BioBlocks

YOUR INNOVATIVE CHEMISTRY PARTNER IN DRUG DISCOVERY

Design and Synthesis of the Comprehensive Fragment Library

A 3D Enabled Library for Medicinal Chemistry Discovery

Warren S Wade[1], Kuei-Lin Chang[1], Todd Meyer[1], Peter Pallai[1], Paolo Tosco[2], James Zapf[3], Laura Lingardo[3], Gordon Alton[3]

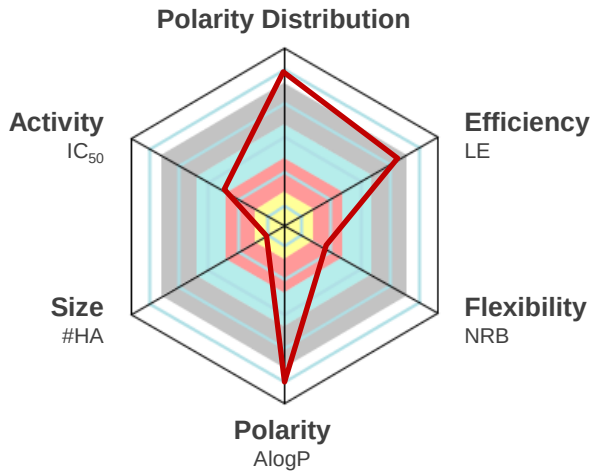
[1]BioBlocks, Inc., [2]Cresset, [3]Visionary Pharmaceuticals

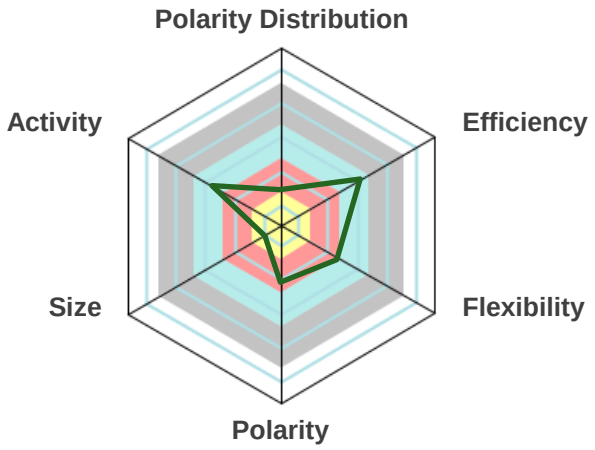
Our Leap-to-Lead™ Platform

BioBlocks

Developed to Improve Lead Generation

Leads for a current internal project





Literature Lead	
• Potent	• Not Permeable
• Not efficient	• Failed in vivo
• Hydrophobic	

BioBlocks Early Lead	
• Less Potent	• Balanced polarity
• More efficient	• Permeable

ProperType plots illustrate Lead quality

A good drug candidate hits the bullseye

Enhancing Drug Discovery

BioBlocks

Leap-to-Lead™

Tools for property based optimization

Comprehensive Fragment Library

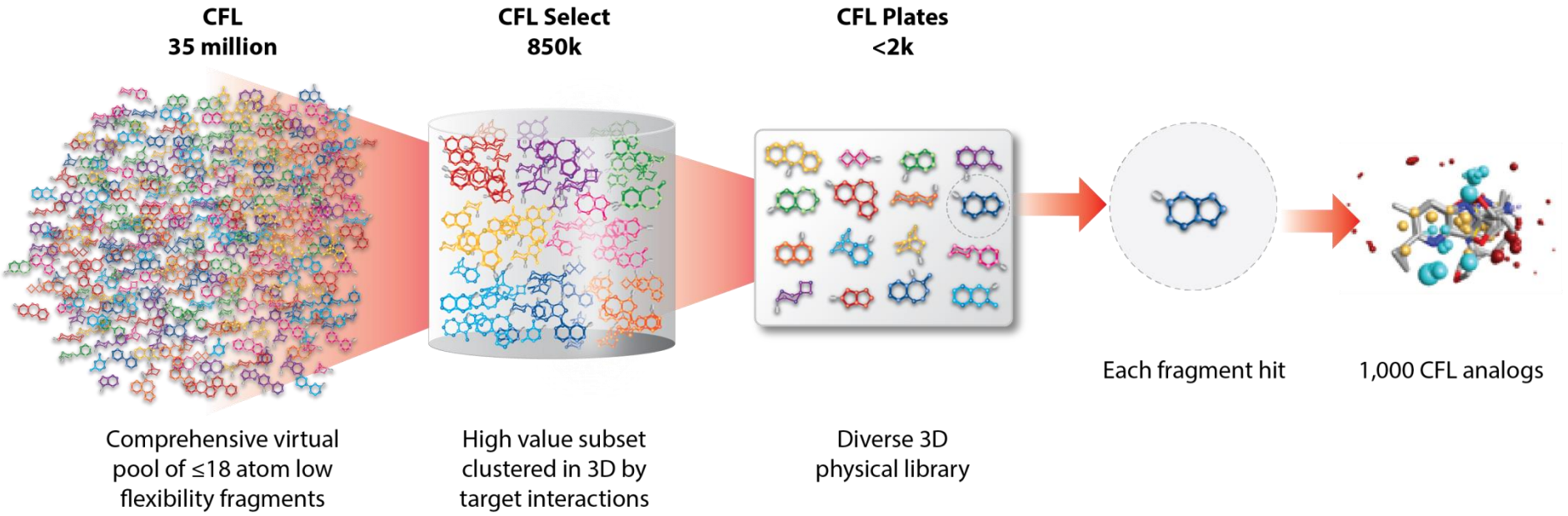
BioBlocks' Proprietary Next Generation 3D Fragment Library

Syntheverse™

Reaction-based access to currently synthesizable compounds

Improved Fragment Screening

Leap-to-Lead™ provides access to new chemical matter through its 3D enabled fragment screening set, the Comprehensive Fragment Library



CFL designed for maximum diversity with medchem friendly properties

Even commercial hits give access to novel, 3D structures

Enables a head start on Hit to Lead and IP

Building The Virtual Fragment Set

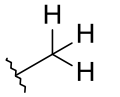
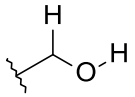
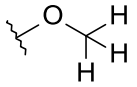
An exhaustive set of ring structures containing ≤18 heavy atoms with 1 handle atom was built in BIOVIA Pipeline Pilot with several inclusion criteria:

- ≥1 ring, including bridged, spiro and fused ring connections
- ≤ 3 unique rings
- ≤ 2 ring assemblies
- ≤ 2 rotatable bonds
- 1 atom has all rotatable bonds
- C,N,S,O atoms only
- ≤ 1 S

Rigid, partially aromatic structures are the highest value 3D enabling subset

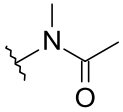
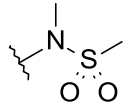
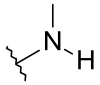
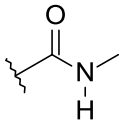
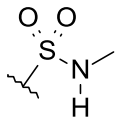
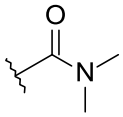
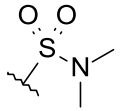
Core structures decorated with handles chosen for synthesis potential

High Value

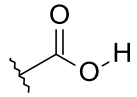
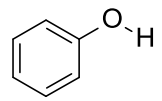




Moderate Value


Follow up only

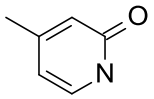
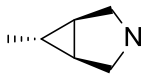
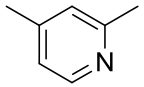
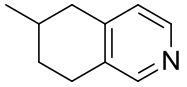
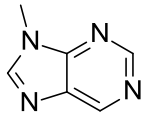
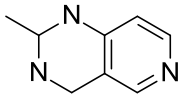
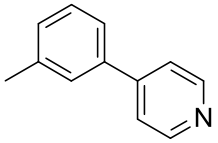
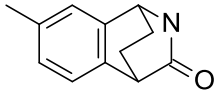
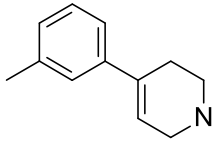
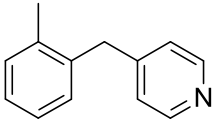
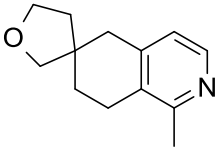
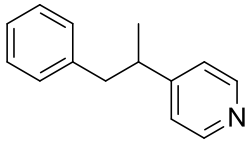









Not chosen

Fragment Handle Criteria

The raw set of >35 million structures with Me handles was filtered by medicinal chemistry criteria to produce a 7 million candidate fragment set

Example CFL Fragments	Excluded	Reason
 		Two handles*
 		Unstable
 		Isolated double bond*
 		Too many rotatable bonds*

*Potential follow up for a CFL fragment hit

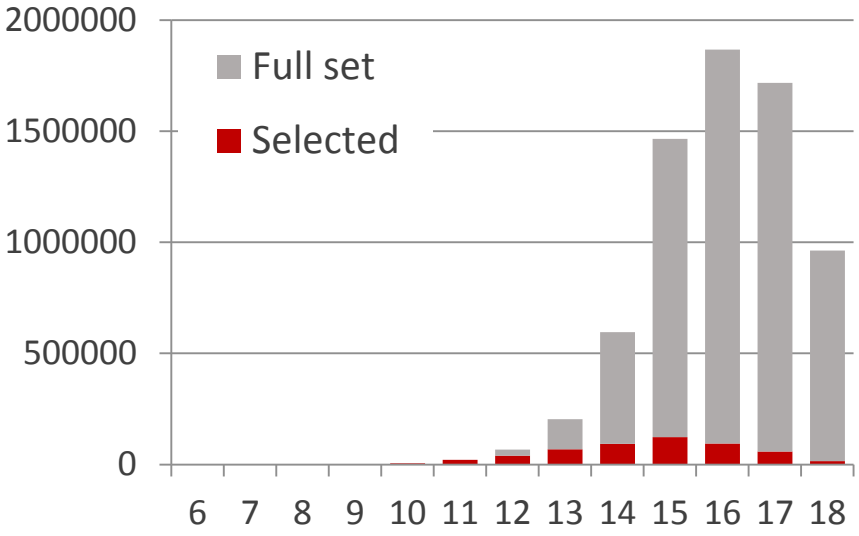
CFL Select set

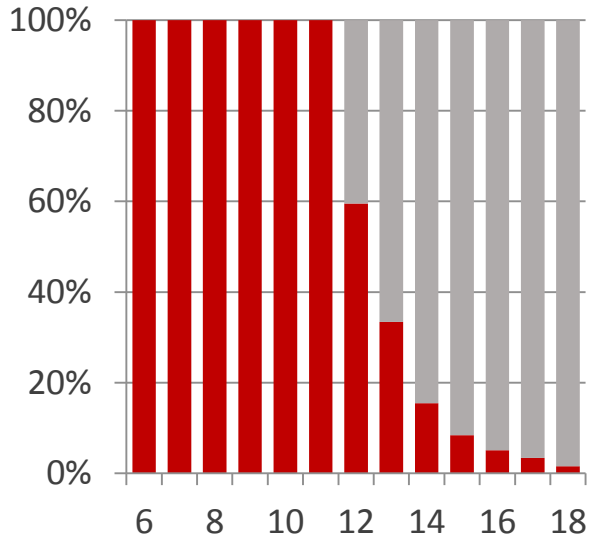
Practical computational limit is <1000k diastereomers

Requires >600 CPU-years for full 3D similarity matrix

A 580k subset was selected for clustering:

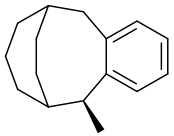
- Reduced mean of 16 heavy atoms to 14.6
- Includes representatives of all 4500 ring types
- Includes 100% of <12 atom fragments to 1.5% of 18 atom fragments



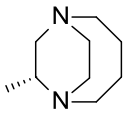


Clustering Workflow

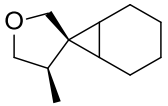
Difficult examples



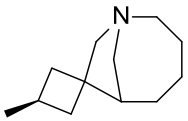
Bridgehead Diastereomers



Nonchiral atom Diastereomers

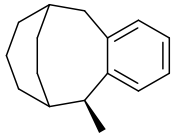


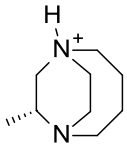
Ring size limited Diastereomers

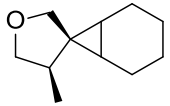


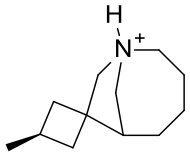
Cis/trans ring Diastereomers

Rule-based protonation

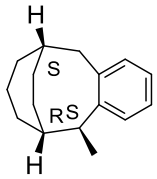




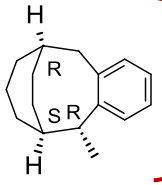


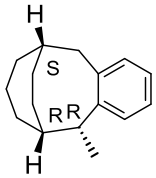


Racemic diastereomer generation

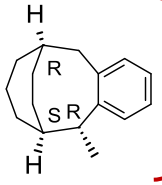


and



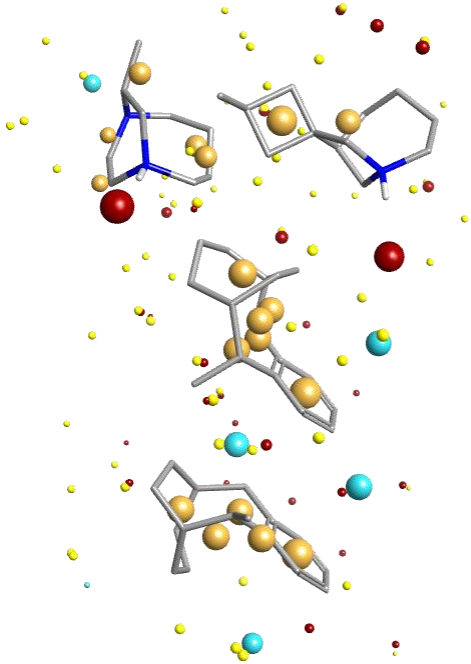


and

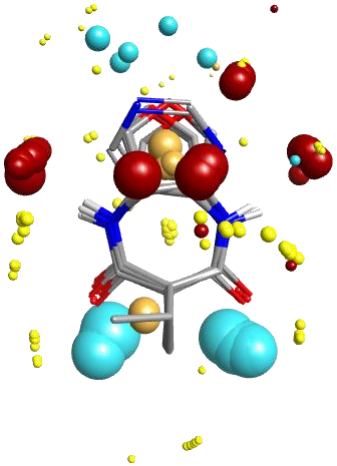


Generation from 2D failed in difficult cases

Conformer and field point generation



3D similarity based clustering



3D Clustering

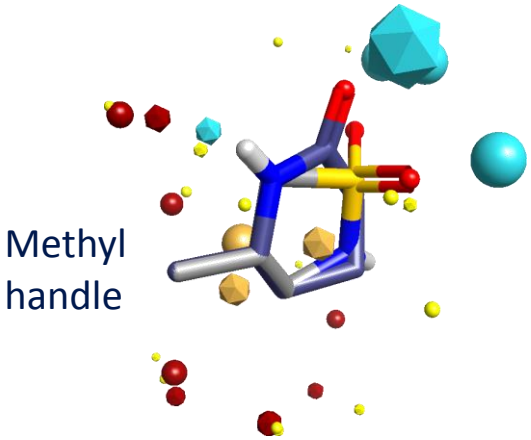
Alignment to methyl handle:

- Rotate structures around the handle
- Determine maximum similarity

Similar compounds:

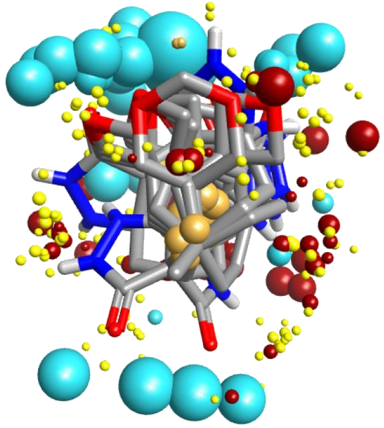
- Have common vectors
- Represent alternative sprouting choices

Comparison of two 6-member clusters

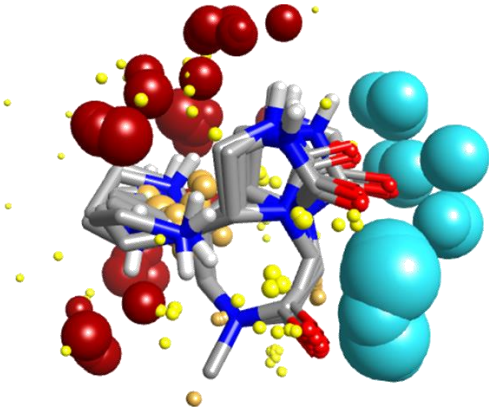


Maximum similarity: 0.78

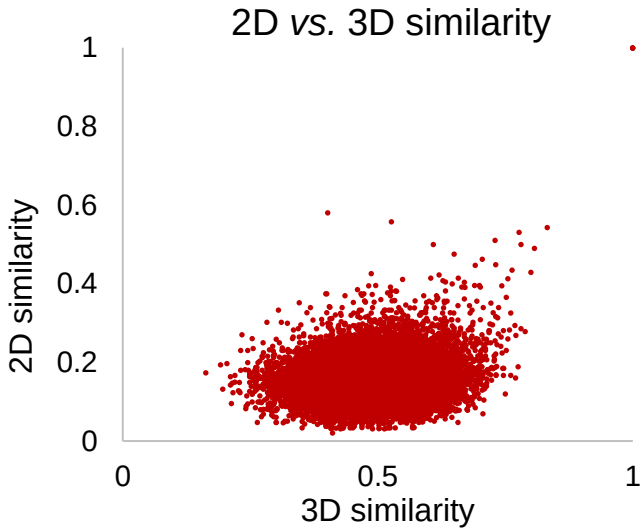
Silhouette: 0.509



2D Cluster



3D Cluster



3D Clustering

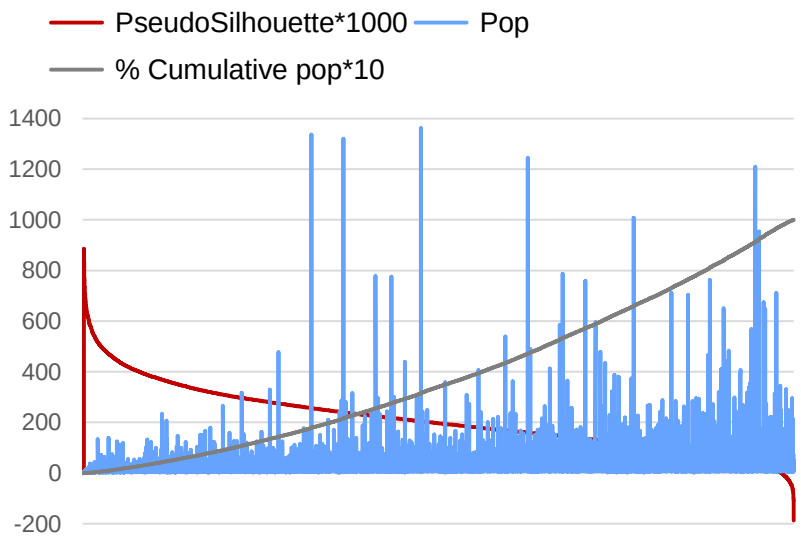
Initial clusters from a full similarity matrix of 150k compounds — 14 CPU-years

Examined 10k, 25k and 50k clusters

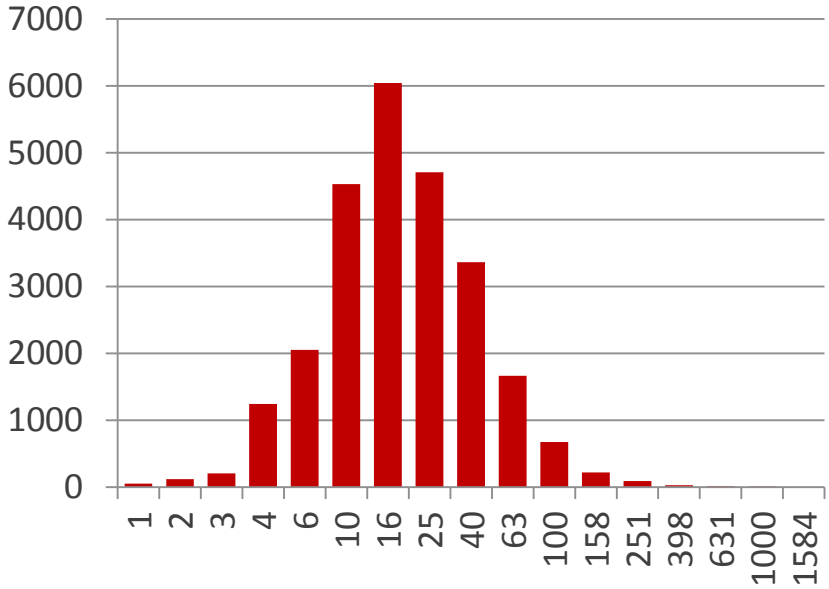
- 10k gave a well-distributed 150k compound set; not scalable to the entire library
- 50k gave cluster sizes too small to be stable

Remaining compounds placed in 25k clusters of most similar medoid — 20 CPU-years

- PseudoSilhouette calculated because full matrix unavailable



Mean intra-cluster similarity: 0.77



Binned Cluster Size, Mean 33

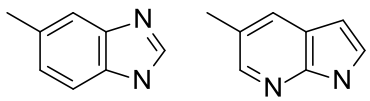
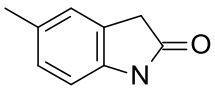
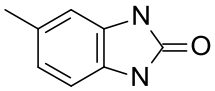
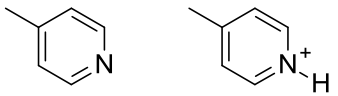
HBond Patterns

HBond patterns were generated for all compounds

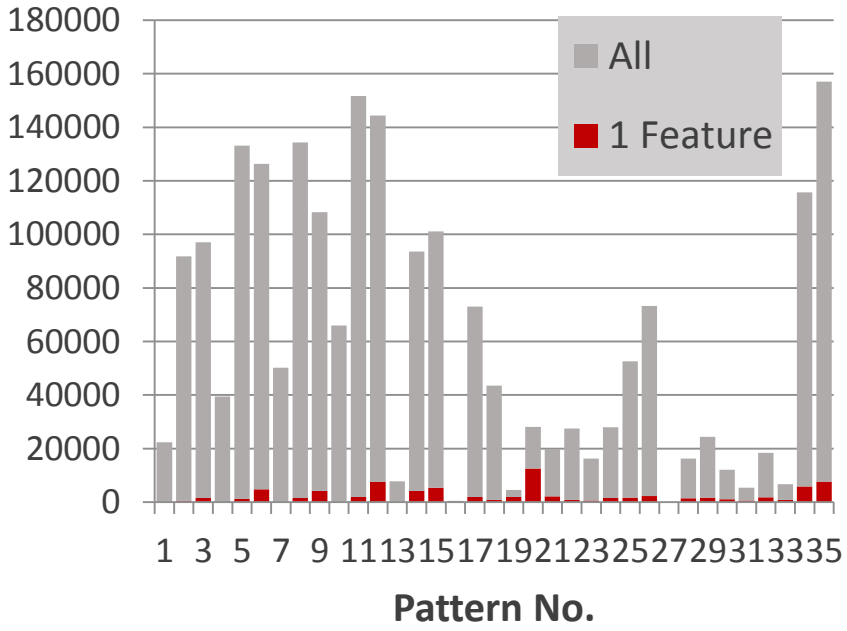
Pipeline Pilot Molecular Pharmacophore Fingerprints Component

HBA, HBD, pairs when separated by 1-6 bonds

35 core patterns including Named Functional Groups

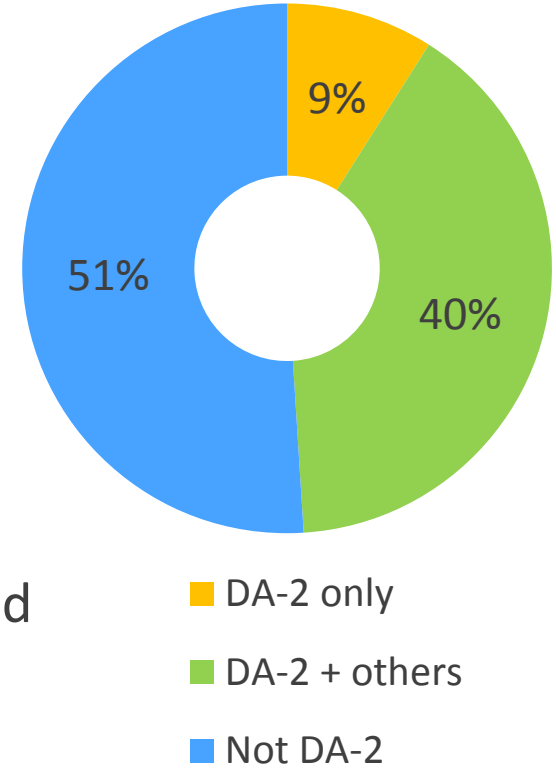
Structure	Pattern Name
	DA-2
	Amide Special case of DA-2
	Urea DA-2,DA-2,DD-2
	Pyridine Protonation state is handle dependent

Pattern Frequencies



H Bonds are independent 2D clusters

- 2D similarity measure
  - Same patterns are distributed over clusters
  - ~50% of clusters have simple, valuable patterns
- Compounds prioritized by target interactions
  - H Bonds important for:
    - Fragment orientation
    - Selectivity
  - Fragments with 1 likely interaction motif preferred



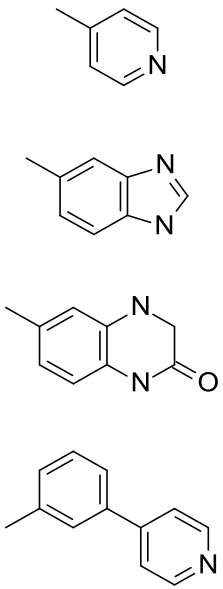
Picking Commercial Compounds

Commercial Compounds found in 1395 clusters (5.6% of total)

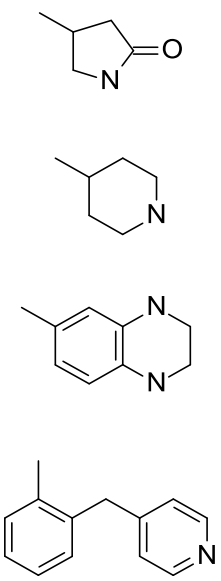
- 905 clusters have only 1 commercial member, 81% are 3D enabled
- Others evenly distributed 2D and 3D by PMI*
- On rod/disk axis, ≤ 2nd percentile of PMI area

Examples

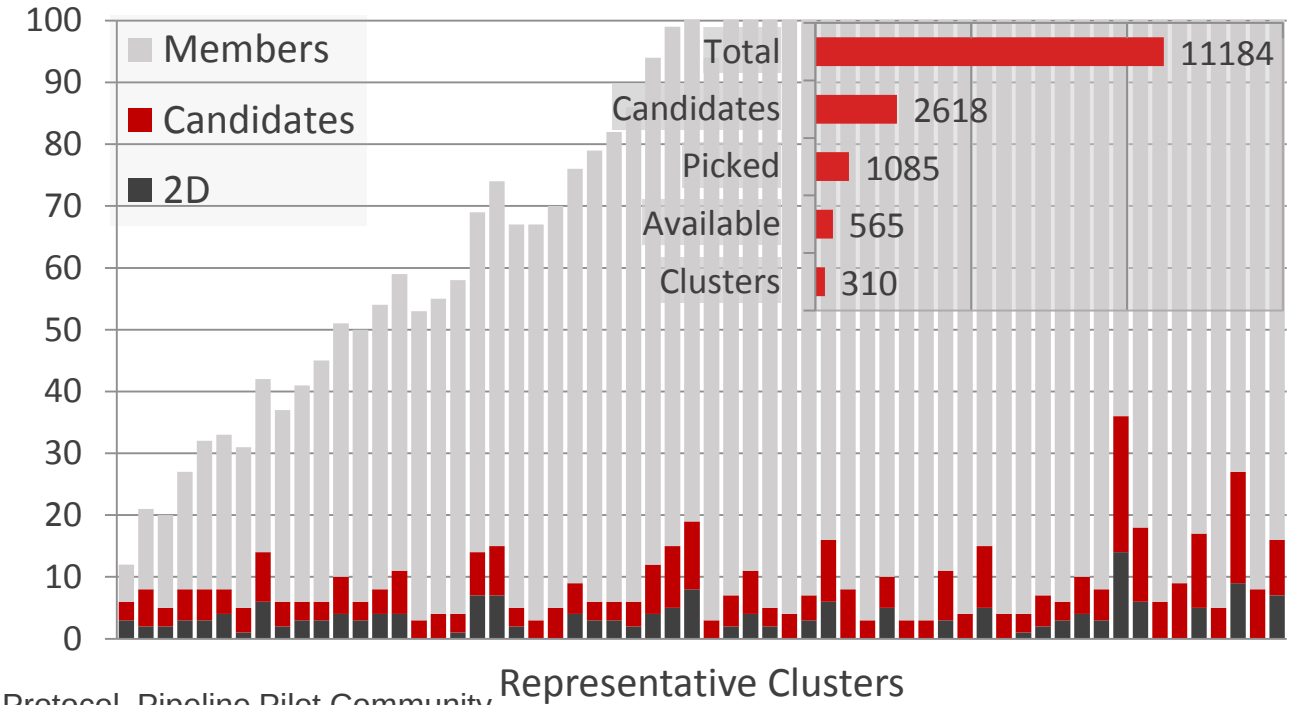
2D



3D

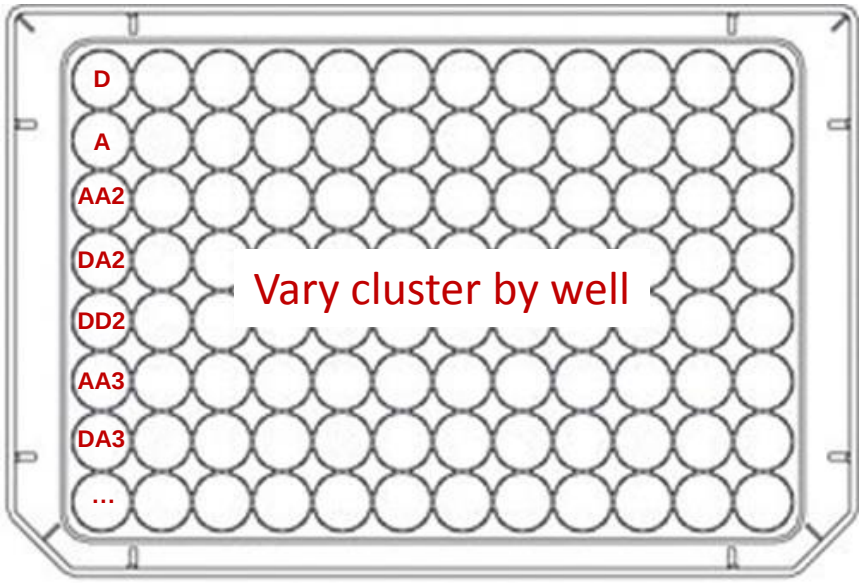


Available Cluster Distribution



*Metz, J.T. Principal Moments of Inertia Protocol, Pipeline Pilot Community.

Building Plate 1

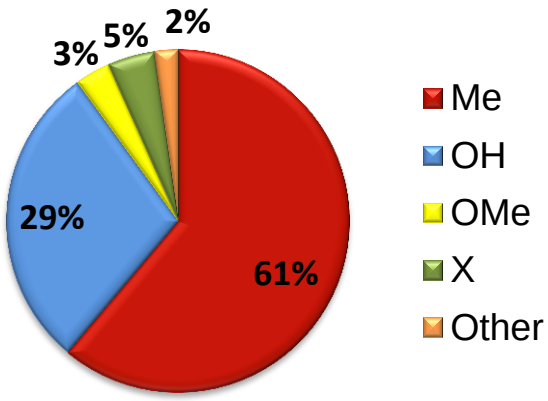


Vary Hydrogen bond patterns by row

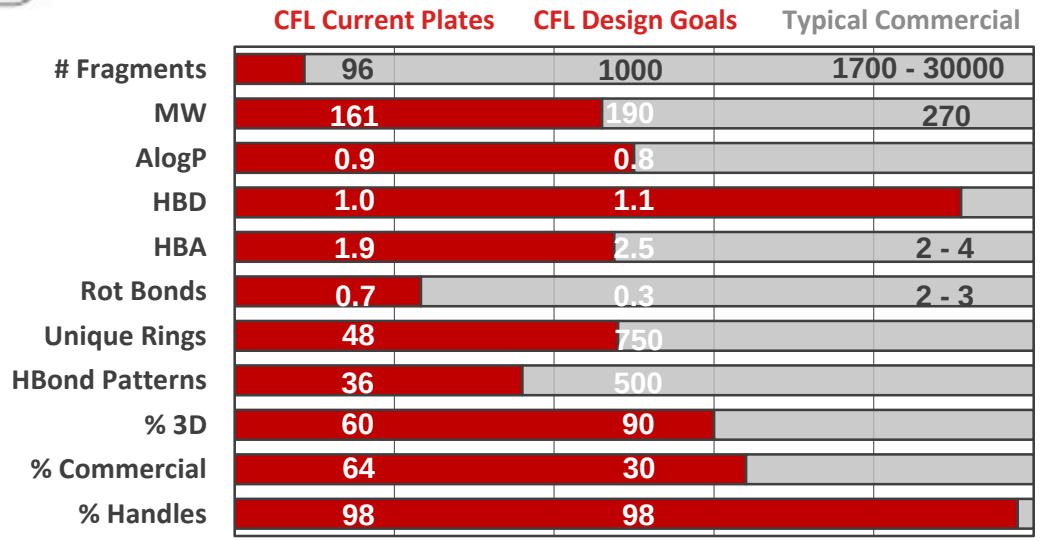
Vary core shape over column

Vary Synthetic Accessibility by plate

Handle Type

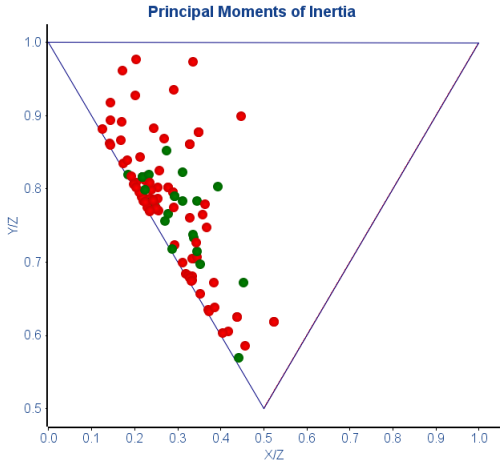


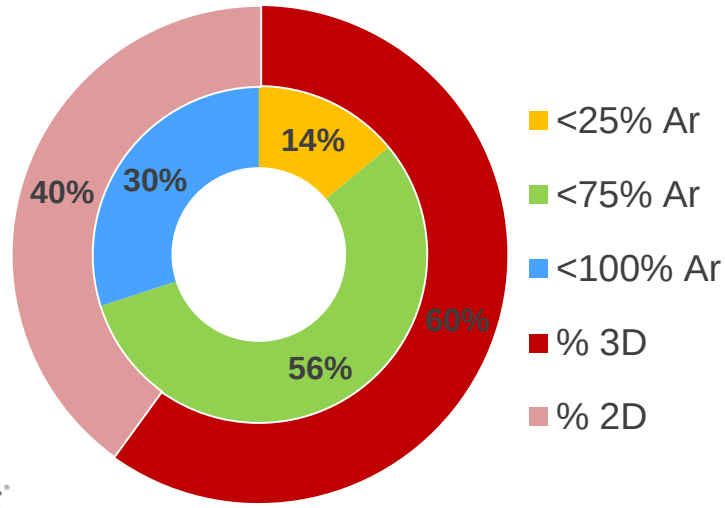
Average CFL Properties



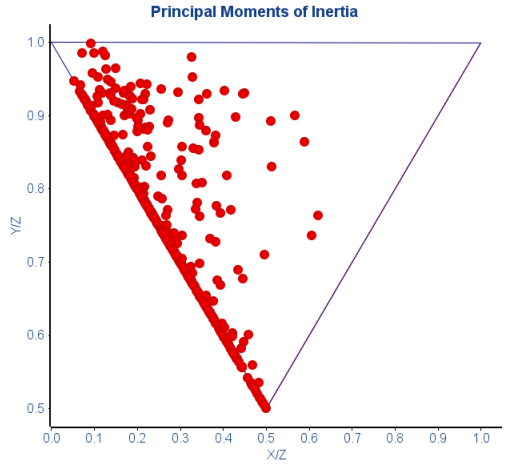
Library comparison

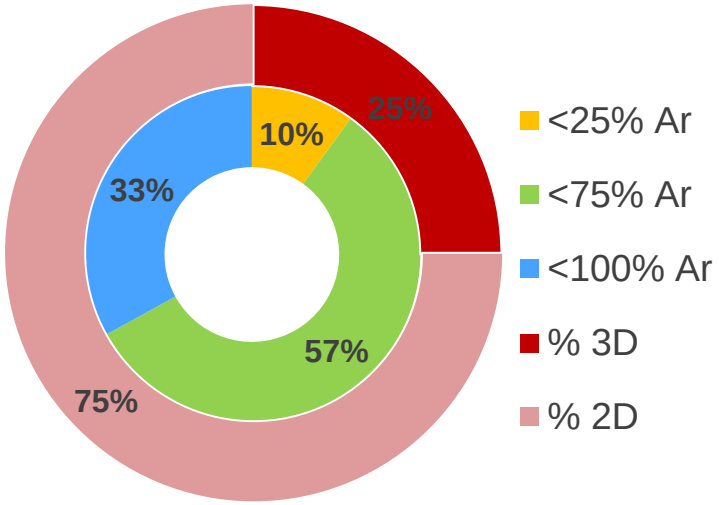
CFL Current Plates





Commercial Library





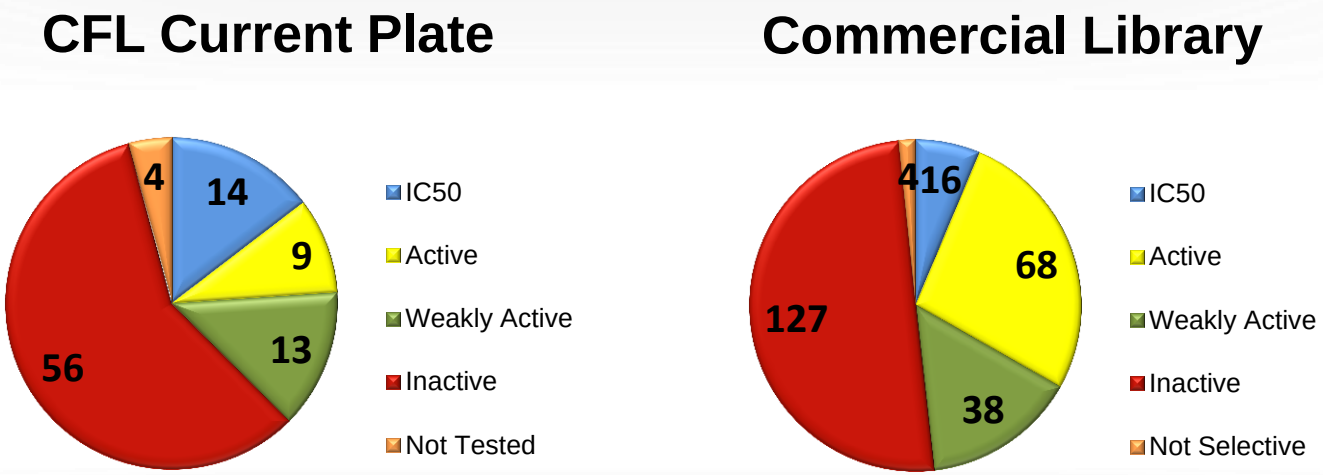
cresset discovery services

Screening comparison

Visionary Pharmaceuticals

BioBlocks

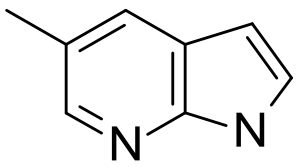
Same kinase target



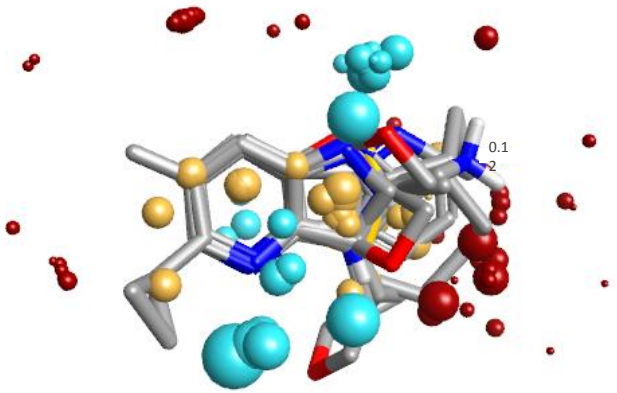
Compared with the previous library:

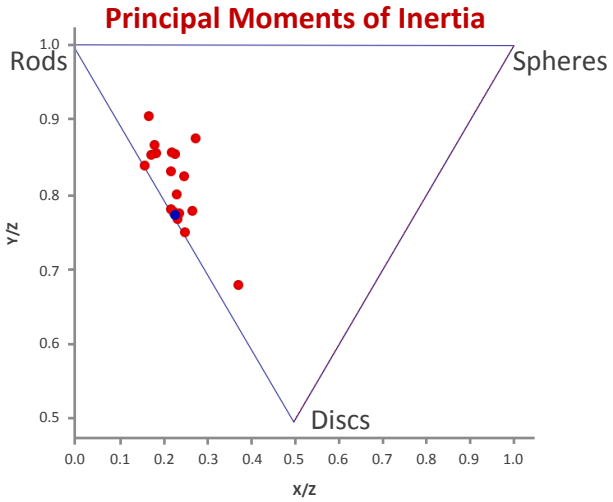
- Plate 1 has a lower hit rate, but more hits generate dose response curves
  - More scaffold variation
  - Allows follow-up on weak but interesting compounds
- Original Chemotypes still recognized

Connecting Fragment Hits to New IP



CFL000025





CFL000025 is a positive control for kinases

- ≥300 µM IC_{50}
- Allows comparison of other hits to known hinge binding motif

CFL provides 3D cluster of related compounds

- 47 member cluster
- Overlay of closest 20 shown

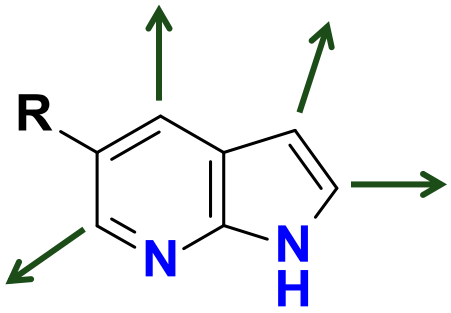
Contains both known and novel alternatives

- Multiple 3D options

Representative Hit

Visionary Pharmaceuticals BioBlocks

CFL clustering reveals many related compounds for rapid fragment SAR

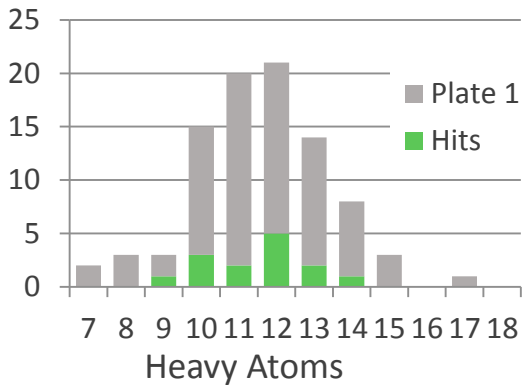


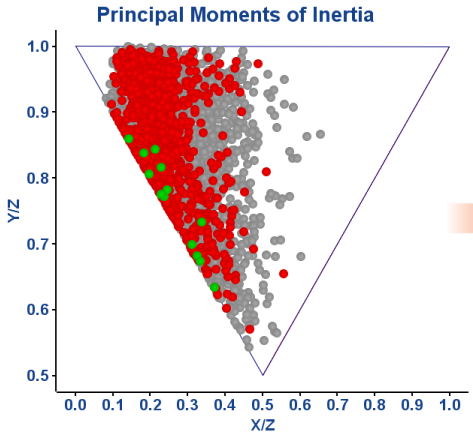
Independent Analogs	CFL000025	14 Fragment Hits
Handle Analogs	**14**	196
Handle Vectors	4	67
Same 2D Scaffold	37	2181
Same HBond Pattern	274	27370
Same 3D Cluster	47	1374
Total CFL Compounds	873	37984
% Unique to Compound	100%	80%
CFL Analogs*	>25,000	>450,000

*Estimated count of 3D clusters, all vectors

Fragment Screening in Action

CFL provides 3D hits for a kinase target





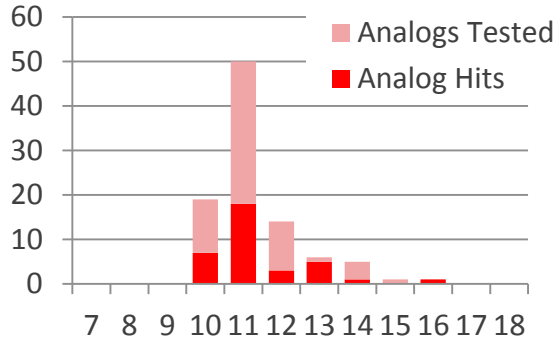


Plate 1 Screened

- 14 hits, 36% 3D
- Mean IC_{50} 400 µM (50-2000)

3D Cluster Coverage

- 14 hit clusters, 83% 3D
- 2500 CFL structures
- Covers valuable 3D space

CFL library follow up

- 35 hits
- Mean IC_{50} 220 µM (20 – 950)

By minimizing overlap, a small library gave wide coverage of chemical space

Conclusion

Available commercial compounds are disappointingly similar to each other.

The CFL defines a valuable area of new chemistry space adjacent to known biologically active fragments.

Our initial CFL screen generated high quality hits through its coverage of chemistry space.

Enhanced 3D coverage revealed new, underexplored chemotypes useful for important drug targets.

The CFL gives high quality entry points to FBLD such as our proprietary Leap-to-Lead™ discovery platform.

Contact

BioBlocks is developing Leap-to-Lead™ for use in Lead Discovery Collaborations. While the CFL is not available for independent purchase, we welcome new collaborators.

For additional information or to discuss using the CFL and the Leap-to-Lead™ platform in a drug discovery effort, please contact wwade@bioblocks.com or visit our website: www.bioblocks.com