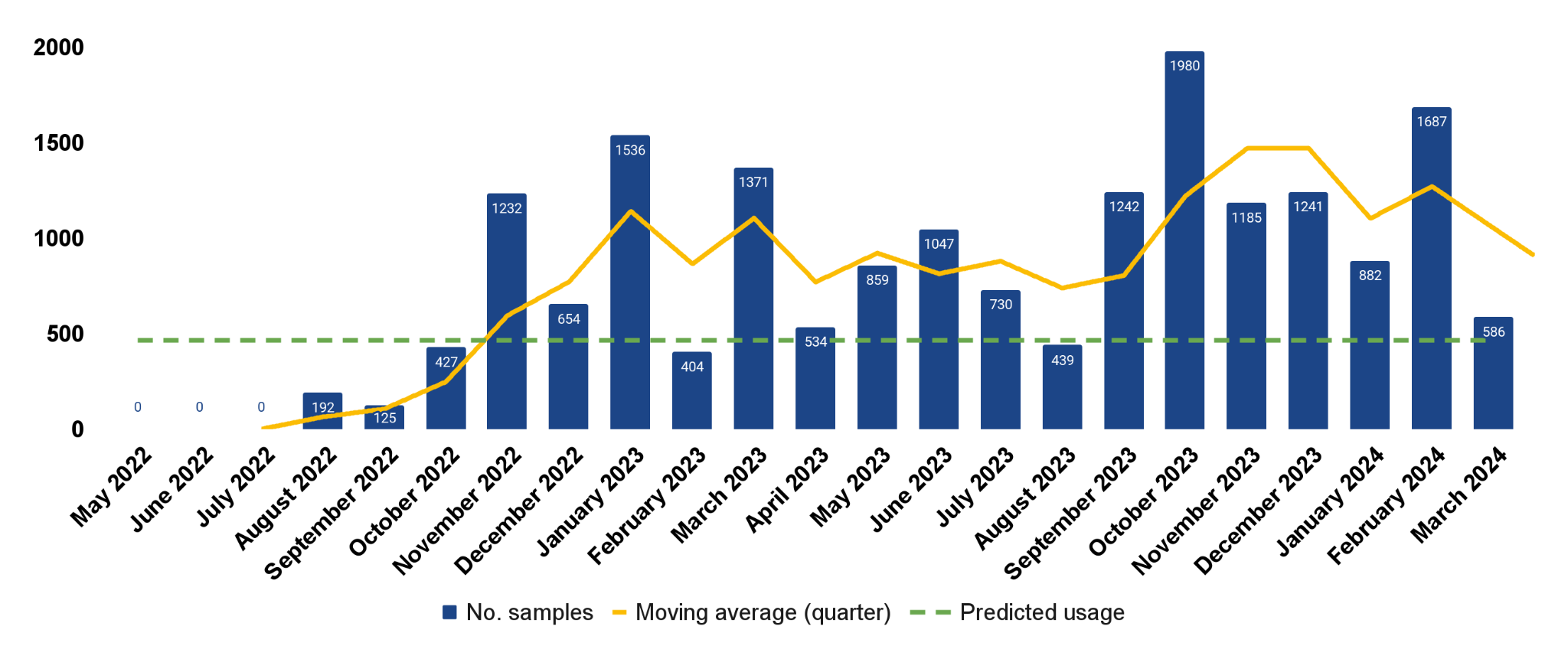
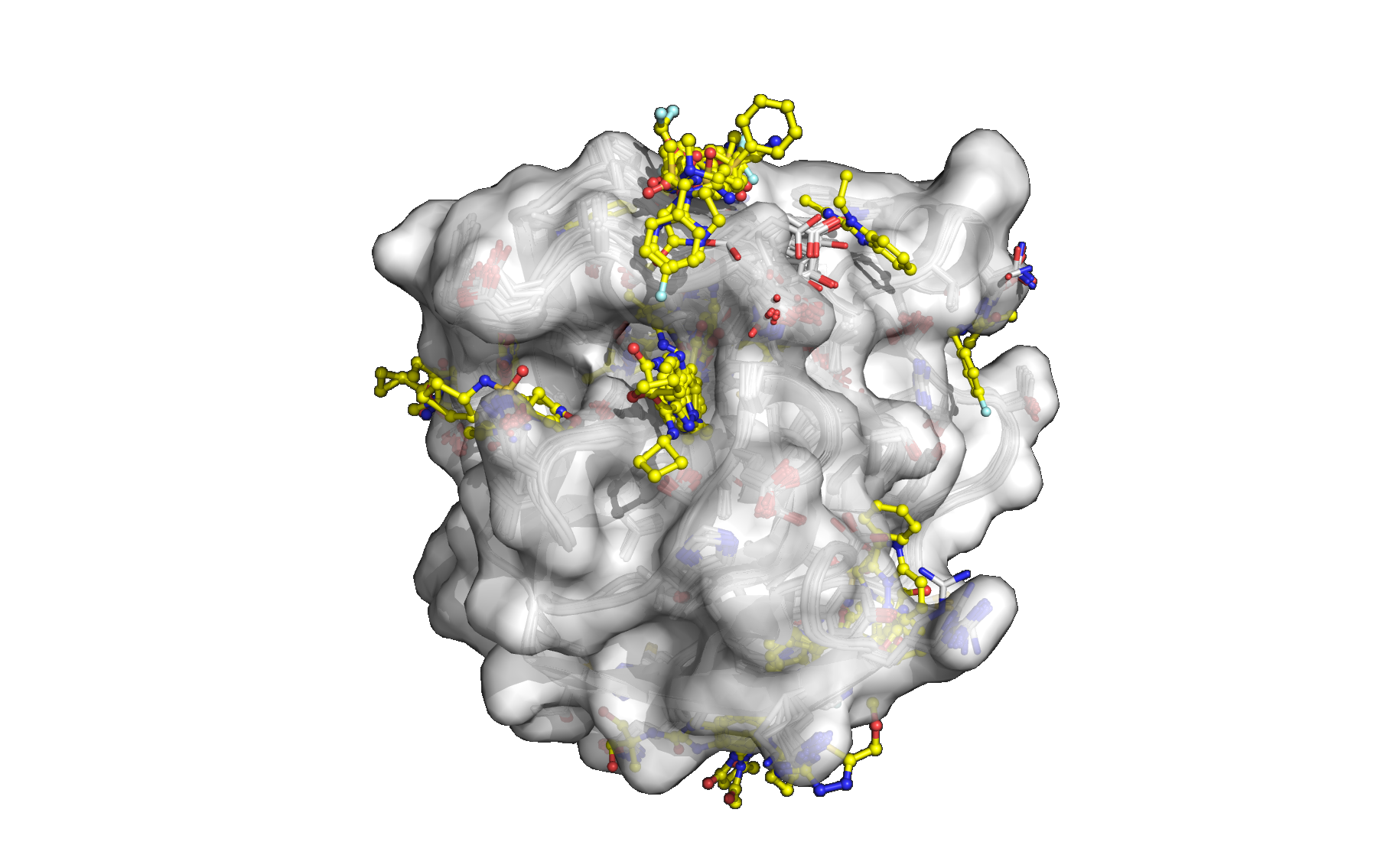
**B. Studies and Results**

Significant accomplishments to date include:

* Unprecedented productivity - sample throughput has been ramped up to match scientific requirements with 12,568 datasets collected in 2023 alone.
* In support of Project 2: completed fragment screens for 5 TEPs, optimized crystals for 16 TEPs, and solved initial structures for 15 TEPs.
* In support of Project 3: performed extensive experimental exploration in order to generate co-crystal structures of lead compounds.
* In support of Project 4: performed extensive experimental exploration in order to generate co-crystal structures of covalent lead compounds.
* In support of Project 5: provided structural enablement for 2 protein targets during lead optimization.
* Deposited 601 structures PDB in 2023, accounting for over 6% of all X-ray structures deposited.
* 91 further structures have been deposited in Q1 of 2024 with 190 pending deposition.
* Implementation of software pipelines to incorporate full ligand information and restraints during X-ray structure dissemination in the PDB.
* Integration of structural biology core data into CDD with protocols in place for all active projects.
* Fast and automated fragment finding and building with PanDDA2.
* Auto-analysis of sites and binding modes with XChemAlign.
* New release of Fragalysis Cloud - a web platform for collaboration and progression.
* Experimental tracking with automated import of all SBCore data to Scarab LIMS and ELN.
* Presented talks & posters describing ASAP work at >20 international locations.

****

**Support of P2**

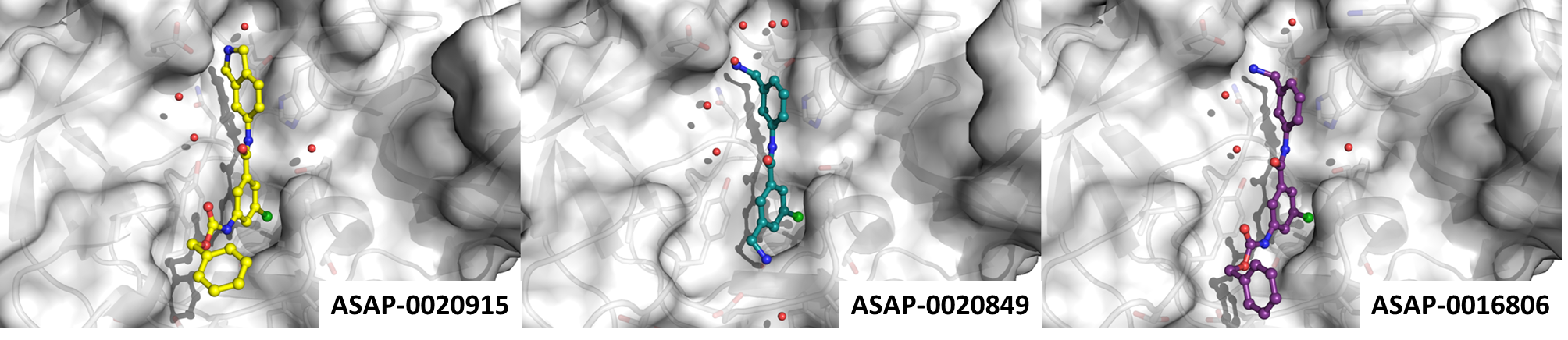
The crystallographic fragment screen of enterovirus 2A protease identified 

a total of 102 fragments with 44 fragments directly binding in the protease active site. The crystal structures of the bound fragments have been openly disseminated via the web platform Fragalysis (<https://fragalysis.diamond.ac.uk/viewer/react/preview/target/A71EV2A>) and the Protein Data Bank (group deposition ID: G\_1002288). Preliminary follow-up work has been initiated with algorithmically designed compounds from P2.

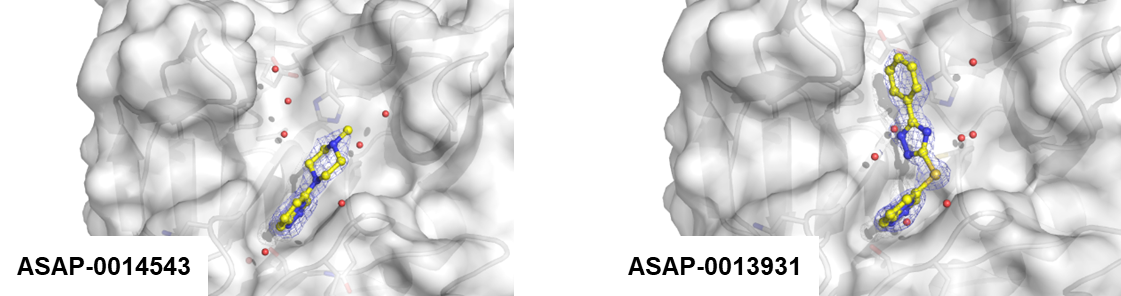
Screening of crystallization conditions for the Zika NS5 RdRp catalytic domain identified robust crystals which diffracted better than 2.5 Å resolution. Large-scale crystallization was carried out and crystallographic fragment screening performed with data analysis currently in progress.

**Support of P3**

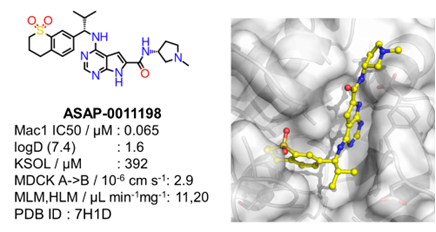
Following the successful crystallographic fragment screens completed in Year 1, the Structural Biology Core supports P3 drive the hit-to-lead process forward for Flavi-virus NS2B3-NS3 and Enterovirus D68-3C protease work by routinely releasing new ligand-bound crystal structures.



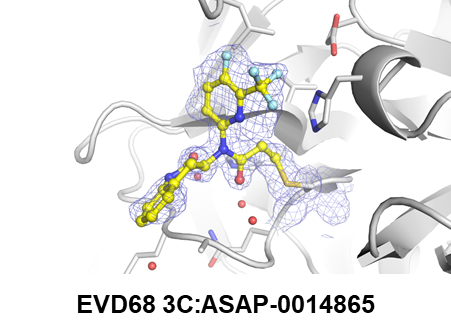
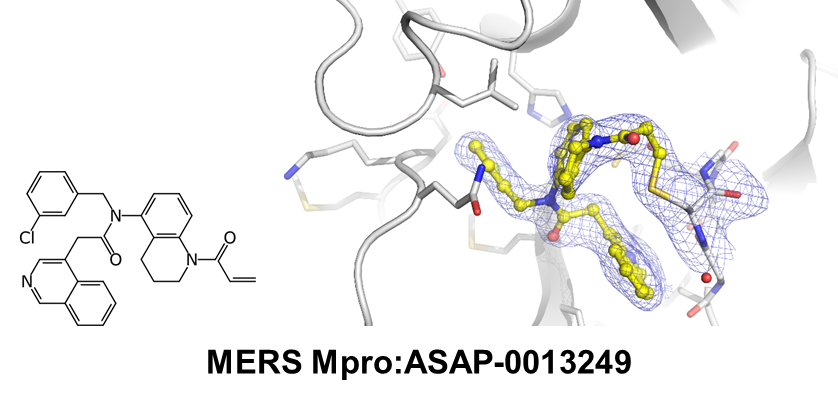
To date, the SB core has tested 717 unique follow-up compounds in Zika NS2B3-NS3 and collected over 1400 datasets, producing around 60 ligand-bound structures. Extra effort is being applied to achieve co-crystallization of DENV2 NS2B3-NS3 with potent small molecules due to lack of suitable soaking crystal system for this homologue. Structural data has been disseminated via Fragalysis (https://fragalysis.diamond.ac.uk/viewer/react/preview/direct/target/XX01ZVNS2B/) and deposited in the PDB (group deposition).

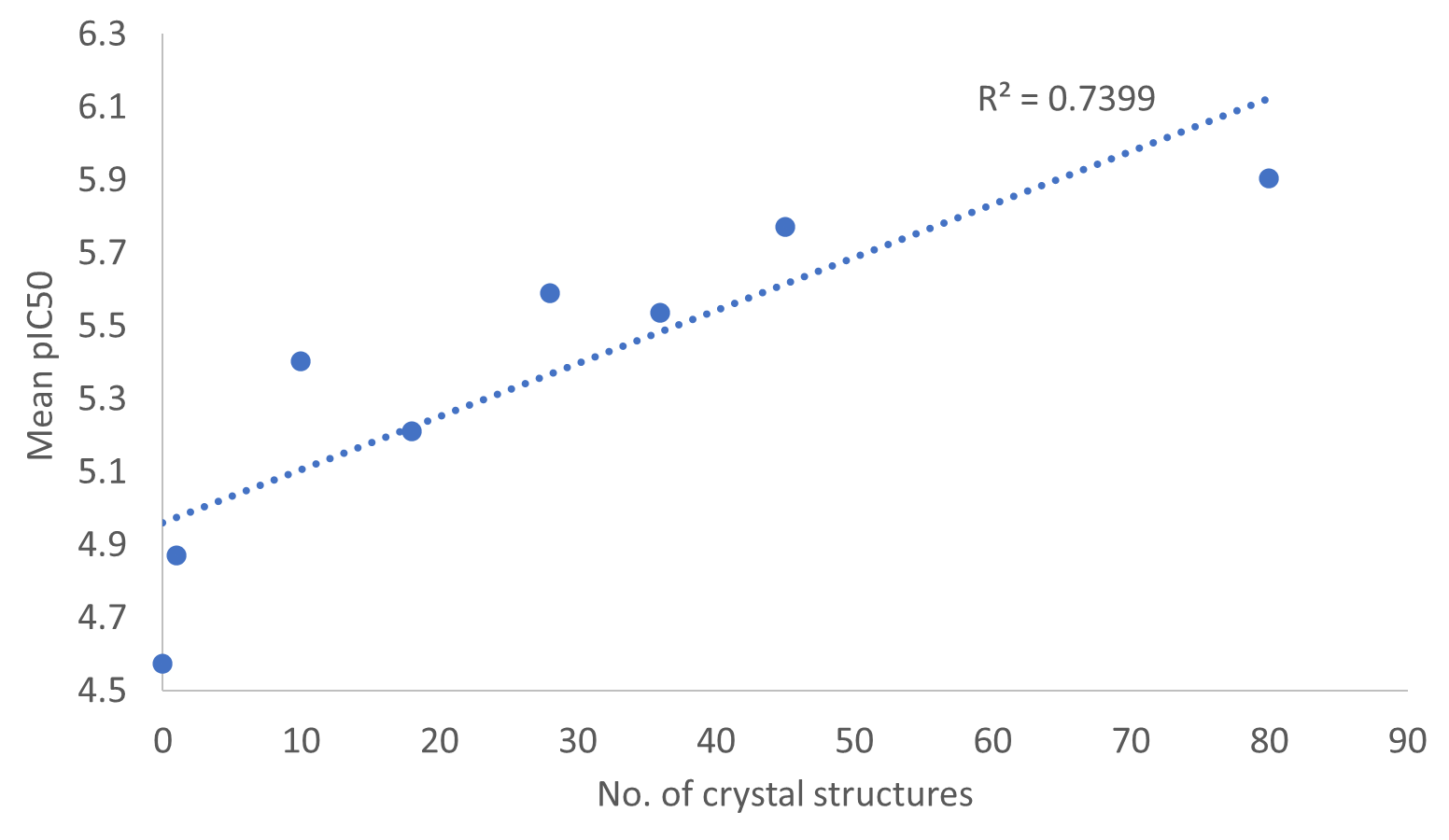
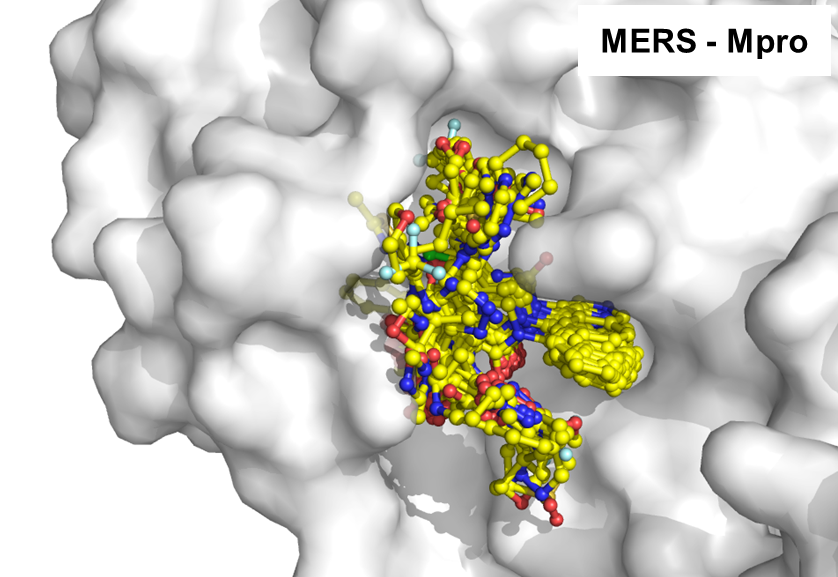


For Enterovirus D68-3C the SB core has identified multiple crystal forms and provided an orthogonal readout to the biochemical assays and structural conformation of P3 and P4 compounds through collection of 1,989 datasets and 1517 unique follow-up compounds. All fragment-bound structures have been deposited in the PDB and structural data is regularly being disseminated via Fragalysis (<https://fragalysis-legacy.xchem.diamond.ac.uk/viewer/react/preview/target/D68EV3CPROA>).

Regarding the development of SARS-CoV-2 nsp3 macrodomain antivirals, the Structural Biology Core was crucial in supporting P3 efforts by weekly releasing new liganded structures. A new crystal system (space group P21) was successfully established that allowed co-crystallisation of the most potent compounds from P3. These compounds were shown to be occluded from binding to the adenosine binding site of the previously used crystal system (space group P43) due to clashes with crystal contacts. Overall, the SB core was soaking and co-crystallising 871 unique compounds and collected 2,151 datasets leading to 90 liganded crystal structures being deposited in the PDB (Group Deposition: G\_1002283).

**Support of P4**

This core generates X-ray structures of new protein-ligand complexes for ligands synthesized by Project 4 to aid in structure-based design, which also aids in determining whether successful covalent targeting is achieved. **  
Support of P5**



The Structural Biology Core support for P5 has primarily focused on assisting the development of a novel series of compounds that demonstrate potent, broad spectrum activity against SARS-CoV-2 and MERS-CoV. Our expertise in MERS-CoV Mpro structural biology in particular is driving the program towards in vivo efficacy testing. 466 SARS-CoV-2 ligand bound crystal structures from the COVID Moonshot were deposited in the PDB in 2023 with ASAP Mpro structure deposition pending disclosure of key compounds.

**Technologies**

**A collage of images of a cell

Description automatically generated**

We have revisited the original PanDDA formalism, firstly to understand and correct its failures and weaknesses, and secondly to implement the fully automated density modelling that had appeared feasible once its power for signal improvement had become evident.This work led to PanDDA2 (https://github.com/ConorFWild/pandda\_2\_gemmi), a reimplementation of the original concepts, refactored to (1) bring about dramatic speed-ups; (2) implement pre-clustering of datasets according to real-space variation, as an alternative to the reciprocal space clustering method developed by Ginn (<https://doi.org/10.1107/S2059798320012619>); (3) introduce fast ranking of putative hits using shape recognition through machine-learning of historic XChem data; (4) achieve fast auto-building using a new statistically robust scoring function; and (5) robustly test the package on historic XChem data.

A close-up of a diagram

Description automatically generated

To fully understand the various binding sites revealed by a fragment screen, and to understand packing or model quality artefacts, requires the ensemble of fragments and literature structures to be placed in common reference frames, spanning diverse asymmetric units, crystal forms and binding site conformations. This includes moving electron density, protein neighbourhood and other packing artefacts to the various reference frames. This task is algorithmically complex, but now implemented in XChemAlign, which heuristically infers the various relationships and variants, preparing them for curation and finally upload to Fragalysis.

We have released a new version of Fragalysis Cloud (http://fragalysis.diamond.ac.uk), which publicly hosts all XChem’s collaborative fragment results. It is fragment-specific, encompassing analysis, progression, collaboration and dissemination. Important features include:

* Fragment structures are pre-curated, so that fragment binding can be directly analyzed and compared.
* Views of the data can be shared as URLs (snapshots) that precisely reconstitute what the user sees, and thus streamline collaboration. This feature was hardened in the COVID Moonshot, where it was heavily used to assist cycles of compound design in a globally distributed team.
* Compute-heavy algorithms can be simply executed, and results reviewed and shared, by pre-configured GUI tools. This provides a vehicle for hosting cutting-edge published algorithms, deployed and hardened in collaboration with their authors.
* Primary and secondary data is disseminated by FAIR mechanisms, in downloadable zip archives that include self-documenting PDFs and snapshot URLs to explain the context and history of the download.

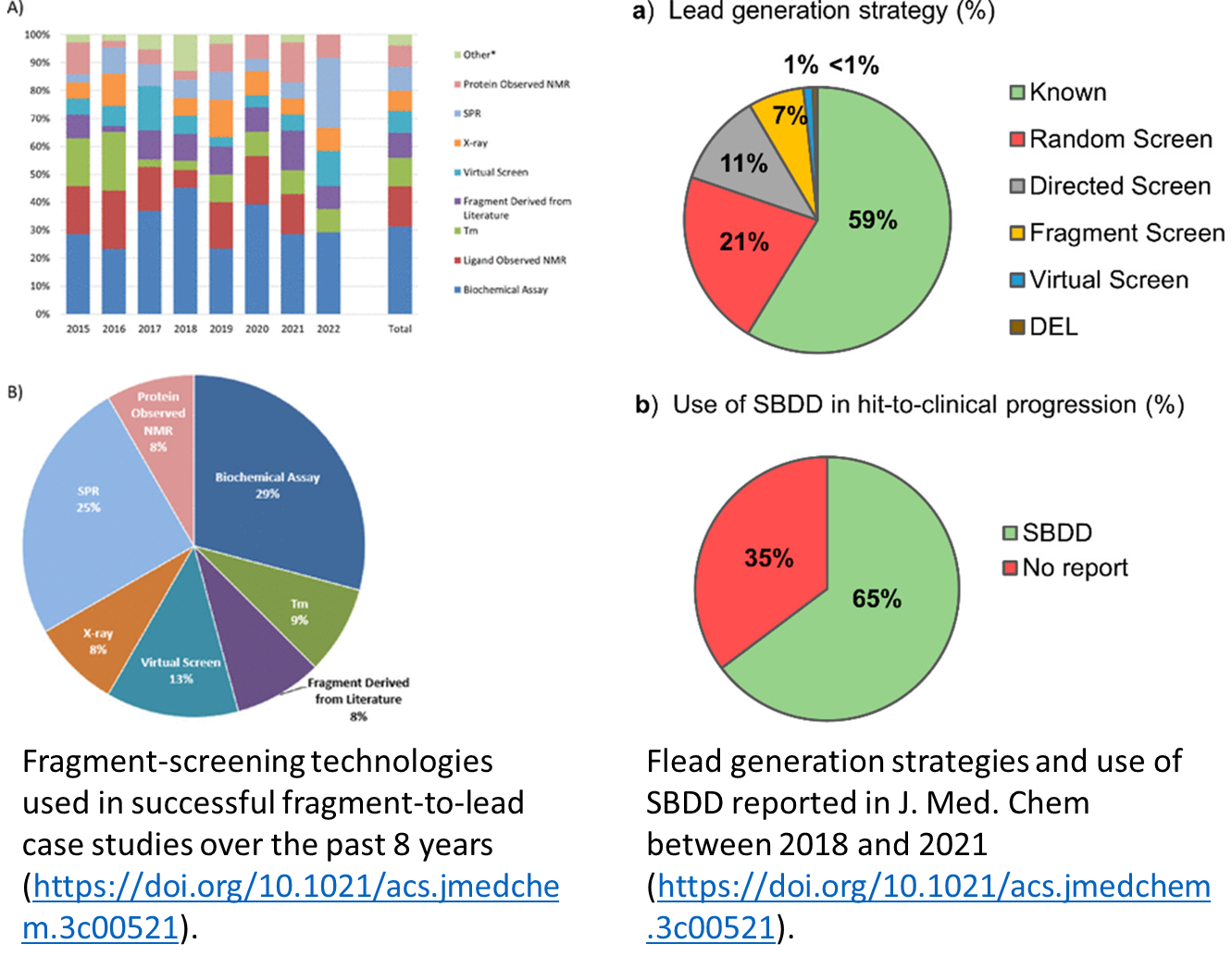
**Other achievements**

Other major results and outputs from this Core are listed in Significant Project Generated Resources and have been posted online.

| **Target** | **Project supported** | **No. crystals mounted** | **No. datasets collected** | **No. compounds soaked** | **No. structures released** | **No. PDB depositions** |
| --- | --- | --- | --- | --- | --- | --- |
| **MERS Mpro** | P4 & P5 | 2,357 | 1,696 | 446 | 187 | 0 |
| **SARS-CoV-2 Mpro** | P4 & P5 | 1,490 | 1,085 | 698 | 198 | 466 \* |
| **SARS-CoV-2 Mac1** | P3 | 2,373 | 2,151 | 871 | 91 | 90 |
| **EV 3C protease** | P3 & P4 | 3,838 | 3,511 | 3,063 | 179 | 98 |
| **EV 2A protease** | P2 | 1,509 | 1,260 | 1,308 | 85 | 85 |
| **Flavi NS2BNS3** | P3 | 3,014 | 2,541 | 2,395 | 131 | 104 |
| **SARS-CoV-2 N protein** | P3 | 679 | 583 | 316 | 0 | 0 |
| **Zika NS5 RdRp** | P2 | 1,262 | 1,104 | 918 | Analysis ongoing | 0 |
| **Zika Ns3 helicase** | P2 | 1,567 | 1,204 | 931 | 46 | 46 |
| **Total** | P2/P3/P4/P5 | 18,089 | 15,135 | 6,686 | 917 | 895 |

*Summary of all data collected by Structural Biology Core to date (19th April 2024) (\*Including structures deposited from COVID Moonshot)*

**C. Significance**

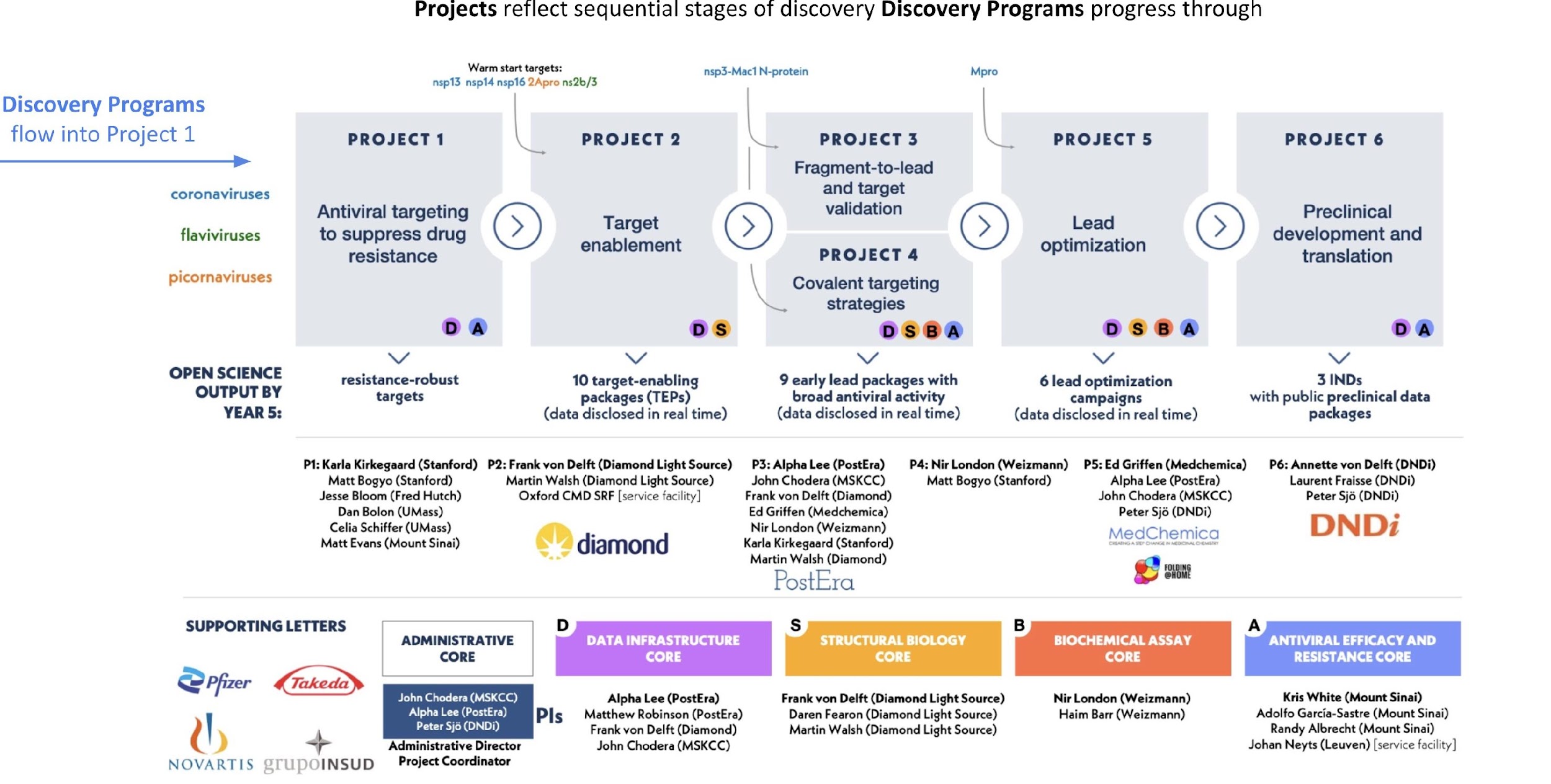


Crystallographic fragment screening is one of many techniques used for fragment hit identification, accounting for 8% of examples of hit-to-lead projects published in J. Med. Chem. in 2022 (<https://doi.org/10.1021/acs.jmedchem.3c02070>). A recent analysis of 18 successful fragment-to-lead case studies in 2022 reported that 83% (15/18) generated a structure for the fragment, 78% (14/18) generated a structure for the lead and 100% (18/18) of entries used structural information during lead optimization.

Furthermore, in 156 hit-to-clinical campaigns published in J. Med. Chem. between 2018 and 2021 65% of campaigns reported used Structure-Based Drug Design as a key enabling technology. This increased to 100% when starting with fragment hits (<https://doi.org/10.1021/acs.jmedchem.3c00521>).

This illustrates that structural enablement is critical for efficient hit-to-lead development and lead optimization.

ASAP utilizes structure-enabled drug discovery processes and the structural biology core is essential in providing the capability of using high-throughput X-ray crystallography to structurally enable antiviral targets pursued by its discovery programs, providing X-ray structures of apo and liganded targets to support Projects 2, 3, 4, and 5.

**

Additionally, rapid public structural data release acts as a public force multiplier to reduce costs to discovery as was demonstrated by the discovery of the SARS-CoV-2 Main protease inhibitor Ensitrelvir (<https://doi.org/10.1021/acs.jmedchem.2c00117>)).

**D. Plans**

Plans for the next project period include:

* Finalize software infrastructure and Fragalysis upgrades for rapid data-to-dissemination
* For Project 2: Target Enablement, deliver fragment screens for 2 TEPs
* For Project 3: Fragment-to-lead, deliver fast turn-around of protein-ligand structures for 5 hit-to-lead projects
* For Project 5: Lead Optimization, deliver fast turn-around of protein-inhibitor structures for 3 hit-to-lead projects