SPECIFIC AIMS

Therapies for COVID-19—which has claimed over 700K lives and infected more than 19M just nine months after the first cases—are urgently needed. The SARS-CoV-2 main viral protease (Mpro) is an attractive target given its distinctiveness from host proteases, essentiality in the viral life cycle, and high sequence conservation to SARS-CoV-1 Mpro. Unfortunately, we found the most potent small molecule inhibitor against SARS-CoV-1 Mpro ($IC_{50} \sim$ 5 nM) is inactive against SARS-CoV-2 ($IC_{50} \sim 5$ μM), motivating the need for a new rapid drug discovery effort capable of leveraging available SARS-CoV-1 inhibitor scaffolds and data and modern technological advances.

In Mar 2020, a a high-throughput fragment screen against SARS-CoV-2 Mpro was conducted at the DiamondMX/XChem automated beamline, revealing an abundance of chemical matter targeting different regions of the active site, providing a roadmap for the development of potent inhibitors. **To rapidly develop novel hits and existing scaffolds into potent and safe inhibitors, we formed the COVID Moonshot open science collaboration, which brings together technologies and partners capable of realizing this goal:** high-throughput crystallography (DiamondMX/XChem) and biochemical assays (London lab) with rapid alchemical free energy calculations exploiting the world's largest computing resource (Folding@home/MSKCC), automated synthetic route planning (PostEra), medicinal chemistry expertise (MedChemica), pharmacology expertise (UCSD), and synthesis and in vivo assay contract research organizations operating at near cost (Enamine, Sai Life Sciences, WuXi, and others). The Moonshot aims to produce a low-complexity compound unburdened by patents that is inexpensive to produce at multiple sites, can be orally administered, and is effective for both treatment of severe COVID-19 disease and prophylactic use by at-risk populations. Multiple partners, including LifeArc and the Drugs for Neglected Diseases Initiative, have expressed their support in aiding subsequent development following our generation of a preclinical package as part of this proposal.

Since its inception four months ago, the COVID Moonshot has synthesized and assayed $>$900 compounds, solved $>$200 X-ray structures of novel Mpro inhibitors, identified numerous sub-micromolar lead series with viable paths to rapid therapeutic development, and demonstrated antiviral activity for multiple lead series. We propose to progress the most promising of these lead series via the following:

Aim 1: Progress novel aminopyridine scaffolds discovered via fragment merger strategies. A merger of fragment hits produced an aminopyridine scaffold compound that crystallographically recapitulates interactions from multiple ligands and exhibits an IC_{50} of 600 nM in a fluorescence inhibition assay. Using exascale free energy calculations to exhaustively explore the exit vector that engages the P3 pocket, we will iterate synthesis, crystallography, assays, and free energy calculations to enable medicinal chemistry to rapidly progress this compound series toward potency.

Aim 2: Progress novel quinolone hits discovered via high-throughput screening. A high-throughput screen conducted by the London lab (Weizmann) identified a quinolone compound with high similarity to a SARS-CoV-1 Mpro hit, and the Moonshot subsequently explored chemical space around these molecules. Starting with quinolone scaffold inhibitors with X-ray structural data with IC_{50} $<$10 μM that bind P1-P2 pockets, we will explore a strategy of expanding these molecules into nearby pockets by merging these molecules with fragments that bind in spatially proximal locations, prioritizing compounds via synthetic accessibility and free energy calculations.

Aim 3: Explore exhaustive computational enumerations of benzotriazoles discovered for SARS-CoV-1 Mpro. Triazole compounds previously developed against SARS-CoV-1—but terminated before a therapy was produced—also show activity against SARS-CoV-2, though with 100x poorer potency (5 μM for SARS-CoV-2 compared to 50 nM for SARS-CoV-1). Building on data that optimization of the three exit vectors of the benzotriazole scaffold can improve potency in SARS-CoV-2, we will use free energy methods and iterative med chem to exhaustively explore three exit vectors to produce potent inhibitors of SARS-CoV-2 Mpro.

Pursuit of these Aims will employ an assay cascade that includes assessment of inhibition of SARS-CoV-2 replication and plaque formation, *in vivo* PK, toxicology, and in an animal infection model, enabling candidate compounds that progress through the cascade to be handed off for subsequent human PK/PD modeling to carry out clinical trials. All compound structures, assay and structural data, and computational predictions will be released in real time, free of IP claims online via the COVID Moonshot portal [`https://postera.ai/covid`] enabling other researchers or pharmaceutical companies to immediately build on these results to develop new potential therapeutics or chemical tools for probing SARS-CoV-2 biology.

SIGNIFICANCE

Therapies for COVID-19—which has claimed over 700K lives and infected more than 19M—are urgently needed. While multiple therapeutic modalities are currently being developed, only non-covalent small molecule inhibitors have the potential to be produced rapidly and inexpensively at scale. Even following vaccine deployment, limitations on the population that can be vaccinated, the need to treat infected patients, and the potential for the emergence of resistance or novel strains motivates the need to develop small molecule therapies against conserved essential viral targets. Moreover, small molecules are more likely to fit the ideal target chemical profile for a direct-acting antiviral—patients with mild to moderate symptoms or pre/post-exposure prophylactics.

The SARS-CoV-2 main viral protease (Mpro, also known as 3C-like, 3CL, or nsp5) is an attractive target for the development of new therapies given its distinctiveness from host proteases, essentiality in the viral life cycle, and high sequence conservation to SARS-CoV-1 Mpro.[2;3] In the veterinary context, the feline coronavirus Mpro inhibitor GC-376[4] has been shown effective in reversing pathology in cats, bolstering our hypothesis that Mpro inhibitor therapies can succeed in mammals. However, many Mpro inhibitors to date, including GC-376,[4] are either peptidomimetic and/or inhibit Mpro via a covalent mechanism.[2;3] Peptidomimetics are challenging to develop into drugs due to the difficulty in achieving useful bioavailability. The development of the HIV-protease inhibitor crixivan, for example, adopted a peptidomimetic strategy that required restarting the development process three times to produce an orally bioavailable compound.[5] Covalent inhibitors carry increased off-target toxicity risk, requiring extended synthetic campaigns to tune warhead reactivity.[6] While previous efforts initiated during the SARS epidemic produced small molecule inhibitors that were not subsequently developed into therapies, we have found the most potent non-covalent small molecule inhibitor against SARS-CoV-1 Mpro (50 nM) is inactive against SARS-CoV-2 (5 μM, 100x weaker; data online at `http://postera.ai/covid`), motivating the need for a new drug discovery effort.

Here, we focus on the fastest route to an inexpensive, orally available therapy: Non-covalent small molecule inhibitors. Since March, we have developed three promising noncovalent chemical series, demonstrating members of these series exhibit significant potency, show activity in viral assays well before toxicity, and could be further optimized to produce compounds likely to achieve the requisite bioavailability whilst maintaining potency.

We have already aligned partners in drug development—the Drugs for Neglected Diseases Initiative (DNDi) and LifeArc—which are non-profits with a proven track record in clinical development and bringing therapeutics to market. Our Consortium (the **COVID Moonshot**, see below), as well as our drug development partners, share the vision of developing a therapeutic that is accessible to all. As such, we have performed all development to date—and plan to continue under this proposal—under an open science model, with all molecules developed free of patents in order to ensure the resulting therapies can be manufactured at low cost by multiple manufacturers.

Although this proposal aims to develop therapeutics for SARS-CoV-2, lead compounds and preclinical candidates we generate will add to the library of repurposable therapeutics for future viruses, significantly decreasing the response time to future pandemics and accelerating the development of pan-coronavirus inhibitors.

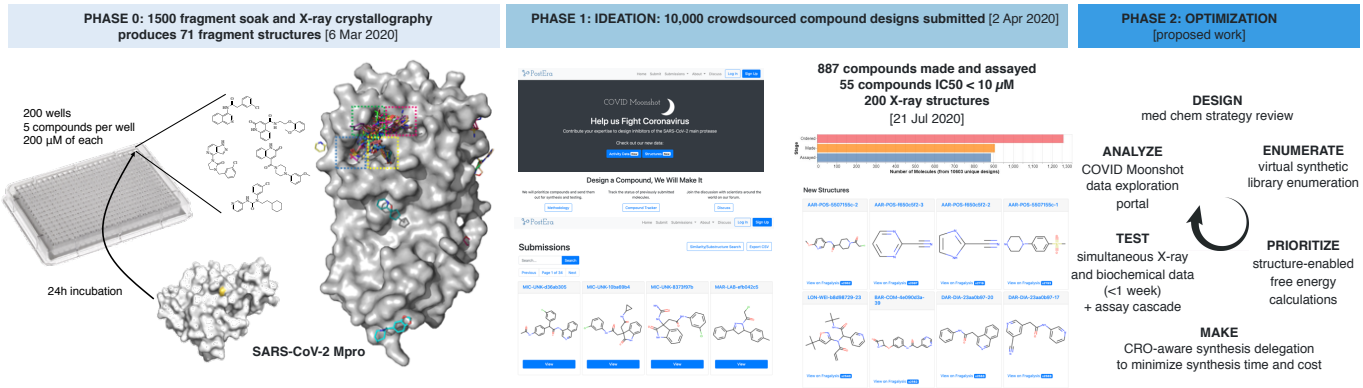

Figure 1. **A high-throughput X-ray fragment soaking experiment with the SARS-CoV-2 main viral protease (Mpro) fueled the COVID Moonshot open science drug discovery effort.** *Phase 0:* The high-throughput crystallography experiment carried out at Diamond Light Source produced 71 small molecule fragment structures that densely sampled the Mpro active site.[1] *Phase 1:* The COVID Moonshot rapidly generated multiple lead series by crowdsourcing strategies for rapidly progressing fragments into potent lead compounds. *Phase 2:* Here, we will execute rapid design-make-test cycles augmented by rapid crystallography and large-scale free energy calculations to develop a preclinical candidate for COVID-19 therapy.

INNOVATION

This proposal combines several technological advances to enable the rapid discovery of potent small molecule therapeutics against SARS-CoV-2 Mpro: (1) high throughput structural biology, (2) automated synthetic route generation and virtual library enumeration, and (3) exascale alchemical free energy calculations for structure-enabled compound prioritization. These technologies, and the whole consortium, have driven the successful hit-to-lead phase of this campaign—since March, we have gone from fragment hits to multiple lead series with sub-micromolar IC_{50} and measurable viral activities, producing multiple distinct promising chemical series with significant potential for cross-pollination between series.

Our campaign is driven by the high volume and throughput of structural data—in less than four months, we have obtained >200 high resolution protein-ligand structures, including over 70 fragments that densely populate the binding site. The Diamond Light Source synchrotron facility at Oxford (United Kingdom) has pioneered an automated workflow of crystal preparation, ligand soaking via acoustic dispensation, crystal mounting and data-analysis.[7] This enables the screening of over 1500 crystals in a single day, and turnaround times from compound receipt to uploading fully solved structures as short as one week.[1] This robust workflow has been applied to the successful discovery chemical probes targeting deubiquitinase OTUB2 and pyrophosphatase NUDT7[8]—proteins with no known chemical probes before—and generated the preliminary data for this proposal. Our preliminary fragment hits populate the entire active site, revealing salient design ideas such as interacting with His163 through a pyridine ring or similar nitrogen containing heterocycle, Glu166 through a carbonyl group in an amide or urea moiety, and aromatic ring forming hydrophobic interactions with Met49 or $\pi - \pi$ stacking with His41. Our primary strategy is to arrive at potent inhibitors via merging/elaborating those fragments, with a fallback strategy of optimising known SARS-CoV-1 inhibitors using interactions inferred from these new fragment hits.

The COVID Moonshot is a collaborative open science effort to rapidly develop a small molecule therapy for COVID-19 based on inhibiting SARS-CoV-2 Mpro. Starting with a dense set of fragment hits (**Figure 1**), we launched the COVID Moonshot. Moonshot comprises two stages: ideation and optimisation. The ideation phase (now complete) utilised an innovative crowdsourcing approach to ensure rapid and broad sampling of chemical space using multiple design methodologies for rapidly progressing fragments to potent leads. The optimisation phase utilises a novel computational approach to rapidly execute the hit-to-lead phase. This proposal leverages the fruits of the ideation stage in terms of diverse leads, and employs the computational approach that has been derisked by the successful hit-to-lead campaigns.

Crowdsourcing generated over 10,000 fragment progression ideas submitted by over 350 chemists, yielding dozens of novel chemical series with low-micromolar potency for subsequent development in. After launching a crowdsourcing campaign in Mar 2020, we triaged crowdsourced designs using synthetic accessibility and physical docking filters to generate rapid assay results and low synthetic complexity for potent lead series starting points. As the total number of commercially available virtual compounds exceeds 10 billion, ensuring rapid synthesizability is a non-trivial task. Custom synthesis prioritization tools developed by PostEra allowed for rapid prioritization of compounds, generation of synthetic routes, and targeted outsourcing of synthesis to partner CROs based on in-stock building blocks, synthetic capabilities, and available capacity. These approaches will also be key to minimizing cycle time in the subsequent lead optimization work proposed here.

Compounds synthesized at partner CROs are injected into the COVID Moonshot assay cascade (**Figure 6**), with initial in vitro screening steps including simultaneous assessment by two independent biochemical SARS-CoV-2 inhibition assays (fluorescence [Weizmann] and RapidFire mass spectrometry (RF-MS) [Oxford]), high throughput X-ray crystallography at DiamondMX, and solubility screening at Enamine (see **Assay Cascade**). Reproducibility of data in two independent, complementary protease assays was deemed critical for making rapid progress by

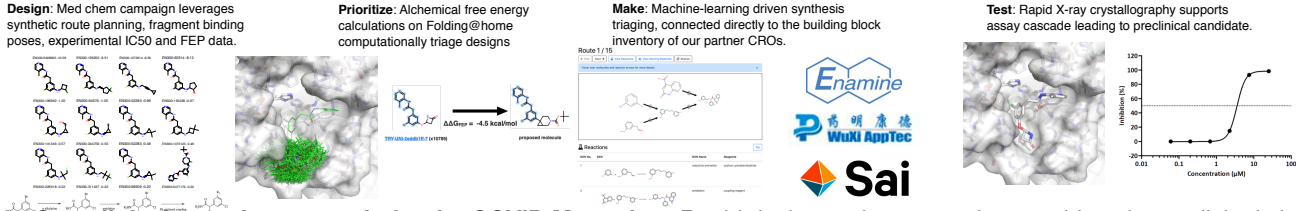

Figure 2. **Overall design-make-test cycle for the COVID Moonshot.** Rapid design-make-test cycles are driven by medicinal chemistry, assisted by ML methods (PostEra) and large-scale alchemical free energy calculations (Folding@home). Compound synthesis times are reduced by optimal matching of synthetic routes to in-house building blocks at each partner CRO to minimize time for compound delivery. Rapid simultaneous biochemical assays and X-ray crystallography (∼1 week) close the rapid design-make-test loop to allow rapid cycle times for concurrent pursuit of all three lead series.

ensuring reliable and robust data is used to guide design iterations. High throughput crystallography has provided extensive new structural data for lead compounds with as little as one week turnaround time, producing >200 new Mpro inhiitor structures since Feb 2020. Rapid structure-enabled design cycles have allowed for rapid generation of design ideas that merge compounds from the initial fragment screen (which cover nearly the entire active site) with more recent active compounds generated early on in the project that focus on P1-P2 (**Figure 3**).

We have now identified over 106 compounds with $IC_{50} < 50\ \mu$M in inhibiting SARS-CoV-2 Mpro (55 with $IC_{50} < 10\ \mu$M; 7 with $IC_{50} < 1\ \mu$M). Those compounds will be further optimised using a rapid rounds of design-make-test cycles that heavily leverages medicinal chemistry expertise, generative machine learning, free energy calculations, and synthesis design tools. Our generative machine learning model is used by the medicinal chemistry team to propose new compounds based on synthetic strategies likely to improve potency to exploring chemical space around the hit clusters. Free energy calculations run on an exascale distributed computing platform enable rapid and reliable prioritization of these designs. Finally, machine learning for synthesis triaging selects compounds our CRO partners can rapidly synthesize. Our one-month hit-to-lead sprint generated multiple noncovalent lead series; here, we focus on rapidly developing the three most promising series.

Our leading chemical series are based around 3-aminopyridines (Aim 1), quinolones (Aim 2), and benzotriazoles (Aim 3) scaffolds. We have obtained extensive crystallographic data on all three series, and our methodology (crystallography and free energy perturbation for structure-enabled design) is optimally suited for optimizing these. (**Aim 1**) The 3-aminopyridine hit was discovered by crowdsourced fragment merges, and further developed by our hit-to-lead sprint. The rapid optimisation of this series from fragment hit to a nM inhibitor derisks our proposed approach. (**Aim 2**) The quinolone series was discovered by a HTS screen in Nir London's lab. It was also one of the hits in a HTS against SARS-CoV MPro (PubChem AID 1879), although to our knowledge no ligand-bound crystal structure has been obtained and this series was not further developed. Our structural studies elucidated the binding mode of quinolones, showing that they have a different interactions pattern with the protein backbone compared to 3-aminopyridines. (**Aim 2**) The benzotriazole chemotype has been developed as a chemical probe for SARS-CoV-1 Mpro. Our preliminary work determined the structure of MPro-bound benzotriazoles, and showed that benzotriazoles with nanomolar potency against SARS-CoV-1 MPro have potency in the micromolar range for SARS-CoV-2 MPro. Nonetheless, the series has exit vectors well-positioned to interact with the different binding pockets. We believe there is significant scope to optimise the benzotriazole series, especially given our knowledge of fragment binding modes and the structure-activity relationship of 3-aminopyridines.

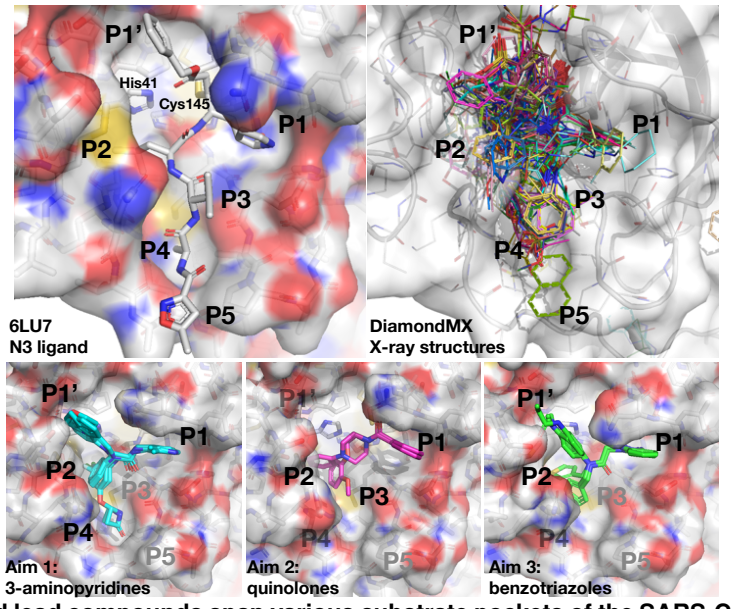

Figure 3. **X-ray fragment hits and lead compounds span various substrate pockets of the SARS-CoV-2 Mpro active site, providing a detailed map of how to rapidly progress toward potency.** *Upper left:* Peptidomimetic N3 ligand covalently bound to SARS-CoV-2 Mpro (PDBID: 6LU7[2]) with substrate peptide pockets P1' through P5 (corresponding to pockets occupied by substrate sidechains) labeled. *Upper right:* DiamondMX/XChem initial fragment hits and new X-ray structures span all pockets, providing a rich set of specific chemical design ideas to draw from for expanding current lead series in a manner that rapidly acquires potency by accessing pockets P1' through P5. *Bottom:* DiamondMX X-ray structures illustrating key compounds from each of the lead series pursued in Aims 1–3. All three series contain potent scaffolds spanning P1-P2 pockets, linkers that engage the protein with different geometries and interactions, and a diverse set of vectors that can be functionalized to acquire potency via access to other pockets. Potent pocket-binding substituents from each series can also transferred to other series with compatible geometries.

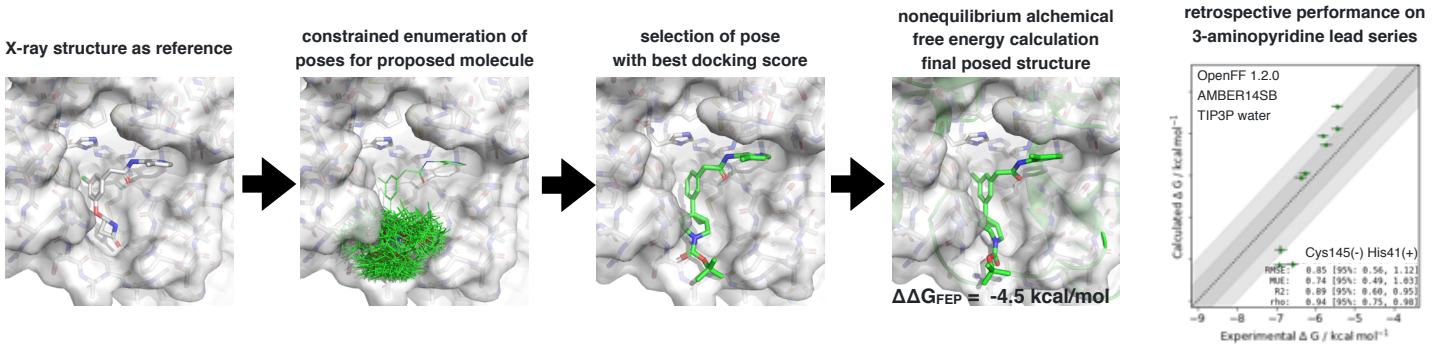

Figure 4. **Massively parallel alchemical free energy calculations on Folding@home will be used to prioritize compound expansions or chemical replcements.** First, one or more X-ray structures that share a common substructure are selected for modeling. Second, the proposed compound is built using a dense sampling of poses where heavy atoms in common are preserved. Next, the pose with the best Chemgauss4 docking score is selected. Finally, at $\geq$100 x 4 ns nonequilibrium switching trajectory cycles (1 ns equilibration followed by 1 ns nonequilibrium alchemical switching) are performed in parallel on Folding@home. Nonequilibrium work values are used to estimate $\Delta\Delta G$ of binding using the Bennett acceptance ratio.[12;13] Retrospective benchmarks on the 3-aminopyridine lead series (Aim 1) shows an RMSE of 0.85 kcal/mol [95% CI: 0.56, 1.12] and correlation coefficient 0.94 [95%: 0.75, 0.98], demonstrating that this approach will be highly useful in prioritizing compounds to achieve rapid potency gains. Folding@home can perform >10K calculations/week.

Structure-enabled alchemical free energy calculations will support rapid design cycles involving substituent replacement and lead expansion to access new pockets. A significant number of design cycles will focus on either (1) replacing a substituent to optimize potency or eliminate liabilities or (2) expanding a lead compound into nearby pockets (often guided by fragment structures highlighting specific interactions) to gain affinity. The abundance of structural data available for our lead compounds (>200 structures have been solved by DiamondMX, all posted publicly online at `https://postera.ai/covid/structures`) and rapid turnaround time for new X-ray structures (as short as seven days from compound receipt) enables us to make rapid use of relative alchemical free energy calculations—sometimes referred to as free energy perturbation (FEP)[9]—within these design cycles to prioritize compounds for synthesis. These methods have demonstrated their value in reducing the number of compounds that must be synthesized to achieve potency goals in structure-enabled programs.[10;11]

Our strategy focuses on conservative but massively parallel application to maximize potency gains in rapid design cycles: PostEra's machine learning platform for synthesis planning selects strategies that permit rapid parallel synthesis of multiple products by our CROs, where choices among in-stock starting materials realize replacement of a portion of the compound or installation of a new substituent to engage a new pocket. A virtual library of hundreds to thousands of possible products is generated and docked using an appropriate X-ray structure which shares most of the compound scaffold, and relative free energy calculations predict the change in affinity for all potential designs. Predicted poses from the best-scoring designs will be visually inspected within Fragalysis [`https://fragalysis.diamond.ac.uk/`] and a set of design compounds prioritized for parallel synthesis by our CRO chemists (via PostEra). Newly synthesized compounds are injected into the assay cascade, which includes the rapid solution of new X-ray structures to permit rapid design cycle iterations.

We utilize a highly scalable nonequilibrium alchemical free energy approach[14] implemented within the OpenMM GPU-accelerated simulation package[15] in the open source software package perses [`http://github.com/choderalab/perses`]. These calculations utilize public force fields (AMBER14SB[16] with TIP3P water[17] and small molecule parameters assigned using either the Open Force Field Initiative Parsley small molecule force field[18] or GAFF[19;20]/AM1-BCC,[21;22] both of which have recently been shown to provide accuracies statistically indistinguishable from proprietary force fields (better than 1 kcal/mol).[23;24] These calculations will be carried out on Folding@home, the world's largest computing resource (>2.5 exaflops, >25M CPU cores, >600K GPUs, now in its twentieth year of providing the largest computing resource available for computational biophysics),[25] which provides <48-hour turnaround times and capacity to run more than 10,000 compounds/week.

Our use of PostEra to manage CROs enables rapid and inexpensive delivery of compounds. PostEra (A. Lee, see **Letter**) has developed a machine learning engine that finds divergent synthetic routes that realize several lead compounds from common intermediates in the vast building block inventories of contract research organizations (CROs). This approach uses Molecular Transformer, a machine learning framework that predicts the outcomes of organic reactions[26] and plans retrosynthetic routes.[27] The algorithm infers the rules of chemistry using a machine translation approach, mining >10M chemical reactions reported in patents and the literature. This method achieves >90% accuracy in predicting the correct product of unseen chemical reactions given reactants and reagents—making it the state-of-the-art—as well as correctly predicted the outcomes of challenging regio- and chemoselective transformations, and has been validated by Pfizer[27] in medicinal chemistry contexts.

Target Product Profile (TPP) for oral SARS-CoV-2 main viral protease (Mpro) inhibitor

Property	Target range	Rationale
protease assay	IC_{50} < 10 nM	Extrapolation from other anti-viral programs
viral replication assay	EC_{50} < 5 μM	Suppression of virus at achievable blood levels
plaque reduction assay	EC_{50} < 5 μM	Suppression of virus at achievable blood levels
route of administration	oral	bid/tid - compromise PK for potency if pharmacodynamic effect achieved
solubility	> 5 mg/mL	Aim for biopharmaceutical class 1 assuming <= 750 mg dose
half-life	> 8 h (human) est from rat and dog	Assume PK/PD requires continuous cover over plaque inhibition for 24 h max bid dosing
safety	Only reversible and monitorable toxicities No significant DDI - clean in 5 CYP450 isoforms hERG and NaV1.5 IC_{50} > 50 μM No significant change in QTc Ames negative No mutagenicity or teratogenicity risk	No significant toxicological delays to development DDI aims to deal with co-morbidities / therapies, cardiac safety for COVID-19 risk profile cardiac safety for COVID-19 risk profile Low carcinogenicity risk reduces delays in manufacturing Patient group will include significant proportion of women of childbearing age

Figure 5. **Target product profile (TPP):** We aim for a target product profile that would support oral use for treatment of severe COVID-19 disease, as well as prophylactic use by health-care workers or other at-risk members of the population.

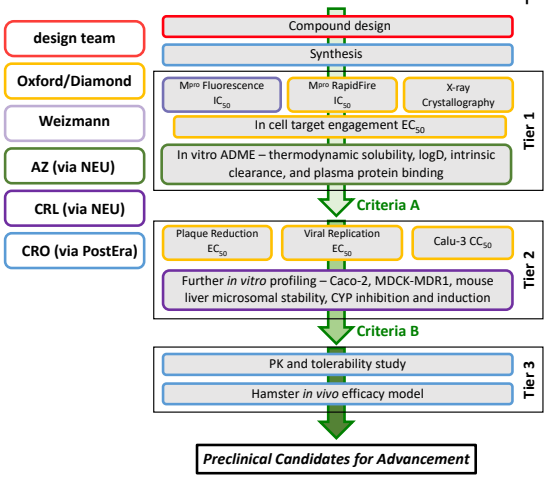

Figure 6: **Assay cascade schematic.** Assay details appear in **Assay Cascade Details**; tier advancement criteria are described below.

PostEra has developed a scalable platform, based on the Molecular Transformer technology,[26] that plans synthetic routes at scale using over 10 billion building blocks available from partner CROs. This platform estimates synthetic cost, designing divergent synthetic strategies where distinct compounds with optimal modifications suggested by FEP can be executed via common intermediates. The platform is directly connected to the inventory of partner CROs with access to different building blocks—Enamine in Ukraine, Wuxi in China, and Sai Life Sciences in India—allowing us to optimally allocate synthesis requests to minimize delivery time and cost. In the initial hit-to-lead phase of our campaign, our platform is applied at scale to triage synthetic accessibility of compounds and design the medicinal chemistry strategy, synthesizing over 1000 compounds at an average cost of $300/molecule.

Underlying these technological innovations is a core team of experienced drug designers driving the design decisions; the core team has already successfully executed the hit-the-lead phase. Design decisions integrate free energy calculations (John Chodera), structure biology (Frank von Delft), biochemical assays (Nir London), and machine learning (Alpha Lee), and industrial drug development expertise (Ed Griffen); see **Letters**.

Approach

We aim to produce an oral SARS-CoV-2 main viral protease (Mpro) inhibitor (bf Figure 5), and have constructed an assay cascade (**Figure 6**) to enable rapid lead optimization to produce a preclinical package for handoff to another entity for development (see **Clinical Handoff Strategy**). Below, we describe the components of the assay cascade and logistics in detail before describing the medicinal chemistry approach.

Assay cascade advancement criteria

Criteria A: <1 μM in both rapidfire and fluorescence IC_{50}; >10 μM aqueous solubility; PPB >1% free (<99% PPB). First few iterations will let compounds through Critiera A to collect clearance etc data on new chemotypes.

Criteria B: <1 μM EC_{50} in plaque reduction and viral replication; Calu-3 CC_{50} >100 μM; >50% remaining after 60 min in hepatocytes and microsomes.

Criteria for preclinical candidates: *Prior to in vivo efficacy:* Half-life in rat ≥ 4 h (equiv 2x/day humans); minimum free blood levels 2x viral replication EC_{50} for 6 h in rat; no irreversible or unmonitorable toxicity in rat. *Progression to preclinical candidates:* Efficacy (via suppression of viral load) in Syrian hamster disease model[28] at <100 mg/kg (ideally <10 mg/kg). Human PK modeling studies indicate probable potential for b.i.d. dosing in human.

Compound and assay logistics: Enamine will handle all compound logistics, including replating and shipping. Initial synthesis will be performed at a 15 mg scale with $\geq$95% purity. Compounds will be plated in DMSO and shipped weekly from Enamine to multiple sites: X-ray crystallography and RF-MS biochemical assays (Frank von Delft, Oxford/Diamond Light Source) and fluorescence biochemical assays (Nir London, Weizmann) will be performed within two days of compound receipt. Enamine will handle shipment of compounds directly to subsequent assay cascade steps once decision points are triggered, and will be responsible for large-scale compound resynthesis to support subsequent assay steps as needed.

During our hit-to-lead phase, our design-make-test cycle time has been 3–4 weeks, and we have seamlessly pushed more than a 1,000 compounds thorough synthesis and assays in four months. As the project progresses to virology and in vivo PK studies, we anticipate cycle time will extend to 4–6 weeks, especially as PK studies require resynthesis at larger scales.

Data management: All data management will be handled by the COVID Moonshot open data platform (A. Lee, PostEra; **Figure 11**). All data is released immediately for both internal and public use.

Design and synthesis decisions: The design team (John Chodera, Frank von Delft, Nir London, Alpha Lee and Ed Griffen) will hold weekly meetings to review new data and design follow-up compounds. Ed Griffen (MedChemica) will be responsible for providing drug design expertise and coordinating input from *ad hoc* contributors from the Moonshot platform; over the past four months, chemists from UCB Pharma and numerous experienced drug-hunters (e.g. Bruce Lefker, Ralph Robinson, Robert Glen, all with experience of launching multiple marketed drugs) have formed an ad hoc design advisory team providing valuable input.

All proposed designs (including virtual library generation) will be captured by the Moonshot platform, which will initiate evaluation of potential for rapid synthesis by CRO partners (Alpha Lee, PostEra) and relative alchemical free energy calculations to closest X-ray structures (John Chodera, MSKCC/Folding@home). Within 48 hours of the design team meeting, a final list will be compiled, and synthesis will be initiated at partner CROs in a manner that minimizes compound turnaround time by accounting for in-stock building blocks (Alpha Lee, PostEra). In addition to the design meeting, Alpha Lee will have weekly reviews with each CRO to monitor and evaluate progress, aided by PostEra's synthesis planning and tracking platform.

As the project progresses to virology and in vivo PK studies, the design team will continue to optimise compound series based on biochemical assay readouts, as well as progress other chemical series. Virology and in PK data will be discussed in weekly design meetings as they become available, and the design team will be augmented by a virologist (co-I Annette von Delft) and pharmacologist (consultant Jeremiah Momper, see **Letter**) as appropriate to aid in interpretation of virology and in vivo data and progression of assay tiers according to **Criteria**.

Quarterly evaluations and selection of in vivo efficacy study candidates: Investigators and design team members will hold quarterly strategic meetings to evaluate the priority of each chemical series.

Clinical development strategy: Successful completion of this project will result in the delivery of a novel MPro protease small molecule inhibitor. The inhibitor will have sub-micromolar efficacy in cellular antiviral assays and antiviral efficacy in animal models at non-toxic doses at doses $<$100 mg/kg (ideally 10 mg/kg; see **Criteria**). All lead molecules will be accompanied by a fully comprehensive safety data package, enabling the translation into first-in-human clinical trials. We have already established collaborative interest from LifeArc and DNDi (see **Letters**). Both have offered to act as advisors for the current project, and downstream to enable the seamless translation of our pre-clinical candidates into clinical trials through their network of partners. Regulatory authorities including the MHRA Innovation Office (UK) will be consulted at an early stage of the project to maximise the likelihood of regulatory approval of a potential drug candidate. To support future development of the candidate therapeutic entity in early phase clinical studies funding may be sought from a range of funding bodies, including the Medical Research Council (MRC) Development Pathway Funding Scheme (DPFS) (UK), the Innovative Medicines Initiative (EU), or the NIH. Extensive clinical expertise in infectious disease and virology is well established at the University of Oxford; our clinical colleagues Prof. Peter Horby and Prof. Peter Landray have the specific expertise and infrastructure in place to carry out multicentre SARS-CoV-2 trials. Indeed, they are currently running the RECOVERY trial repositioning a range of molecules against SARS-CoV-2 (`clinicalTrials.gov`, NCT04381936). Oxford has recently established a first-in man clinical trials centre (Oxford Clinical Trials Research Unit) led by Prof Duncan Richards, with which the BRC translational science team (A. von Delft, see **Letter**) closely collaborates.

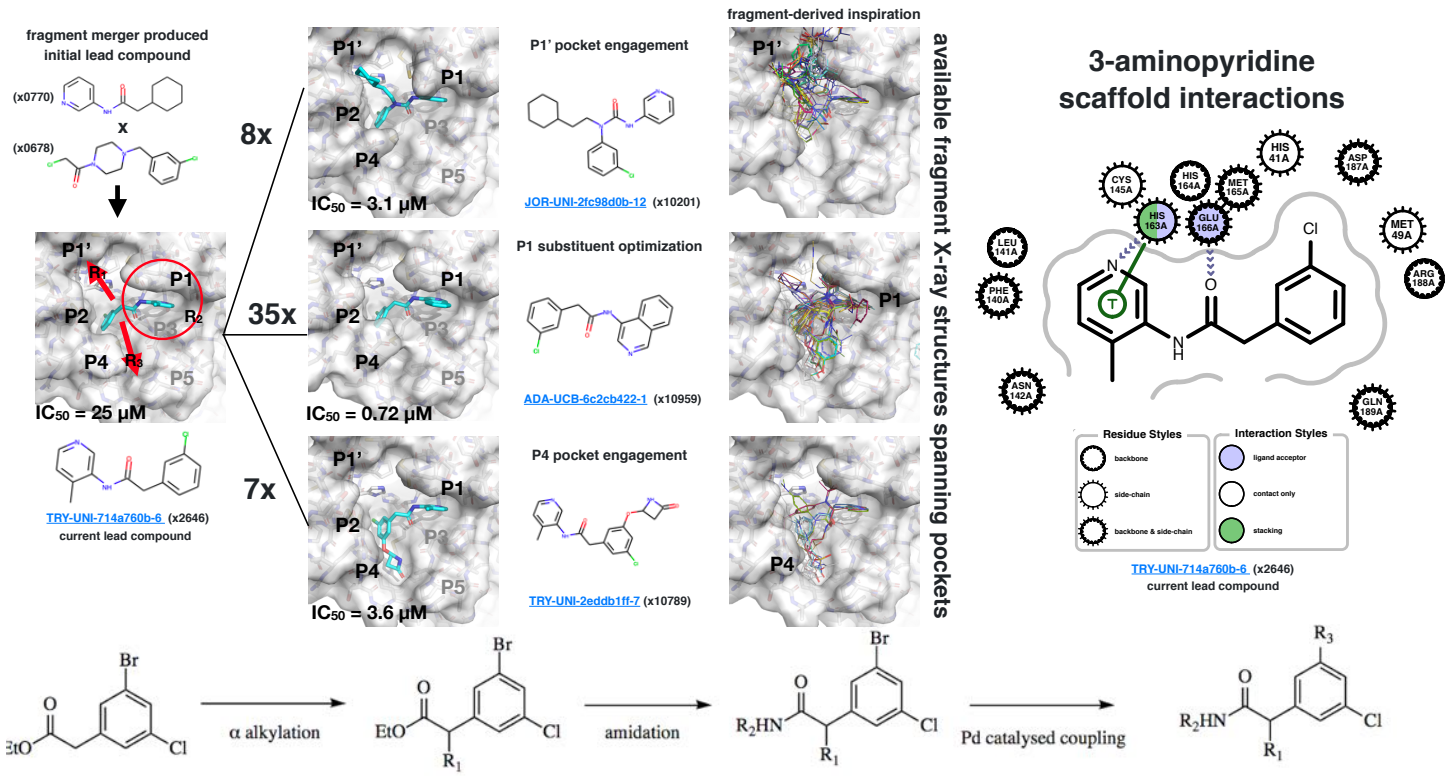

Figure 7. **Optimization of the 3-aminopyridine lead compound (Aim 1) will seek to optimize existing components of the structure and gain affinity by engaging new interactions in pockets P1' and P4.** *Top left:* Current 3-aminopyridine lead compound and selected compounds (of 256 compounds in this series assayed) illustrating SAR around current design. *Top right:* Fragment structures in P1' and P4 pockets that will be used to guide ideation process for engaging these pockets. Bottom: One possible synthetic route that permits the installation of substituents to access P1' (R1), vary engagement in P1 (R2), and grow into P4 (R3).

Aim 1: Progress novel aminopyridine scaffolds discovered via fragment merger strategies.

The 3-aminopyridine lead series originates from a merger of initial X-ray fragments x0770 and x0678 (**Figure 7**), producing a scaffold that engages Mpro His163 and Glu166 via hydrogen bonds and aromatic stacking interactions. This scaffold provides three exit vectors tractable through validated synthetic routes, enabling extensive and rapid optimization of interactions with the P1 pocket (via R2) as well as extensions into P1' (via R1) and P4 pockets (via R3). Our proposed lead optimisation campaign aims to utilize all three exit vectors, leveraging structure-activity relationships uncovered in the 264 compounds already assayed in this series as well as the large number of fragment interactions exploring these three pockets from our initial fragment screen to inform compound ideation (**Figure 7**). Using our synthetic strategy, large virtual libraries of rapidly synthesizable compounds will be enumerated given in-stock building blocks and prioritized with large-scale, rapid-turnaround free energy calculations to identify compounds that can be synthesized in parallel. Rapid rounds of structure-enabled optimization will quickly enable potency optimization from the current best lead in this series (IC_{50} = 0.72 μM).

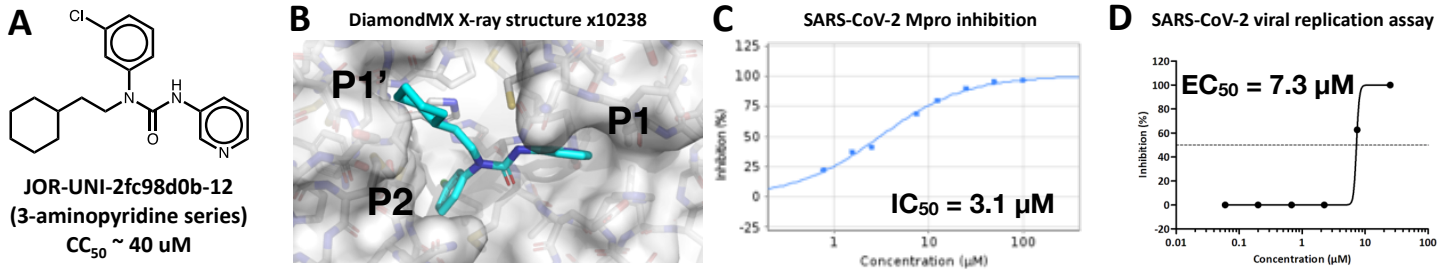

Figure 8. **An early 3-aminopyridine series inhibitor (Aim 1) shows sub-micromolar efficacy in a SARS-CoV-2 viral assay at concentrations well below cytotoxicity. (A)** The 3-aminopyridine inhibitor JOR-UNI-2fc98d0b-12 derives from merging multiple fragments from the initial DiamondMX screen. (B) The X-ray structure of this compound shows engagement of the P1, P2, and P1' pockets. (C) Concentration-dependent inhibition shows this compound inhibits SARS-CoV-2 Mpro with an IC_{50} of 3.1 μM (details in **Assay Cascade Details**). (D) This compound demonstrates an EC_{50} of 7.3 μM in a SARS-CoV-2 viral replication assay performed at the Radboud Institute for Molecular Life Sciences, Nijmegen, Netherlands. Note this assay differs from protocols in **Assay Cascade Details**. Briefly: Vero E6 cells were seeded in 24-well plates were infected with SARS-CoV-2 Bavarian isolate in the presence of 6 different concentrations of the inhibitors (25 μM–0.06 μM). After 1 h incubation at 37°C, virus inoculum was discarded, cells were washed with PBS and replaced with infection medium containing the same concentration of inhibitors. 0.1% DMSO served as a control. 24 h post-infection, 100 μL of cell culture supernatant was added to TRIzol reagent for RNA isolation. EC_{50}s were computed via fits to qPCR readouts.

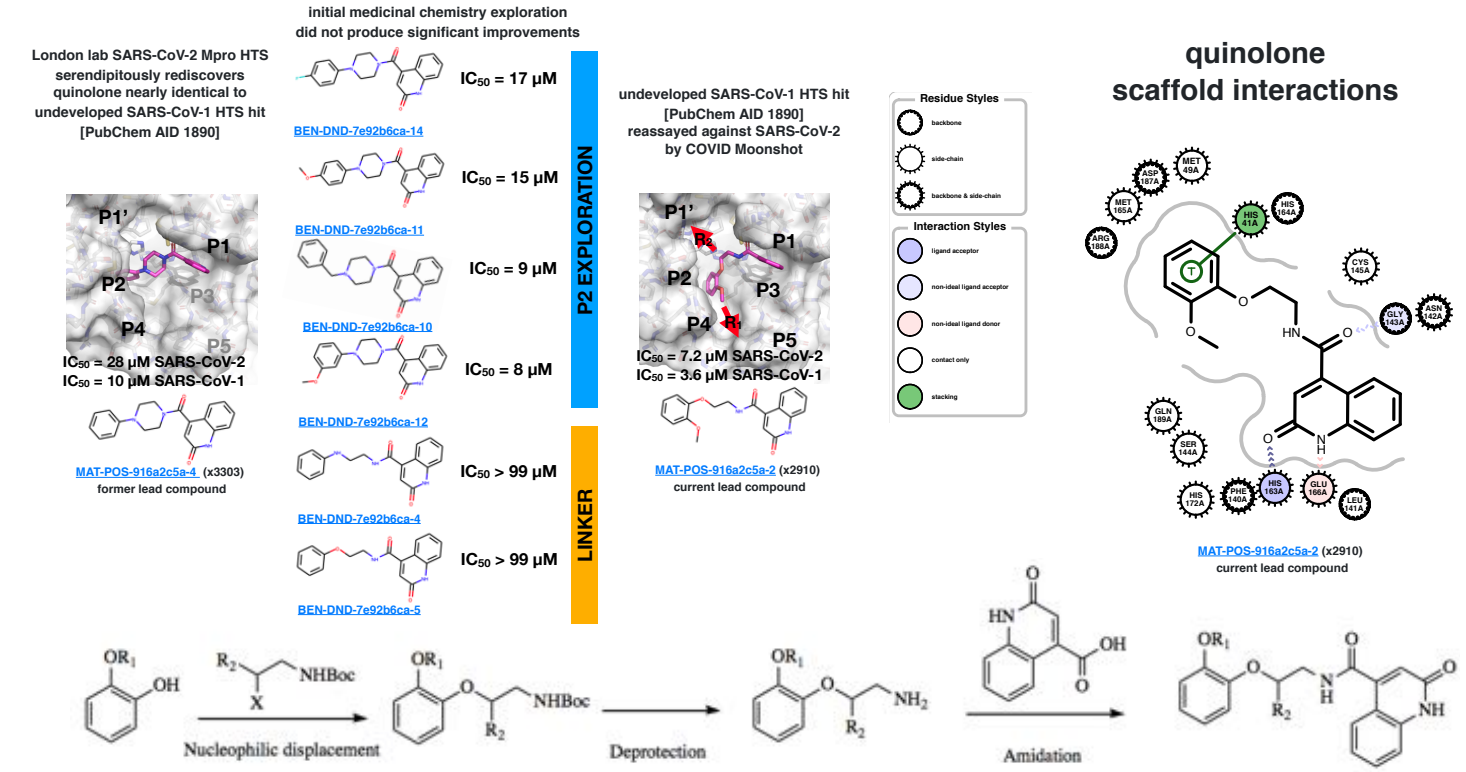

Figure 9. **Optimization of the quinolone lead compound (Aim 2) enables an alternative linker and exploration of distinct exit vectors from the 3-aminopyridine series to gain affinity by engaging interactions in pockets P1' and P3/P4/P5.** *Top left:* Original quinolone lead compound and selected compounds (of 57 compounds in this series assayed) illustrating SAR around current design, demonstrating difficulty in optimizing original scaffold. *Middle:* Current quinolone lead compound which exhibits improved potency and simple positioning of exit vectors to access P1' and P4. *Right:* Ligand interaction diagram illustrating scaffold engagement with active site residues. *Bottom:* One possible synthetic route that permits the installation of substituents to access P1' (R2) and grow into P3/P4/P5 (R1).

First, we aim to maximize the potency of the minimal P1-P2 spanning fragment as an anchor for further expansion and exploration of neighboring pockets, where existing SAR has suggested the potential for large affinity gains. Second, we will expand the fragments into the P1' and P4 pockets by replacing substituents in this series that have already demonstrated the ability to access these pockets and provide 7–8x affinity gains prior to more extensive optimization. Our aim is to keep the compound small (in heavy atom count) and minimize the number of hydrogen bond donors to avoid bioavailability issues commonly associated with peptidomimetics.

One compound in the 3-aminopyridine series (JOR-UNI-2fc98d0b-12, x10201, IC_{50} 3.1 μM) has already been included in early viral assays, where it demonstrated low cytotoxicity (CC_{50} $\sim$40 μM) and inhibited viral killing of Vero E6 cells at concentrations below 10 μM (**Figure 8**), demonstrating the viability of developing this scaffold for potent antiviral activity with further optimization.

Aim 2: Progress novel quinolone hits discovered via high throughput screening.

Starting with quinolone scaffold inhibitors with X-ray structural data with IC_{50} $<$10 uM that bind P1-P2 pockets, we will explore a strategy of expanding these molecules into nearby pockets by merging these molecules with fragments that bind in spatially proximal locations, prioritizing compounds via synthetic accessibility and free energy calculations. The quinolone scaffold offers the potential for a different toxicological and pharmacokinetic profile compared to 3-aminopyridines, with the opportunity to engage a different set binding pockets via clear exit vectors, thus pursuing those aims derisks this campaign.

We will first optimise the quinolone core, as this will serve as a potent anchor, allowing us to reveal SAR trends when varying multiple exit vectors. We will introduce substituents on the quinolone to enhance hydrogen bonding with His163, introduce heteroatoms on the ring to modulate electron density, and introduce a protonated amine proximal to His163.

We will then optimise the series by tuning the linker, which preliminary data suggests is tolerant to modification. Distinct from 3-aminopyridines (Aim 1), the amide carbonyl interacts with Asn142 rather than Glu166. We will carry out further optimisation by varying linker length, rigidity, and patterns of hydrogen bond donors/acceptors (e.g. ether and amine linkages). As a backup, we will investigate bioisosteric replacement of the amide to modulate hydrogen bond strength with Asn142. Designs will be prioritised by synthetic accessibility and building block

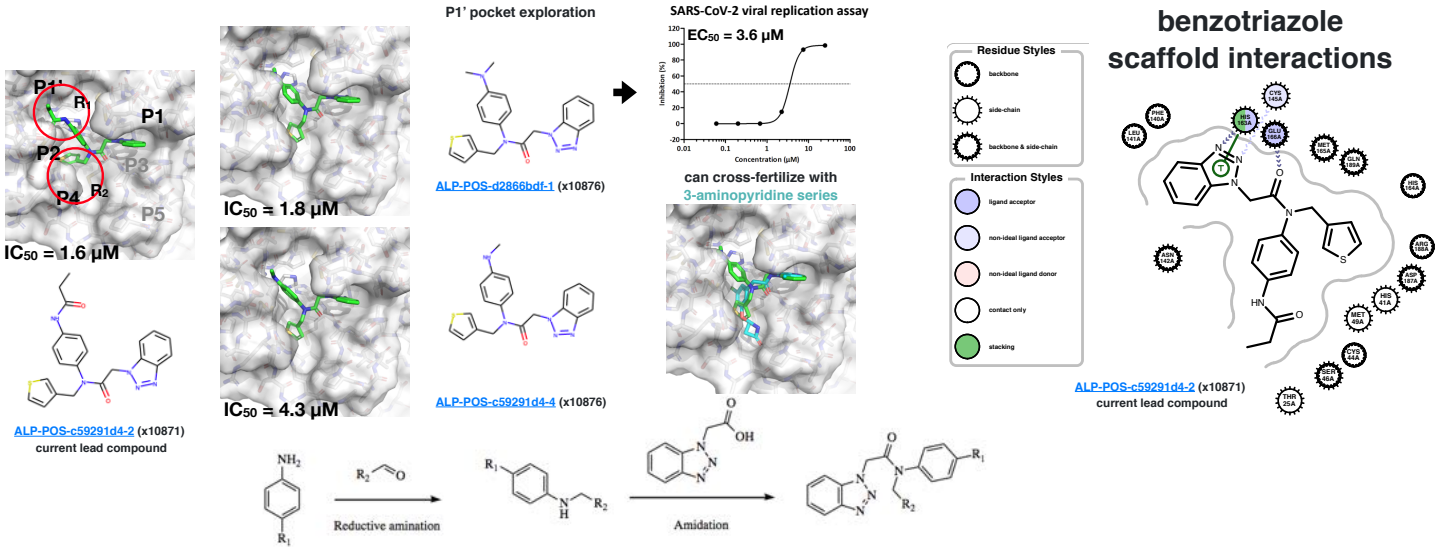

Figure 10. **The benzotriazole lead series (Aim 3) enables an optimization of a scaffold which already accesses P1' and P2/P4 pockets through distinct exit vectors.** *Left:* Current benzotriazole lead compound, with selected compounds illustrating exploration of the P1' pocket is well-tolerated and could lead to potency gains. *Middle:* X-ray and 2D structures of 3-aminopyyridine series show substituents that could be grafted onto benzotriazole scaffold (and vice-versa) to permit cross-fertilization between lead series. Compound ALP-POS-d2866bdf-1 (x10876) already shows activity in a SARS-CoV-2 viral replication assay (same protocol as Figure 8D). *Right:* Interaction schematic illustrating key interactions with benzotriazole scaffold. Bottom: One possible synthetic route that permits the optimization of substituents that access P1' (R1) and P2/P4 (R2). Specifically, to optimize binding with the P1 site, we will employ the position analogue scan technique[29] of moving nitrogen around the 5-membered aromatic ring and replacing N with C whilst keeping the crucial N–His163 interaction. This technique has previously been applied fruitfully to optimise P1 binding[30] in another series of SARS-CoV-1 Mpro inhibitors, demonstrating that P1 subpocket is sensitively dependent on the pattern of hydrogen bond acceptor and ring electron density. Inspired by the binding mode of the 3-aminopyridines, we will also investigate the retroamide and urea—the synthesis of both are facile. FEP is a powerful technique to rapidly evaluate all synthetically feasible possibilities with high confidence.

availability.

We will then turn to the ortho dialkoxy groups in MAT-POS-916a2c5a-2. Optimisation of hydrogen bonding with water forces the alkyl substituents to orient opposite to each other. This creates poised exit vectors that can penetrate into P3 and P1' pockets, by optimising R1 and R2 respectively. This optimization will be done using free energy perturbation—we will first benchmark our method on the R1 exit vector using free energy perturbations to screen an enumerated synthetic library, collecting structural data that probes the conformational plasticity of P3, a pocket that we have thus far not engaged. Armed with the revised structural knowledge, we will employ FEP to screen a larger library of compounds where the P1' and P3 substituents are varied. SAR and knowledge on P1' substituent optimisation learnt from executing Aim 1 will be used to guide the selection of P1' substituents.

Aim 3: Explore exhaustive computational enumerations of benzotriazoles discovered for SARS-CoV-1 Mpro. Benzotriazole inhibitors of SARS-CoV-1 Mpro previously discovered following SARS epidemic—but terminated before a therapy was produced—also show activity against SARS-CoV-2, though with 100x poorer potency (IC_{50} of 5 μM for SARS-CoV-2 compared to 50 nM for SARS-CoV-1).[3] The scaffold is already well-positioned to engage with the P1, P1', P4 and P2 binding sites, as evidenced by the wealth of SAR literature. We believe significant potency gains can be unlocked by the simultaneous tuning of the three exit vectors via introducing small functional group modifications. The rigidity of the binding mode, as evidenced by our structural data for the series, makes our high throughput FEP platform the ideal tool.

In terms of P1', we anticipate a slight tuning of R1 to have a significant impact, as evidenced by our preliminary data of a 2.3x increase in potency going from methylamine to dimethylamine, although our structural data show that the binding mode is the same. The same strategy can be applied to the P4 exit vector, where our structural data for 3-aminopyridines suggests that the 2-position on the thiophene is ripe for exploration. In both cases, FEP is again the method of choice, by scanning through hundreds of functional group interconversions with $<$4 heavy atom change. Our priority access to Folding@Home, the first and (to date) only exascale supercomputer, gives us a unique edge in terms of tackling this combinatorial challenge of simultaneously optimising all three exit vectors with large-scale free energy calculations.

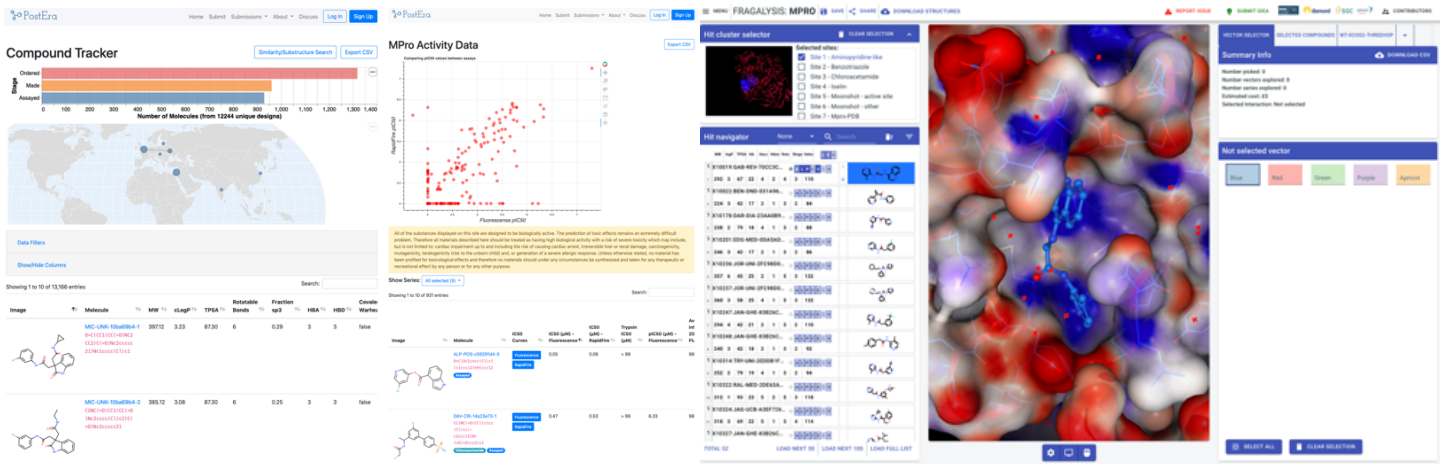

Figure 11. **The COVID Moonshot data platform integrates compound design and rationale capture, compound tracking, activity data, data analytics, and structural data in an open science data sharing platform.** The open science COVID Moonshot [`http://postera.ai/covid`] developed and operated by PostEra integrates all data collected for the project and provides a common platform for internal and external data access, as well as facilitating public discussion of data.

Pitfalls and alternative approaches

We identify three salient risks associated with our proposal: (1) Our proposed chemical series falling in in vivo studies; (2) Hamster animal model falling short of capturing salient human pathology; (3) Escape mutations in SARS-CoV-2 decreasing the therapeutic efficacy of our candidate.

1. **We identified three additional avenues to generate additional lead series should all three fail.** First, Nir London's lab has run a high throughput screen with a library of 200,000 compounds, identifying over 100 hits to date that are distinct from our three proposed series, providing ample material for new lead series should all three fail. Second, our Moonshot crowdsourcing platform has solicited over 10,000 submissions with well-described design rationales; most of these designs are potentially developable, and discarded at this stage only because of relative synthetic complexity. Third, our structure-enabled design paradigm allows us to easily combine design ideas from all three lead series, arriving at a new chemical space. We anticipate the three series to have different potency, ADMET and in vivo activity profiles, thus even if none of the series can individually progress to preclinical candidate, we anticipate the merge space of the series to be a fertile ground.

2. **We identified alternative animal models and partners.** Our scientific understanding of the symptoms and effects of COVID-19 is still incomplete, and as such we acknowledge that the correlation between any animal model and human pathology is still a matter of active scientific research. Although we believe the available evidence to date suggests that a hamster model is the most promising, we have also obtained positive responses from collaborators at Imperial College London (Prof Wendy Barclay) and Charles River Laboratories to run ferret models. As in vivo infection study will be executed only in the later years of the proposal, we will constantly appraise the literature and adjust our in vivo strategy if needed. Annette von Delft, a co-investigator on our proposal, will be responsible for evaluating and updating in vivo strategy in consultation with our clinical development partners, LifeArc and DNDi.

3. **We have the expertise to perform virology experiments to assess the mutation risk of our preclinical candidates.** The virology of SARS-CoV-2 and the mutational fitness landscape of Mpro is still a rapidly evolving field. More broadly, there has been no FDA-approved Mpro inhibitor against human coronavirus. Promisingly, the literature on feline coronavirus suggests that the best-in-class Mpro inhibitor is robust to mutations with no significant increase in EC_{50} after 20 rounds of serial passaging.[4] We will proactively monitor strains reported in the literature and global sequencing databases to track evidence of relevant Mpro mutations, assessing whether to add these to our assay panels. Although currently not part of our research plan, our virology laboratory has the expertise to perform serial passaging experiments to assess the emergence of mutations against our lead candidate and its impact on viral fitness, which can be used as an additional metric to select the final preclinical candidate. We will adapt our virology experiments as the understanding of SARS-CoV-2 evolves

Assay Cascade Details

Fluorescence MPro assay: (London lab, Weizmann; see **Letter**) Compounds are dispensed into 384-well plates using an Echo acoustic dispenser. Mpro assay reagents are dispensed into the assay plate in 10 μL volumes for a final working volume of 20 μL. Final reaction concentrations are 20 mM HEPES pH 7.3, 1 mM TCEP, 50 mM NaCl, 0.01% Tween-20, 10% glycerol, 5 nM Mpro, 375 nM fluorogenic peptide substrate ([5-FAM]-AVLQSGFR-[Lys(Dabcyl)]-K-amide). Mpro is pre-incubated for 15 minutes at room temperature with compound before addition of substrate. Protease reaction is measured continuously in a BMG Pherastar FS with a 480/520 ex/em filter set. Screening assays are performed in duplicate at 20 μM and 50 μM. Compounds showing $>$50% inhibition at either 20 or 50 μM are selected for 12-point dose response curves with 100–0.05 μM compound to obtain IC_{50}s.

RapidFire MPro assay: (F von Delft, Oxford; see **Letter**)) Compounds are dispensed into 384-well plates using an Echo acoustic dispenser ($<$1% DMSO). Enzyme solution, containing 20 nM Mpro in 20 mM HEPES pH 7.5 50 mM NaCl, is added to each well and incubated with inhibitor for 15 min at RT. 2.0 μM substrate (TSAVLQSGFRK) is added, and the reaction quenched after 10 min with 10% formic acid and injected into an Agilent RapidFire LC-MS. Initial screens are performed in duplicate at 20 μM and 50 μM; compounds showing $>$50% inhibition at either are selected for 11-point dose response curves with 100–0.0017 μM compound to obtain IC_{50}s.

High-throughput X-ray crystallography: (F. von Delta, Oxford; see **Letter**) The high-throughput crystallography protocol closely follows that previously reported for the initial fragment screen.[1] Compounds received at Oxford/Diamond are dissolved in 100% DMSO at varying stock concentrations and soaked into Mpro crystals using an Echo 555 acoustic dispenser. X-ray reflection data for looped crystals is collected at beamline I04-1 at 100 K using automated instrumentation. Ligand-binding events are identified using PanDDA, solved using Coot, and refined with Refmac and Buster as described in.[1] So far, Diamond has produced $>$200 X-ray structures (average resolution 1.8 Å) for novel compounds bound to SARS-CoV-2 Mpro since Feb 2020.

Cellular target inhibition assays: (A. von Delft; see **Letter**) To assess intracellular target inhibition of small molecule inhibitors of the 3-CL main protease of SARS-CoV-2 (MPro), a NanoBRET assay is being established based on an adapted protein:protein interaction protocol by Promega. In brief, two vectors containing MPro NSP5 3-CL protease (pVLX-Nsp5 WT, plus inactive control pVLX-Nsp5 C145A), and a protease cleavage site flanked by NanoLuc (donor) and Halo (acceptor) tag (pCI3P-Nsp7-8) are co-transfected into HEK293 cells. After 16 hours, transfected cells are re-seeded into 96-well plates, and a HaloTag ligand added to enable the bioluminescence resonance energy transfer (BRET) transfer between the two tags of the protease cleavage vector. After 24 hours, NanoLuc substrate is added, and donor and acceptor signals measured on a GloMax Discover System to calculate NanoBRET ratios. MPro 3-CL protease activity is associated with a reduction in Halotag fluorescence at 618 nm, whereas inhibition of the protease through small molecule inhibitors is associated with increased emission.

Thermodynamic aqueous solubility: (AstraZeneca via Northeastern; see **Letter**) This assay, using established methods at AZ, is performed by adding drug to buffer, agitating for 16 h, and measuring soluble drug concentration.

Microsomal and hepatocyte stability: (AstraZeneca and Charles River Labs via Northeastern; see **Letter**) In vitro intrinsic clearance is determined from human and mouse liver microsomes and rat hepatocytes. Following incubation of NADPH with microsomes or hepatocytes for 0 and 60 min, samples are analyzed with LC/MS to determine percentage of drug remaining.

Log $D_{7.4}$: (AstraZeneca via Northeastern; see **Letter**) Shake-flask octanol-water distribution coefficients are measured in 10 mM sodium phosphate pH 7.4 buffer using a method validated for Log $D_{7.4}$ ranging from -2 to 5.0.

Human Plasma Protein Binding (PPB): (AstraZeneca via Northeastern; see **Letter**) PPB is determined using equilibrium dialysis (RED device) to separate free from bound compound. The amount of compound in plasma (10 μM initial concentration) and in dialysis buffer (pH 7.4 phosphate buffer) is measured by LC-MS/MS after equilibration at 37°C in a dialysis chamber. The fraction unbound (f_u) is reported.

Choice of cell lines for antiviral assays: The widely used African green monkey kidney epithelial Vero E6 cell line (Vero 76, clone E6, ATCC CRL-1586) was used for SARS-CoV-2 viral stock production, as this cell line propagates virus fast and can be readily infected with multiple viral strains, including SARS-CoV-2.[31] Subsequently, a lung cell line was chosen to assess for compound toxicity and antiviral activity in plaque assays with the aim of provide a more physiological readout for antiviral activity of small molecule inhibitors, as lung and intestine are the main anatomical sites where SARS-CoV-2 RNA has been isolated in patients.[32] Commonly used lung cell lines (such as A549 cells) were not susceptible to SARS-CoV-2 infection, whereas Calu-3 lung cells showed high rates of SARS-CoV-2 replication and were used for subsequent experiments (ATCC HTB-55).[31] All used cell lines were verified using short tandem repeat (STR) profiling (DNA fingerprinting) and screened for mycoplasma before use.

Compound toxicity: (A. von Delft; see **Letter**) Cell viability will be measured using the CellTiter 96R AQueous One Solution Cell Proliferation MTA (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H - 15 tetrazolium, inner salt) Assay (Promega) according to the manufacturer's instruction. Briefly, Calu-3 lung cells are treated in 9 half-log concentrations in triplicate for 2 d. As controls, wells with 200 μL growth media with and without cells are included in triplicate. After the incubation, 20 μL of MTS reagent is added to the 200 μL of media in each well. After a further 1–2 h incubation, absorbance at 490 nm is measured on a Molecular Devices SpectraMax M5 microplate reader. The evaluation endpoint is taken when untreated cells reach absorbance $\sim$1.

SARS-CoV-2 antiviral assays: (A. von Delft; see **Letter**) Antiviral efficacy assays of novel small molecules will be defined through viral replication and plaque assay performed in BSL-3 containment facilities at the University of Oxford. For antiviral assays, SARS-CoV-2 viral isolate England02/2020 is propagated in African green monkey kidney epithelial Vero E6 cells. Experiments are performed using passage 3 virus that had been sequence-confirmed by the Oxford Genomics Centre. Calu-3 cells are infected with SARS-CoV-2 MOI = 0.5, and incubated with novel antiviral inhibitors, positive and negative controls in triplicate for 72 hours at 37°C. For viral replication assays, after 24 hours incubation time cell culture supernatant are added to TRIzol reagent for RNA isolation and EC50 curves generated based on inhibition levels observed with qPCR. For plaque-forming assay readouts, supernatant is collected after 72 hours incubation time from individual virus samples, and Vero E6 cells infected with tenfold serial dilutions. Overlay medium is added after an incubation time of 1 h, and incubated at 37°C for 3 days before fixation with 4% formalin. Formed viral plaques are visualized using crystal violet solution and measured using an AID systems EliSPOT reader.

Caco-2 permeability: (Charles River Labs via Northeastern; see **Letter**) Caco-2 permeability of test compounds is measured using an established protocol at Charles River Labs by measuring peak area ratios (to assess test compound concentration) at T=0 and 2 h on the apical and basolateral sides of the basal plate.

CYP induction: (Charles River Labs via Northeastern; see **Letter**) The assay is performed using established protocols at Charles River Labs by adding the test compound to the positive or negative controls, incubating for 3 h at 37°C, following determination of the fold-induction relative to the vehicle control by LC/MS/MS analysis. (Providing party: Charles River Labs via Northeastern University).

CYP inhibition: (Charles River Labs via Northeastern; see **Letter**) CYP inhibition is measured using established protocols at Charles River Labs. A 50 mM stock solution was prepared and serially diluted (1.5-fold and 4-fold), this was then combined with a buffer/cofactor/substrate solution before combining with the human liver microsomes. The resulting mixture was incubated at 37°C for 30 mins with gentle shaking before the percentage inhibition was calculated relative to zero inhibition by LC/MS/MS analysis.

In vivo pharmacokinetic (PK) studies (Jeremiah Momper, UCSD; see **Letter** and **Vertebrate Animals** section): Single-dose pharmacokinetic studies in both rats and hamsters will be performed for investigational agents that meet Criteria B (expected $\leq$5 compounds/year). Six rodents will be used per compound (3 male, 3 female). Each rodent will receive a 20 mg/kg dose of the test compound via oral administration. PK samples (50 μL blood) will be taken via retro-orbital bleeds at T=0 (pre-dose) and 0.5, 1, 2, 4, 8, and 24 hr. Plasma concentrations will be determined by LC-MS/MS. Non-compartmental PK parameters (AUC, CL, Vss, T1/2, Cmax, Tmax, Clast, Tlast) will be estimated using Phoenix WinNonlin 6.3. Studies will be conducted under existing approved IACUC protocols at UC San Diego (Approval S17023, Date: 12/19/19; PI: Momper) at the UCSD Drug Development Pipeline Recharge Facility under supervision of Dr. Momper, a Consultant to this proposal, via a fee-for-service model.

In vivo efficacy: NIAIDs Preclinical Services has stated their in vivo SARS-CoV-2 animal models are a potential source to serve our needs; other opportunities (e.g. Syrian hamster disease model[28] offered by Charles Rivers Labs) are being continually explored. In addition, we plan to conduct tolerability studies in healthy animals (CRL or Momper lab, contingent on the site of the efficacy model), controlling for sex as a biological variable.

Serial passages of SARS-CoV-2 to identify resistance mutations to lead compounds: (A. von Delft, Oxford; see **Letter**) Sequential in vitro passage experiments using wild-type SARS-CoV-2 in the presence of lead compounds can be performed to select resistant viruses. Briefly, Calu-3 cells are infected with SARS-CoV-2 at a MOI of 0.5 in the presence of lead compounds. At each passage, supernatants containing viruses are passed on to fresh cells in the presence of lead compounds, with control mock virus passaged in the absence of drug. Virus titers at certain passage numbers are determined by the 50% tissue culture infective dose assay and the fold changes in EC50 values relative to the wild-type virus were determined. After 10 passages in the presence or absence of lead compound, total viral RNA is isolated using the RNeasy Mini kit (Invitrogen) and the 3CLpro sequenced following amplification by RT-PCR and analyzed for the presence of mutations.

References

[1] **Douangamath A**, Fearon D, Gehrtz P, Krojer T, Lukacik P, Owen CD, Resnick E, Strain-Damerell C, Ábrányi-Balogh P, Brandaõ-Neto J, et al. Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease. bioRxiv. 2020; .

[2] **Jin Z**, Du X, Xu Y, Deng Y, Liu M, Zhao Y, Zhang B, Li X, Zhang L, Peng C, et al. Structure of M pro from SARS-CoV-2 and discovery of its inhibitors. Nature. 2020; p. 1–5.

[3] **Liu Y**, Liang C, Xin L, Ren X, Tian L, Ju X, Li H, Wang Y, Zhao Q, Liu H, et al. The Development of Coronavirus 3C-Like Protease (3CLpro) Inhibitors from 2010 to 2020. European Journal of Medicinal Chemistry. 2020; p. 112711.

[4] **Kim Y**, Liu H, Galasiti Kankanamalage AC, Weerasekara S, Hua DH, Groutas WC, Chang KO, Pedersen NC. Reversal of the progression of fatal coronavirus infection in cats by a broad-spectrum coronavirus protease inhibitor. PLoS pathogens. 2016; 12(3):e1005531.

[5] **Lin JH**, Ostovic D, Vacca JP. The integration of medicinal chemistry, drug metabolism, and pharmaceutical research and development in drug discovery and development. In: *Integration of Pharmaceutical Discovery and Development* Springer; 2002.p. 233–255.

[6] **Bauer RA**. Covalent inhibitors in drug discovery: from accidental discoveries to avoided liabilities and designed therapies. Drug discovery today. 2015; 20(9):1061–1073.

[7] **Collins PM**, Ng JT, Talon R, Nekrosiute K, Krojer T, Douangamath A, Brandao-Neto J, Wright N, Pearce NM, Von Delft F. Gentle, fast and effective crystal soaking by acoustic dispensing. Acta Crystallographica Section D: Structural Biology. 2017; 73(3):246–255.

[8] **Resnick E**, Bradley A, Gan J, Douangamath A, Krojer T, Sethi R, Geurink PP, Aimon A, Amitai G, Bellini D, et al. Rapid covalent-probe discovery by electrophile-fragment screening. Journal of the American Chemical Society. 2019; 141(22):8951–8968.

[9] **Cournia Z**, Allen B, Sherman W. Relative binding free energy calculations in drug discovery: recent advances and practical considerations. Journal of chemical information and modeling. 2017; 57(12):2911–2937.

[10] **Abel R**, Wang L, Harder ED, Berne B, Friesner RA. Advancing drug discovery through enhanced free energy calculations. Accounts of chemical research. 2017; 50(7):1625–1632.

[11] **Wang L**, Wu Y, Deng Y, Kim B, Pierce L, Krilov G, Lupyan D, Robinson S, Dahlgren MK, Greenwood J, et al. Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. Journal of the American Chemical Society. 2015; 137(7):2695–2703.

[12] **Bennett CH**. Efficient estimation of free energy differences from Monte Carlo data. Journal of Computational Physics. 1976; 22(2):245–268.

[13] **Crooks GE**. Excursions in statistical dynamics. PhD thesis, Citeseer; 1999.

[14] **Gapsys V**, Michielssens S, Peters JH, de Groot BL, Leonov H. Calculation of binding free energies. In: *Molecular Modeling of Proteins* Springer; 2015.p. 173–209.

[15] **Eastman P**, Swails J, Chodera JD, McGibbon RT, Zhao Y, Beauchamp KA, Wang LP, Simmonett AC, Harrigan MP, Stern CD, et al. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. PLoS computational biology. 2017; 13(7):e1005659.

[16] **Maier JA**, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. Journal of chemical theory and computation. 2015; 11(8):3696–3713.

[17] **Jorgensen WL**, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. The Journal of chemical physics. 1983; 79(2):926–935.

[18] **Jang H**, Maat J, Qiu Y, Smith DGA, Boothroyd S, Wagner J, Bannan CC, Gokey T, Lim VT, Lucas X, Tjanaka B, Shirts MR, Gilson MK, Chodera JD, Bayly CI, Mobley DL, Wang LP, openforcefield/openforcefields: Version 1.2.0 "Parsley" Update. Zenodo; 2020. https://doi.org/10.5281/zenodo.3872244.

[19] **Wang J**, Wang W, Kollman PA, Case DA. Automatic atom type and bond type perception in molecular mechanical calculations. Journal of molecular graphics and modelling. 2006; 25(2):247–260.

[20] **Wang J**, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. Journal of computational chemistry. 2004; 25(9):1157–1174.

[21] **Jakalian A**, Bush BL, Jack DB, Bayly CI. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. Journal of computational chemistry. 2000; 21(2):132–146.

[22] **Jakalian A**, Jack DB, Bayly CI. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. Journal of computational chemistry. 2002; 23(16):1623–1641.

[23] **Gapsys V**, Pérez-Benito L, Aldeghi M, Seeliger D, Van Vlijmen H, Tresadern G, de Groot BL. Large scale relative protein ligand binding affinities using non-equilibrium alchemy. Chemical Science. 2020; 11(4):1140–1152.

[24] **Rufa DA**, Macdonald HEB, Fass J, Wieder M, Grinaway PB, Roitberg AE, Isayev O, Chodera JD. Towards chemical accuracy for alchemical free energy calculations with hybrid physics-based machine learning/molecular mechanics potentials. bioRxiv. 2020; .

[25] **Zimmerman MI**, Porter JR, Ward MD, Singh S, Vithani N, Meller A, Mallimadugula UL, Kuhn CE, Borowsky JH, Wiewiora RP, et al. Citizen Scientists Create an Exascale Computer to Combat COVID-19. BioRxiv. 2020; .

[26] **Schwaller P**, Laino T, Gaudin T, Bolgar P, Hunter CA, Bekas C, Lee AA. Molecular transformer: A model for uncertainty-calibrated chemical reaction prediction. ACS central science. 2019; 5(9):1572–1583.

[27] **Yang Q**, Sresht V, Bolgar P, Hou X, Klug-McLeod JL, Butler CR, et al. Molecular transformer unifies reaction prediction and retrosynthesis across pharma chemical space. Chemical Communications. 2019; 55(81):12152–12155.

[28] **Imai M**, Iwatsuki-Horimoto K, Hatta M, Loeber S, Halfmann PJ, Nakajima N, Watanabe T, Ujie M, Takahashi K, Ito M, et al. Syrian hamsters as a small animal model for SARS-CoV-2 infection and countermeasure development. Proceedings of the National Academy of Sciences. 2020; 117(28):16587–16595.

[29] **Pennington LD**, Aquila BM, Choi Y, Valiulin RA, Muegge I. Positional Analogue Scanning: An Effective Strategy for Multiparameter Optimization in Drug Design. Journal of medicinal chemistry. 2020; .

[30] **Jacobs J**, Grum-Tokars V, Zhou Y, Turlington M, Saldanha SA, Chase P, Eggler A, Dawson ES, Baez-Santos YM, Tomar S, et al. Discovery, synthesis, and structure-based optimization of a series of N-(tert-butyl)-2-(N-arylamido)-2-(pyridin-3-yl) acetamides (ML188) as potent noncovalent small molecule inhibitors of the severe acute respiratory syndrome coronavirus (SARS-CoV) 3CL protease. Journal of medicinal chemistry. 2013; 56(2):534–546.

[31] **Harcourt J**, Tamin A, Lu X, Kamili S, Sakthivel SK, Murray J, Queen K, Tao Y, Paden CR, Zhang J, et al. Isolation and characterization of SARS-CoV-2 from the first US COVID-19 patient. BioRxiv. 2020; .

[32] **Wu Y**, Guo C, Tang L, Hong Z, Zhou J, Dong X, Yin H, Xiao Q, Tang Y, Qu X, et al. Prolonged presence of SARS-CoV-2 viral RNA in faecal samples. The lancet Gastroenterology & hepatology. 2020; 5(5):434–435.