BIONET 2nd Generation Premium Fragment Library

Building a Diverse and Experimentally-Curated Fragment Library

Andrew Lowerson[1], Patrick McCarren[2], Steven Laplante[3], and Michael Serrano-Wu[2]

1. Key Organics Ltd, Highfield Road Industrial Estate, Camelford, Cornwall, PL32 9RA, UK
2. Center for the Development of Therapeutics, Broad Institute, 415 Main Street, Cambridge, MA 02142,
3. NMX Research and Solutions Inc., 500 Cartier Blvd., Suite 6000, Laval, Québec, Canada H7V 5B7

Key Organics
Chemistry | Innovation | Quality

BROAD INSTITUTE

NMX Research & Solutions

Introduction

Fragment libraries are commonly assembled by Rule of 3 filtering followed by manual curation, however robust experimental data that ensures the proper physicochemical attributes needed for high-concentration screening is often lacking and replaced instead by *in silico* calculations of uncertain predictive value. A fragment collection with experimentally-determined aqueous solubility will address a major source of false positives and attrition in fragment screening libraries: aggregation, stability, and solubility. ^{1}H NMR spectral data in aqueous buffer will further enable practitioners to rapidly build fragment pools and initiate screening.

Diversity selection methods in shape, scaffold, fingerprint, and predicted property space combined with industry-standard substructure filtering were used to select over 2,500 Key Organics compounds for experimental profiling. NMR and LCMS analysis allowed the careful selection of highly-soluble fragments with desirable physicochemical and stability characteristics. Importantly, the curated molecules are enriched in cyclic scaffolds commonly found in drug candidates, and spans chemical space that minimally overlaps with existing commercial collections. **This poster will summarize the design, cheminformatic and experimental features of this next generation Key Organics fragment library.**

Build Strategy

Rule of 3 Compliant
- Heavy atom count ≤ 16
- MW ≤ 300 clogP ≤ 3 HBA/HBD ≤ 3 PSA ≤ 60 Rot. Bonds ≤ 3

Remove the Undesirables
- BMS, PAINS, FAFDrugs2, Kazius and Bursius toxicophores, Lilly Med Chem Rules
- Med Chem curation

Enrich for Desirables
- Privileged scaffolds
- BioCores
- 3-dimensional shape

Maximize Diversity
- Calculated properties
- Chemical fingerprints
- Scaffolds
- 3D shape

QC by NMR
- Structure Verification
- Purity
- Solubility
- Stability
- Aggregation
- Chemical Shift Encoding (pooling)

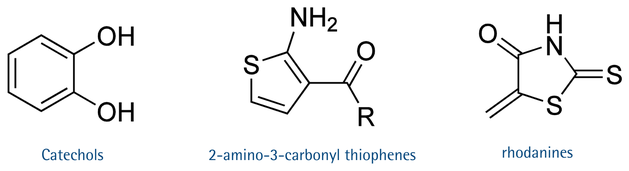
Remove the Undesirables

Substructure Filtering

As part of our Fragment selection process, industry-standard substructure filtering - including PAINS filtering - was implemented and as a result the BIONET 2nd Generation Premium Fragment Library does not include substructures identified as promiscuous or reactive by empirically determined rejection rules. A fragment was rejected if it failed any one of five rejection rules: BMS[1], PAINS[2], FAFDrugs2[3], Kazius and Bursi toxicophores[4], and Lilly MedChem Rules[5].

Focus on Pan Assay Interference Compounds (PAINS) substructure filtering – a deciding factor in the quality of a fragment library. PAINS are compounds that frequently show up as screening hits, but that act through non-specific mechanisms such as covalent attachment to proteins or generation of hydrogen peroxide. The problem with PAINS is that they may show convincing biochemical and even cell based activity, but mechanistically be useless for further advancement to drugs or even chemical probes. PAINS remain common in many vendors Fragment Libraries. PAINS compounds have been identified and substructure filters constructed that recognise these compounds[1].

Examples of PAINS

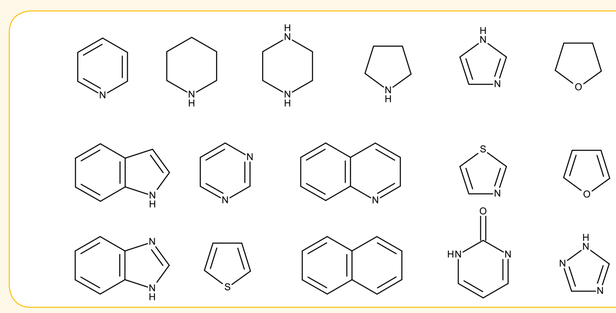


Catechols | 2-amino-3-carbonyl thiophenes | rhodanines

Enrich for Desirables

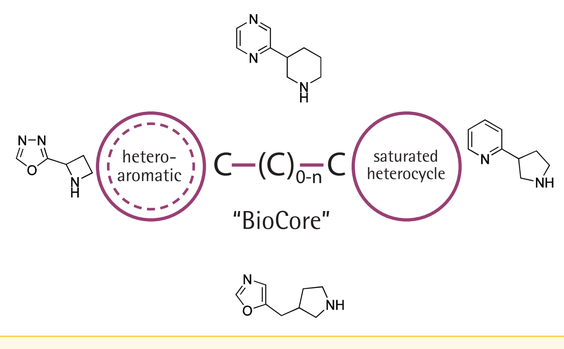
Privileged scaffolds

Rings, ring systems and frameworks in drugs have been analyzed[6,7] to understand their frequency in bioactive compounds. Our strategy involved enrichment of the fragment library with privileged scaffolds typically found in bioactive compounds.



BioCores

BioCores are defined as a heteroaromatic ring connected by a carbon linker to a saturated heterocyclic molecule in Kombarov et al. Molecular Diversity 2010, 14, 193-200. Incorporation of BioCores into the scaffolds of a molecular library enhances their natural/drug like properties. BioCores are underrepresented in commercially available fragment libraries.



Maximize Diversity

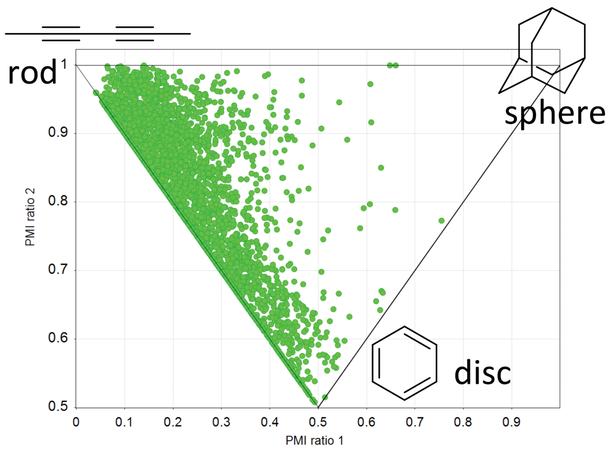
General approach

A diverse selection was made initially using an iterative selection (preferring BioCores or privileged scaffolds) adding compounds that filled in the most chemical space as determined by 3D shape, 2D structure, and calculated properties. The resulting selection was then examined in the context of commercial fragment space in two spaces: cells of a Self-Organizing Map of 2D fingerprint space and in bins of Principal Component Analysis of property space.

2-D structure: Scaffold and Fingerprint

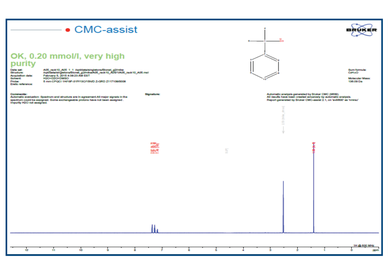
Scaffold | Fingerprint

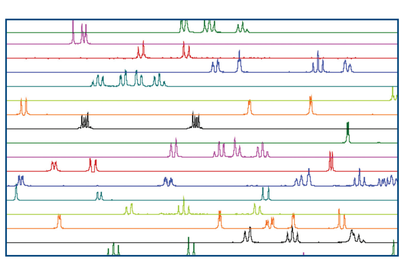
3D shape: Principle Moments of Inertia



QC by NMR

All fragments in the 2nd Generation Premium Fragment Library have been analyzed by ^{1}H NMR at 600 MHz for structure verification, purity, solubility, and lack of aggregation. Spectra are available to customers for Chemical Shift Encoding (CSE), thus allowing custom pools to be built with significant time and cost savings.





NMR Spectra Allows for Assessment of:

- Structure Verification
- Purity
- Solubility
- Stability
- Aggregation
- Chemical Shift Encoding (pooling)

In Summary

A 2nd generation BIONET Premium Fragment Library has been constructed employing Rule of Three and industry standard substructure filtering including PAINS analysis. Diversity selection utilized methods in shape, scaffold, fingerprint and predicted property space. All fragments in the 2nd Generation Premium Fragment Library have been analyzed by 1H NMR for structure verification, purity, solubility, and lack of aggregation. Spectra are available to customers for Chemical Shift Encoding (CSE), thus allowing custom pools to be built with significant time and cost savings.

- ☑ Rule of 3 compliant: MW ≤300, cLogP ≤ 3, number of HBA/HBD ≤ 3, PSA ≤ 60 and Number rotatable bonds ≤3
- ☑ Heavy atom count (HAC) ≤ 16
- ☑ Does not include substructures identified as promiscuous or reactive by empirically determined rejection rules
- ☑ Inclusion of diverse scaffolds that are present in bioactive compounds and that have 3-dimensionality
- ☑ Clustering and Diversity analysis
- ☑ Passes chemist visual inspection
- ☑ Solubility in PBS buffer and signs of aggregation determined by 1H NMR spectra

References

1. Bradley C. Pearce , Michael J. Sofia, Andrew C. Good, Dieter M. Drexler, and David A. Stock. An Empirical Process for the Design of High-Throughput Screening Deck Filters. *Journal of Chemical Information and Modeling* **2006**, 46, 1060-1068.
2. Jonathan B. Baell and Georgina A. Holloway. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *Journal of Medicinal Chemistry* **2010**, 53, 2719-2740.
3. David Lagorce, Olivier Sperandio, Hervé Galons, Maria A Miteva, and Bruno O Villoutreix. FAF-Drugs2: Free ADME/tox filtering tool to assist drug discovery and chemical biology projects. BMC Bioinformatics 2008, 9:396.
4. Jeroen Kazius, Ross McGuire, and Roberta Bursi. Derivation and Validation of Toxicophores for Mutagenicity Prediction. *Journal of Medicinal Chemistry* 2005, 48, 312-320.
5. Robert F. Bruns and Ian A. Watson. Rules for identifying potentially reactive or promiscuous compounds. *Journal of Medicinal Chemistry* **2012**, 55, 9763-9772
6. Lisurek M, Rupp B, Wichard J, Neuenschwander M, von Kries JP, Frank R, Rademann J, Kühne R. Design of chemical libraries with potentially bioactive molecules applying a maximum common substructure concept. Mol Divers. 2010 May;14(2):401-8. doi: 10.1007/s11030-009-9187-z. Epub 2009 Aug 15
7. Richard D. Taylor, Malcolm MacCoss, and Alastair D. G. Lawson. Rings in Drugs. J. Med. Chem., 2014, 57 (14), pp 5845–5859.

BIONET Fragment Libraries

For more information, please contact us at: **andrewl@keyorganics.net**