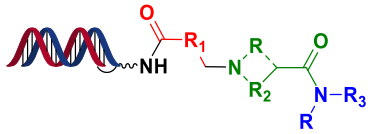
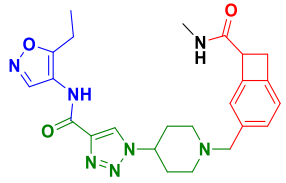
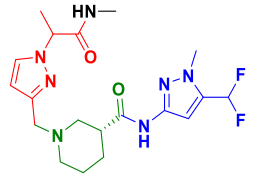




Org. Lett. **2020**, 22, 16, 6277–6282

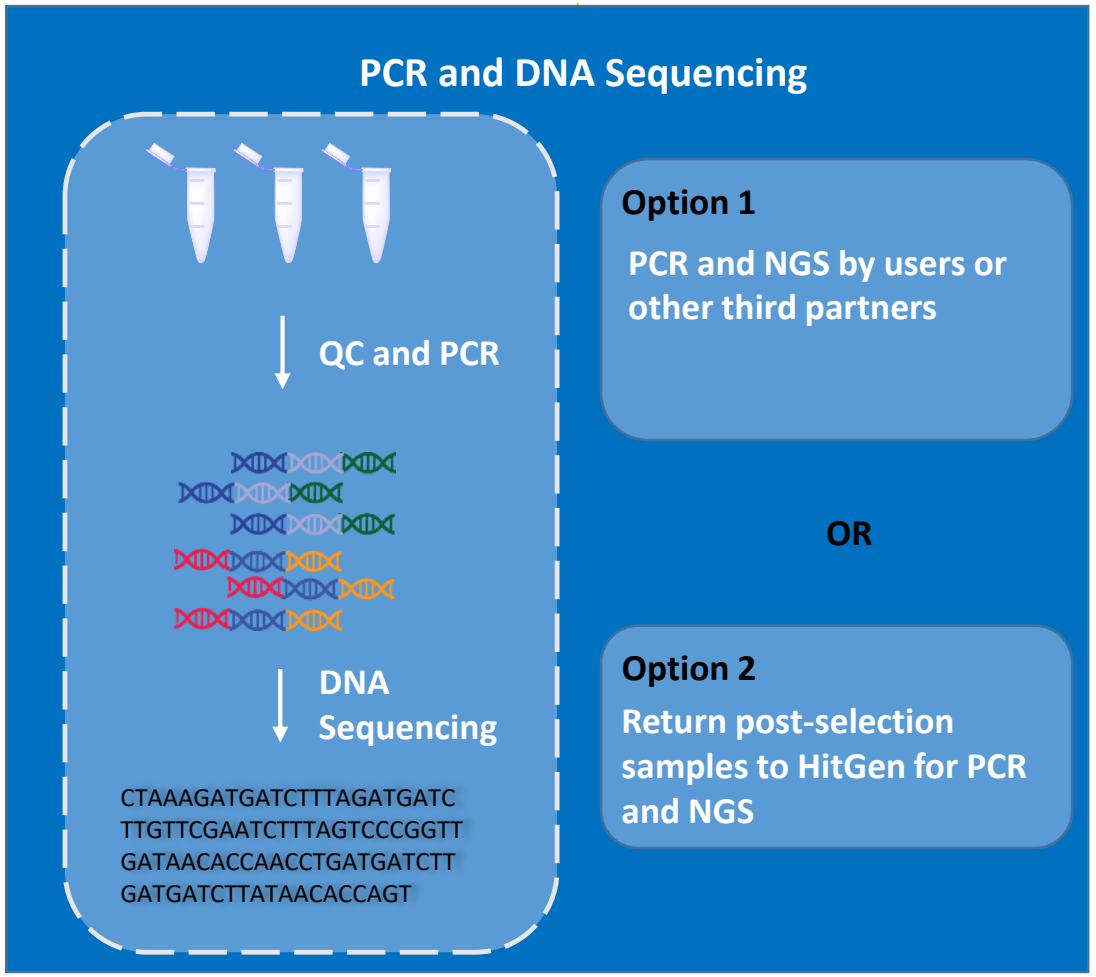
HITGEN

Examples of Libraries and Molecule Structures

Examples of DEL	Examples of Molecule Structures	Examples of DEL	Examples of Molecule Structures
	 Molecular Weight: 393.4406 CLogP: 3.93881  Molecular Weight: 350.43 CLogP: 2.90339		 Molecular Weight: 298.35 CLogP: 0.844995  Molecular Weight: 374.44 CLogP: 3.50217
	 Molecular Weight: 421.4756 CLogP: -0.309298  Molecular Weight: 461.5593 CLogP: 1.77574		 Molecular Weight: 444.5040 CLogP: 2.1628  Molecular Weight: 508.68 CLogP: 3.59632
	 Molecular Weight: 427.4552 CLogP: 2.45652  Molecular Weight: 485.54 CLogP: 2.12716		 Molecular Weight: 448.5606 CLogP: 1.1413  Molecular Weight: 496.61 CLogP: 3.37025
	 Molecular Weight: 446.5480 CLogP: 0.68044  Molecular Weight: 494.66 CLogP: 3.73032		 Molecular Weight: 463.54 CLogP: -0.659856  Molecular Weight: 423.47 CLogP: 0.0573396

Workflow of OpenDEL™

(OpenDEL™ selection manual is provided)



★ Users have options to conduct experiments in-house or to commission HitGen for services

Full Journey with HitGen OpenDEL™ Team

OpenDEL™ Kit

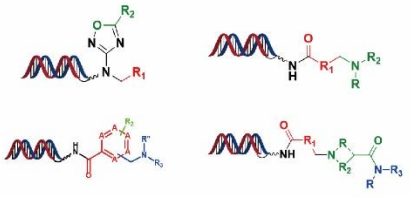
OpenDEL™

- 59 Encoded Libraries
- 4 Bn Compounds

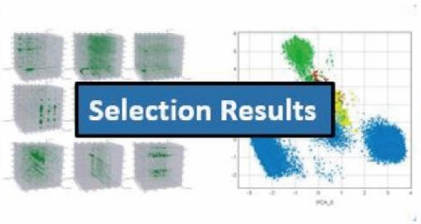
Users only need to prepare biological targets

- Fully open access to building blocks/codons
- No structure disclosure fee
- No compound IP license fee
- DEL selection manual

Examples of Libraries



Screening & NGS



HitGen's GOLD sequencing & data analysis service

- **Global sample shipment** — Average 4-7 days to receive DEL samples
- **Outstanding sequencing quality** — Average Q30 > 92%
- **Lightning-speed result delivery** — Average Sequencing + Data analysis < 7 days
- **Diverse sequencing options** — Choose from a throughput range of 500 million to up to 10,000 million reads

- Be carried out by our team of experienced professionals
- Complete the screening experiments within 1-2 weeks

Off-DNA Synthesis

- Traditional Off-DNA Synthesis
- High Throughput Chemical Synthesis
  - State-of-the-art automation platform
  - High throughput with different scales
  - Applicable to SAR studies

OpenDEL™ Kit → **On-DNA Reference Compound** → **Screening & Sequencing (NGS)** → **Data Analysis & Hit Proposal** → **Off-DNA Synthesis** → **Binding/ Functional Confirmation** → **Hit Expansion** → **ADMET &Other Profiling**

On-DNA Reference Compound

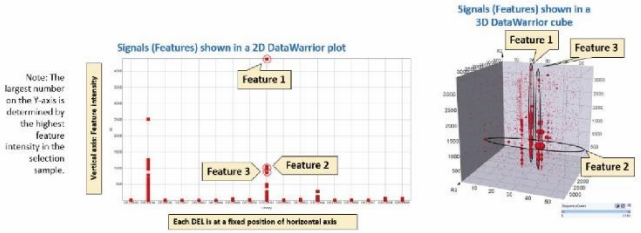
- Tool Compound Synthesis and Modification

For example

- Conjugate Synthesis

For example

Data Analysis & Hit Proposal



- The DELs with Features will be plotted in a set of 3D DW Cubes to allow visual SAR analysis

OpenDEL™ Sales

Academic Groups, AI-Tech, Biotech, Big Pharma

- OpenDEL™ kits are sold to ***15 +*** countries and regions
- ***50+*** OpenDEL™ kits were purchased by a single customer
- ***60+%*** of customers have made multiple orders

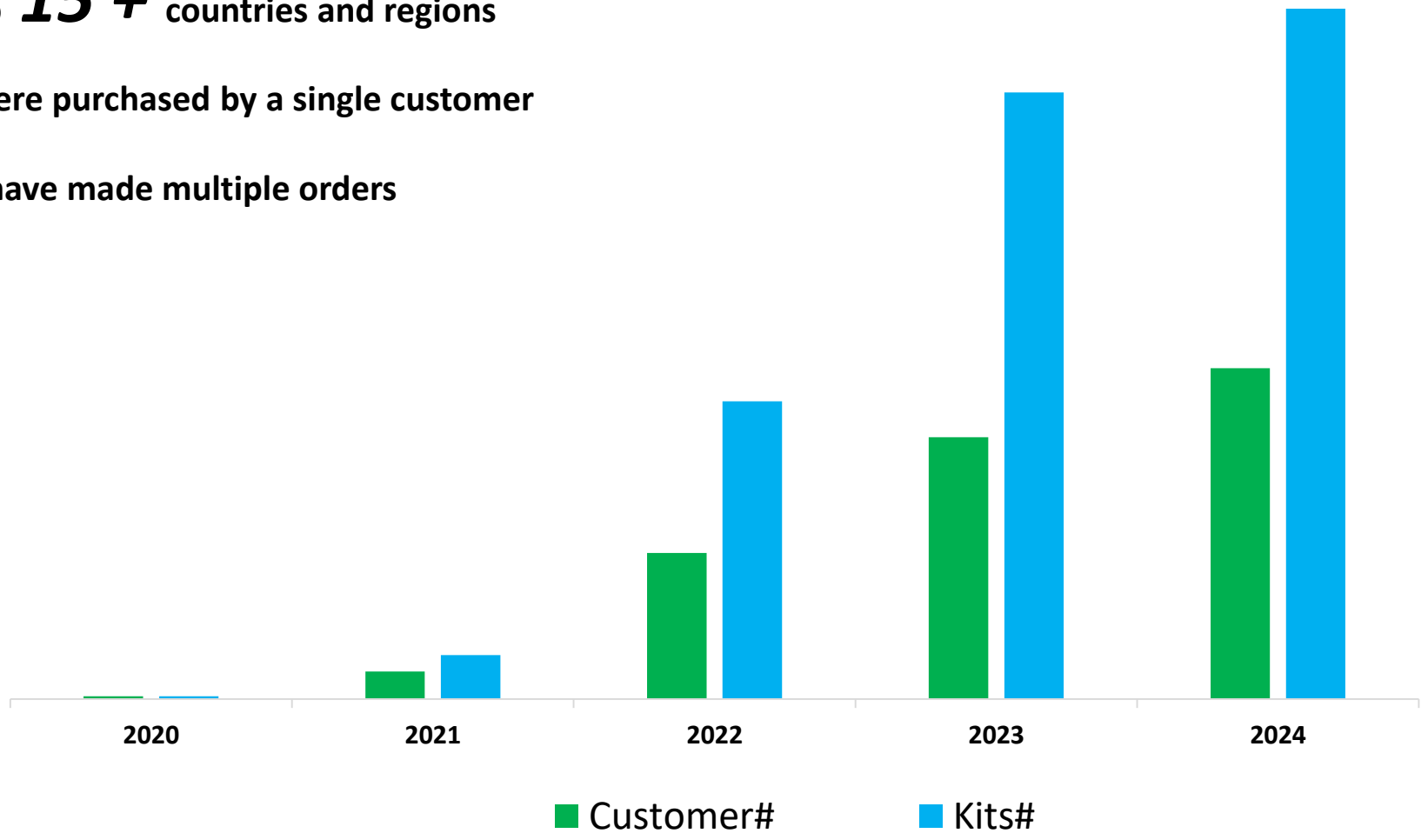
2021-2024

# of Customers:

Achieved 9 fold growth

# of Kit sales:

Achieved 10 fold growth



HITGEN

Third-Party Validation: HitGen OpenDEL™ Shows Superior Performance in Drug Discovery

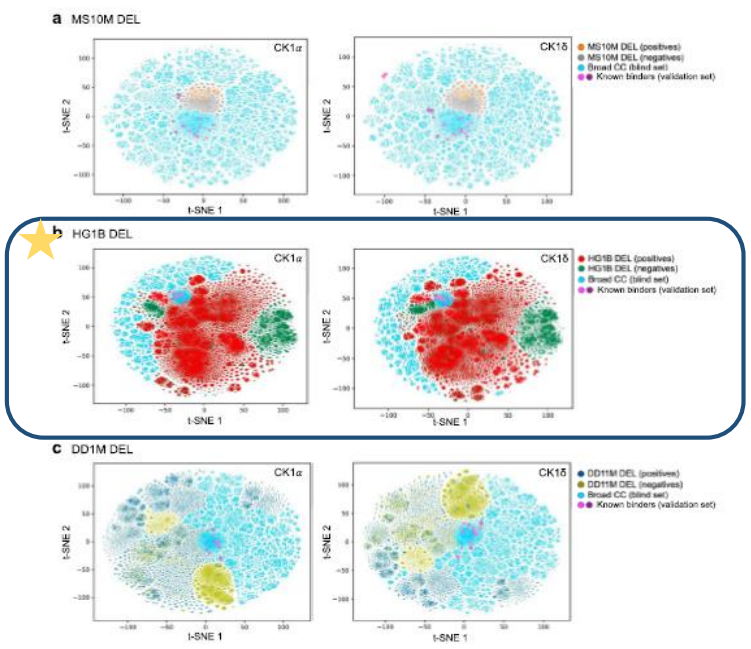


Fig.1 Chemical space comparison

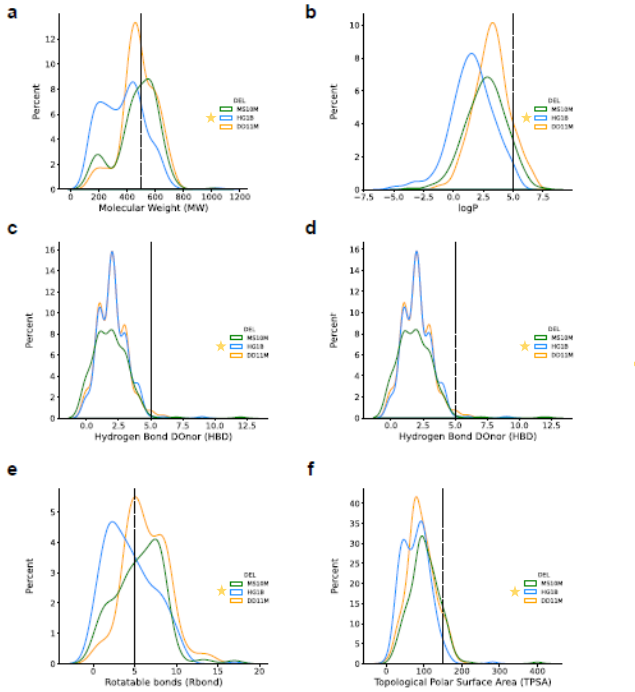


Fig.2 Distribution of physicochemical properties

(a) Confirmed hit count and hit rate per DEL			
DEL	Number of compounds selected for experimental validation	Number of compounds identified as confirmed binders	Hit Rate
MilliporeSigma (MS10M)	237	23	10%
HitGen (HG1B)	283	43	**15%**
DOS-DEL (DD11M)	288	14	5%

Bold indicates the best performance.

Table 1. Confirmed hit (i.e., binder) count from different DEL+ML combinations.

Iqbal, S., Jiang, W., Hansen, E. et al. Evaluation of DNA encoded library and machine learning model combinations for hit discovery. *npj Drug Discov* 2, 5 (2025). https://doi.org/10.1038/s44386-025-00007-4

✓ **Superior Predictive Power**:

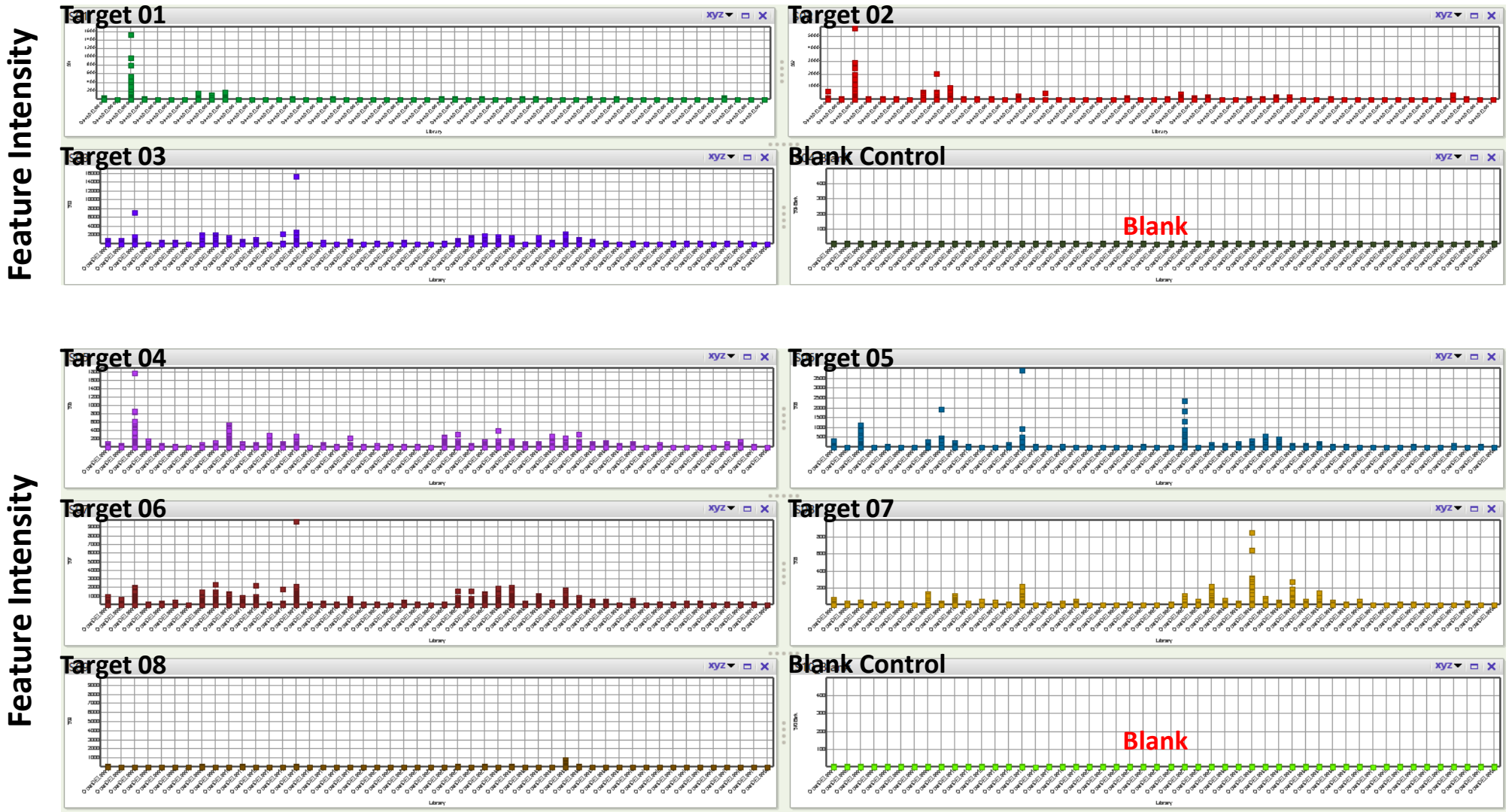
- Models trained with **HitGen OpenDEL™** data outperformed others in identifying binders beyond training chemical spaces.

✓ **Validated Success**: OpenDEL-derived compounds showed:

- Highest confirmation rates (**15% hit rate**)
- Optimal drug-like properties (**48% Lipinski compliance**)
- Nanomolar binders (e.g., **187 nM KD**)

HITGEN

OpenDEL™ Application --- Ligandability of New Target

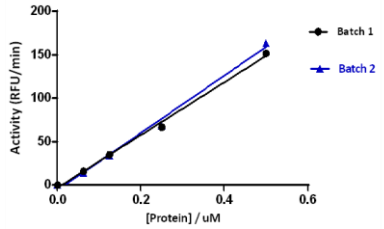


✓ Abundant signals could be identified across multiple types of new targets during screening with the OpenDEL™ kit

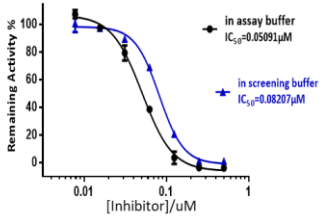
Quick Hit identification process with OpenDEL™

Pre-selection

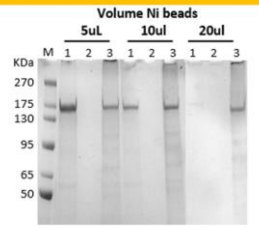
Step 1: Target activity confirmation



Step 2: Selection buffer confirmation



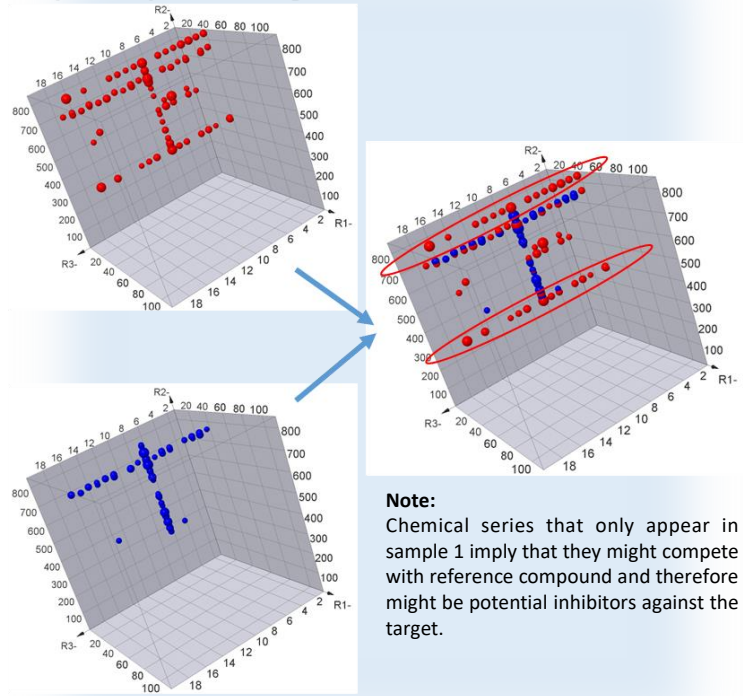
Step 3: Immobilization confirmation



M: Marker, 1: Flow through
2: Wash 3: Retained on bead

OpenDEL® selection

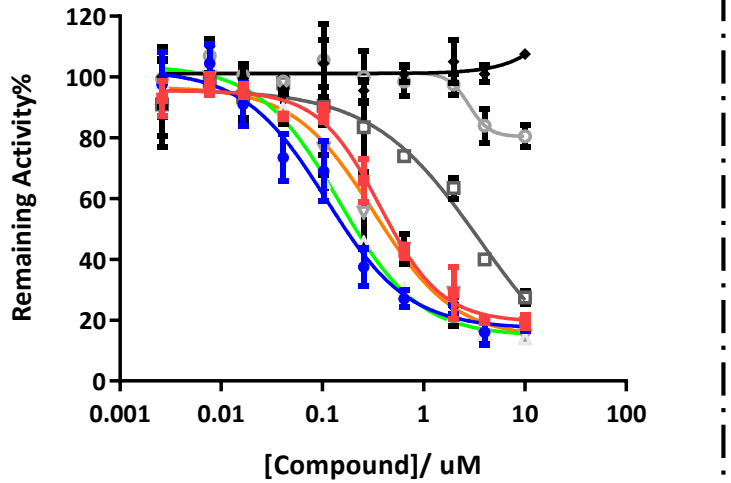
Sample 1: Apo form target



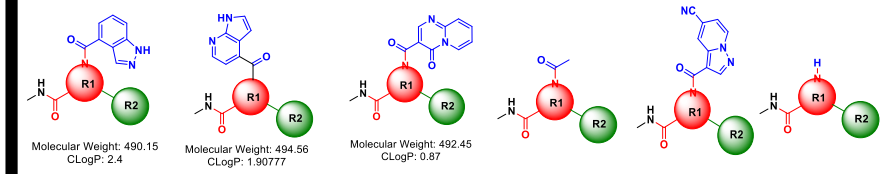
Sample 2: Target + Ref cpd

Note:
Chemical series that only appear in sample 1 imply that they might compete with reference compound and therefore might be potential inhibitors against the target.

Validation of re-synthesized hits



Note: red curve: reference cpd



IC_{50}=0.1067μM | IC_{50}=0.1557μM | IC_{50}=0.3230μM | IC_{50}=3.398μM | IC_{50}>10μM | IC_{50}>10μM

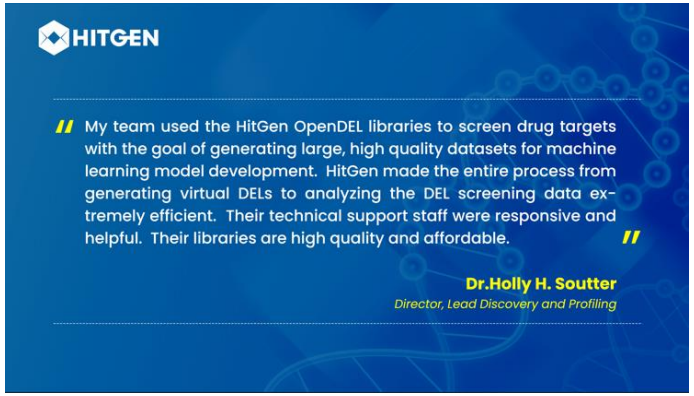
✓ Rapid discovery of nM activity hits with OpenDEL™

Customer Witness & Collaboration

Septerna

"HitGen has created one of the most exciting new DEL products. Septerna was initially attracted to the openness and flexibility with the OpenDEL model. And we have now successfully identified functionally validated hits for multiple ***GPCR targets*** from the OpenDEL libraries. Most importantly, the HitGen team has been a very open and collaborative partner throughout the process."

Testimonial from Dr.Holly H. Soutter



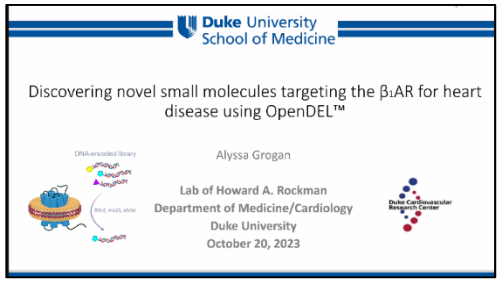
Collaborate with SGC- Providing AI/ML Community with High-quality and Well-curated Data

HitGen will utilize its DNA-encoded library (DEL) technology platform, specifically OpenDEL™, to screen under-represented targets chosen by SGC. ***The screening datasets, curated in a ML-ready format, will be posted to a publicly accessible portal to facilitate drug discovery and ML experts from around the world to model the data and make predictions about new active molecules*** that would be experimentally tested at SGC as part of the Target 2035 initiative.

Structural Genomics Consortium and HitGen Announce Research Collaboration Focused on DNA-Encoded Library Based Drug Discovery | SGC (thesgc.org)

Customer Witness & Collaboration

Duke University



Abstract P2154: A Novel Allosteric Modulator Of The β_1AR Identified By DNA-encoded Small Molecule Library Screening Demonstrates Unique Pharmacology And Function

Alyssa Grogan, Seungkirl Ahn, David Israel, Alex Shaginian, Qiuxia Chen, Jian Liu, Max Wilkey, Jialu Wang, Alem Kahsai, Robert J Lefkowitz and Howard A Rockman

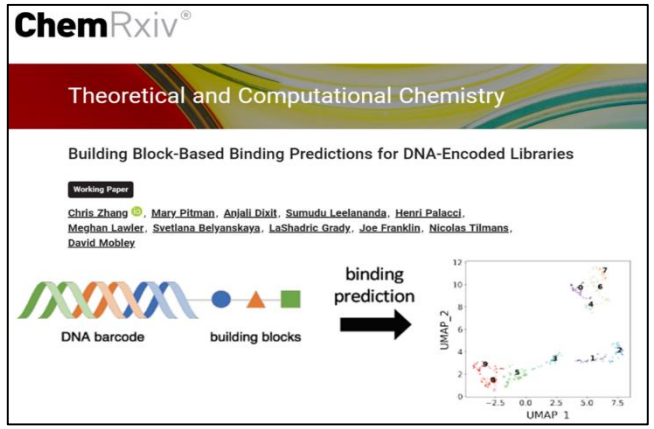
Originally published 13 Oct 2023 | https://doi.org/10.1161/res.133.suppl_1.P2154 | Circulation Research. 2023;133:AP2154

Abstract

While traditional β-blockers (i.e. competitive orthosteric antagonists of β-adrenergic receptors; βARs) are widely used as cardiovascular therapeutics, adverse effects such as fatigue and the nonselective inhibition of multiple receptor subtypes often limit maximal effectiveness. An emerging approach to enhance therapeutic targeting is to identify allosteric modulators that act cooperatively with orthosteric ligands. In contrast to orthosteric ligands which bind the endogenous ligand binding site, allosteric modulators bind to regions that are topographically distinct from the orthosteric pocket and can enhance (positive allosteric modulator; PAM) or reduce (negative allosteric modulator; NAM) the activities of orthosteric agonists/antagonists. Since allosteric regions exhibit greater sequence and structural diversity among receptor subtypes relative to the more highly conserved orthosteric pocket, allosteric modulators are more likely to be subtype specific and/or generate less adverse effects. We therefore embarked on a DNA-encoded small molecule library screen to identify novel allosteric modulators of the β_1AR. Following multiple rounds of affinity selection using purified, functional, β_1ARs reconstituted in lipid nanodiscs and HitGen's OpenDEL® small molecule library containing more than 1 billion unique compounds, we identified Compound 11 (C11) as an allosteric modulator with unique pharmacological properties. Notably, C11 binds to the β_1AR with micromolar affinity and enhances the binding affinity of orthosteric agonists and certain antagonists to the β_1AR. In contrast to its positive cooperative effect on ligand binding, cell signaling assays showed C11 potently inhibits G protein and β-arrestin signaling downstream of the β_1AR. Importantly, C11 showed high β_1AR specificity with no effect on β_2AR or AT_1R signaling. These results suggest that C11 is a β_1AR-specific PAM in terms of ligand binding but a NAM in terms of agonist efficacy, belonging to a largely under-characterized class of allosteric modulators termed PAM-antagonists. With an extremely unique pharmacological profile, C11 is a promising potential therapeutic and experiments evaluating its ability to modulate β_1AR signaling *in vivo* are ongoing.

https://doi.org/10.1161/res.133.suppl_1.P2154

Anagenex & UC Irvine



DOI: 10.1021/acs.jcim.3c00588

SGC

Abstract

Recent advances in DNA-encoded library (DEL) screening have created bioactivity datasets containing billions of molecules, unlocking new opportunities for machine learning (ML) in drug discovery. However, most ultra-large DEL libraries are proprietary, limiting the advancement of ML tools for big chemical data analytics and hindering the democratization of DEL-ML technology. We address this gap by developing an open, end-to-end DEL-ML framework using public datasets, where enriched binders are represented by common chemical fingerprints, ensuring proprietary data protection. We demonstrate that ML models can be built and validated on fingerprinted DEL data and then applied to virtual screening (VS) of billion-sized, publicly accessible chemical libraries. As a proof-of-concept, we screened the human protein WDR91 using the HitGen OpenDEL library (3 billion molecules) and trained ML models, which were used to screen the Enamine REAL Space library (37 billion molecules). Fifty potential binders were identified, 48 of which were tested, and seven were confirmed as novel binders with dissociation constants (KD) from 2.7 to 21 μM that were successfully co-crystalized with WDR91. This fully automated, open-source workflow demonstrates the potential of DEL-ML models in discovering novel binders and promotes the use of open chemical bioactivity datasets and ML to accelerate drug discovery.

DOI: 10.26434/chemrxiv-2024-xd385

Broad Institute of MIT and Harvard

Building on the above-mentioned advances and applications of ML to DELs, we sought to understand better how the composition of different DELs and different ML models trained using these DEL data impact the outcome of DEL + ML paradigm for hit discovery. We chose to screen two well-characterized drug targets[31], *CSNK1A1*(CK1α) and *CSNK1D* (CK1δ), against three DELs of different sizes and chemical compositions: MilliporeSigma DEL, HitGen OpenDEL®, and DOS-DEL[32]. The resulting DEL screening data were then used to train five different ML models that included both traditional models, such as Random Forest[33], and Deep Neural Network models, such as Multi-Layer Perceptron[34] and ChemProp[35]. The developed ML models were applied to a blind (i.e., unseen by the models and with unknown labels) assessment set of 140,000 compounds. Predicted binders from the blind assessment set were tested in a biophysical binding assay to confirm if they were correctly predicted as binders. We further tested molecules that were predicted not to bind to the screened targets, to understand the potential DEL + ML pipeline for filtering out true negatives. As far as the authors are aware, this work is the first such analysis of its kind. In total, 80 (10%, 80 out of 808) and 83 (94%, 83 out of 88) compounds were confirmed as binders and not-binders, respectively, in the biophysical assay. Our cross-DEL and cross-ML results analyses highlight the influence of DEL data quality, chemical space overlap between training and test datasets, ML algorithms on the outcome of a DEL + ML paradigm for hit discovery. Finally, we released the developed DEL + ML pipeline with trained models in an open-source GitHub repositories (https://github.com/broadinstitute/DEL-ML-Refactor), to foster data sharing and community usage and refinement of the developed models for hit identification.

https://doi.org/10.1038/s44386-025-00007-4

THANKS

Welcome to visit us by

www.hitgen.com

or scan 2D barcode.

HITGEN