# Practical Aspects of Designing the BIONET Fluorine Fragment Library

Steven R. LaPlante[1], Patrick McCarren[2], Francois Bilodeau[1], Andrew Lowerson[3].

[1]NMX Research and Solutions, Laval, Quebec, Canada; University of Quebec, INRS-IAF, Laval, Quebec, Canada.
[2]Broad Institute of MIT and Harvard, Cambridge, MA, USA
[3]Key Organics, Camelford, Cornwall, United Kingdom

A BIONET Fluorine 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.

## Key Features and Benefits:

- Ro3 compliant
- Measured solubility in PBS buffer ≥ 200µM
- 97% of Fragments soluble in DMSO at 200mM
- Purity ≥ 95%
- Filtered to remove toxic and reactive groups
- Filtered for PAINS substructures
- Diversity coefficient = 0.68

### All 461 fragments in the Fluorine Fragment Library have been analyzed by $^{19}$F NMR and $^{1}$H NMR for:

- Structure verification
- Purity
- Measured solubility in PBS buffer ≥ 200µM
- Lack of aggregation.

### All customers will be supplied with the following data package for each aqueous soluble fluorine fragment purchased:

- Aqueous buffer $^{19}$F NMR pdf
- Aqueous buffer $^{19}$F NMR raw data file
- $^{19}$F NMR chemical shifts supplied in an excel file

NMR files also include analyses data derived from the CMC Assist module which automatically picked resonance peaks and tabulated the data within the spectral files, these can then be used for chemical shift encoding purposes.

The process began with the virtual assessment of the Key Organics Fragment Collection (> 20K fragments) that would be physically available in large quantities then filtering this to a fragment-like space. The "Rule of Three" was used to bias towards fragments likely to be fragment hits[1]. Further, a design element of the Astex screening library was adopted that further reduces the size of fragments ≤ 16 heavy atoms, above which much lower hit rates were observed[2]. All properties were calculated using Pipeline Pilot[3].

## The parameters used in the design of Fluorine Fragment Library:

- Heavy atoms ≤16
- logP ≤ 3,
- Hydrogen bond donors ≤3
- Hydrogen bond acceptors ≤3
- Polar surface area ≤ 60
- Rotatable bonds ≤ 3

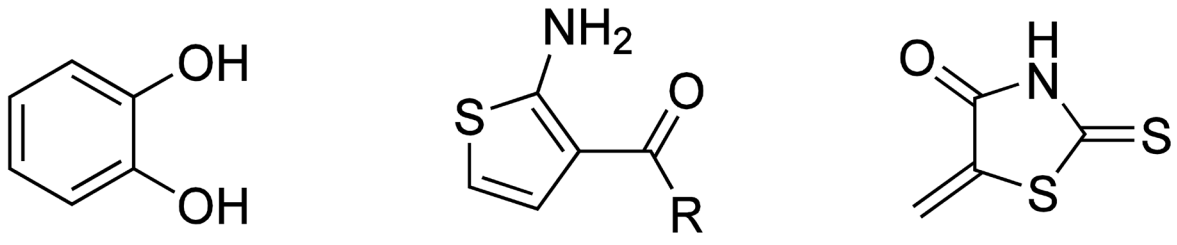
As part of our Fluorine Fragment selection process, industry-standard substructure filtering - including PAINS filtering - was implemented and as a result the BIONET Fluorine Fragment Library does not include substructures identified as promiscuous or reactive by the following empirically determined rejection rules:

- Lilly MedChem Rules[4]
- PAINS[5]
- BMS[6]

## 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 fragments have been identified and substructure filters constructed that recognise these fragments

**Figure 1: Example of PAINS**



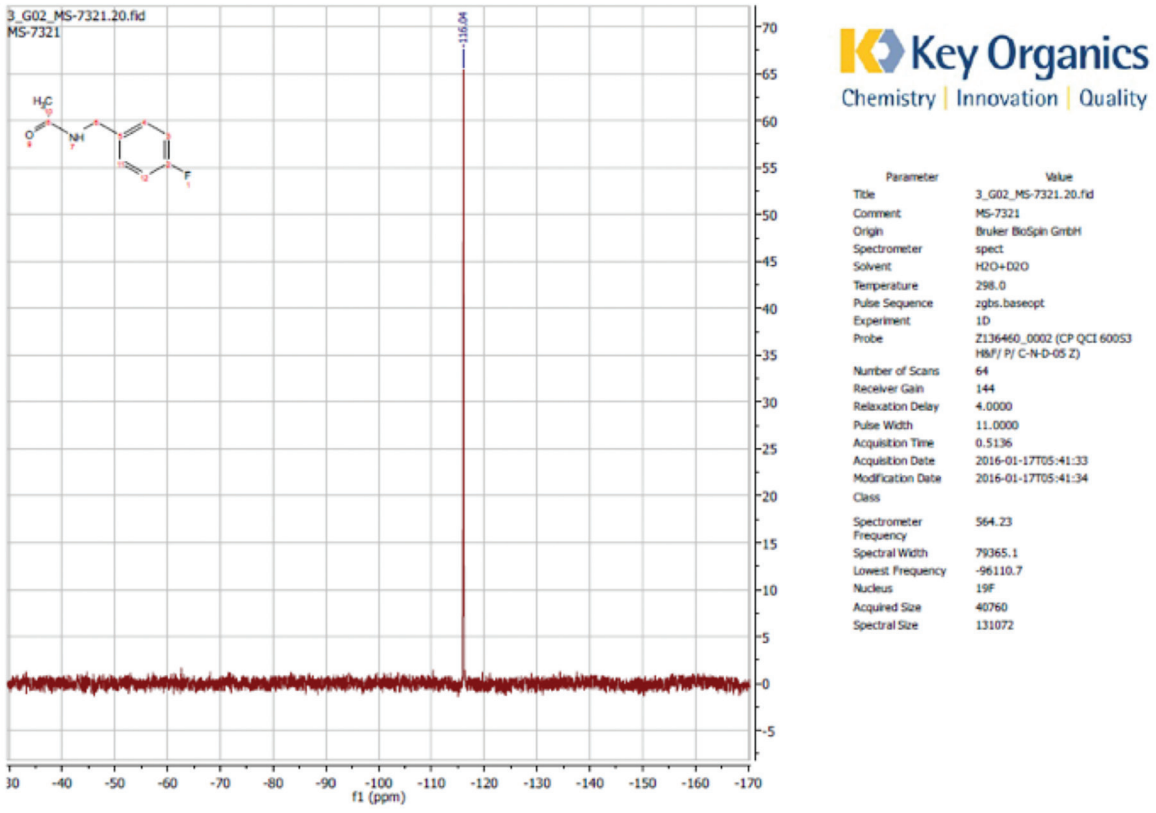
## $^{19}$F and $^{1}$H NMR curation for fragment prioritisation and library characterization

$^{19}$F and $^{1}$H NMR were then employed to select compounds with the appropriate solution behaviour and be amenable to rigorous biophysical analysis in physiologically-relevant aqueous solution. Each singleton sample consisted of nominal 300 µM compound in buffer (50 mM sodium phosphate pH 7.4, 100 mM NaCl). $^{1}$H NMR spectra were acquired at 600 MHz spectrometer equipped with a helium cryoprobe that significantly increased signal-to-noise. Simple 1D $^{19}$F and $^{1}$H NMR spectra were acquired. Compounds with solubility at 200 µM and higher were prioritized given that most fragment screens and assays require high concentrations.

The compounds and respective spectra were then re-subjected to visual inspection by medicinal chemists experienced in the art of fragment-based drug discovery. Compounds that were inconsistent with the NMR spectra were removed, along with those that showed signs of insolubility, instability or aggregation.

Moreover, the NMR curation also allowed us to evaluate and characterize the final library. This experimental data served to verify structural integrity, purity, solubility, stability, aggregation and chemical shift positions. The final library consisted of 461 fragments that had solubility's of at least 200 µM in PBS aqueous buffer. The library included 64 fragments found in approved drugs and all 461 contain rings found in drugs.

**Figure 2:** Example $^{19}$F NMR spectra.



## Salient statistics of the final library.

A visualization of the final Fluorine Fragment Library properties and diversity is provided in the following Figures 3, 4 and 5:

**Figure 3:** Overview of calculated properties of the final library

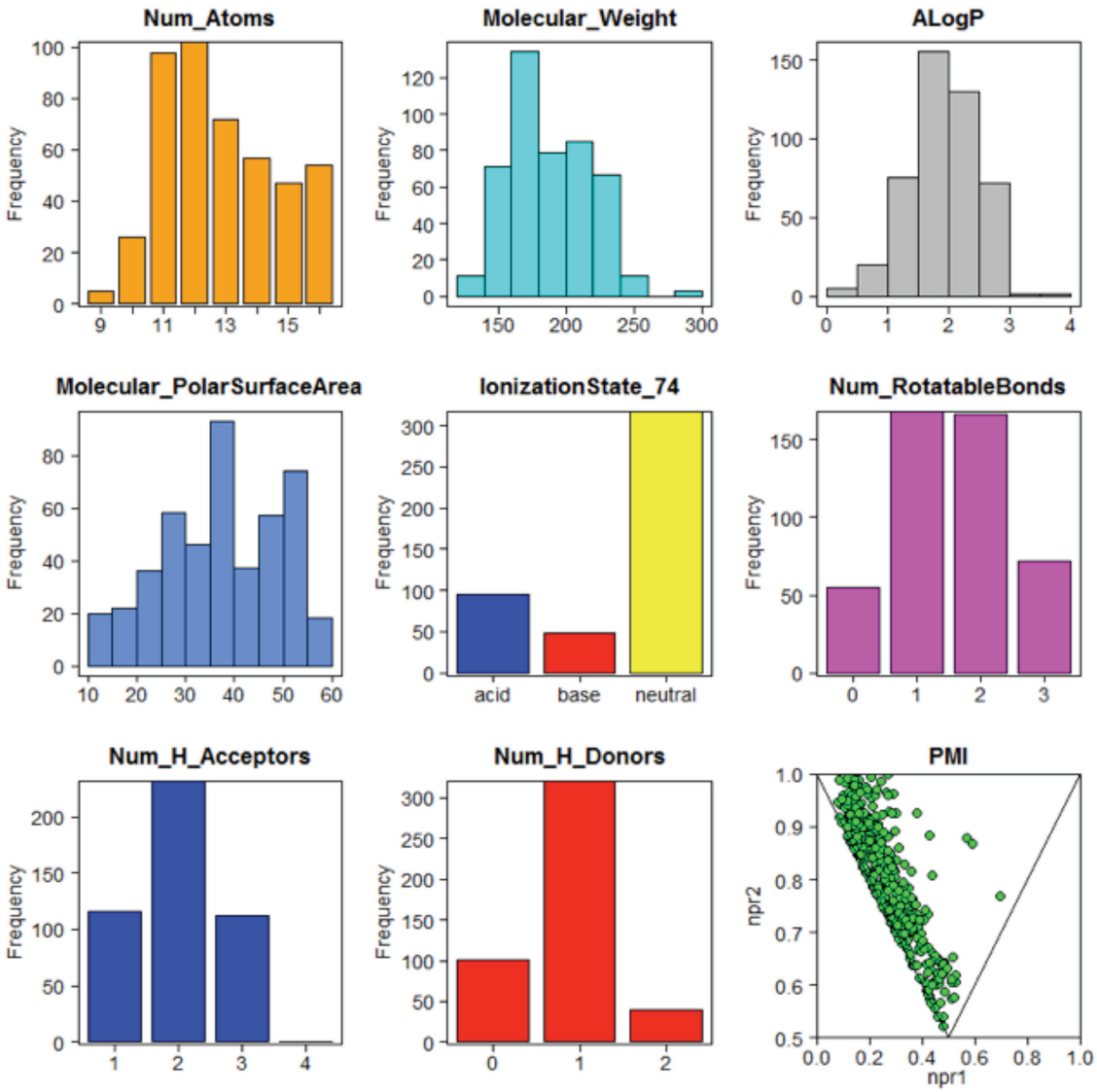
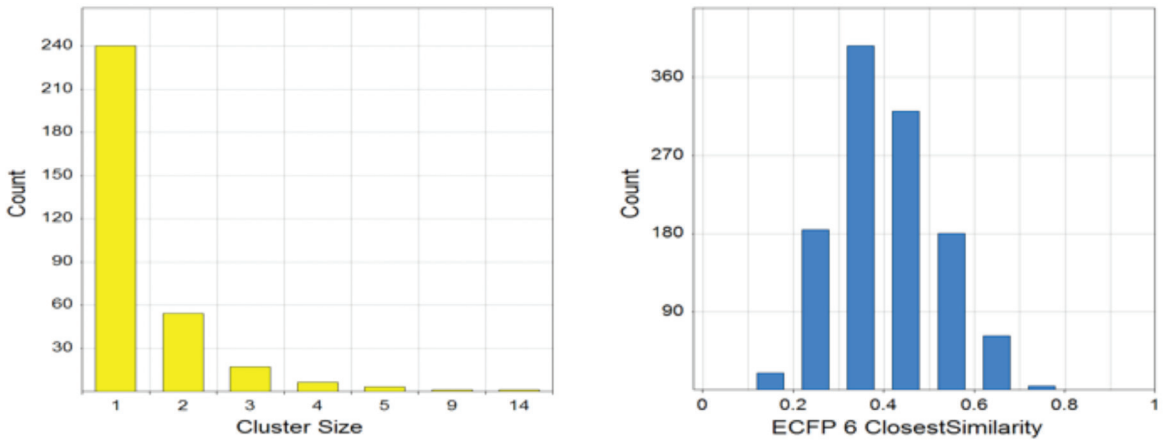



**Figure 4:** Overview of the range of properties of the final library

| Property | Average | Min | Max |
|---|---|---|---|
| Molecular Weight | 187 | 122 | 282 |
| Heavy Atom Count | 12.8 | 9 | 16 |
| Polar Surface Area | 37 | 12 | 58 |
| ALogP | 1.9 | 0.2 | 3.5 |
| # H-bond acceptors | 2 | 1 | 4 |
| # H-bond donors | 0.87 | 0 | 2 |
| # rotatable bonds | 1.6 | 0 | 3 |

**Figure 5:** Diversity measure summary.

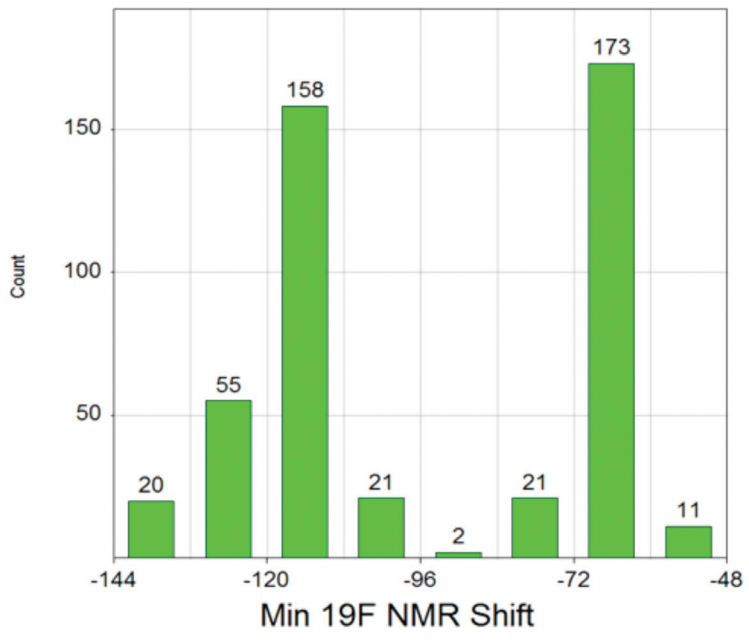



## Diversity Statistics:

- Diversity coefficient (average distance using MACCS 166-bit keys) = 0.68
- # clusters at 0.85 Tanimoto similarity (MACCS 166-bit MOE) = 240 singletons, 322 clusters / 461 fragments 75%

## $^{19}$F NMR shifts in the set

**Figure 6:**
$^{19}$F NMR shift statistics. 199 CF3 (-84.7 to -52.9 ppm), 2 CF2 (-104 to -88 ppm), 3 OCF3 (-57 ppm), 9 OCF2H (-81.0 to -81.7 ppm), 252 F (-141 to -96 ppm).



# Summary

All 461 fragments in the Fluorine Fragment Library have been analysed by $^{1}H$ NMR and $^{19}F$ NMR for:

- Structure verification
- Purity
- Measured solubility in PBS buffer ≥ 200μM
- Lack of aggregation.

## Physiochemical Properties of the library:

- Heavy atoms ≤16
- AlogP ≤ 3.5 (ONLY 4 FRAGMENTS ≥ 3)
- Hydrogen bond donors ≤3
- Hydrogen bond acceptors ≤3
- Polar surface area ≤ 60
- Rotatable bonds ≤ 3

## Substructure Filtering:

- Lilly MedChem Rules[4]
- PAINS[5]
- BMS[6]

## Diversity Statistics:

- Diversity coefficient (average distance using MACCS 166-bit keys) = 0.68
- # clusters at 0.85 Tanimoto similarity (MACCS 166-bit MOE) = 240 singletons, 322 clusters / 461 fragments 75%

## DMSO Solubility:

447 Fragments are soluble in DMSO at 200mM

## Data Provision:

All customers will be supplied with the following data package for each aqueous soluble fluorine fragment purchased:

- Aqueous buffer $^{19}F$ NMR pdf
- Aqueous buffer $^{19}F$ NMR raw data file
- $^{19}F$ NMR chemical shifts supplied in an excel file

# References


1. Congreve, M.; Carr, R.; Murray, C.; Jhoti, H. *Drug Discov. Today* **2003**, 8 (19), 876–877.
2. Hall, R. J.; Mortenson, P. N.; Murray, C. W. *Prog. Biophys. Mol. Biol.* **2014**, 116 (2-3), 82–91.
3. Pipeline Pilot v 8.5; note H-bond acceptors and donors used the default Accelrys Num_H_Acceptors and Num_H_Donors which are different from the Lipinski definition.
4. Bruns, R. F.; Watson, I. A. *J. Med. Chem.* **2012**, 55 (22), 9763–9772.
5. Baell, J. B.; Holloway, G. A. *J. Med. Chem.* **2010**, 53 (7), 2719–2740.

Key Organics

Chemistry | Innovation | Quality

BIONET

Fragment Libraries

For more information please contact us at:

T: +44 (0)1840 212137 F: +44 (0)1840 213712 E: fragments@keyorganics.net www.keyorganics.net

Key Organics Ltd., Highfield Road Industrial Estate, Camelford, Cornwall PL32 9RA, United Kingdom