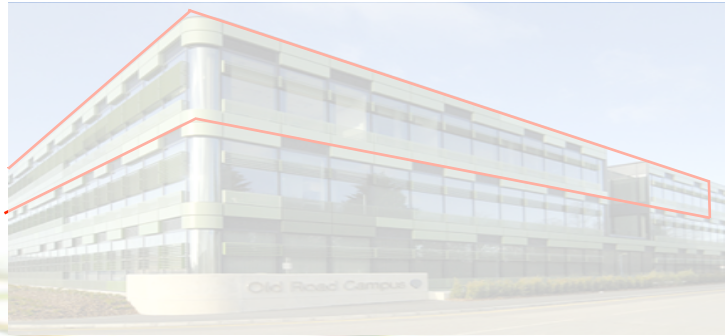
CCP CMC – SQUONK update

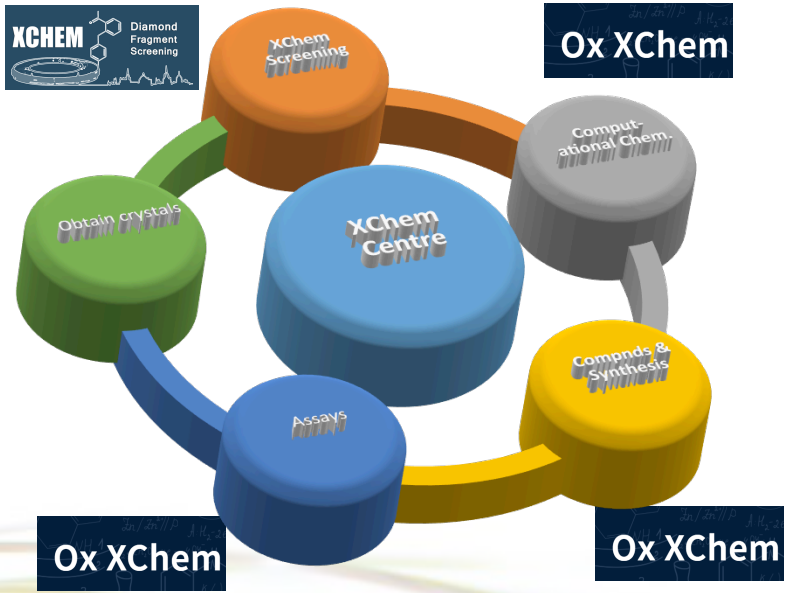
Anthony Bradley
Project Leader, Ox XChem
Chemistry and SGC Oxford, Diamond

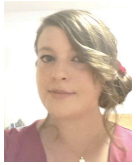


Current project

- **Aim:** Routine medicinal chemistry on hits from FBDD

Two year seed funding project to establish routine follow-up chemistry on X-ray structure hits





Rachael
Comp-chem

Anthony
Chemistry

Kannan
Biology

Computational platform
CCP CompMedChem: >20 people workshop 10th March (pharma,partners academics)
£1.3 M proposal for InnovateUK to build infrastructure

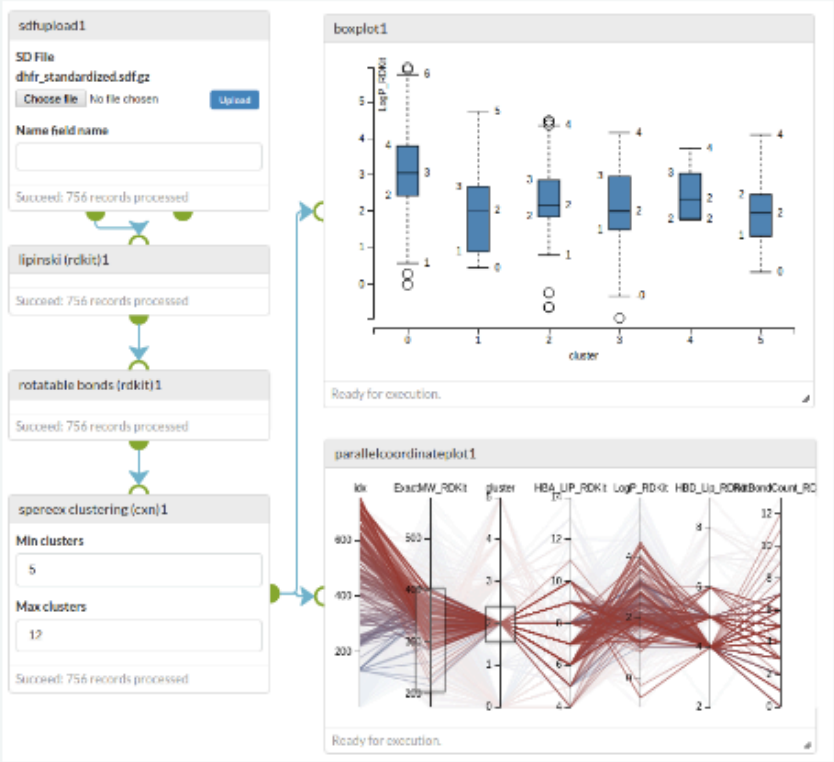
Follow-up chemistry
Oxford: 250 compounds collected from groups in University
Harwell: Lab-space at Harwell for synthesis
Collaborations: Working with academic groups (UK and China) to provide rapid follow-up chemistry
CROs: Working with CROs to improve the availability of follow-up chemistry to our users

Biology and crystallography
Oxford compounds: Screened against 4 targets
AMR: Two AMR targets screened
Dementia: Aim to screen 4 novel dementia targets – starting summer

Purpose

- Demonstrate (rapid) progress made for SQUONK project
- Encourage others to engage with open-source elements
- Source ideas for future features – ways that would encourage additions

How Squonk works



Get your work done

Squonk provides an easy to use, browser based environment that allows you to execute complex computational workflows and analyse the results.

Reproducibility and Traceability

Squonk automatically records details of your operations providing a version controlled record of what you did and allows to re-run the process in a reproducible manner.

Collaboration

Squonk allows to securely share workflows with colleages facilitating team working and collaboration.

Integration

Squonk integrates best of breed tools and data, both commercial and open source, and provides the interoperability between these allowing you to choose those that are best suited to your needs and budget.

How are we using SQUONK 1 - Projects

DCP2

Name	Owner
Enamine-D3R-DCP2	arbradley
DCP2_SPR	arbradley
DCP2 - round three	arbradley
DCP2B-x0033	arbradley
DCP2B-x0533	arbradley
DCP2B - Napatha	arbradley

NUDT7

Name	Owner
NUDT7-x0438	arbradley
NUDT7-x0399	arbradley
NUDT7-x0389	arbradley
NUDT7-x0384	arbradley
NUDT7-x0374	arbradley
NUDT7-x0129	arbradley

Score crystal poses → Enumerate reactions → Model poses

How are we using SQUONK 2 – Sharing best practice

- Create and share notebook with common process in (e.g. tautomer enumeration)

- We can upload and run through pipeline

1) No need for them to know what to do
2) No need for us to install any code anywhere

How we are adding functionality (and how you can)

Oakley Susan

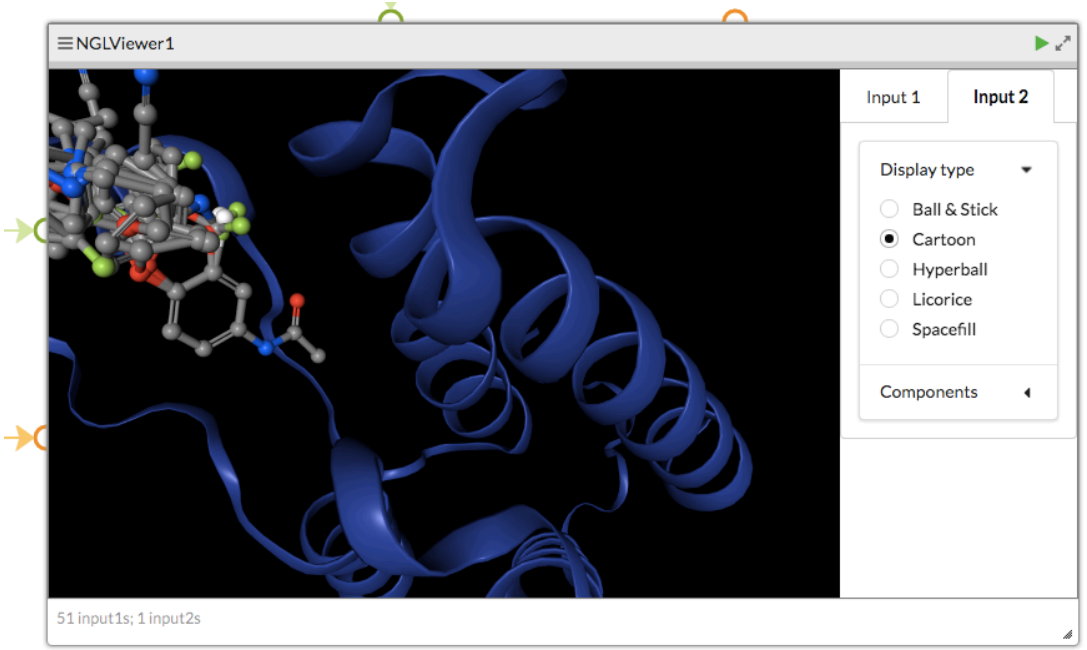
- **Pure RDKit code – from PhD projects**
  - Only need Python Script calling RDKit functionality
  - Input molecules – write out output molecules
  - Constrained conformer generator and reaction enumerator / detector
- **RDKit based packages**
  - Chose standardisers (unmet need)
  - Simple pip install (one line Docker file)
  - Compare the tool outcomes

https://github.com/flatkinson/standardiser and https://github.com/mcs07/MolVS

- **External third-party packages**
  - Open-source Docking/scoring programs (PLI and SMOG2016)
  - More complex install (ten line Docker file..) + Python / bash script
  - Trivial comparison of score values

Visualiser

- NGL viewer (RCSB PDB's viewer
- Supports two data inputs (molecules / PDBs)
- Supports data filtering
- Demonstrates SQUONK capability with JavaScript based components



Pull it together

LIVE DEMO

OR

https://www.youtube.com/watch?v=Ek1ojZiIAG0&feature=youtu.be