Science and Technology Facilities Council

Scientific Computing

diamond

STFC Cloud Speeds-Up Drug Discovery

A project that will benefit pharmaceutical development, clinicians and, ultimately, patients is using the Science and Technology Facilities Council (STFC) Cloud to enable faster, easier collaboration between researchers searching for ways to combat diseases like COVID-19.

The Challenge

On average, it takes at least ten years for a new medicine to complete the journey from initial discovery to the clinic. The average cost to research and develop each successful drug is estimated to be $2.6 billion (USD). Of the thousands and sometimes millions of compounds that may be screened, only a few will ultimately receive approval. The likelihood that a drug entering clinical testing will eventually be approved is estimated to be less than 12%.[1]

Our Approach

Drugs are generally developed iteratively, starting from initial compounds which might be identified by high-throughput screening. While this can identify starting points for the generation of drug leads, they are frequently difficult to optimise, or to rationalise the drivers for their potency.

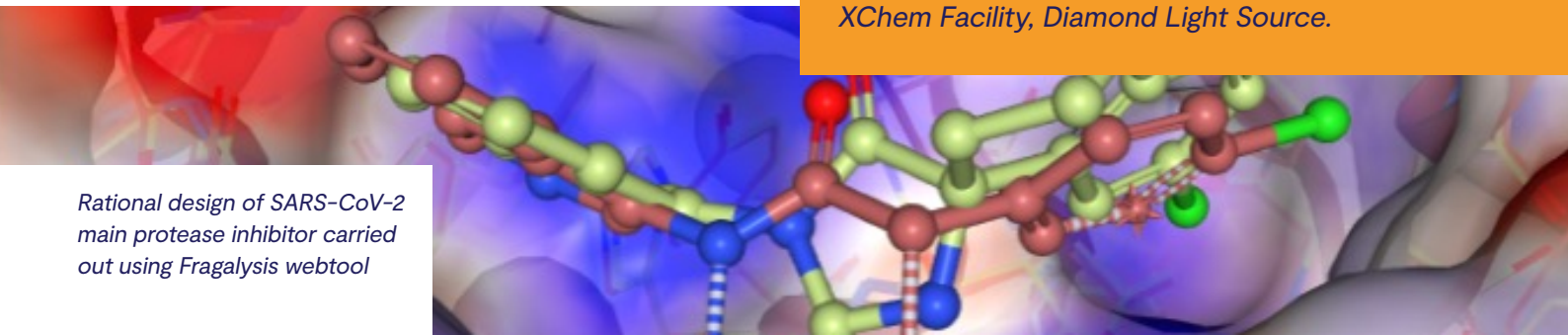
Scientists at the Diamond Light Source's XChem Facility use an alternative approach known as Fragment-Based Drug Discovery (FBDD). This involves screening with compounds much smaller than typical drug molecules (hence 'fragment') to assess which chemical features are useful when interacting with the protein being targeted. These features can then be expanded or combined into more complex molecules that will have the potency required for drug-like action.

XChem uses X-ray crystallography to visualise, as a detailed 3D structure, precisely how the drug molecule binds with the protein. Thousands of fragments can be screened in a week, producing dozens or even hundreds of starting points, each with the potential of reducing the time and cost of getting drugs into clinical trials.

Facilities like XChem have given rise to an unprecedented increase in data describing potential drug molecule and protein combinations. To cope with the vast data output, and to explore it in the context of FBDD, the XChem team designed an opensource web-based application. They called it **Fragalysis**.[2]

"Having Fragalysis served from the STFC cloud has been nothing short of sensational in terms of what it has allowed us to do scientifically. We believe Fragalysis has the potential to provide a step-change in how scientists think about fragment-based drug discovery, and how they collaborate on projects and share data with the rest of the world. If we manage to achieve this, then this will hopefully lead to much shorter and cheaper timescales for drugs getting into the clinic. This will ultimately have individual benefits for patients and collective benefits for society."

Dr Rachael Skyner
XChem Facility, Diamond Light Source.



Rational design of SARS-CoV-2 main protease inhibitor carried out using Fragalysis webtool

STFC Cloud

This is where the STFC Cloud comes in. It is a dedicated cloud infrastructure which provides access to compute resources for users across STFC and partner organisations. Developed and managed by the Scientific Computing Department, it enables users to perform complex data analysis as and when required, without the overheads of running their own computing infrastructure.

The Benefits

The STFC Cloud hosts Fragalysis and has enabled multiple projects to work in a much more collaborative environment, and at a much faster pace. One such project is **COVID Moonshot**[3], an international consortium of scientists drawn from academia, biotech companies, contract research organisations and pharma companies, all working pro-bono or at cost, to rapidly develop easy-tomanufacture antivirals that can treat COVID-19 by inhibiting the main protease, one of the enzymes involved in the viral life cycle. The project was kickstarted from a massive fragment screen of the SARS-CoV-2 main protease carried out at XChem by Diamond researchers in early 2020 as the pandemic was evolving[4].

The Cloud platform has enabled the sharing of data in almost real-time, allowing the structures to be released as they are solved. This has allowed hundreds of contributors from across the globe to rapidly consider new designs for molecules and has contributed to an extremely swift turn-around for optimization cycles. They expect to progress some of the molecules designed using Fragalysis to become preclinical candidates[5]. Instead of taking many years, the drug candidates from the COVID Moonshot have been developed in less than 12 months[6].

(1) www.phrma.org

(2) https://fragalysis.diamond.ac.uk

(3)https://www.nature.com/articles/d41586-021-01571-1

(4)https://pubmed.ncbi.nlm.nih.gov/33028810/

(5) https://www.diamond.ac.uk/Home/News/LatestNews/2021/27-09-21.html

(6)https://www.biorxiv.org/content/10.1101/2020.10.29.339317v3

"We couldn't have carried out this project as effectively without the STFC Cloud. Hosting Fragalysis elsewhere, with the computational resource it needs to keep the cogs turning, would have come at a considerable cost. The STFC Cloud ensures it has a long-term home."

Dr Rachael Skyner
XChem Facility, Diamond Light Source.

The Fragalysis project benefitted from an STFC Cloud allocation through IRIS (www.iris.ac.uk).

IRIS helps develop and grow the digital research infrastructure that allows STFC to continue to play a leading role in world class science.

Additional funding was provided by the Ada Lovelace Centre.