It's a small world.
What can we learn from published hits?
A work in progress……

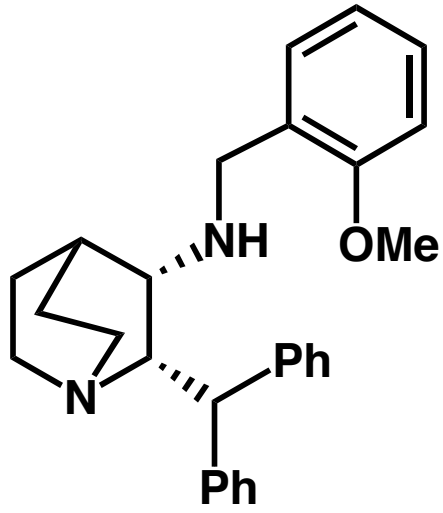
Chris Swain

Cambridge MedChem Consulting

Personal Journey

- How did I become interested in fragment-based drug discovery?

The first Non-Peptide Substance P Antagonist



CP-96,345

hNK_1 IC_{50} 0.6 nM

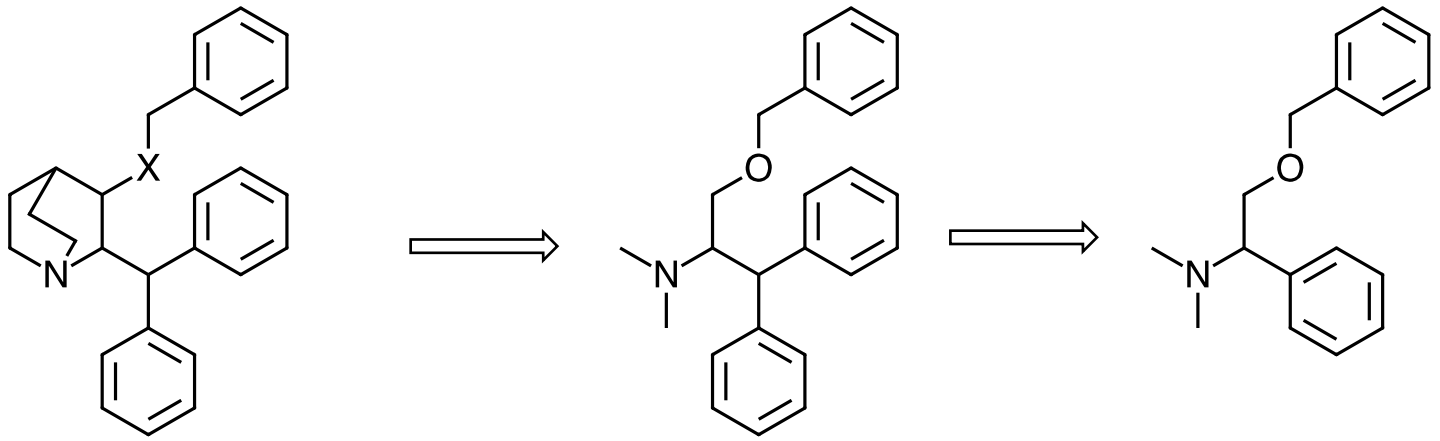
Poor Oral Availability

Cardiovascular side-effects

(Possibly due to Ca^{++} channel blockade)

- Medicinal Chemistry Programme to address :
  - Novelty
  - Selectivity particularly over ion channels
  - Oral Availability and half-life

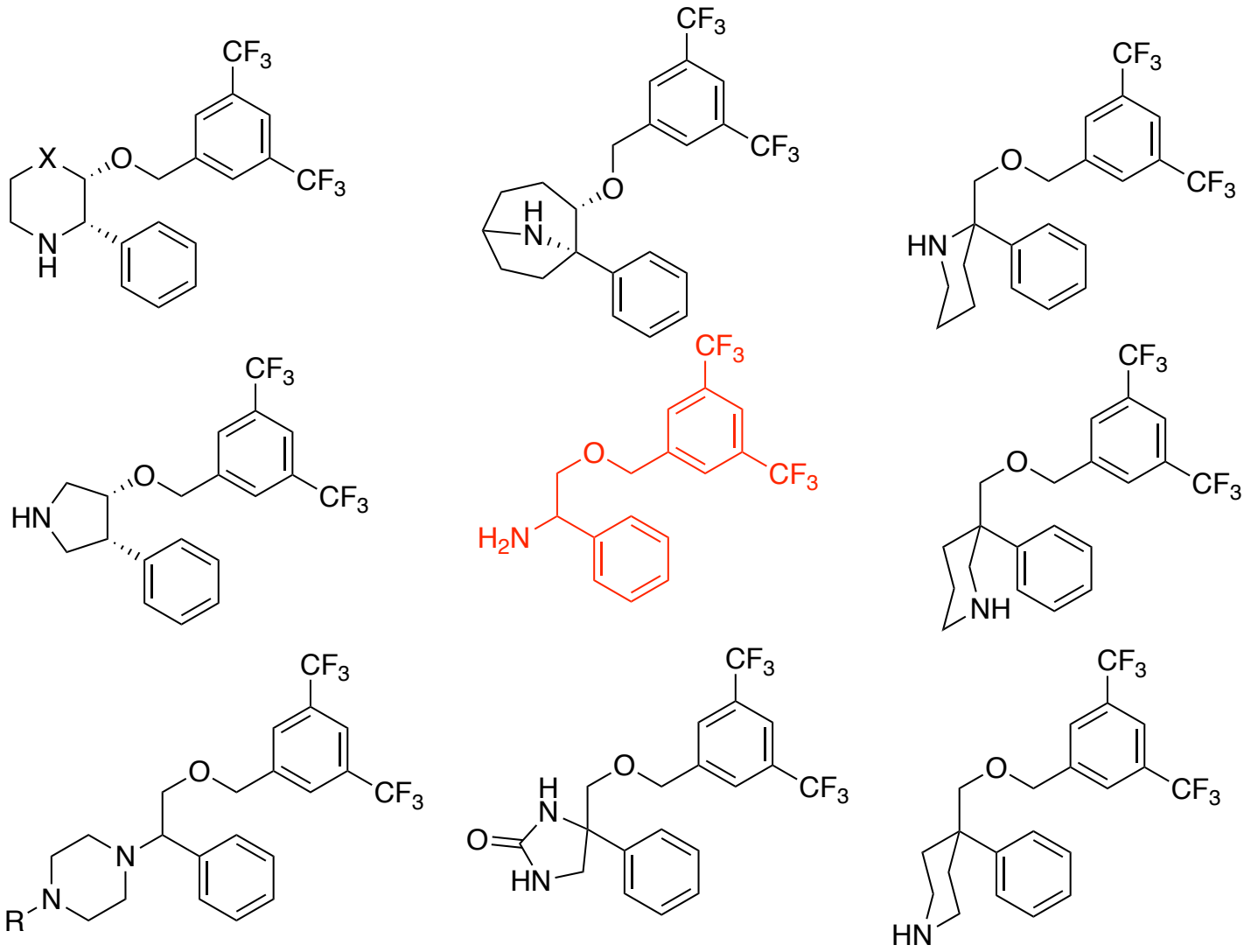
Identification of minimum structure



MWt 383.5
xLogP 6.8

MWt 255
xLogP 3.67

Introducing Conformational Restraints



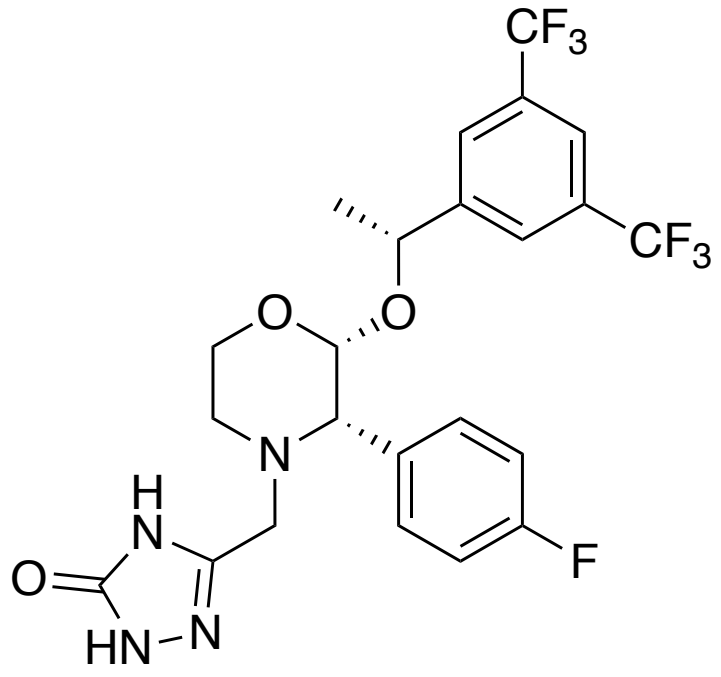
J. Med. Chem. **1998**, *41*, 4623-4635

MK-869: Aprepitant

hNK_1 IC_{50} 0.09 nM

Foot-tapping ID_{50} (iv, 1 h) 0.33 mg/kg

Foot-tapping ID_{50} (iv, 24h) 0.36 mg/kg



J.Med Chem **1998**, *41*, 4607-4614

Identification of novel, low molecular weight 5-HT2A antagonists

- Identify ALL low molecular weight compounds in the HTS deck (MWt 150-200)
- Cluster using maximum common substructure
- Add screening data and select ALL clusters that contain at least one structure > 50% inhibition at 5 μM
- Define SMARTS queries that represent each of the active cluster.

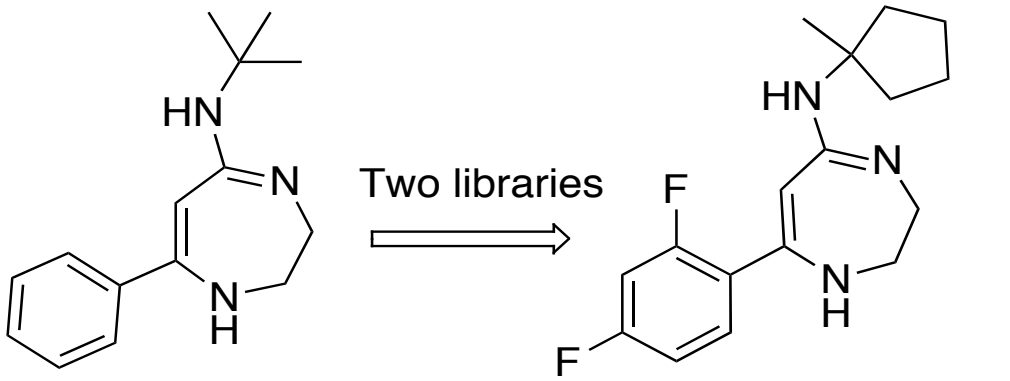
Swain et al Bioorg and Medchem Lett 16, 6058.

Searching

- Each of the individual SMARTS was then used to search:-
  - A file of known 5-HT_{2A} ligands
  - A file of actives from other screens
- Many of the SMARTS queries identified structures in the file of known 5-HT_{2A} ligands, whilst others were found in both 5-HT_{2A} and other screens and possibly represent promiscuous motifs.
- The most interesting SMARTS queries are those representing clusters containing actives in the HTS that are not present in either known 5-HT_{2A} or other screens.

Results

- Seventeen compounds (50-70% inhibition at 5μM in the HTS) were selected for titration and ten were subsequently shown to have useful activity.



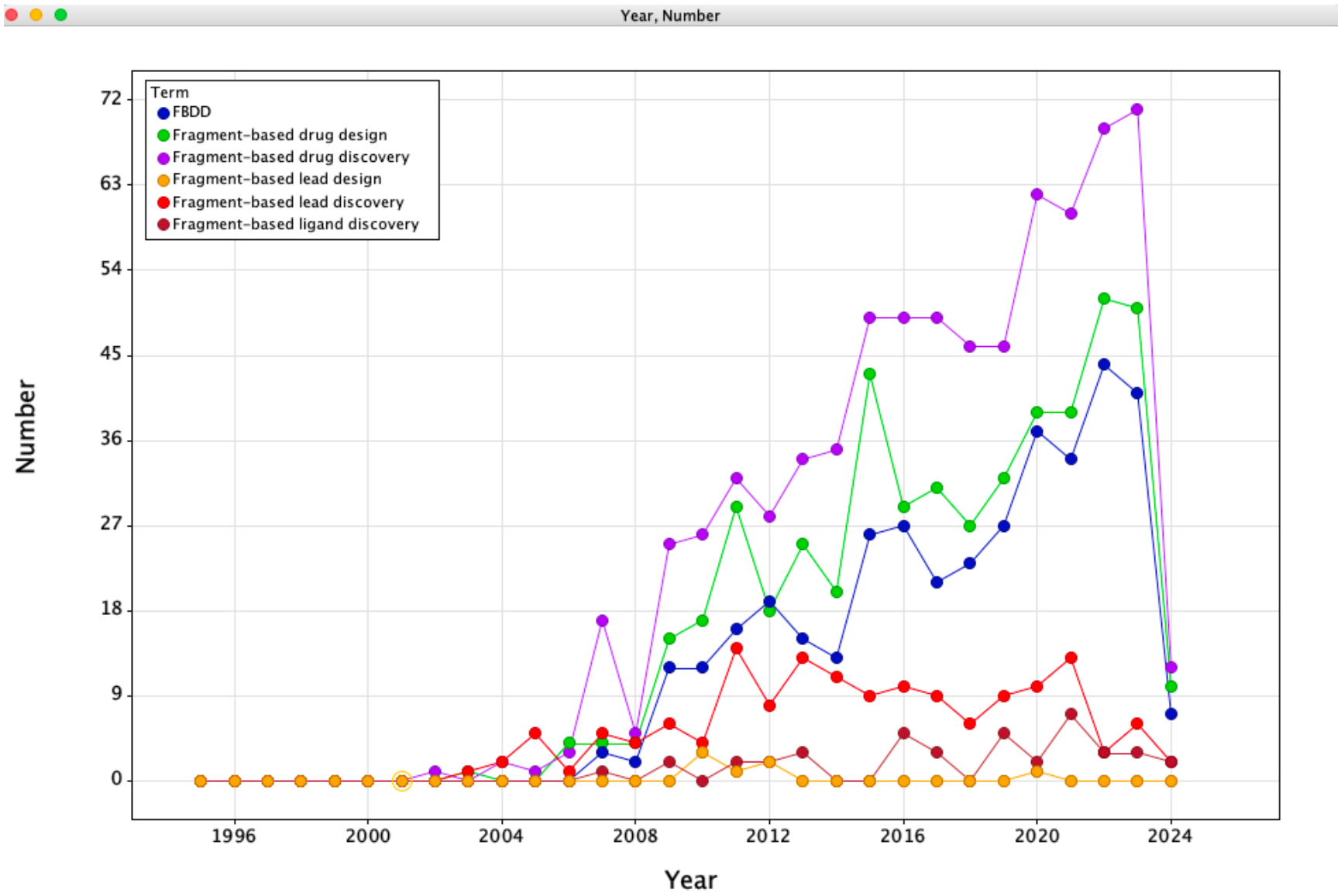
Sub-nanomolar activity

MWt 305

Fragment-based drug discovery

- SB Shuker, PJ Hajduk, RP Meadows, SW Fesik, Discovering high-affinity ligands for proteins: SAR by NMR
  - *SCIENCE* 29 Nov 1996 Vol 274, Issue 5292 pp. 1531-1534. https://doi.org/10.1126/science.274.5292.1531
- Murray, C.W., Verdonk, M.L. The consequences of translational and rotational entropy lost by small molecules on binding to proteins.
  - *J Comput Aided Mol Des* **16**, 741–753 (2002). https://doi.org/10.1023/A:1022446720849

Mentions in PubChem



Fragment-Based Screening

- Fragment-based screening has become increasingly popular and has proven to be a viable alternative to high-throughput screening.
- Fragment space is smaller
  - A million compounds cover only a small fraction of the suggested 10^{60} Chemical Space, whilst 2000 compounds can probe much of the 10^{6} Fragment Space
- Hit rates for Fragment-based screening appear to be higher, typically 3-10%.
- Binding Efficiency for small molecules are often higher.

Design of the Fragment Library

- Several approaches have been described in the design of fragment libraries. Most comply with the commonly accepted Astex "Rule-of-Three"
  - MW <300, H-bond donors/acceptors <=3, cLogP <3.
- Solubility is key requirement since screening carried out at higher concentrations
  - Often overlooked
- Rather than simply cull available molecules there have been recent attempts to design libraries based on known drugs, PDB ligands, natural products, or enhanced 3D structure.

- Can we use the information from fragment hits reported in the literature to help design fragment libraries?

What can we learn from known fragment hits?

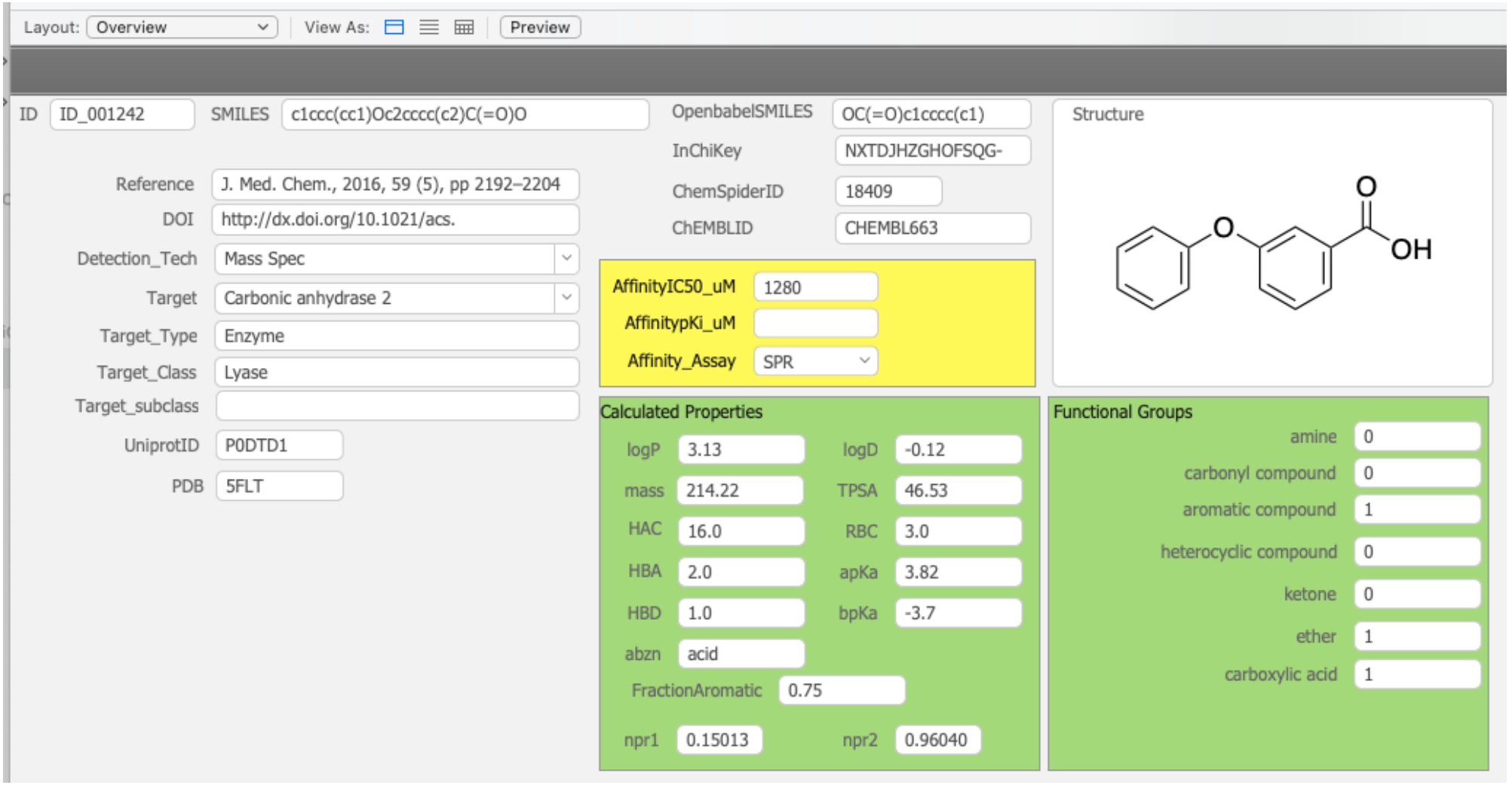
- Compile database of published hits from fragment screens. (Store as SMILES).
- Also include:-
  - Screening technology
  - Target and Uniprot ID, affinity (how measured), PDB code
  - Target type/class, using ChEMBL ontology
- Calculate
  - Physicochemical properties
    - cLogP, cLogD, PSA, HBA, HBD, RotB, pKa, shape descriptors, MR, HAC, fraction aromatic heavy atoms. (ChemAxon, MOE)
  - Functional groups
  - Cluster analysis

Current Status (Jan 2024)

- >300 Publications
- >2500 Published hits
- 265 Different targets
- 32 Detection technologies

- Finding the data is getting more of a challenge, it seems that as fragment screening becomes more mainstream it is often not mentioned in the title or abstract.

Fragments Database

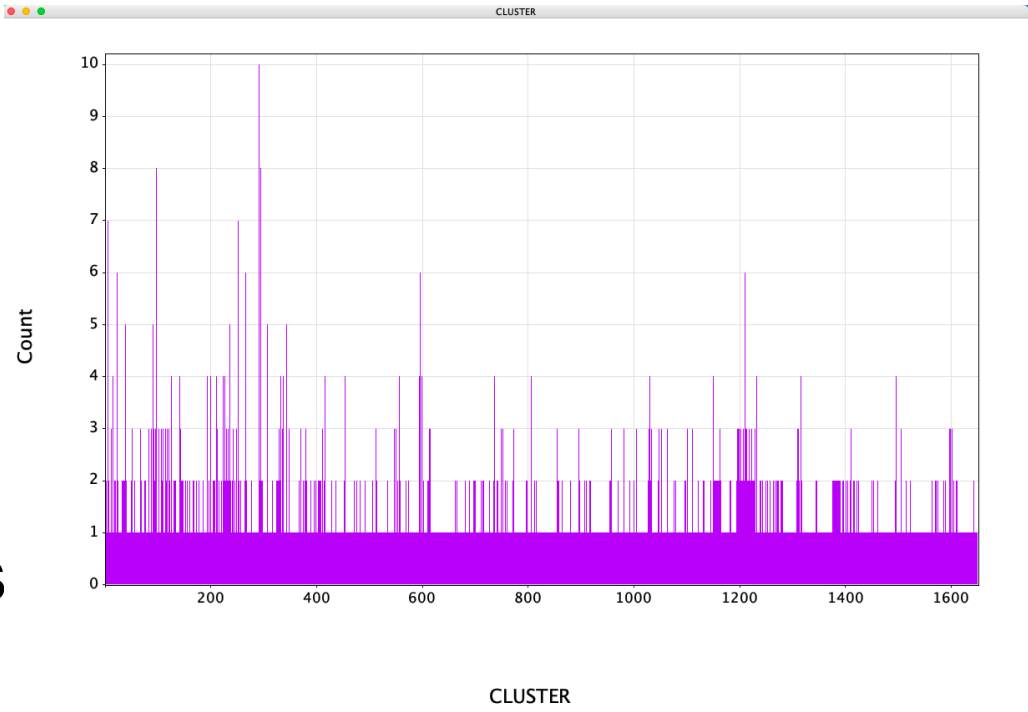


Mining the information

- Where do hits come from?
- Are there “preferred” structural classes?
- How do physicochemical properties impact?
- Does the target influence fragments?

Diversity

- Calculate similarity between each fragment and every other fragment in the database
  - Using 1028 bit circular fingerprints
- Average Tanimoto 0.09
- Clustering
  - Using Tanimoto 0.85
  - Mainly singletons
  - Some small clusters
  - Often duplicate structures

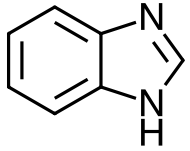


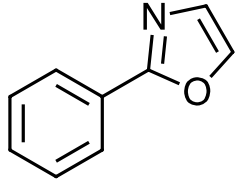
Functional Group Analysis of 2200 hits

- 2075 contain an aromatic ring, 1500 contain heteroaromatic
- 496 contain an arylhalide, 186 contain a phenol
- 298 contain an acidic group, 348 a basic group
- 42 contain a nitro group
- 132 contain a hydroxy, 339 an ether
- 720 contain an amine, 347 “anilines” (mainly on heteroaromatic systems)
- 605 amides, 98 esters, 79 ureas

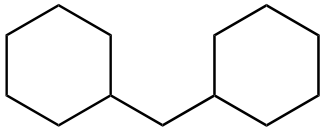
What are the most common scaffolds

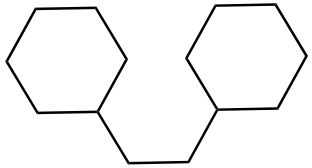
- Identify most common scaffolds



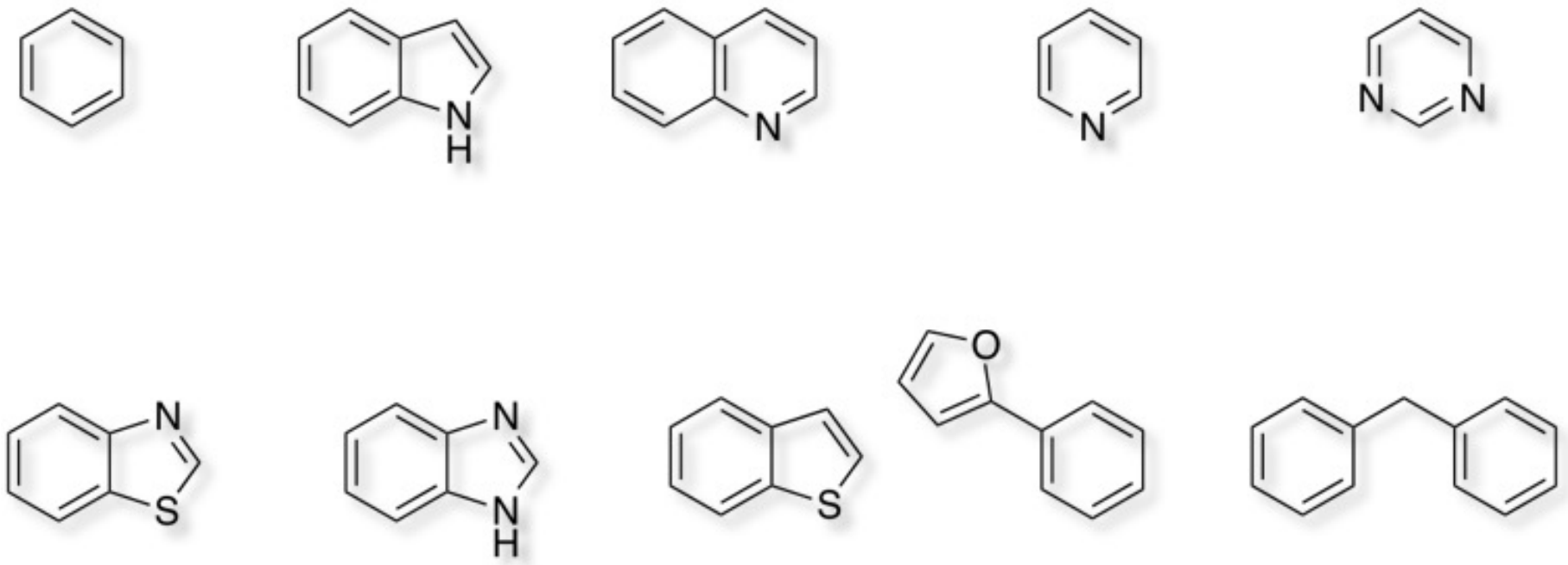


- Graph based scaffold analysis
  - Ignoring atom types

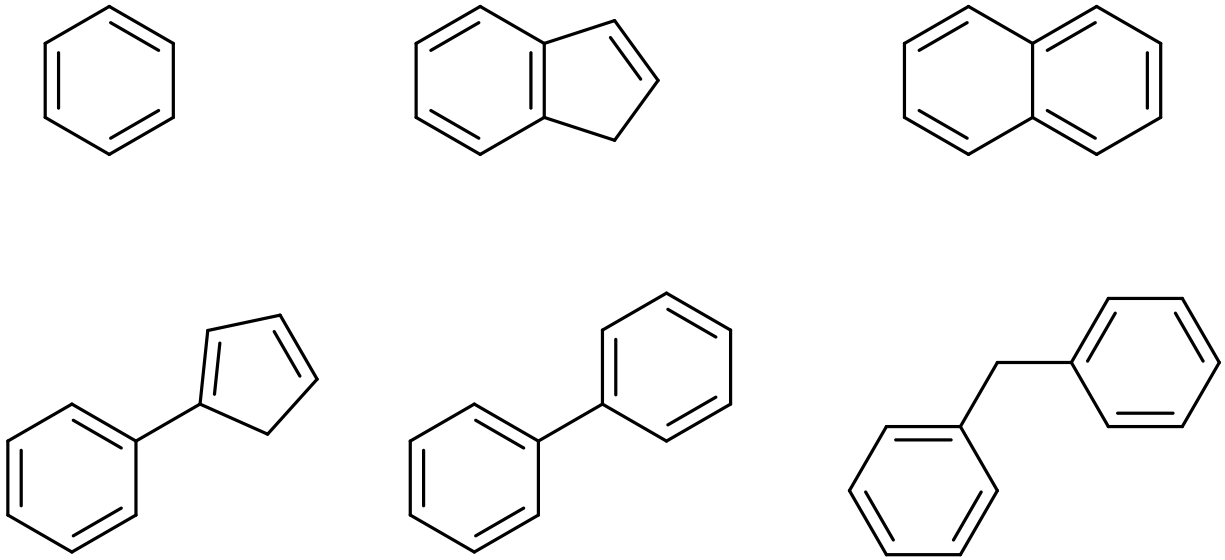




Most common scaffolds



Graph analysis



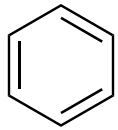
How does this compare with known ligands?

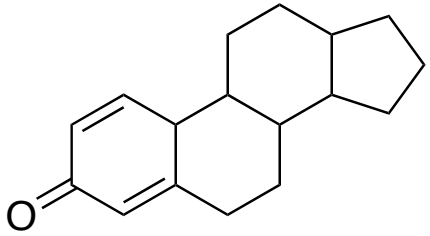
- Compare with
  - DrugBank
  - PDB
  - BindingDB

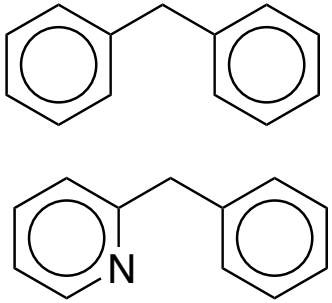
How does this compare with drugs

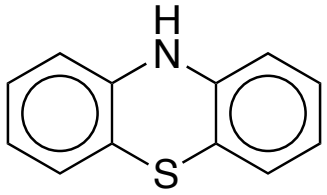
- Search DrugBank (www.drugbank.ca)
  - Approved, small molecule drugs.
- 1474 molecules exported
- Import into MOE database
- Use sca.svl to identify scaffolds
  - The script finds all scaffold in a database, writes them to a separate database
  - A New Approach to Finding Natural Chemical Structure Classes; J. Med. Chem. 2002, 45, 5311-5320
    - http://dx.doi.org/10.1021/jm010520k

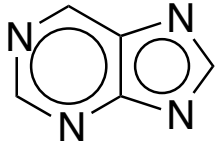
Most common Scaffolds DrugBank

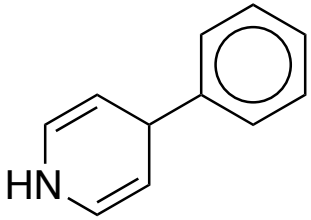


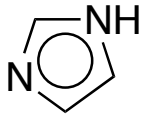


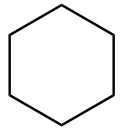


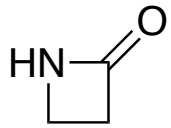


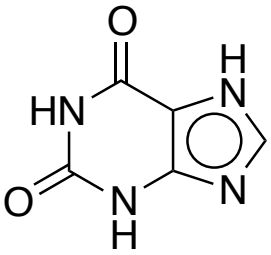




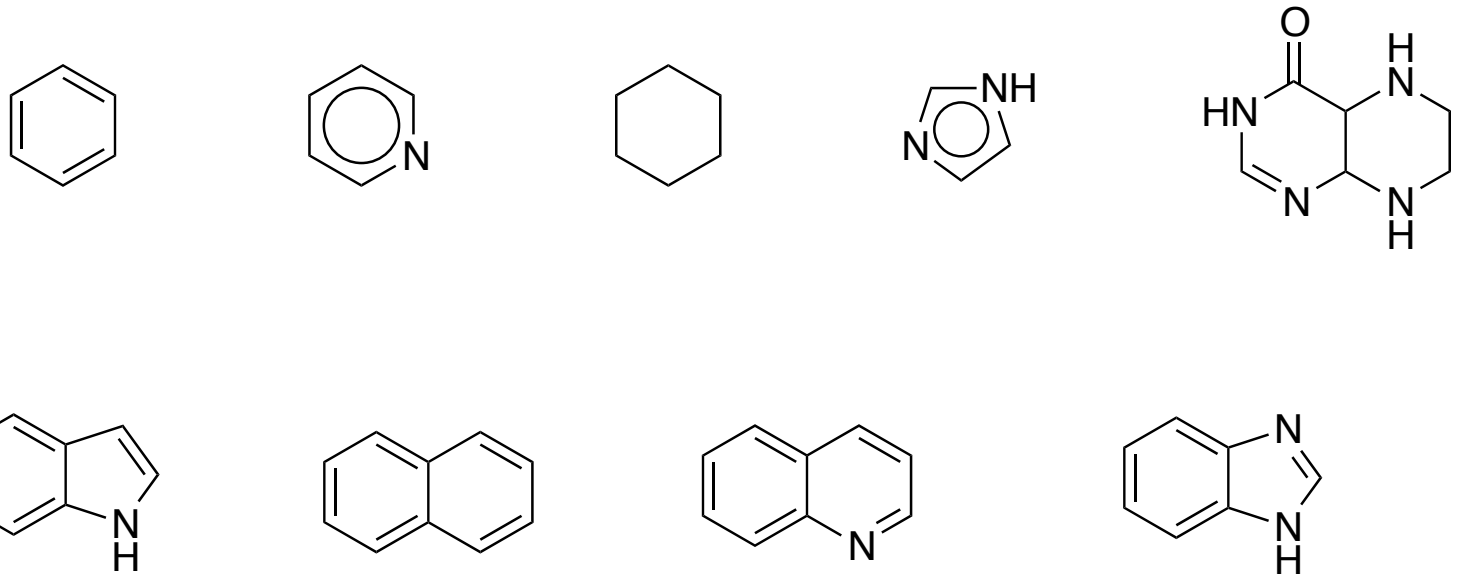




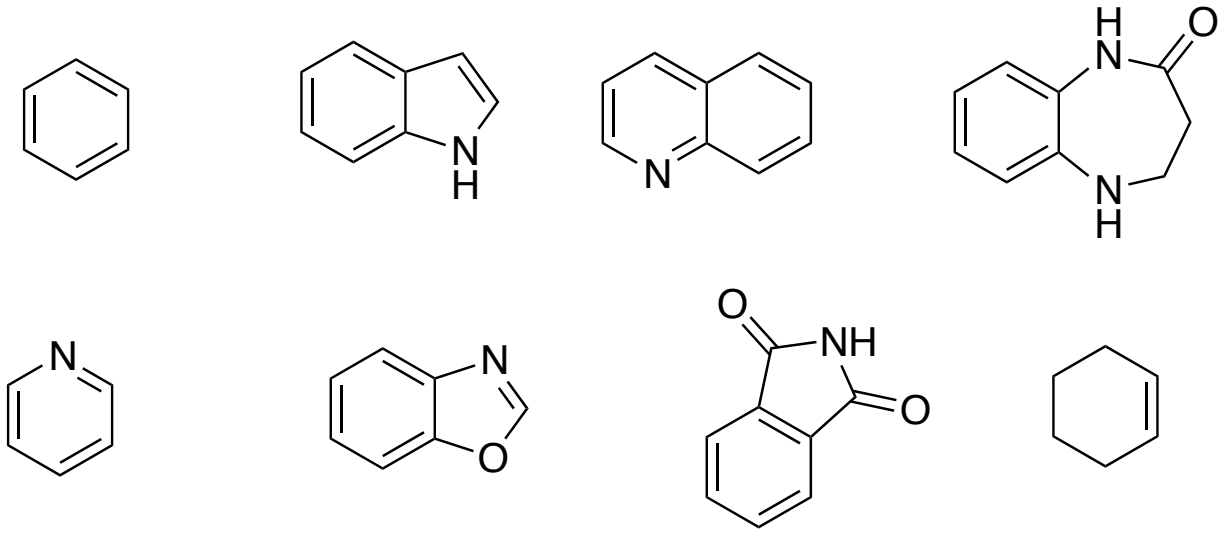




Most common fragments in PDB



Most common scaffolds in BindingDB



Conclusions

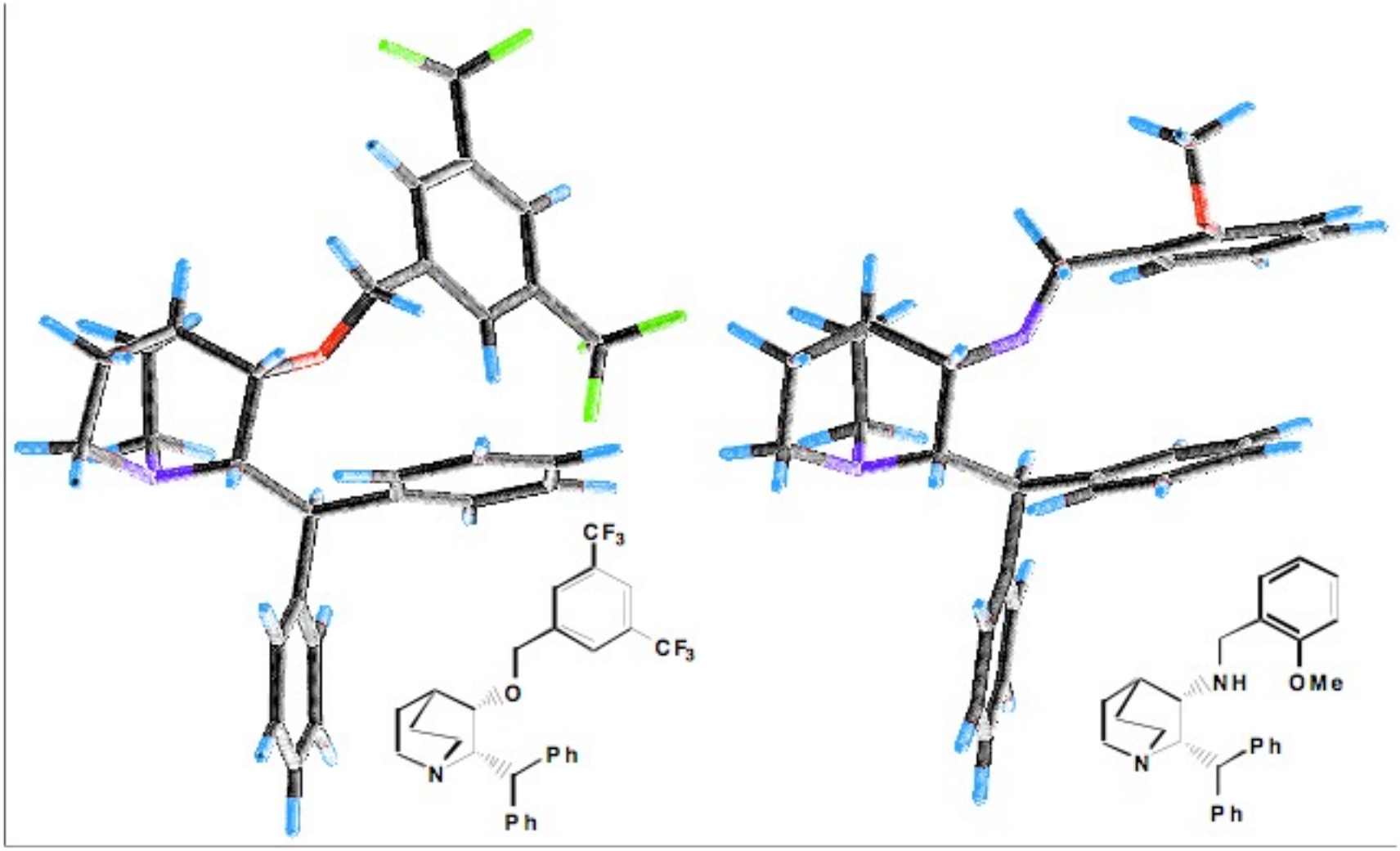
- Analysis of reported fragment hits highlights the preponderance of aromatic systems.
- Exploration of three public data sources of ligands indicates a similar observation.

  - Is there something special about aromatic scaffolds?

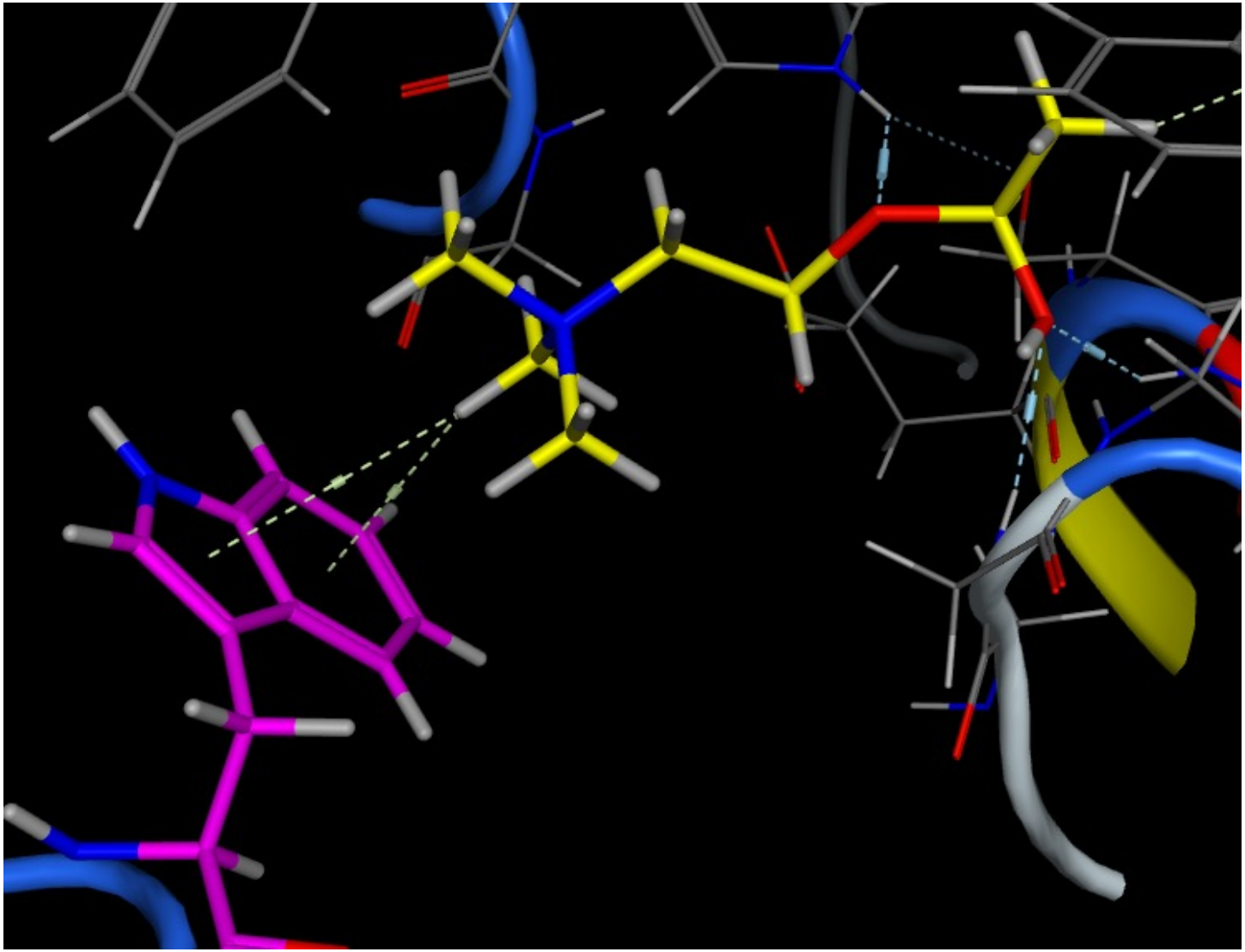
Is there something special about aromatic fragments?

- Simple hydrophobic interactions
- A hydrophobic ligand (or surface on the binding site) disrupts the structure of bulk water and decreases entropy because of stronger bonding and ordering of water molecules around the solute, if ligand and binding site associate then some of the water molecules can be returned to bulk water.

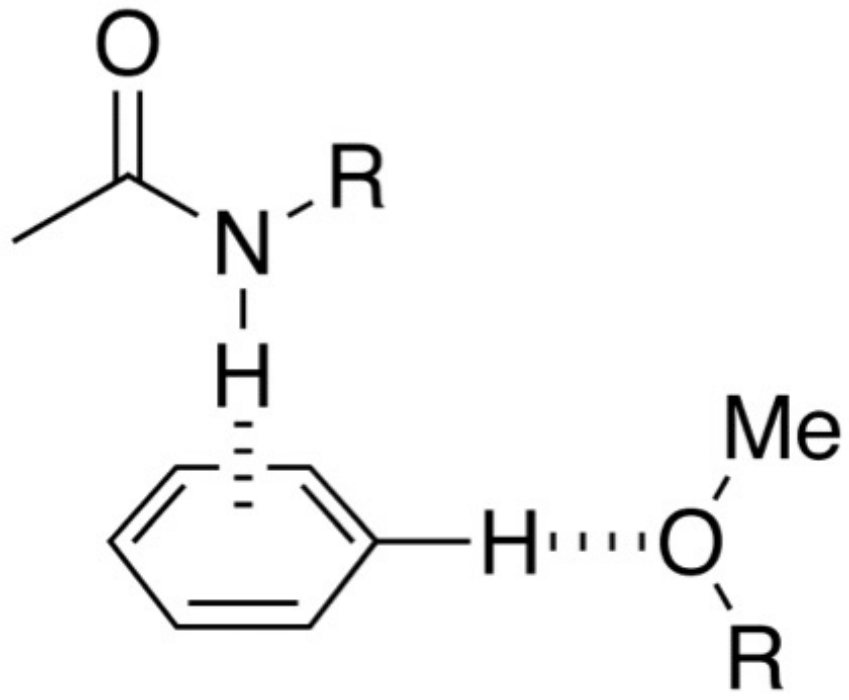
Aryl-Aryl Interactions



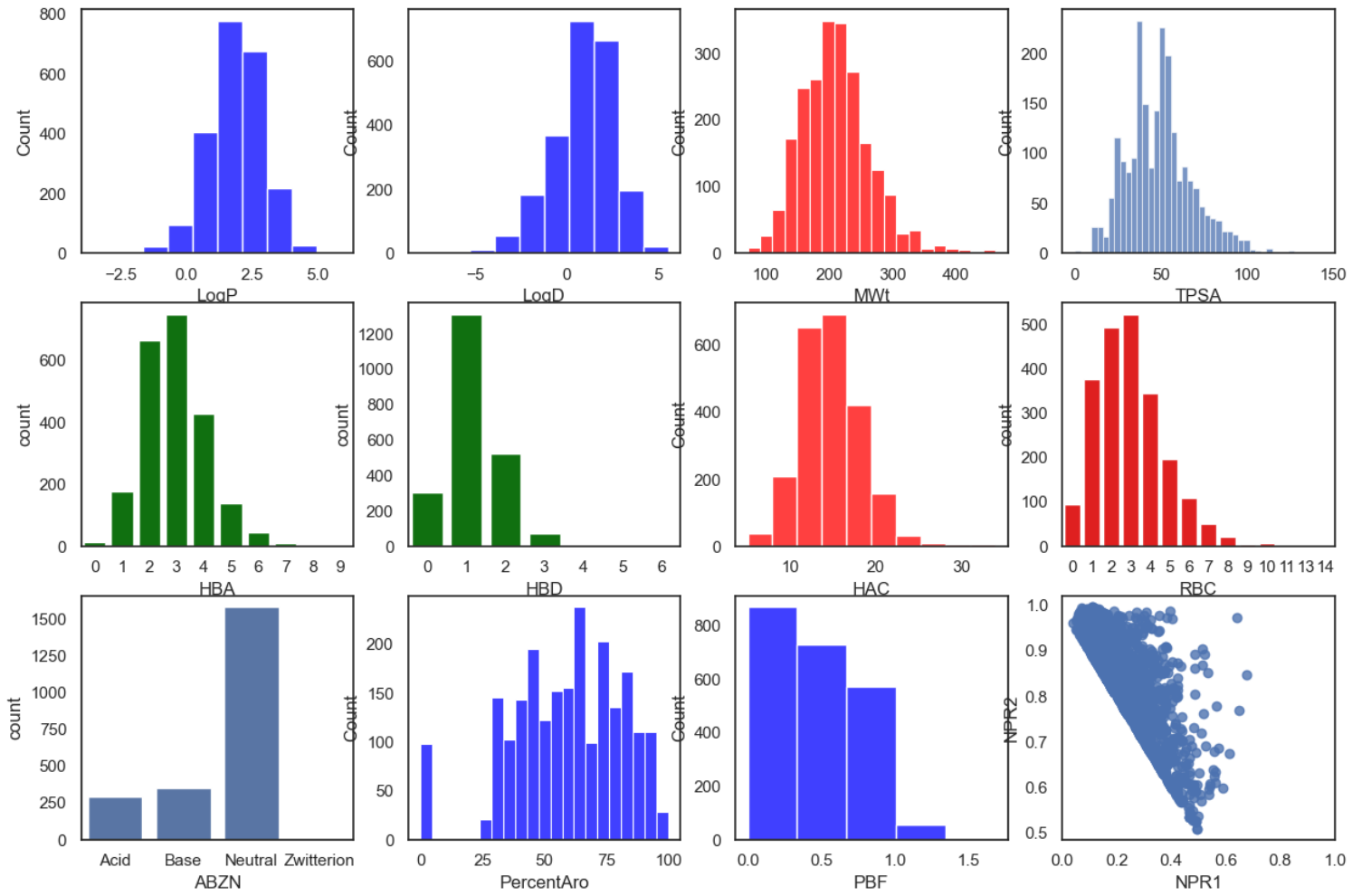
Cation-Aryl Interaction



Aryl O-H or N-H interactions



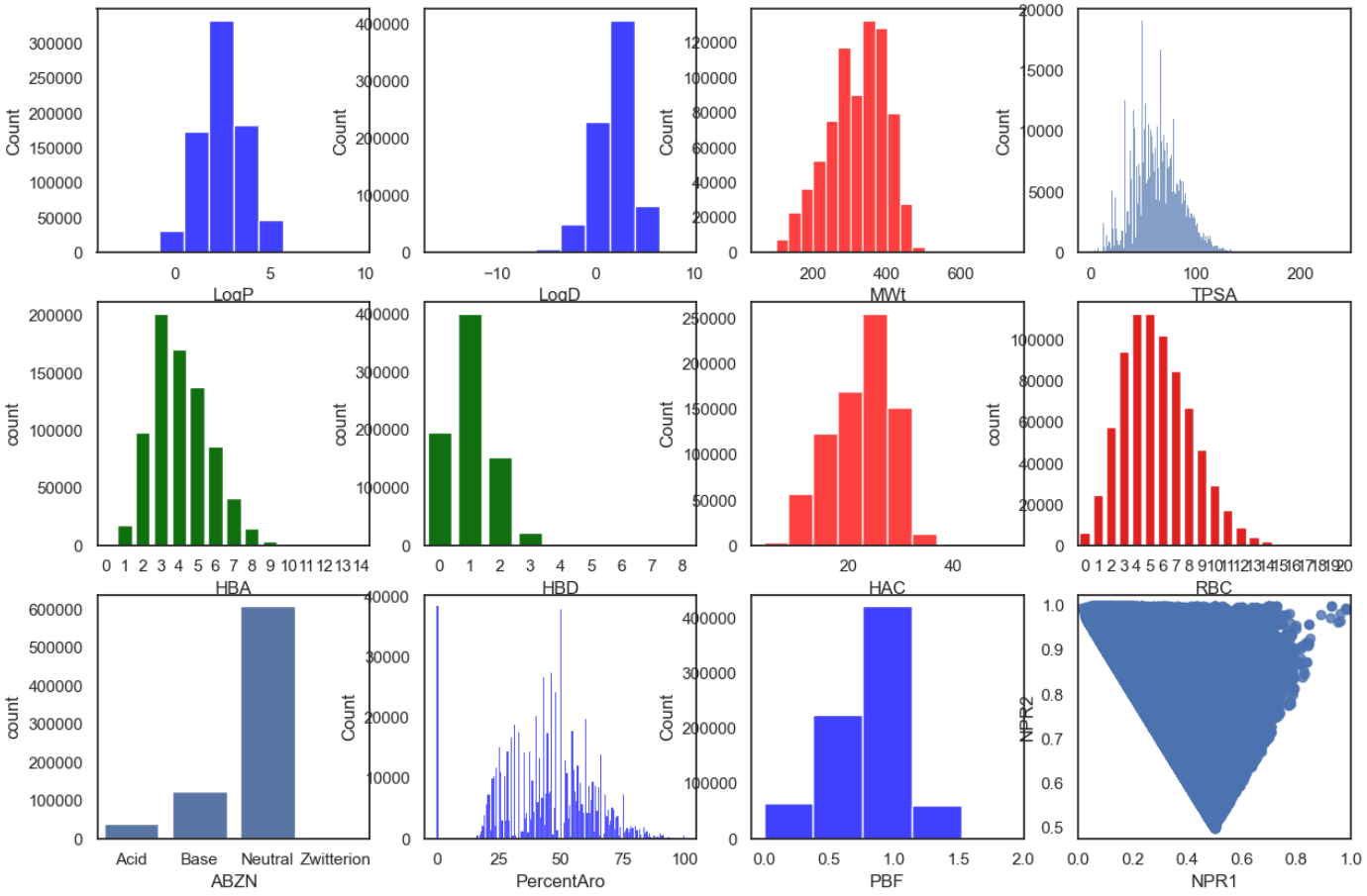
Calculated properties



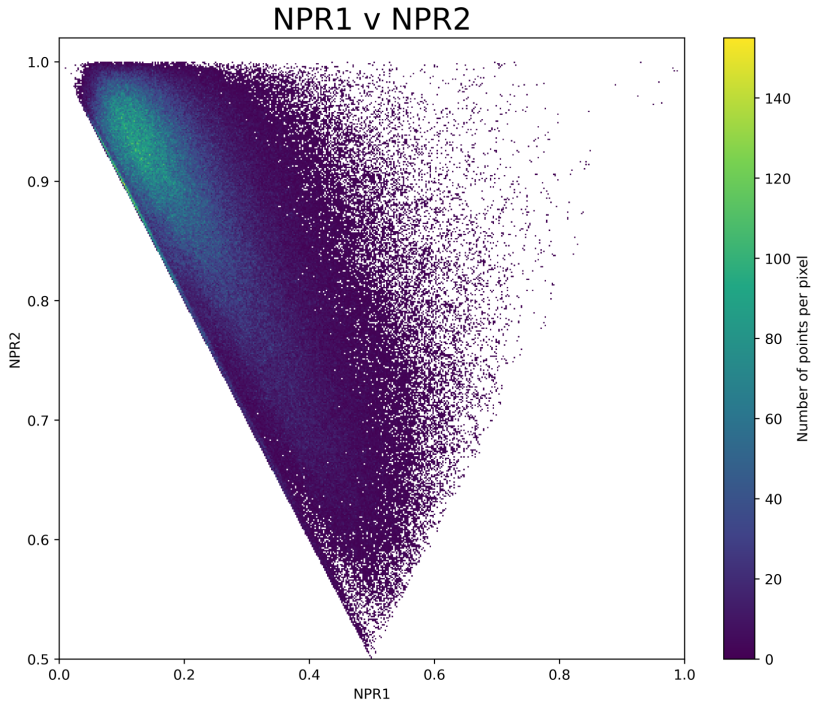
You can only test what is available

- Some papers describe the source of the screening compounds, many do not.
- Compile a database of commercially available fragments

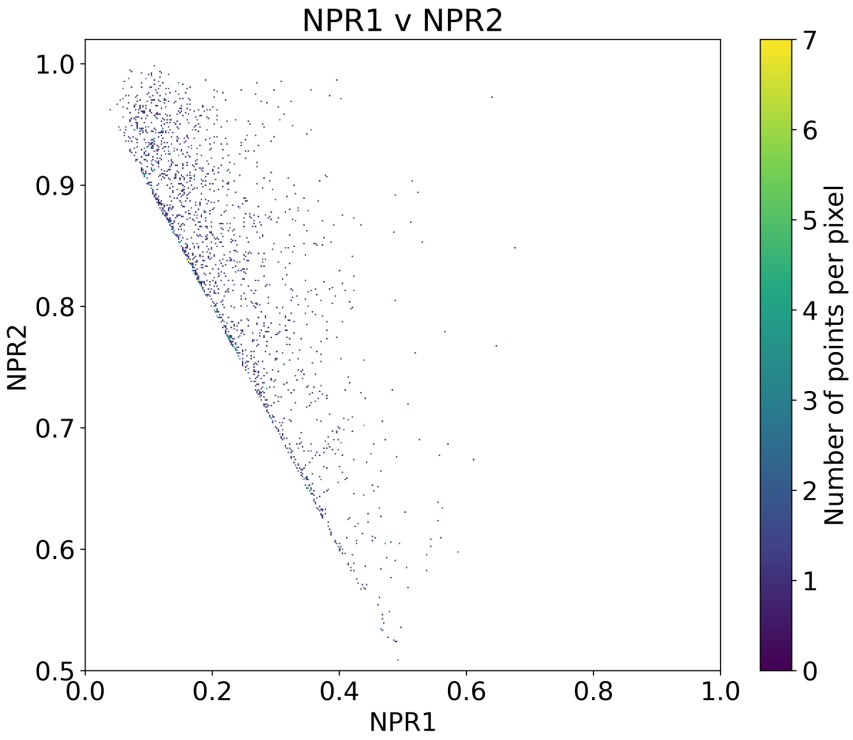
Available Fragments



Plot npr1 v npr2 heatmap



Available Fragments



Published Fragments

Conclusions

- Published fragments are lower molecular weight
- They contain a greater proportion of ionisable groups
- They contain a greater proportion of aromatics rings
- The role of increased 3D shape is unproven.

3D shape and Chirality

- Protein active sites are 3D and chiral
- 43% of phase 3 or greater molecules contain a chiral centre
  - Dataset compiled from ChEMBL, DrugBank, Clinical trials.gov
- Planer systems compromise solubility
  - https://doi.org/10.1021/jm901241e
- Increasing FP3 and chiral carbon count reduces promiscuity and CYP450 inhibition
  - https://doi.org/10.1039/C2MD20347Bcuity and Cyp450 inhibition.

So why not more 3D and chiral fragments?

- Tend to be more complicated structures
- Need to have access to both enantiomers
- Just include racemates
- Need to determine absolute configuration
- Hit expansion more challenging
- Fragment collections tend not to include them

Determination of absolute stereochemistry

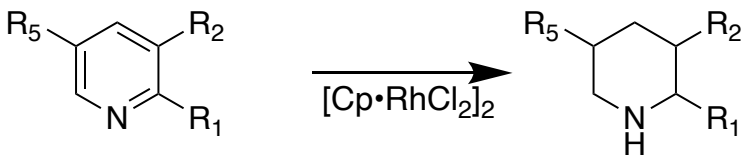
- Traditional methods
  - X-ray crystallography,
  - Comparison with known standard(s) using chiral HPLC, polarimetry
  - NMR and shift reagents.
- More recent
  - Vibrational circular dichroism (VCD)
    - https://doi.org/10.1039/D3CB00082F
    - https://doi.org/10.1366/11-06321

Enantiomeric pair covalent fragments

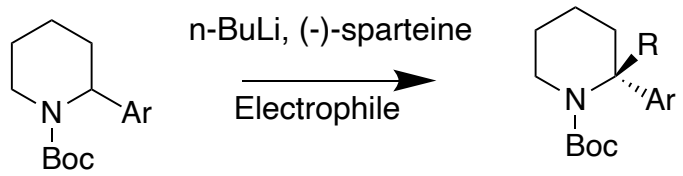
- Expedited mapping of the ligandable proteome using fully functionalized enantiomeric probe pairs
  - … identified >170 stereochemistry-dependent small molecule-protein interactions in human cells
    - https://doi.org/10.1038%2Fs41557-019-0351-5
- Using enantiomeric pairs to identify Helicase inhibitors
  - https://doi.org/10.1021/jacs.3c10581
- Rather than complicating the situation, enantiopair fragments offer an additional method to separate true actives from noise.

Increasing repertoire of asymmetric reactions

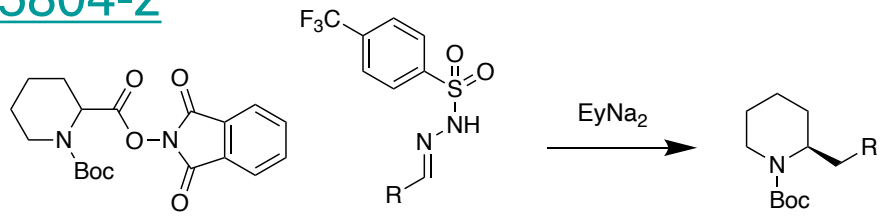
- Asymmetric hydrogenation
  - https://doi.org/10.1002/adsc.201201034



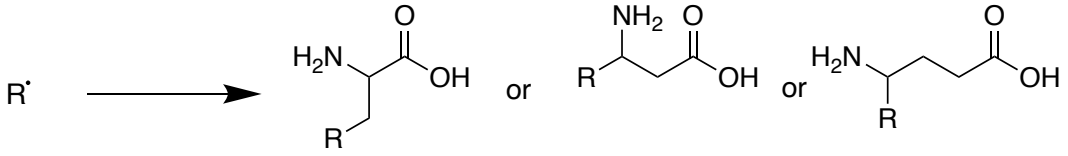
- Kinetic resolution
  - https://doi.org/10.1039/C4CC04576A



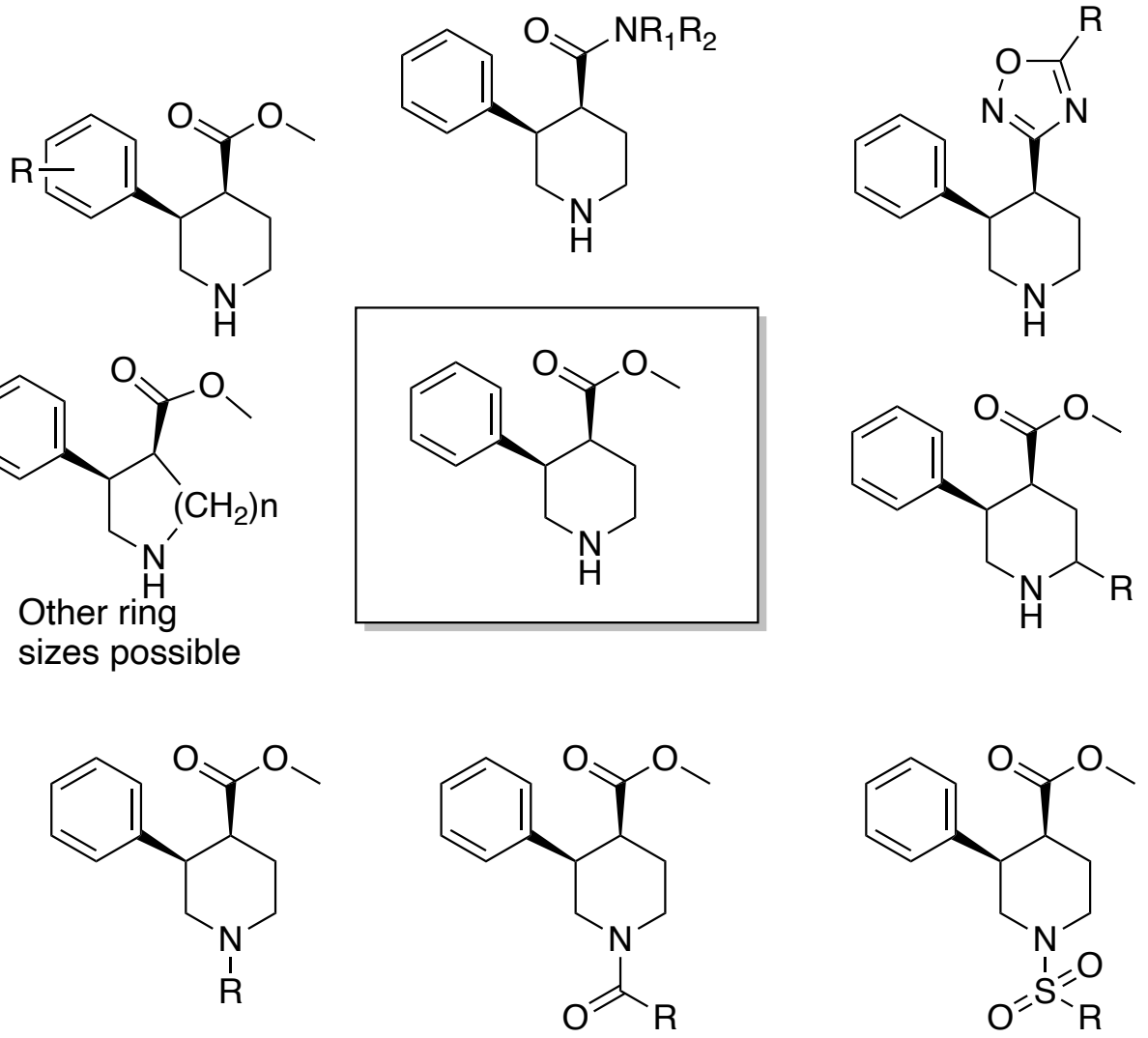
- Photochemical coupling
  - https://doi.org/10.1038/s41467-024-45804-z



- Radical mediated Amino Acid synthesis
  - https://doi.org/10.1002/adsc.202000753



Potential Hit Expansion



Liverpool ChiroChem chemically poised fragment library

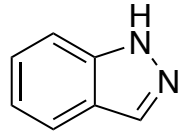
3D and chirality conclusions

- Enantiopure synthesis of chiral fragments well established
  - Robust methods for determining stereochemistry
- Enantiomeric pairs offers potential advantage in analysis of results, particularly for covalent libraries.
- Hit expansion strategies available
  - Liverpool ChiroChem poised library with automated parallel synthesis support.
- Commercial libraries are available

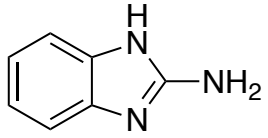
Multiple targets

- Over 240 fragment hits have been shown to be active against multiple targets.
- Whilst a few are active against similar targets (e.g. kinases), many are active against seemingly unrelated proteins.

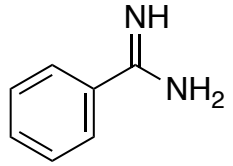
Fragments active against multiple targets



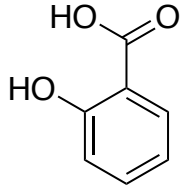
CDK2
DNA Gyrase
Dehaloperoxidase B
RadA



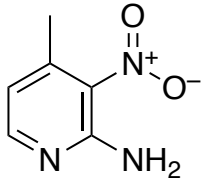
phenylethanolamine N-methyltransferase
Urokinase
Tryptase
DNA Gyrase



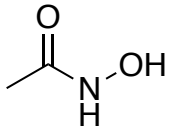
Factor Xa
Urokinase
Tryptase
Thrombin



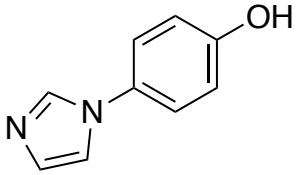
tyrosinase
GlcNAc-PI de-N-acetylase
PB1
PI3Kalpha



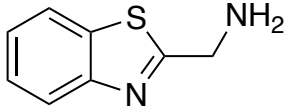
PDE10A
PDE4a
thiM
rmD in Mycobacterium abscessus



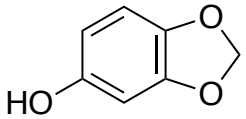
MMP-2
anthrax lethal factor
tyrosinase
Stromelysin



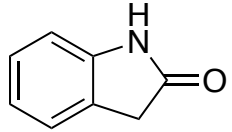
MNK1 and MNK2
UDP-3-O-acyl-N-acetylglucosamine deacetylase
Mycobacterium tuberculosis IMPDH
Inosine-5'-monophosphate dehydrogenase



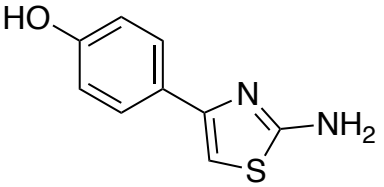
Inositol-3-phosphate synthase
ASIC3
Mycobacterium tuberculosis transaminase



Inositol-3-phosphate synthase
HIV Integrase

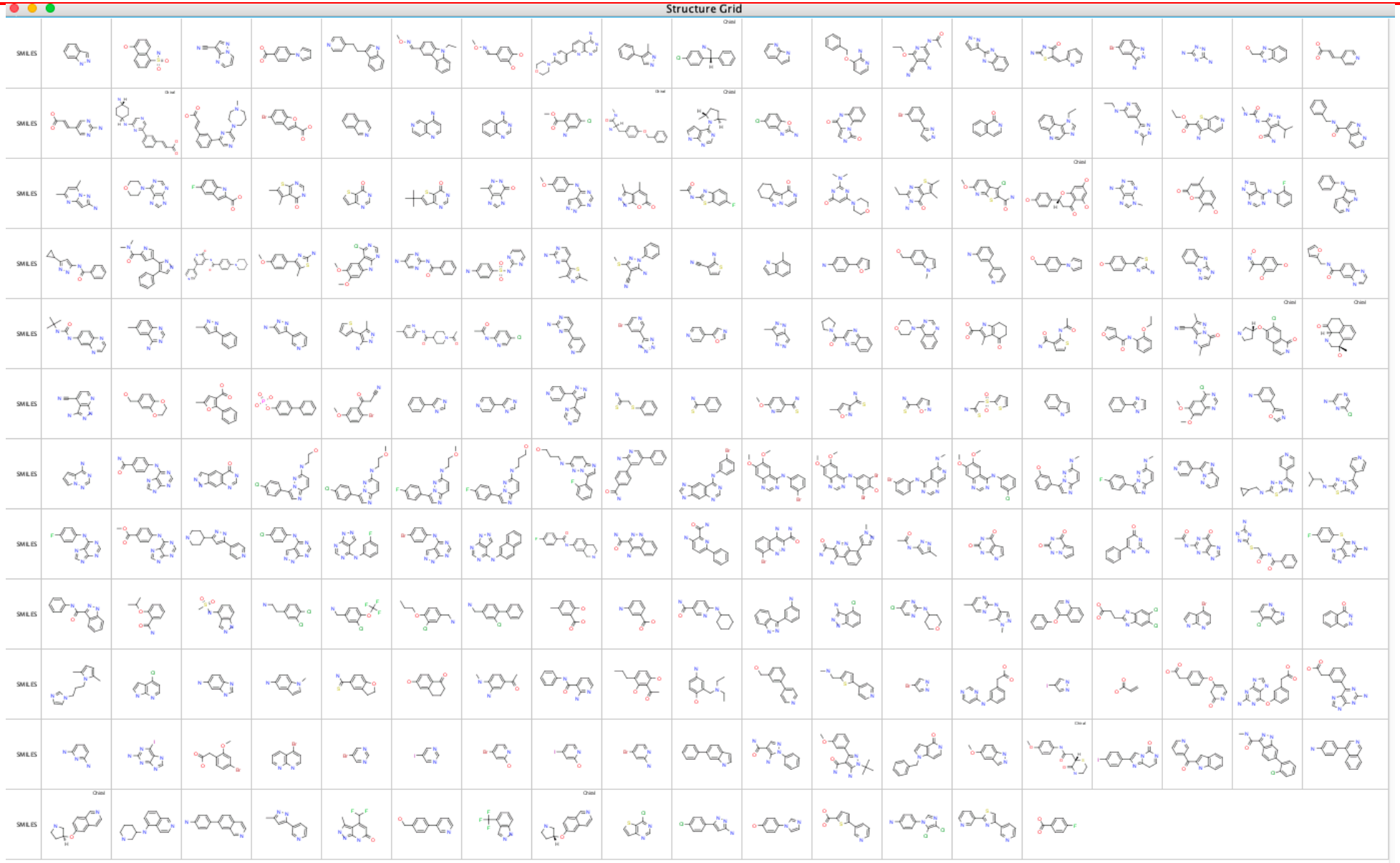


Inositol-3-phosphate synthase
DNA Gyrase
CD54



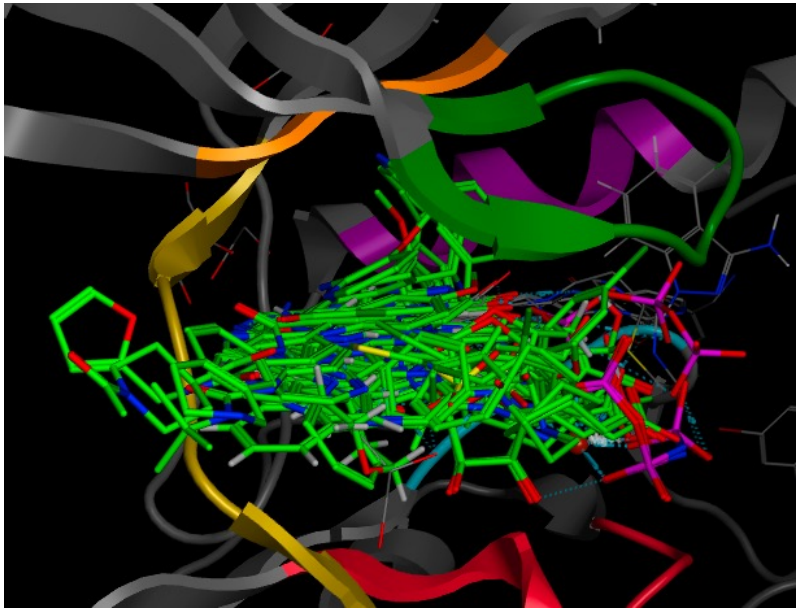
LTA4H
Trypanosoma brucei Choline Kinase
Apical Membrane Antigen 1
MNK1 and MNK2

224 Kinase Fragments



Fragments with structures in PDB

- Of 224 fragment hits identified against kinase targets, 71 are in PDB.
- Most Fragments in PDB are in active site



Conclusions

- Fragment-based screening successful against most target classes
- Technologies that also give 3D structural information dominate
- The same fragment may be a hit against multiple target types
- Opportunity to explore 3D and chiral fragments

Lockdown teaching

Lockdown teaching

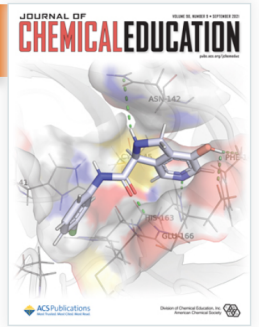
- The COVID pandemic lockdown in 2020 resulted in the closure of the UCL teaching labs
- I was approached by Alethea Tabor to design an online course on computational drug discovery for MRes students.
- Developed course using open-source tools
- Aim to allow students to design their own ligands.

Published

Molecular Docking with Open Access Software: Development of an Online Laboratory Handbook and Remote Workflow for Chemistry and Pharmacy Master's Students to Undertake Computer-Aided Drug Design

Bethanie A. Clent, Yuhang Wang, Hugh C. Britton, Frank Otto, Christopher J. Swain, Matthew H. Todd*, Jonathan D. Wilden*, and Alethea B. Tabor*

Cite this: *J. Chem. Educ.* 2021, 98, 9, 2899–2905
Publication Date: August 18, 2021
https://doi.org/10.1021/acs.jchemed.1c00289



Journal of Chemical Education

PDF (3 MB) SI Supporting Info (3) »

SUBJECTS: Drug discovery, Protein structure, Software, Structural bioinformatics, Students

Issues

- Many students unhappy using command line to install and run packages
- We ended up installing on UCL cluster and wrote Jupyter notebook as interface
- Target choices not always ideal
  - Poorly resolved PDB
  - Complex ligands
    - Difficult to understand key binding interactions
    - Multiple rotatable bonds
  - Not possible to test experimentally

Next version

- Collaborate with CCDC and University of Sussex
- Choose limited number of targets
- Choose PDB with fragment as initial hit
  - Easier to understand binding interactions
  - Limited number of conformations
  - May be possible to test experimentally by synthesis or purchase of proposed analogues
  - Maybe already made and tested?
- Or are there examples where there is already data for many analogues of initial hit already available?

Acknowledgements

- All those who published results

- Liverpool ChiroChem for collaboration on poised fragment library and covalent library.

- UCL students and staff

- And you for your attention!