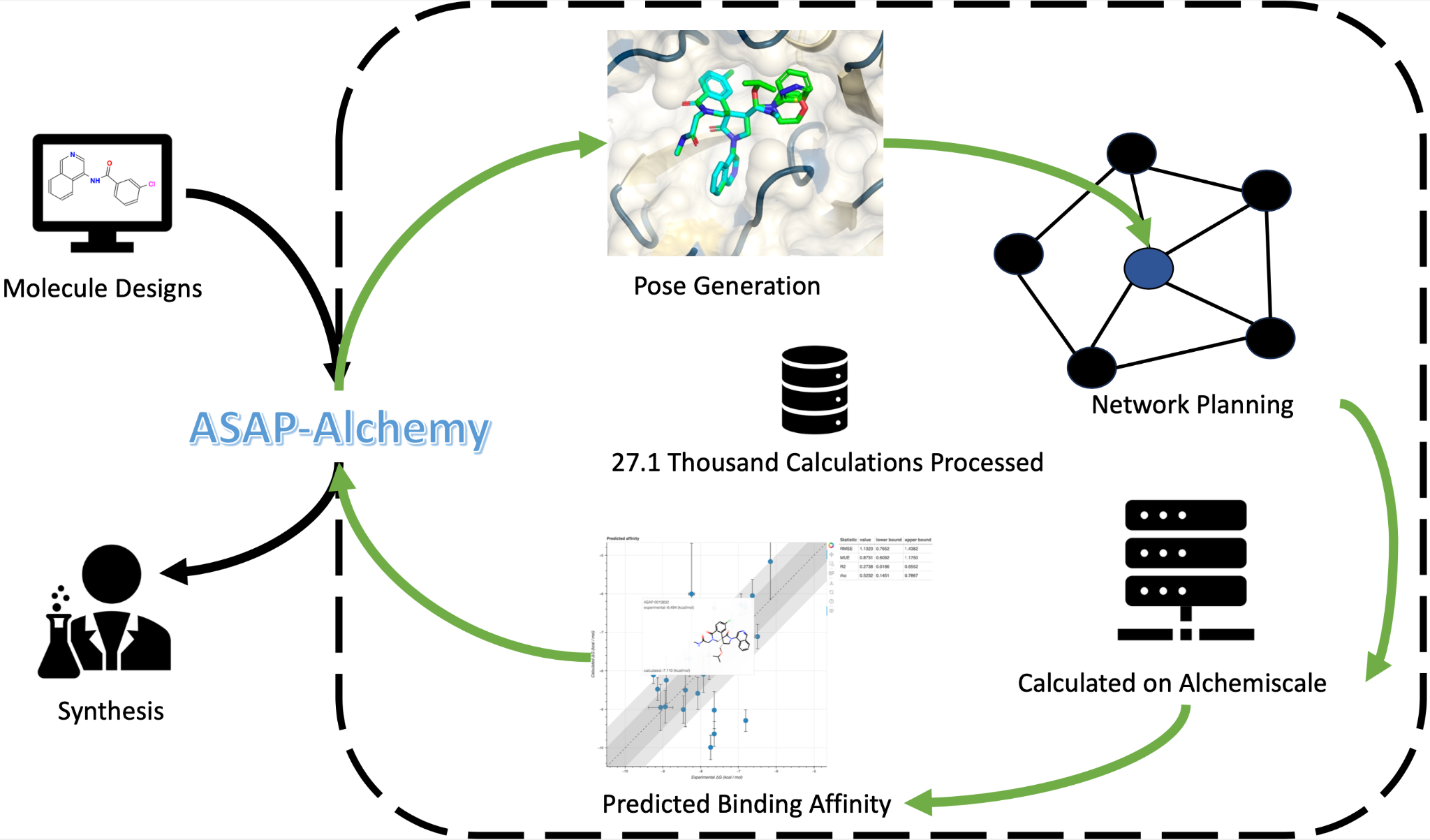
**Significance of the work**

The AI-driven Structure-Enabled Antiviral Platform (ASAP) employs state-of-the-art computational free energy calculations to predict binding affinities between a lead compound and its therapeutic target to efficiently use synthetic resources. While free energy calculations provide a formally rigorous protocol for extracting binding free energy data from the dynamics of candidate molecules in complex with their therapeutic target, they are only useful in guiding synthesis if sufficiently accurate (for example, <1 kcal/mol r.m.s. error). In practice, accuracy is limited by the general molecular mechanics force fields used to describe the dynamics and interactions. This project aims to improve the accuracy of these atomistic models by deriving molecule-specific parameters from high-level quantum mechanical (or machine learning) reference data on the fly providing more robust and accurate estimates of ligand binding affinity.

**Significant project-generated resources**

An automated end-to-end open-source free energy workflow ([ASAP-Alchemy](https://github.com/choderalab/asapdiscovery/tree/main/asapdiscovery-alchemy)) has been created, **enabling 27.1 thousand calculations of protein-ligand binding free energies** to date.



**Figure 1:**  ASAP-Alchemy improves the reliability and scale at which protein-ligand binding free energies can be estimated with its automated end-to-end workflow.

This includes the generation of the initial ligand-protein complex coordinates, planning of the optimal alchemical network, submission and execution on HPCs, estimation of the final binding affinities and automatic distribution of the results with the med chem team. The automated pipeline has significantly increased throughput and will aid adoption by other cores and targets due to ease of use. The workflow is expected to be a valuable addition to the computational chemistry community as its general design and end-to-end nature make it applicable to any drug discovery campaign.

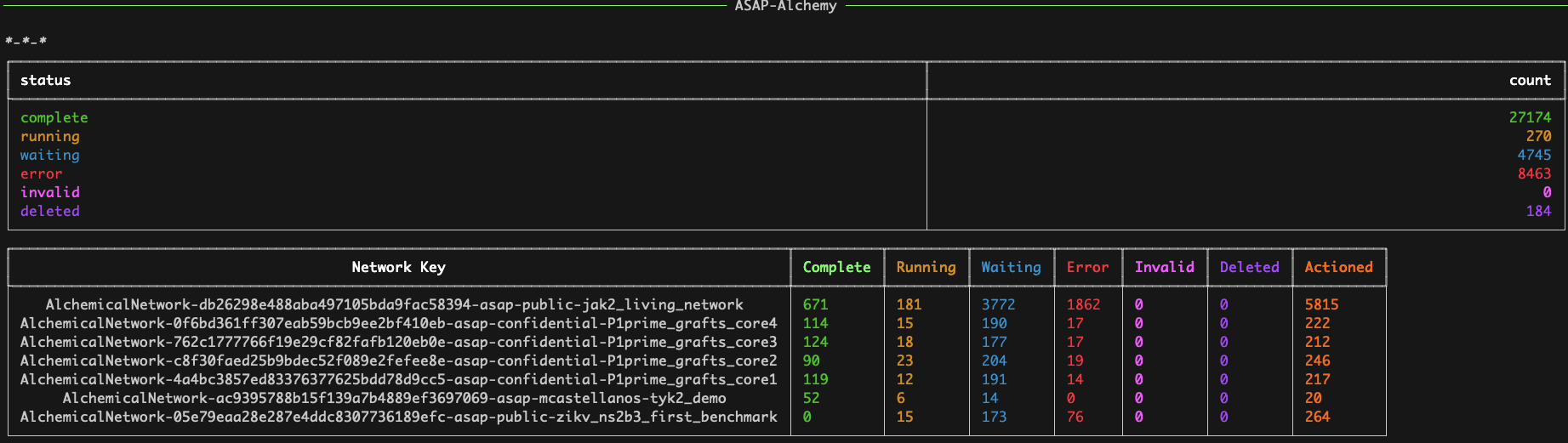
Integration of state-of-the-art virtual screening fragment elaboration pipeline which has been used to screen over **140 thousand ligands** **from Enamine real space against the** **EV71A-2A target**. A subset of the top hits are now prioritized for experimental testing.

**A. Specific Aims for the MP/DRP**

The Specific Aims have not been modified from the original, competing application.

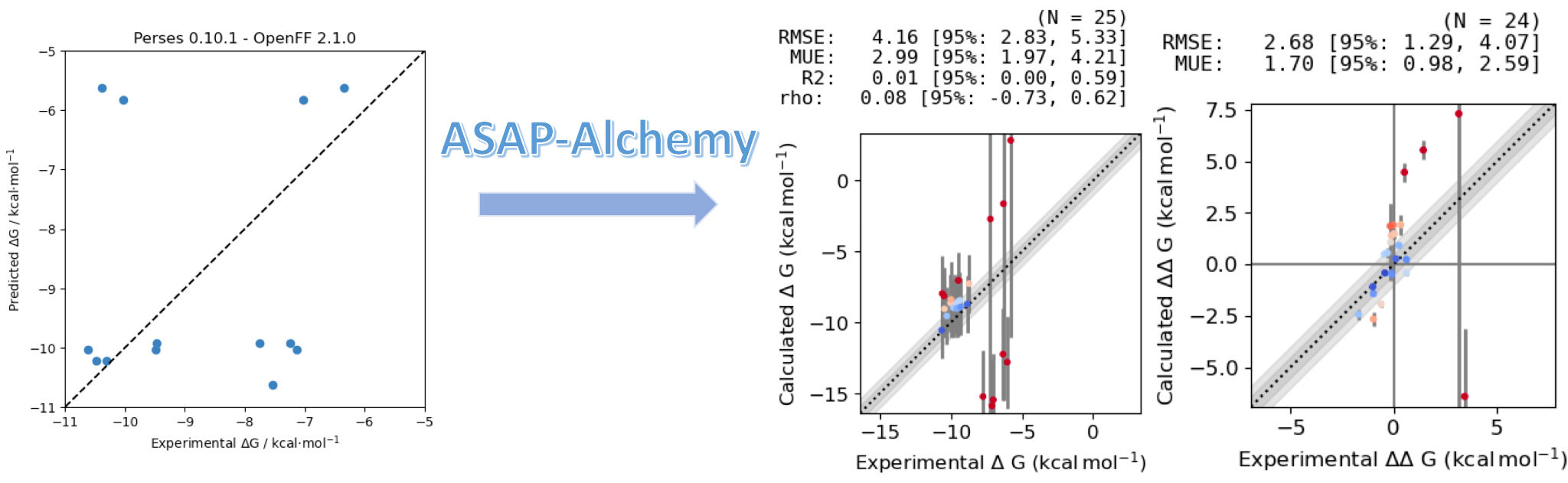
**B. Studies and Results**

**Free energies calculations at scale:** ASAP-Alchemy dramatically expands the scale at which we can perform free energy calculations as its fully automated workflow makes them trivial to prepare, execute and analyse. Interfaces to state-of-art distributed compute platforms such as [Alchemiscale](https://github.com/openforcefield/alchemiscale) enable the execution of thousands of free-energy calculations simultaneously. This powerful combination has **enabled over 27.1 thousand calculations of protein-ligand binding free energies** to date across multiple viral targets, allowing the computational-chemistry core to provide support at a scale relevant to drug discovery.

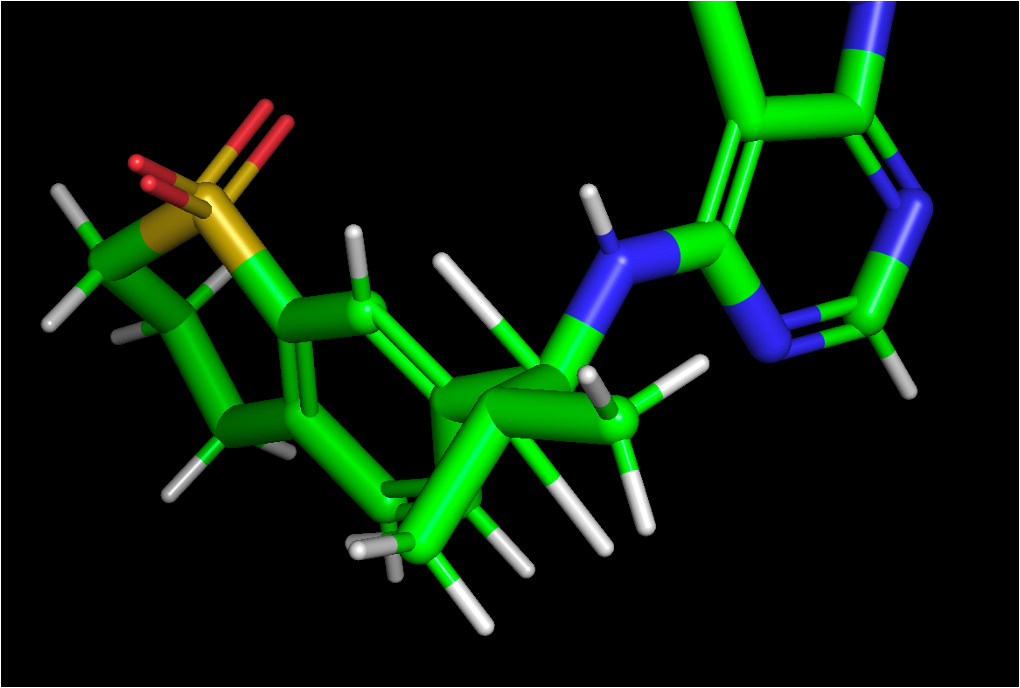


**Figure 2**: ASAP-Alchemy makes it trivial to execute and track thousands of free-energy calculations simultaneously. A screen shot of the dashbord provided by ASAP-Alchemy is shown divided into the overal status of all calculations run to date (top) and a per network breakdown for the currently actioned calculations (bottom).

**Setup validation improves accuracy -** A retrospective benchmark on the MAC1 target compared the old and new automated free energy workflows. This target has proved problematic and predicted binding affinities showed no correlation with experimental data, our new workflow however fixes this and shows **good predictive accuracy**. The latest workflow performs strict checks on the stereochemistry of the modeled ligands to ensure they always match the intended input demonstrating the importance of careful system preparation. This strict system validation will help to ensure the success of free energy calculations applied to current and future drug discovery campaigns.



**Figure 3:** Robust system setup protocols increase the reliability of protein-ligand binding calculations. The previous workflow showed little correlation with experiment (left) this is improved with the application of ASAP-Alchemy (right) which shows correlation with experiment when removing outliers caused by unconverged calculations involving large transformations.



**Figure 4:** Inconsistent stereochemistry reduces the predictive power of binding free energy calculations. A SARS2-Nsp3-MAC1 inhibitor with an inconsistent pose and intended stereochemistry is shown, these cases are now automatically detected and removed from the alchemical networks.

**Diversified starting points:** With access to high-throughput structural biology and carefully designed fragment libraries, scientists in P2 at Diamond are able to generate 10s-100s of fragment hit compounds against a single target. However, selecting follow-on compounds from make-on-demand libraries comprising billions of potential purchasable compounds is non trivial. We have employed the FEgrow software (<https://github.com/cole-group/FEgrow>), which was developed at Newcastle University to grow and score compounds based on crystallographic input structure. The virtual screening pipeline has been applied to the EV71A-2A target (currently P2) to generate fragment elaborations. Over **140 thousand compounds have been screened with many top hits now prioritized for experimental testing**. This has significantly diversified the novelty of the starting points while increasing the volume of designs put forward for testing.

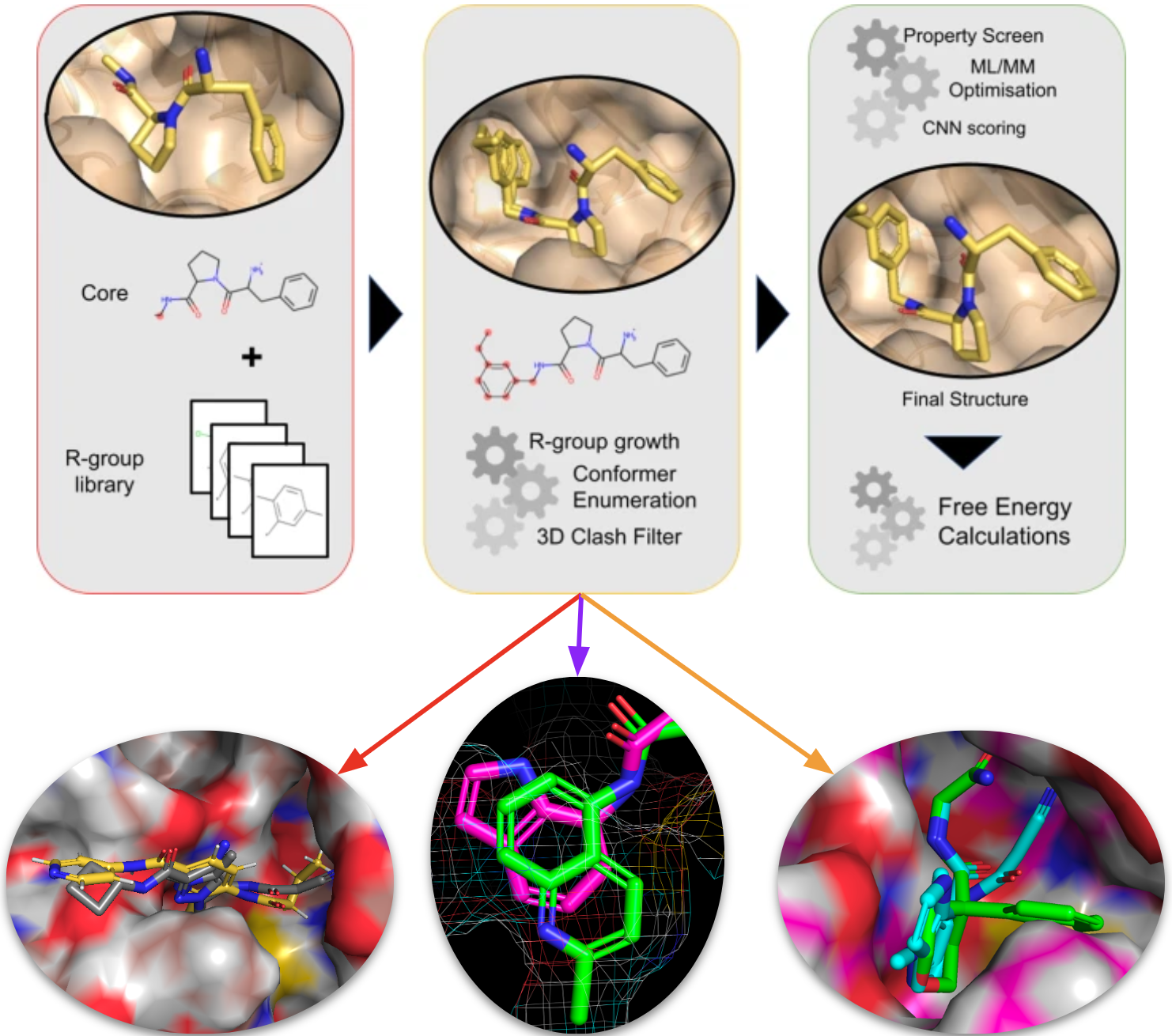


Figure 5: State-of-the-art virtual screening techniques diversify the novelty of designs for P2 starting points. The FEGrow workflow is shown along with 3 diverse starting points identified from make-on-demand libraries as potential binders which have now been prioritised for experimental validation.

**C. Significance**

The project will improve the predictive accuracy of free energy calculations on the exascale, thus aiding the prioritization and filtering of potent antivirals for synthesis. Our open science approach will enable global, equitable access to our workflows and data, and accelerate follow-on development of methods and compounds. The future career development of the PI (Horton) will benefit immensely from interactions with the world-leading target validation, structural biology, AI synthesis technologies, medicinal chemistry, antiviral assay, and preclinical development expertise at ASAP.

**D. Plans**

We will increase the throughput of the bespoke parameterization workflow by integrating state-of-the-art machine learning potentials to generate high-quality reference fitting data rapidly. A retrospective benchmark will then be conducted to assess a hierarchy of robust fitting protocols that offer a trade-off between speed and model accuracy. The workflow will be incorporated into ASAP-alchemy to enable prospective, accurate free energy calculations at scale. This will facilitate tight integration of the workflow with the rapid design-make-test-analyze cycles of the projects.

We will enable the adoption of the workflow by other campaigns outside the consortium via dissemination activities covering journal articles, posters and talks at leading conferences in the field including the alchemistry workshop and ccp-biosim. This alongside comprehensive documentation and examples will provide the community with a robust state-of-the-art open source free energy workflow.

**Training and Professional Development (For Mentored Projects only)**

The mentee (Horton) is mentored throughout the project by Dr John Chodera, who is the ASAP virtual computational chemistry core lead. Chodera leads on several large distributed open science collaborations and is co-founder of the Folding@Home Consortium, the largest computing resource on the planet for biology, and serves on Scientific Advisory Boards of multiple discovery companies. The mentored project has been integrated into the virtual computational chemistry core, and has begun to support active discovery activities in P2 and P5.

Horton contributes a proven track record in open software development, molecular modeling and force field design, but learning more about the application of these in computer-aided drug design from the core lead and integration with the wider target validation, structural biology, AI synthesis technologies, medicinal chemistry, antiviral assay, and preclinical development expertise at ASAP will benefit his career immensely.