slides will be posted at: http://choderalab.org/news

OPEN SCIENCE ANTIVIRAL DISCOVERY WITH THE COVID MOONSHOT 🌙🚀 AND THE OPEN SOURCE DRUG DISCOVERY ECOSYSTEM

John D. Chodera
MSKCC Computational and Systems Biology Program
Sildes will be posted to http://www.choderalab.org/news

DISCLOSURES:
Scientific Advisory Board, OpenEye Scientific, Redesign Science*, Interline Therapeutics*, Ventus Therapeutics
All funding sources: http://choderalab.org/funding
* Denotes equity interests

3 Feb 2022 - NIH BISTI - Cyberspace

Memorial Sloan Kettering Cancer Center

Sloan-Kettering Institute

In more than 100 laboratories, our scientists are conducting innovative research to advance understanding in the biological sciences and improve human health.

Dana Pe'er

Quaid Morris

Christina Leslie

Joao Xavier

John Chodera

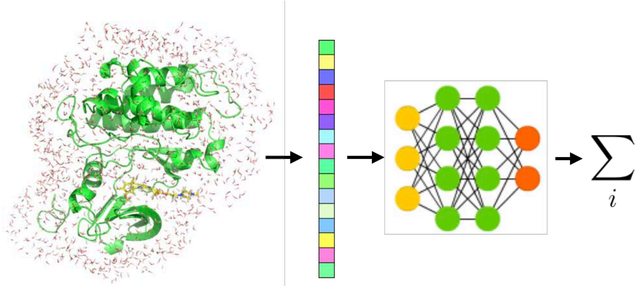
Thomas Norman

csbio@MSKCC

CHODERA LAB

We develop quantitative predictive modeling approaches to frontier problems

MODELING



FOLDING @HOME

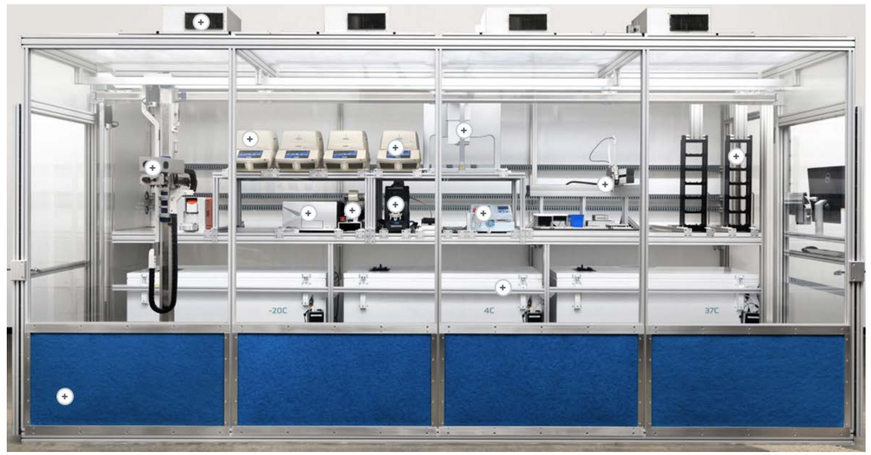
$$V(\mathbf{q}) = \sum_{\text{bonds}} K_r(r - r_{eq})^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_{eq})^2$$

$$+ \sum_{\text{dihedrals}} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] + \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\epsilon R_{ij}} \right]$$

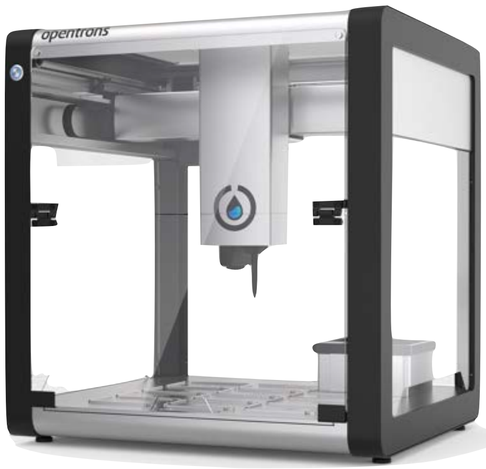
amazon web services EC2

AUTOMATION

CHODERA LAB @ MSKCC



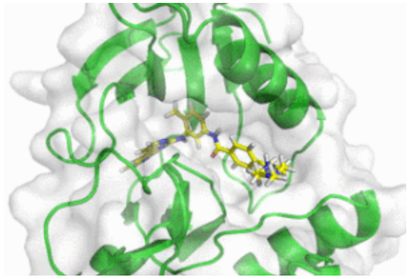
STRATEOS CLOUD WETLAB



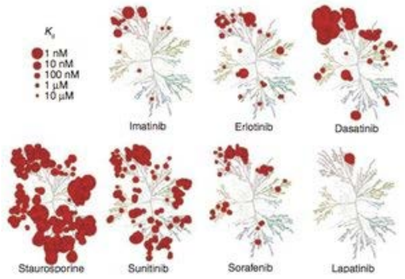
OPENTRONS

CHODERA LAB

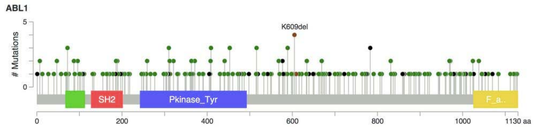
We develop quantitative predictive modeling approaches to frontier problems



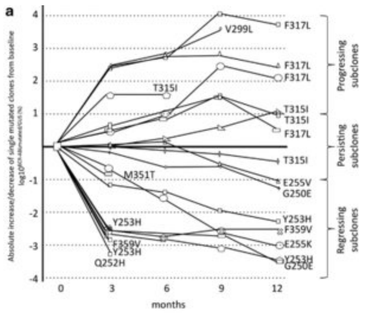
SELECTIVE INHIBITOR DESIGN: TARGETS/ANTITARGETS



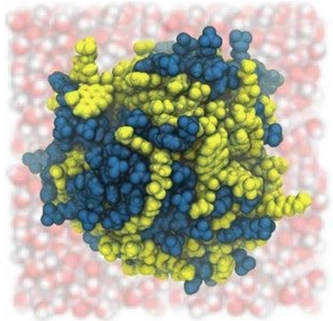
KINASE INHIBITOR SELECTIVITY



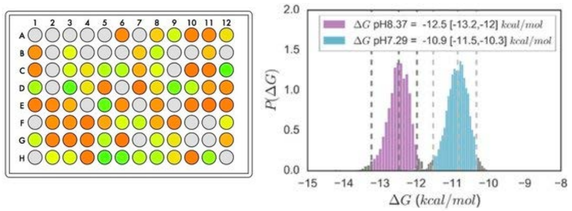
PREDICTING DRUG SENSITIVITY/RESISTANCE



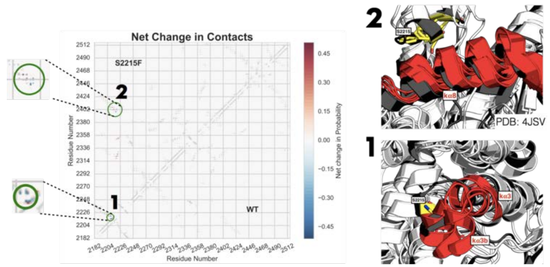
ANTICIPATING DRUG RESISTANCE



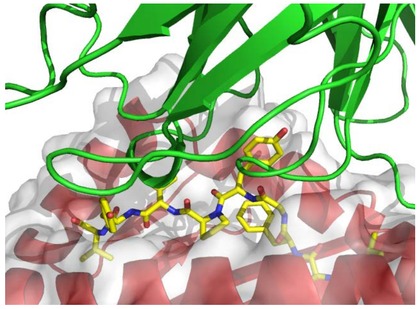
NOVEL DRUG DELIVERY MODALITIES



AUTOMATED BIOPHYSICAL ASSAYS AND INFERENCE



MECHANISMS OF ONCOGENIC ACTIVATION



CANCER IMMUNOTHERAPY

COLLABORATIONS WITH OPEN SCIENCE, OPEN SOURCE SOFTWARE, AND INDUSTRY ARE SYNERGISTIC

BAYER

open forcefield consortium

OpenMM

MolSSI
Molecular Sciences Software Institute

INTERLINE therapeutics

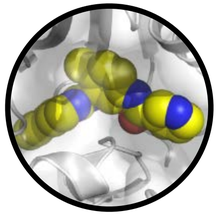
FOLDING @HOME

FORESITE LABS

open source software

COVID Moonshot

SILICON Therapeutics



choderalab
(algorithms, open science, and open source software)

industry

open science

OpenEye SCIENTIFIC

NIH
National Center for Advancing Translational Sciences

academia

VIR

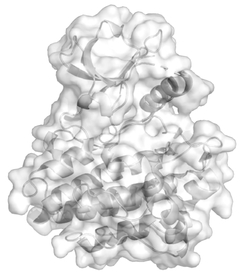
XCHEM Diamond Fragment Screening
Diamond Light Source / XChem

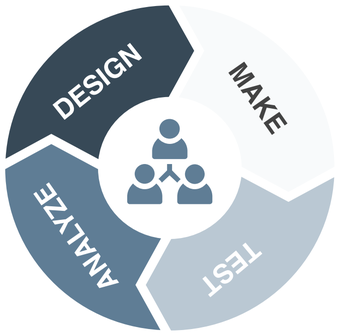
SGC

DRUG DISCOVERY AND DEVELOPMENT IS SLOW, COSTLY, AND PRONE TO FAILURE

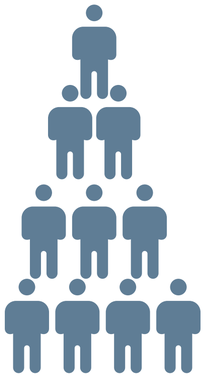
Target Product Profile (TPP)

	Target ID	Discovery	Preclinical	Phase 1 safety	Phase 2 efficacy	Phase 3 effectiveness	NDA new drug approval
% survival		48%	38%	25%	12%	7%	6%
time*		4.5 y	6.8 y	8.9 y	12.0 y	14.5 y	15 y





Human-driven design iterations



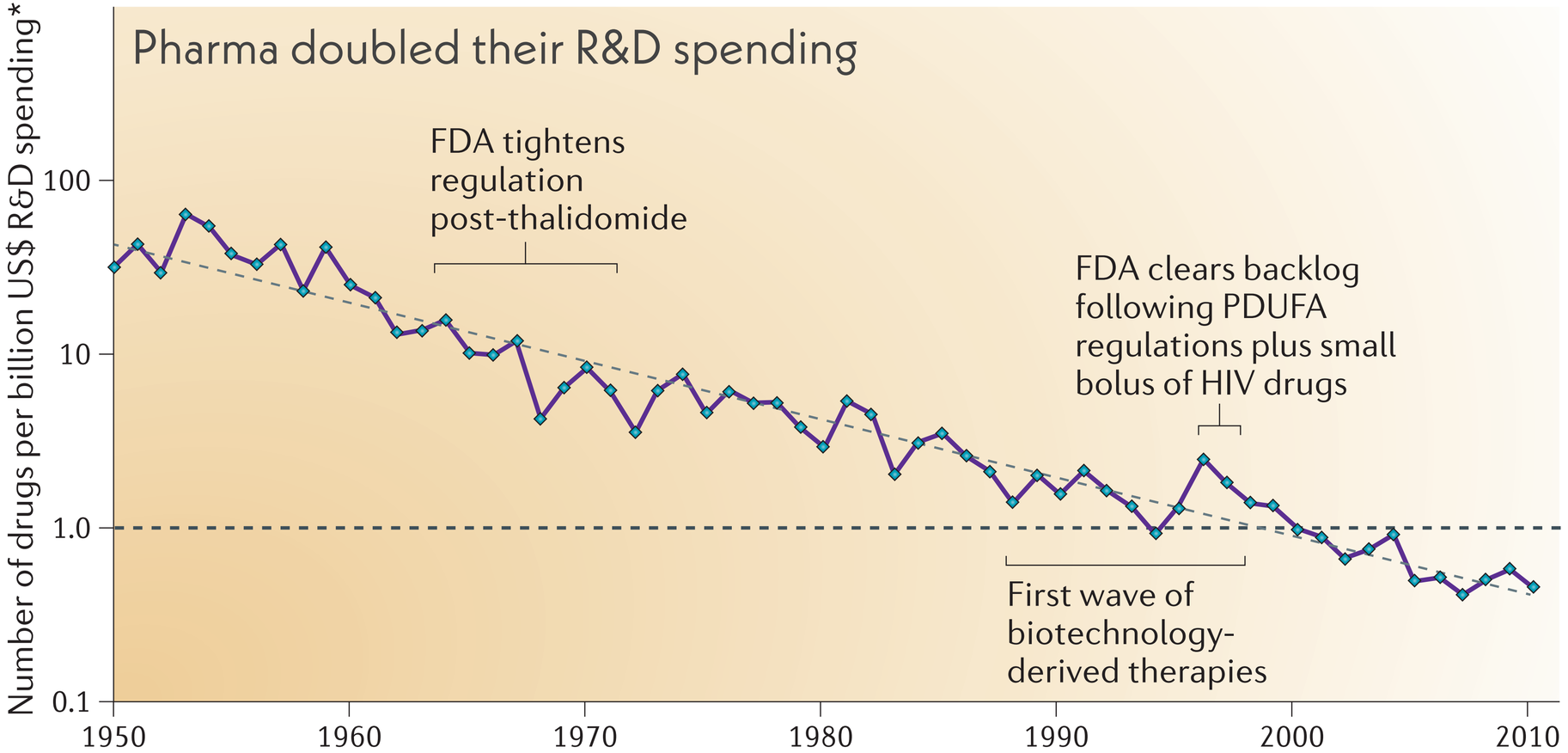
* denotes mean

sources: [1] [2] [3] [4] [5]

Global annual prescription drug **market will reach $1.6T** by 2026 [5]

DRUG DISCOVERY SEEMS TO BE GETTING HARDER

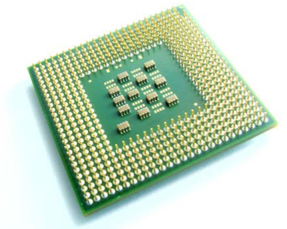
a Overall trend in R&D efficiency (inflation-adjusted)

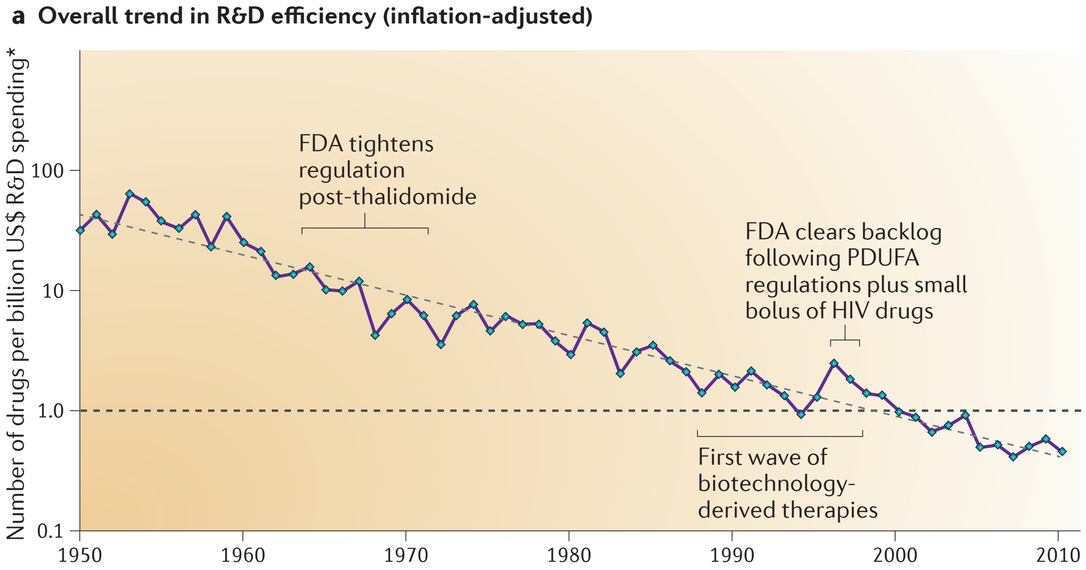


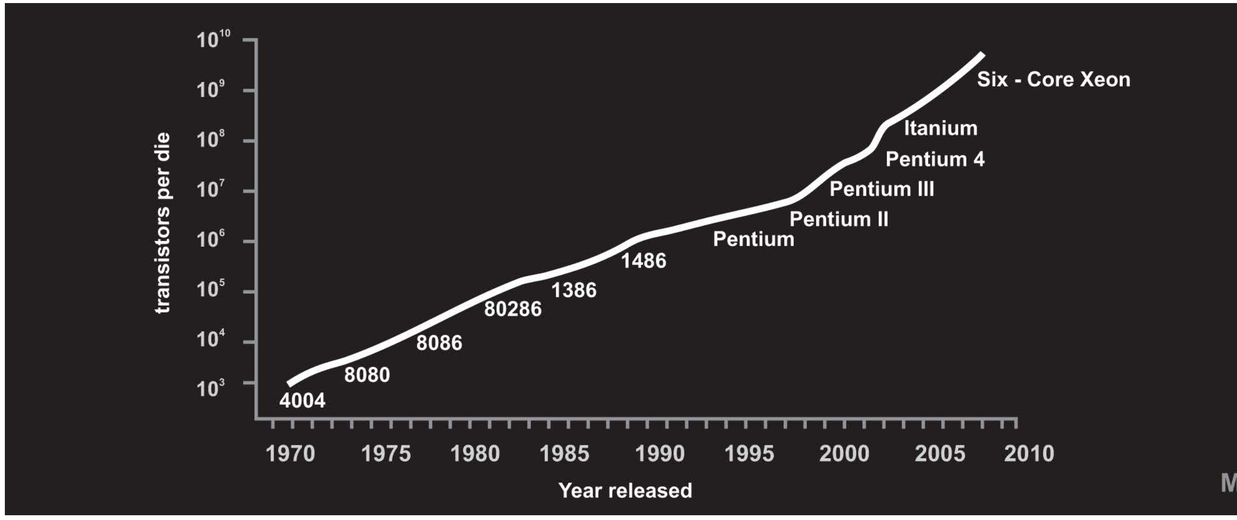
←NOW: >$2.6B/drug

Scannell et al. Nature Rev Drug Disc 11:191, 2012
DiMasi et al. J Health Econ 47:20, 2016

DRUG DISCOVERY SEEMS TO BE GETTING HARDER



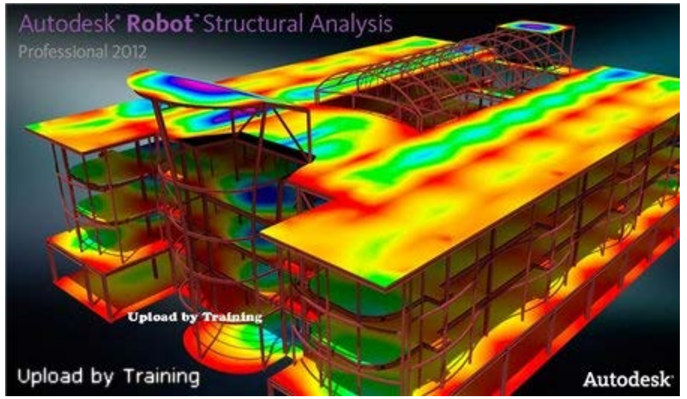




EROOM'S LAW

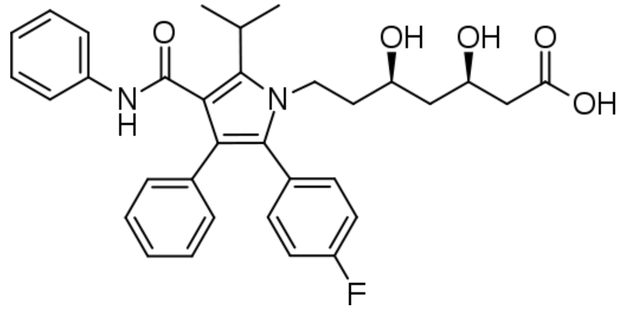
MOORE'S LAW

WE REGULARLY DESIGN PLANES, BRIDGES, AND BUILDINGS ON COMPUTERS



10^3 - 10^6 parts

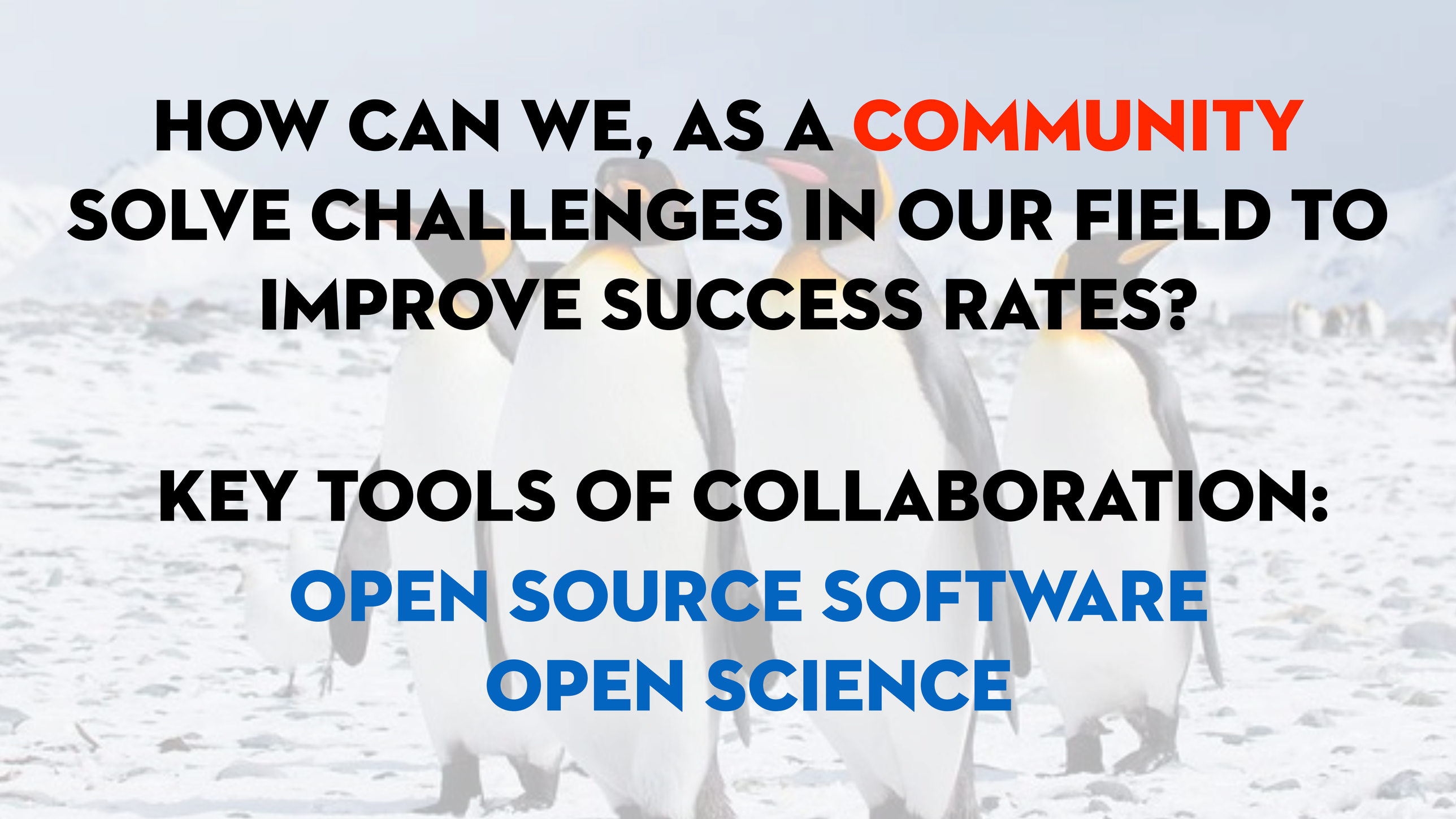
WHY NOT SMALL MOLECULE DRUGS?



< 10^2 atoms

A RISING TIDE LIFTS ALL BOATS

By working together to solve major challenges,
we can improve success rates for everybody



HOW CAN WE, AS A COMMUNITY SOLVE CHALLENGES IN OUR FIELD TO IMPROVE SUCCESS RATES?

KEY TOOLS OF COLLABORATION:

OPEN SOURCE SOFTWARE

OPEN SCIENCE

THOUGHTFUL LICENSING MODELS CAN ENSURE RESEARCH HAS MAXIMUM SCIENTIFIC IMPACT

Goal: Ensure our work has **maximum impact** by allowing use, modification, and redistribution. Aim to explicitly rescind any restrictions that would prevent this.

reproducible research product	**licenses that encourage others to build on work**
paper	**CC-BY 4.0** https://creativecommons.org/licenses/by/4.0/
data	**CC-BY 4.0**
experiment (code)	**MIT, BSD 3-clause** https://opensource.org/licenses/MIT
documentation	**CC-BY 4.0**

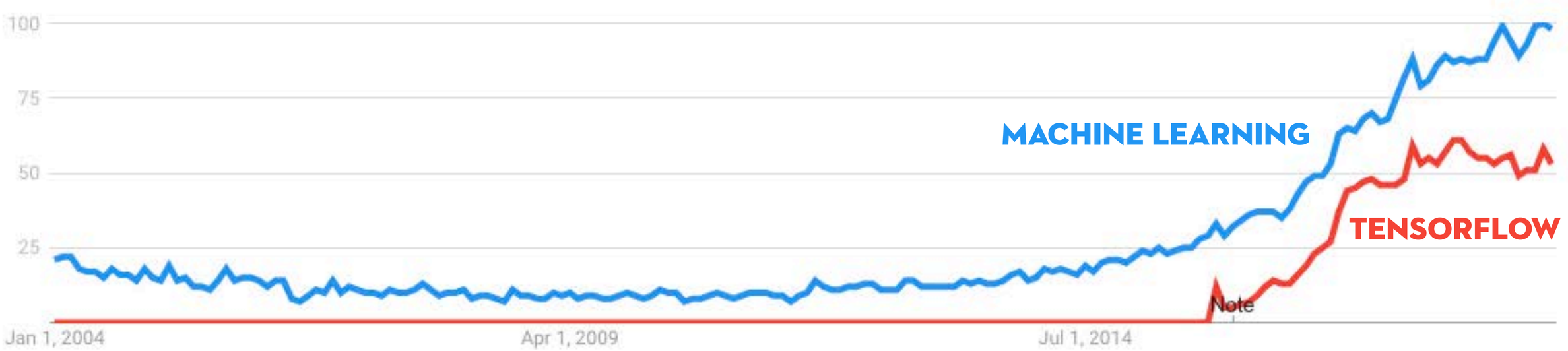
Stodden, Victoria, Enabling Reproducible Research: Open Licensing for Scientific Innovation (March 3, 2009). International Journal of Communications Law and Policy, Forthcoming. Available at SSRN: https://ssrn.com/abstract=1362040

https://web.stanford.edu/~vcs/papers/Licensing08292008.pdf

OPEN SOURCE SOFTWARE ECOSYSTEMS HAVE THE POTENTIAL TO ACCELERATE PROGRESS

WE CAN LOOK TO THE MACHINE LEARNING ECOSYSTEM FOR INSPIRATION

WE CAN LOOK TO THE MACHINE LEARNING ECOSYSTEM FOR INSPIRATION



What did TensorFlow accomplish?

* Created new opportunities
* Accelerated rate of progress

OpenMM

A high performance toolkit for molecular simulation. Use it as a library, or as an application. We include extensive language bindings for Python, C, C++, and even Fortran. The code is open source and actively maintained on Github, licensed under MIT and LGPL. Part of the Omnia suite of tools for predictive biomolecular simulation.

ABOUT FORUM GITHUB

Extreme Flexibility. Extreme Speed.

Extreme flexibility through custom forces and integrators. Extreme performance through GPU Acceleration, with optimizations for AMD, NVIDIA, and Intel Integrated GPUs. It's fast on CPUs too. See the benchmarks.

Install

Install using the conda Python package manager that powers the Omnia ecosystem.

Docs

For more information about the science, the code base, and the API behind OpenMM.

Support

For more information about filing bug reports, requesting new features, and other issues.

Resources

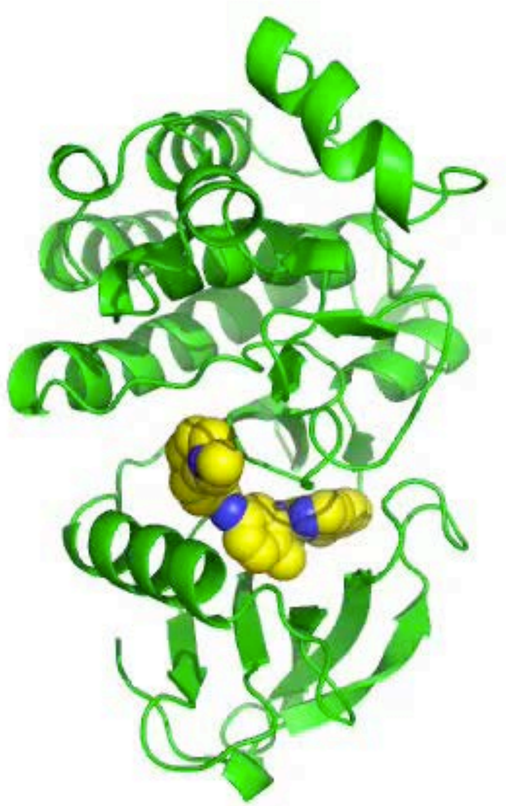
Explore additional libraries and third-party tools built around OpenMM.

Tutorials

Get started right away with OpenMM tutorials.

http://openmm.org/

OPENMM ARCHITECTURE MAKES DEVELOPMENT SIMPLE



Python Scripting	• Simulation protocols • File I/O
C++ API	• Forces • Integrators
Computational Kernels	• Optimized C++/CUDA/OpenCL code

OpenMM also has bindings for **C++**, **C**, and **FORTRAN**

OPENMM IS USED BY RESEARCHERS ALL OVER THE WORLD

Geographic statistics from http://simtk.org

CONDA-FORGE

ANACONDA CLOUD

OpenMM

http://openmm.org

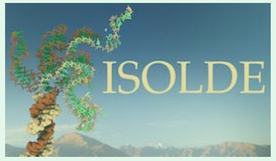
downloads 402k total

OpenMMTools

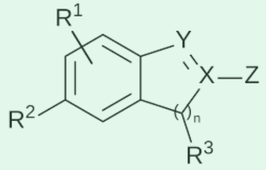
http://github.com/choderalab/openmmtools

downloads 156k total

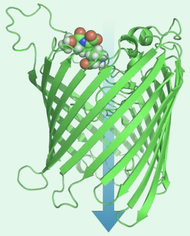
OPENMM CAN BE USED AS A LIBRARY TO ENABLE APPLICATIONS TO INTEGRATE PHYSICAL MODELING



isolde



perses



iapetus

targeted domain-specific applications
(Python, C++, C, or Fortran)

APPLICATIONS

openmmtools

high-level simulation algorithms, alchemical tools
(Python to enable rapid development)

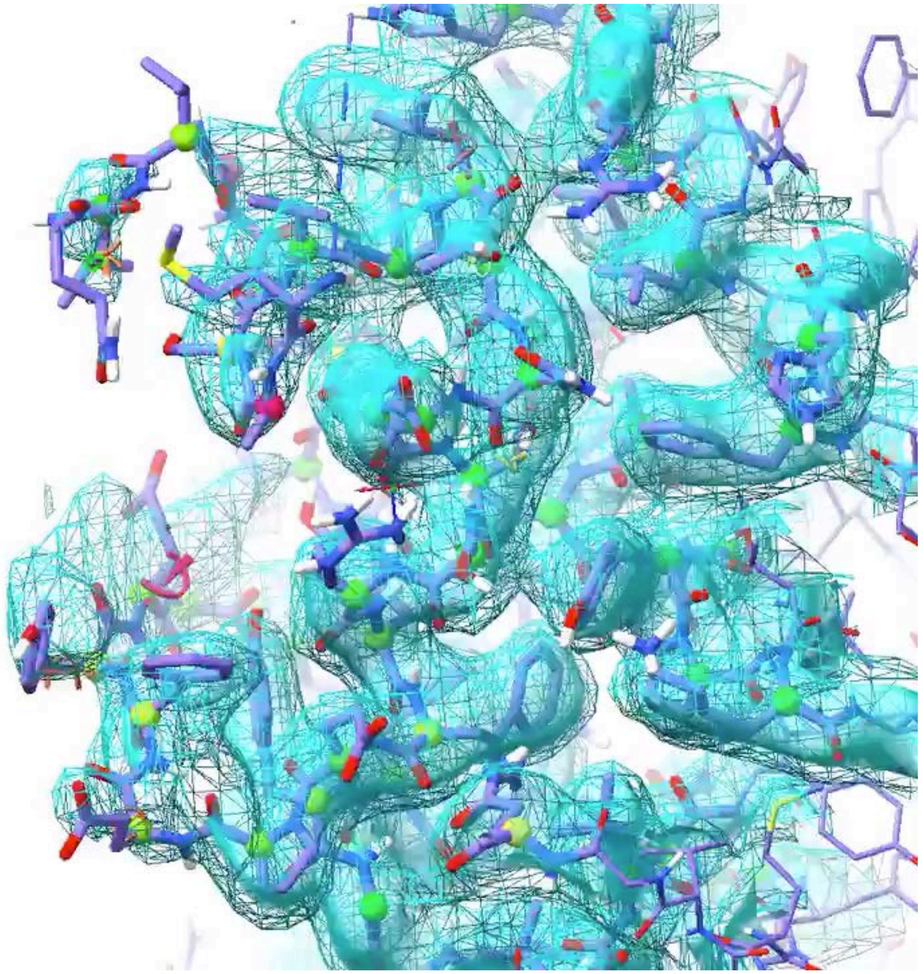
ALGORITHMS

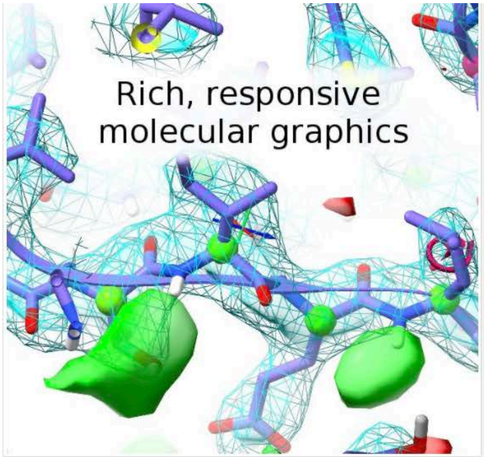
OpenMM

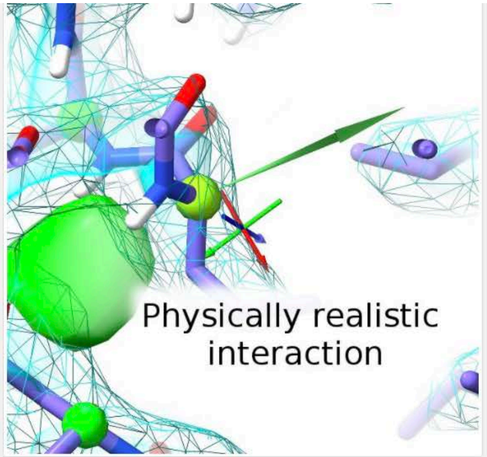
general GPU-accelerated MD simulation engine
(C++/CUDA/OpenCL with Python API)

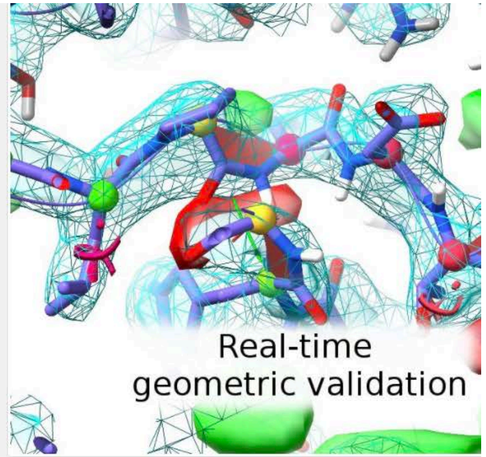
CORE

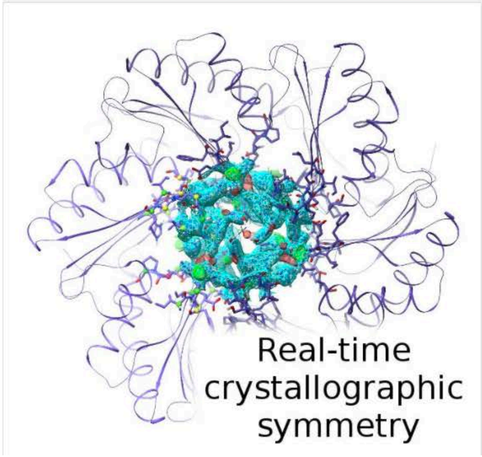
EXAMPLE: REAL-TIME SIMULATION-BASED REFINEMENT IN A VISUALIZATION PROGRAM

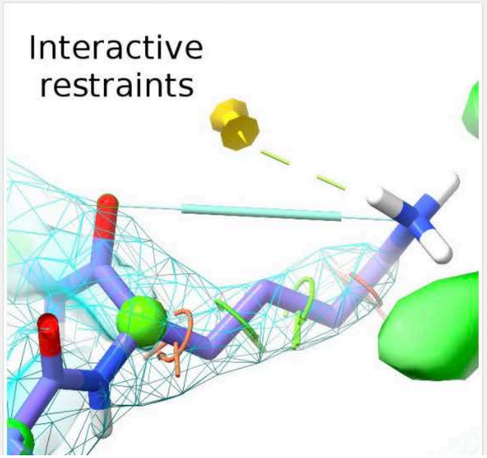


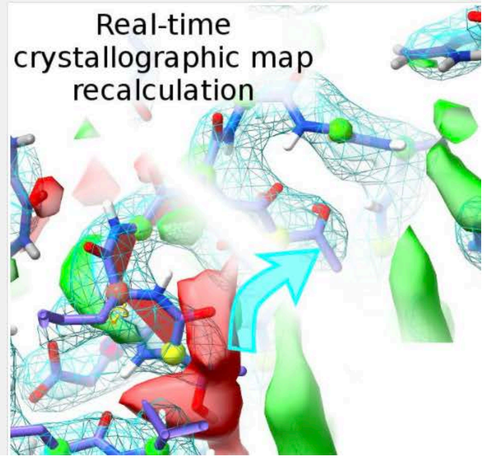












ISOLDE https://isolde.cimr.cam.ac.uk | Tristan Croll, Cambridge University

MODERN ML FRAMEWORKS PROVIDE A LEVEL OF ABSTRACTION THAT ENABLES HIGH PRODUCTIVITY AND INTEROPERABLE ECOSYSTEMS

```python
import tensorflow as tf
mnist = tf.keras.datasets.mnist

(x_train, y_train),(x_test, y_test) = mnist.load_data()
x_train, x_test = x_train / 255.0, x_test / 255.0

model = tf.keras.models.Sequential([
  tf.keras.layers.Flatten(input_shape=(28, 28)),
  tf.keras.layers.Dense(128, activation='relu'),
  tf.keras.layers.Dropout(0.2),
  tf.keras.layers.Dense(10, activation='softmax')
])

model.compile(optimizer='adam',
              loss='sparse_categorical_crossentropy',
              metrics=['accuracy'])

model.fit(x_train, y_train, epochs=5)
model.evaluate(x_test, y_test)
```

Run code now Try in Google's interactive notebook

load your tools

grab a dataset

define a new kind of model

declare your objectives in training it

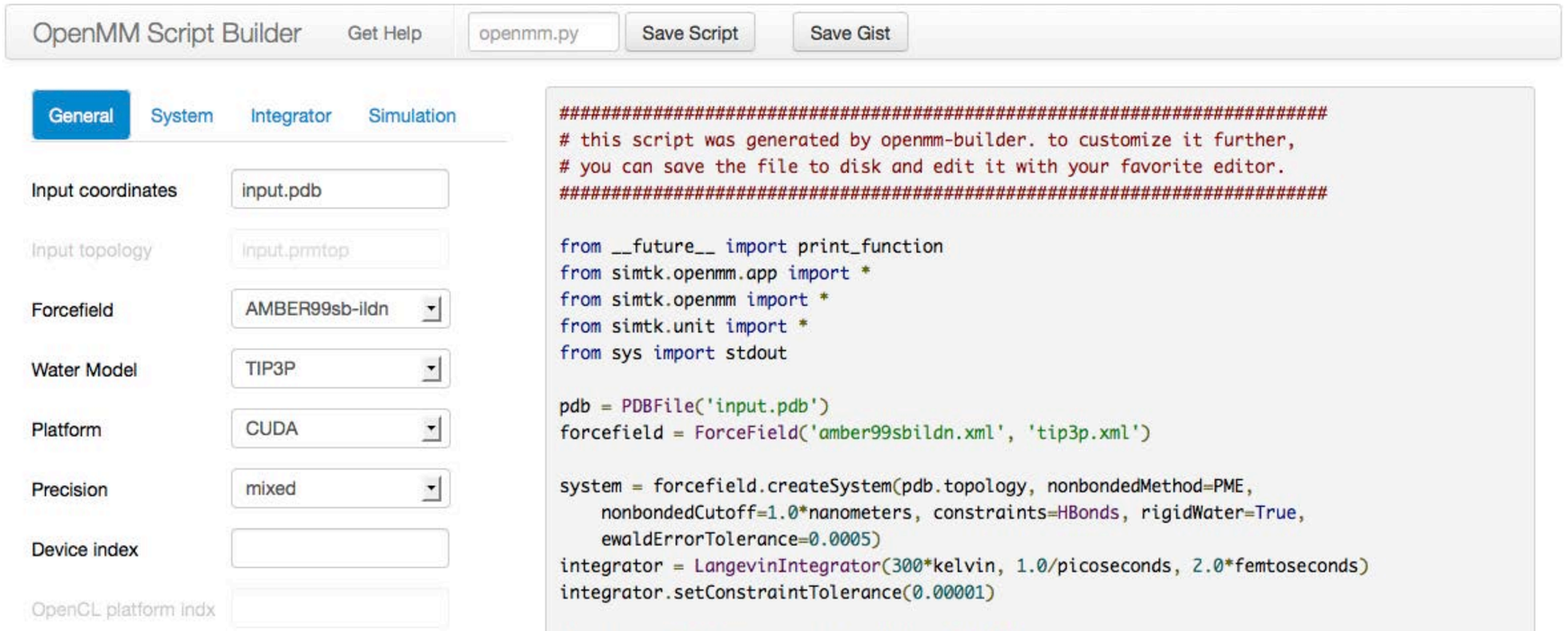
fit it

use it

https://www.tensorflow.org/overview

Why can't we make it this easy to do new things in molecular modeling?

OPENMM AIMS TO PROVIDE A CLEAR, HIGH PRODUCTIVITY API AROUND WHICH AN ECOSYSTEM CAN GROW IN PYTHON




```
#############################################################################
# this script was generated by openmm-builder. to customize it further,
# you can save the file to disk and edit it with your favorite editor.
#############################################################################

from __future__ import print_function
from simtk.openmm.app import *
from simtk.openmm import *
from simtk.unit import *
from sys import stdout

pdb = PDBFile('input.pdb')
forcefield = ForceField('amber99sbildn.xml', 'tip3p.xml')

system = forcefield.createSystem(pdb.topology, nonbondedMethod=PME,
    nonbondedCutoff=1.0*nanometers, constraints=HBonds, rigidWater=True,
    ewaldErrorTolerance=0.0005)
integrator = LangevinIntegrator(300*kelvin, 1.0/picoseconds, 2.0*femtoseconds)
integrator.setConstraintTolerance(0.00001)

platform = Platform.getPlatformByName('CUDA')
properties = {'CudaPrecision': 'mixed'}
simulation = Simulation(pdb.topology, system, integrator, platform, properties)
simulation.context.setPositions(pdb.positions)

print('Minimizing...')
simulation.minimizeEnergy()

simulation.context.setVelocitiesToTemperature(300*kelvin)
print('Equilibrating...')
simulation.step(100)

simulation.reporters.append(DCDReporter('output.dcd', 1000))
simulation.reporters.append(StateDataReporter(stdout, 1000, step=True,
    potentialEnergy=True, temperature=True))

print('Running Production...')
simulation.step(1000)
print('Done!')
```

load a structure
load a force field

define a simulation

target the GPU

minimize the energy

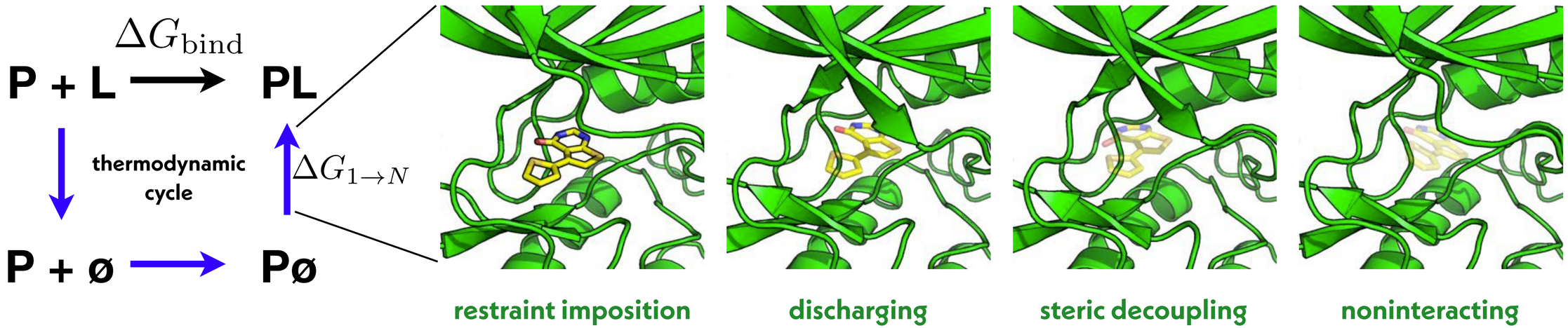
equilibrate the system

define outputs

run a simulation

http://builder.openmm.org/

ALCHEMICAL FREE ENERGY CALCULATIONS PROVIDE A RIGOROUS STRUCTURE-ENABLED WAY TO COMPUTE BINDING AFFINITIES



Includes all contributions from enthalpy and entropy of binding to a flexible receptor

$$\Delta G_{1\to N} = -\beta^{-1} \ln \frac{Z_N}{Z_1} = -\beta^{-1} \ln \frac{Z_2}{Z_1} \cdot \frac{Z_3}{Z_2} \cdots \frac{Z_N}{Z_{N-1}}$$

$$Z_n = \int dx\, e^{-\beta U_n(x)}$$ partition function

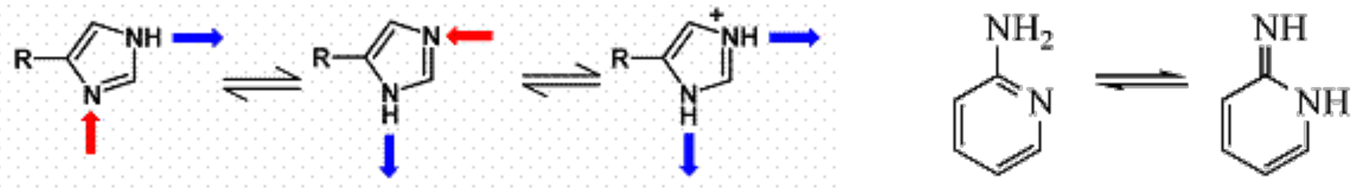
Pioneering work from many: McCammon, van Gunsteren, Kollman, Jorgensen, Chipot, Roux, Boresch, Fujitani, Pande, Shirts, Swope, Christ, Mobley, Schrödinger, and many more

PHYSICAL MODELING FOR DRUG DISCOVERY FACES THREE MAJOR CHALLENGES

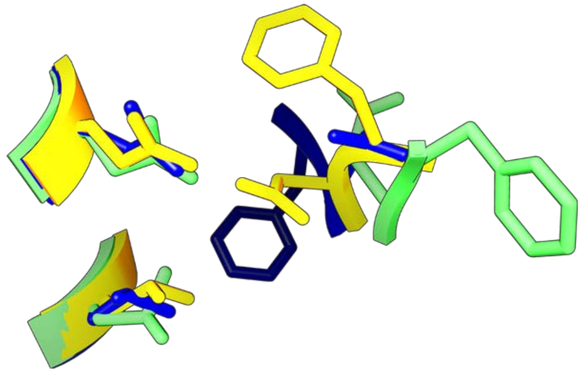
1. The forcefield does a poor job of modeling the physics of our system

$$V(\mathbf{q}) = \sum_{\text{bonds}} K_r(r - r_{eq})^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_{eq})^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] + \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\epsilon R_{ij}} \right]$$

2. We're missing some essential chemical in our simulations
(e.g. protonation states, tautomers, covalent association)



3. We haven't sampled all of the relevant conformations



open forcefield

An open and collaborative approach to better force fields

OPEN SOURCE

Software permissively licensed under the MIT License and developed openly on GitHub.

OPEN SCIENCE

Scientific reports as blog posts, webinars and preprints

OPEN DATA

Curated quantum chemical and experimental datasets used to parameterize and benchmark Open Force Fields.

NEWS TUTORIALS ROADMAP

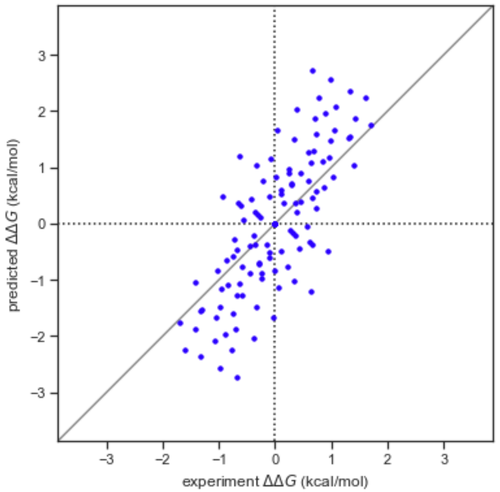
http://openforcefield.org

http://qcarchive.molssi.org

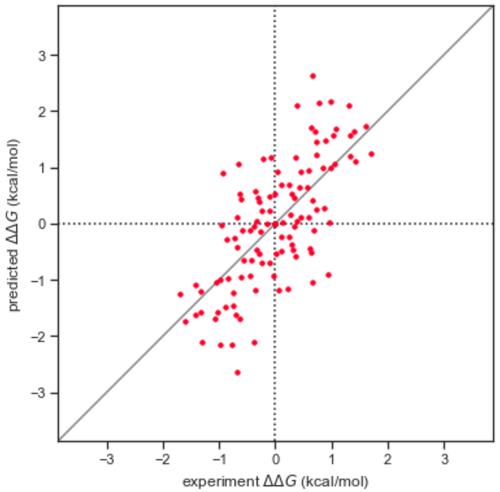
WE'VE MADE RAPID AND SIGNIFICANT PROGRESS

Open Force Field Initiative →

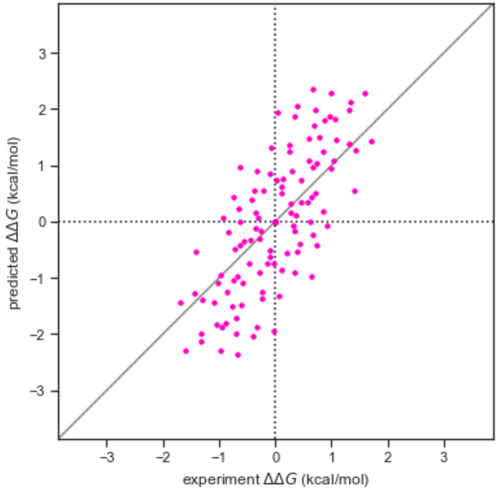
GAFF 1 (1999)



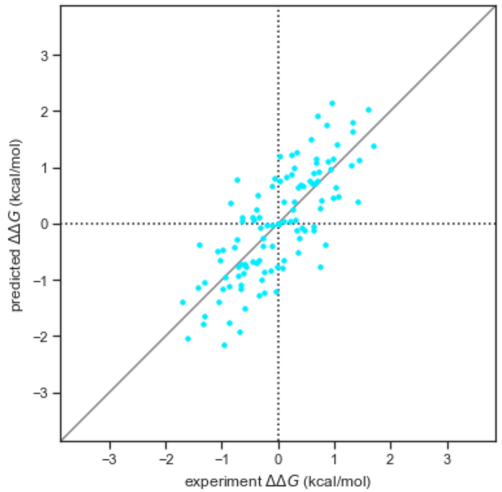
OPLS2.1 (2015)



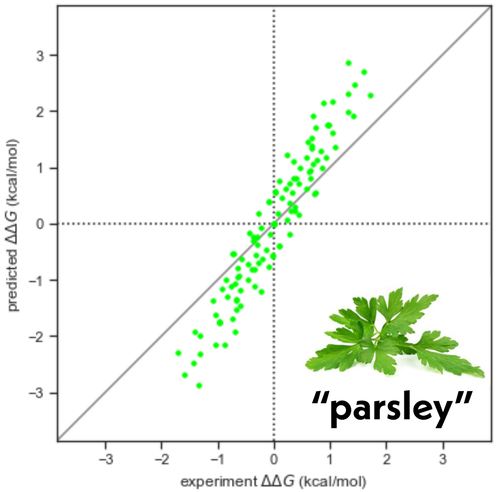
GAFF 2 (2016)



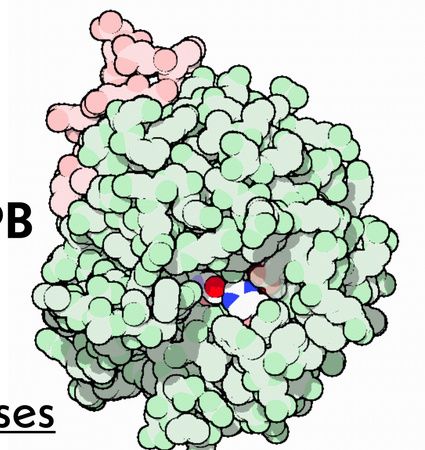
smirnoff99Frosst (2018)



openff 1.0 (2019)



thrombin
PDB101: 1PPB

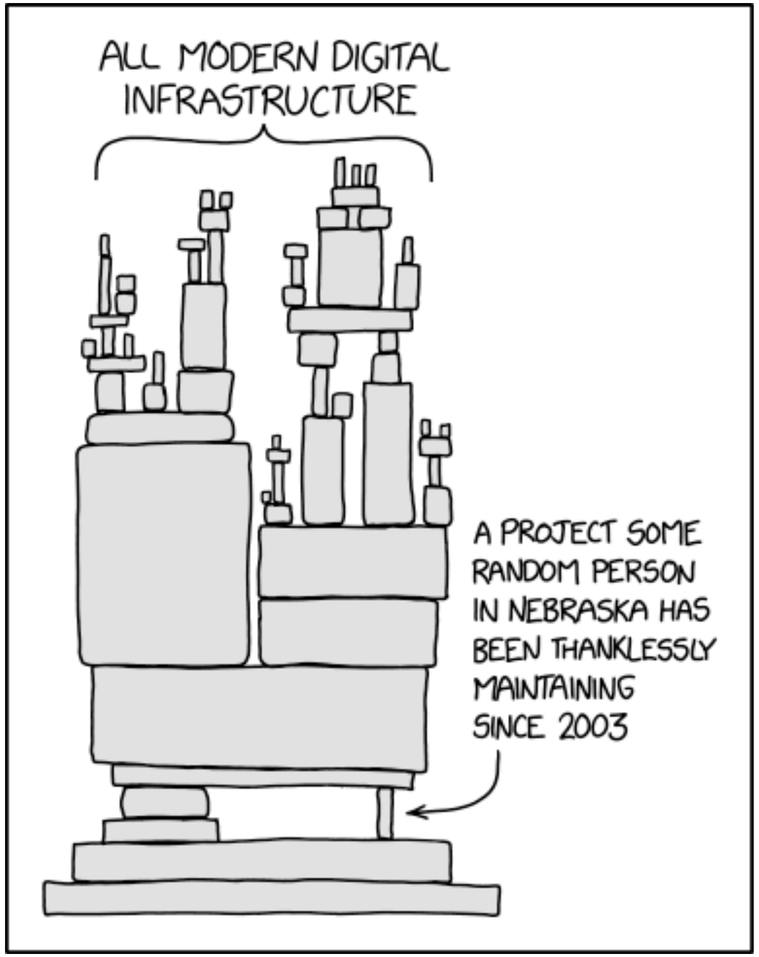


HANNAH BRUCE MACDONALD
MSKCC

http://github.com/choderalab/perses

DOMINIC RUFA

OPEN SOURCE SOFTWARE IS CHRONICALLY UNDERFUNDED—EVEN WHEN ESSENTIAL TO ENTIRE FIELDS



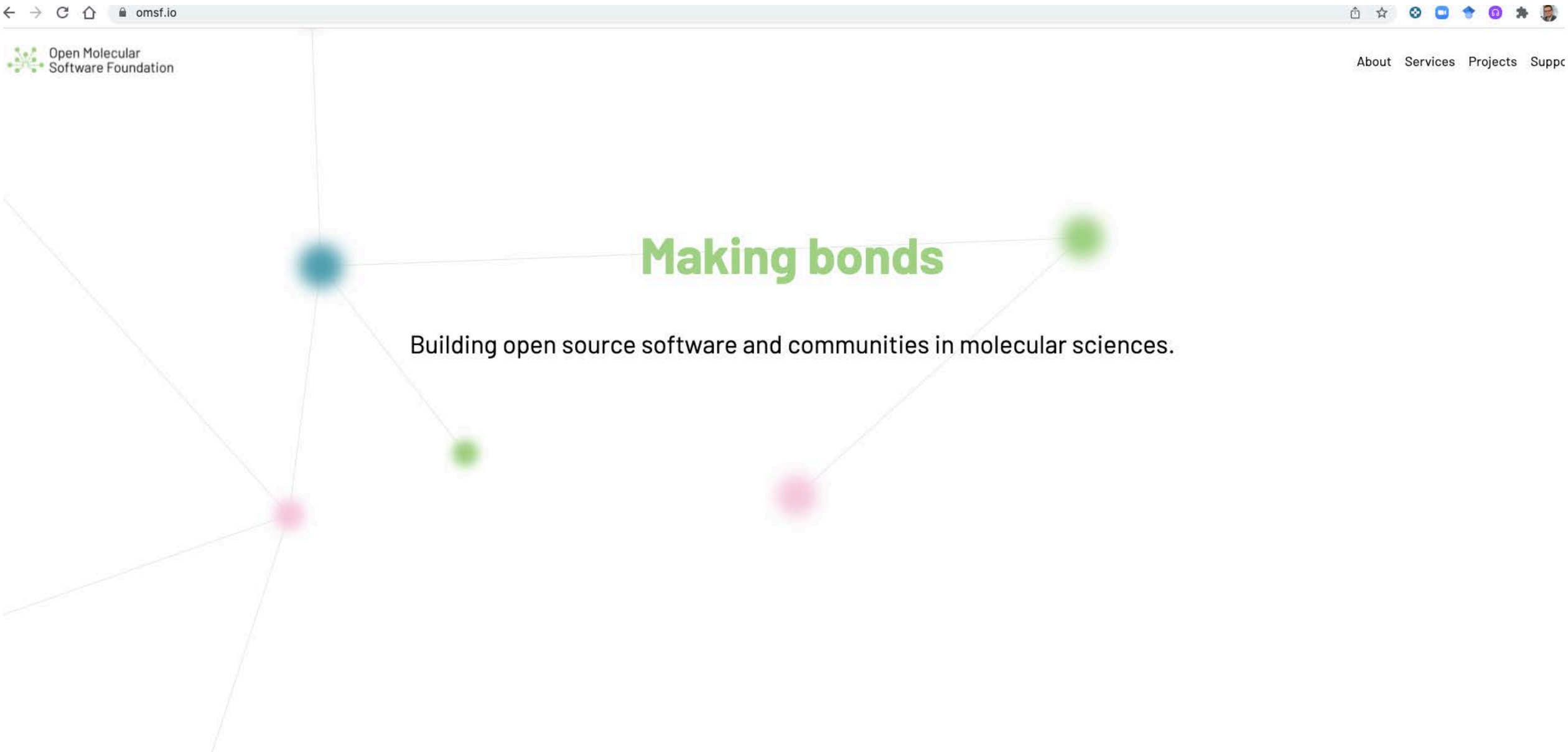
https://github.com/ParmEd/ParmEd

A project someone who works for a lightswitch company has been thanklessly maintaining since 2014

Are there better models for supporting open source software essential to science?

https://xkcd.com/2347/

THE OPEN FORCE FIELD MODEL IS SO SUCCESSFUL, WE HAVE CREATED A FOUNDATION TO HELP REPLICATE IT



COMMUNITY BLIND CHALLENGES ACCELERATE PROGRESS TO SOLUTIONS

SAMPL

Model systems of **intermediate complexity** to focus community on challenges in blind tests

Model protein-ligand systems

Isolate individual physical challenges (e.g. binding of charged ligands)

Physical properties

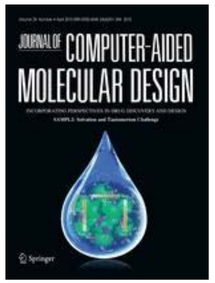
Tests of forcefield accuracy in hydrated or protein-like environments

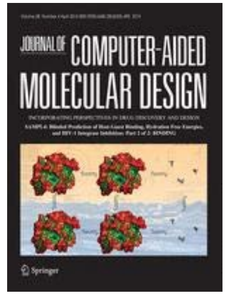
Isolate chemical effects (protonation states, ligand conformations) without slow protein timescales

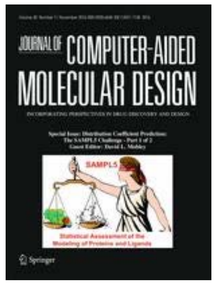
Host-guest systems

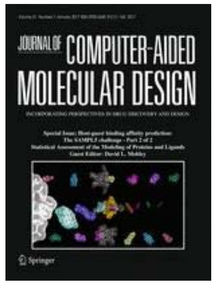
Binding of small drug-like molecules with protein-like affinities, without slow protein timescales

SAMPL0 2007	SAMPL1 2008	SAMPL2 2009	SAMPL3 2011	SAMPL4 2013	SAMPL5 2016
JNK3 kinase inhibitors hydration free energies	CDK2 kinase inhibitors hydration free energies	hydration free energies tautomer ratios	trypsin inhibitors hydration free energies	HIV-1 integrase inhibitors hydration free energies octoacid host-guest CB7 host-guest	distribution coefficients CBClip host-guest CB7 host-guest









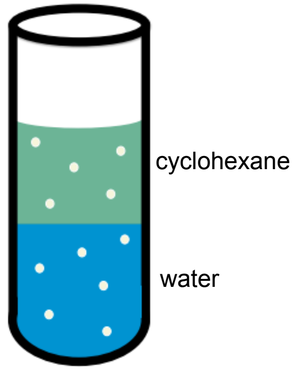
BLIND CHALLENGES CAN DRIVE PROGRESS BY FOCUSING COMMUNITY EFFORT

SAMPL0 2007	SAMPL1 2008	SAMPL2 2009	SAMPL3 2011	SAMPL4 2013	SAMPL5 2016
JNK3 kinase inhibitors	CDK2 kinase inhibitors	hydration free energies	trypsin inhibitors	HIV-1 integrase inhibitors	distribution coefficients
hydration free energies	hydration free energies	tautomer ratios	hydration free energies	hydration free energies	CBClip host-guest
				octoacid host-guest	CB7 host-guest
				CB7 host-guest	

BLIND CHALLENGES CAN DRIVE PROGRESS BY FOCUSING COMMUNITY EFFORT

SAMPL0 2007	SAMPL1 2008	SAMPL2 2009	SAMPL3 2011	SAMPL4 2013	SAMPL5 2016
hydration free energies	hydration free energies	hydration free energies	hydration free energies	hydration free energies	WE RAN OUT OF DATA!

Lots of disagreement in predictions → **Can tell when experiments are wrong** X



ALPHAFOLD2'S SUCCESS AT CASP IS AN EXAMPLE OF THE PAYOFF FOR SUSTAINED INVESTMENT

Article

Highly accurate protein structure prediction with AlphaFold

https://doi.org/10.1038/s41586-021-03819-2
Received: 11 May 2021
Accepted: 12 July 2021
Published online: 15 July 2021
Open access
Check for updates

John Jumper[1,4✉], Richard Evans[1,4], Alexander Pritzel[1,4], Tim Green[1,4], Michael Figurnov[1,4], Olaf Ronneberger[1,4], Kathryn Tunyasuvunakool[1,4], Russ Bates[1,4], Augustin Žídek[1,4], Anna Potapenko[1,4], Alex Bridgland[1,4], Clemens Meyer[1,4], Simon A. A. Kohl[1,4], Andrew J. Ballard[1,4], Andrew Cowie[1,4], Bernardino Romera-Paredes[1,4], Stanislav Nikolov[1,4], Rishub Jain[1,4], Jonas Adler[1], Trevor Back[1], Stig Petersen[1], David Reiman[1], Ellen Clancy[1], Michal Zielinski[1], Martin Steinegger[2,3], Michalina Pacholska[1], Tamas Berghammer[1], Sebastian Bodenstein[1], David Silver[1], Oriol Vinyals[1], Andrew W. Senior[1], Koray Kavukcuoglu[1], Pushmeet Kohli[1] & Demis Hassabis[1,4✉]

Proteins are essential to life, and understanding their structure can facilitate a mechanistic understanding of their function. Through an enormous experimental effort[1–4], the structures of around 100,000 unique proteins have been determined[5], but this represents a small fraction of the billions of known protein sequences[6,7]. Structural coverage is bottlenecked by the months to years of painstaking effort required to determine a single protein structure. Accurate computational approaches are needed to address this gap and to enable large-scale structural bioinformatics. Predicting the three-dimensional structure that a protein will adopt based solely on its amino acid sequence—the structure prediction component of the 'protein folding problem'[8]—has been an important open research problem for more than 50 years[9]. Despite recent progress[10–14], existing methods fall far short of atomic accuracy, especially when no homologous structure is available. Here we provide the first computational method that can regularly predict protein structures with atomic accuracy even in cases in which no similar structure is known. We validated an entirely redesigned version of our neural network-based model, AlphaFold, in the challenging 14th Critical Assessment of protein Structure Prediction (CASP14)[15], demonstrating accuracy competitive with experimental structures in a majority of cases and greatly outperforming other methods. Underpinning the latest version of AlphaFold is a novel machine learning approach that incorporates physical and biological knowledge about protein structure, leveraging multi-sequence alignments, into the design of the deep learning algorithm.

The development of computational methods to predict three-dimensional (3D) protein structures from the protein sequence has proceeded along two complementary paths that focus on either the physical interactions or the evolutionary history. The physical interaction programme heavily integrates our understanding of molecular driving forces into either thermodynamic or kinetic simulation of protein physics[16] or statistical approximations thereof[17]. Although theoretically very appealing, this approach has proved highly challenging for even moderate-sized proteins due to the computational intractability of molecular simulation, the context dependence of protein stability and the difficulty of producing sufficiently accurate models of protein physics. The evolutionary programme has provided an alternative in recent years, in which the constraints on protein structure are derived from bioinformatics analysis of the evolutionary history of proteins, homology to solved structures[18,19] and pairwise evolutionary correlations[20–24]. This bioinformatics approach has benefited greatly from the steady growth of experimental protein structures deposited in the Protein Data Bank (PDB)[5], the explosion of genomic sequencing and the rapid development of deep learning techniques to interpret these correlations. Despite these advances, contemporary physical and evolutionary-history-based approaches produce predictions that are far short of experimental accuracy in the majority of cases in which a close homologue has not been solved experimentally and this has limited their utility for many biological applications.

In this study, we develop the first, to our knowledge, computational approach capable of predicting protein structures to near experimental accuracy in a majority of cases. The neural network AlphaFold that we developed was entered into the CASP14 assessment (May–July 2020; entered under the team name 'AlphaFold2' and a completely different model from our CASP13 AlphaFold system[10]). The CASP assessment is carried out biennially using recently solved structures that have not been deposited in the PDB or publicly disclosed so that it is a blind test

[1]DeepMind, London, UK. [2]School of Biological Sciences, Seoul National University, Seoul, South Korea. [3]Artificial Intelligence Institute, Seoul National University, Seoul, South Korea. [4]These authors contributed equally: John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, Alex Bridgland, Clemens Meyer, Simon A. A. Kohl, Andrew J. Ballard, Andrew Cowie, Bernardino Romera-Paredes, Stanislav Nikolov, Rishub Jain, Demis Hassabis.
✉e-mail: jumper@deepmind.com; dhcontact@deepmind.com

https://www.nature.com/articles/s41586-021-03819-2

everyone else —>

AlphaFold2 error —>

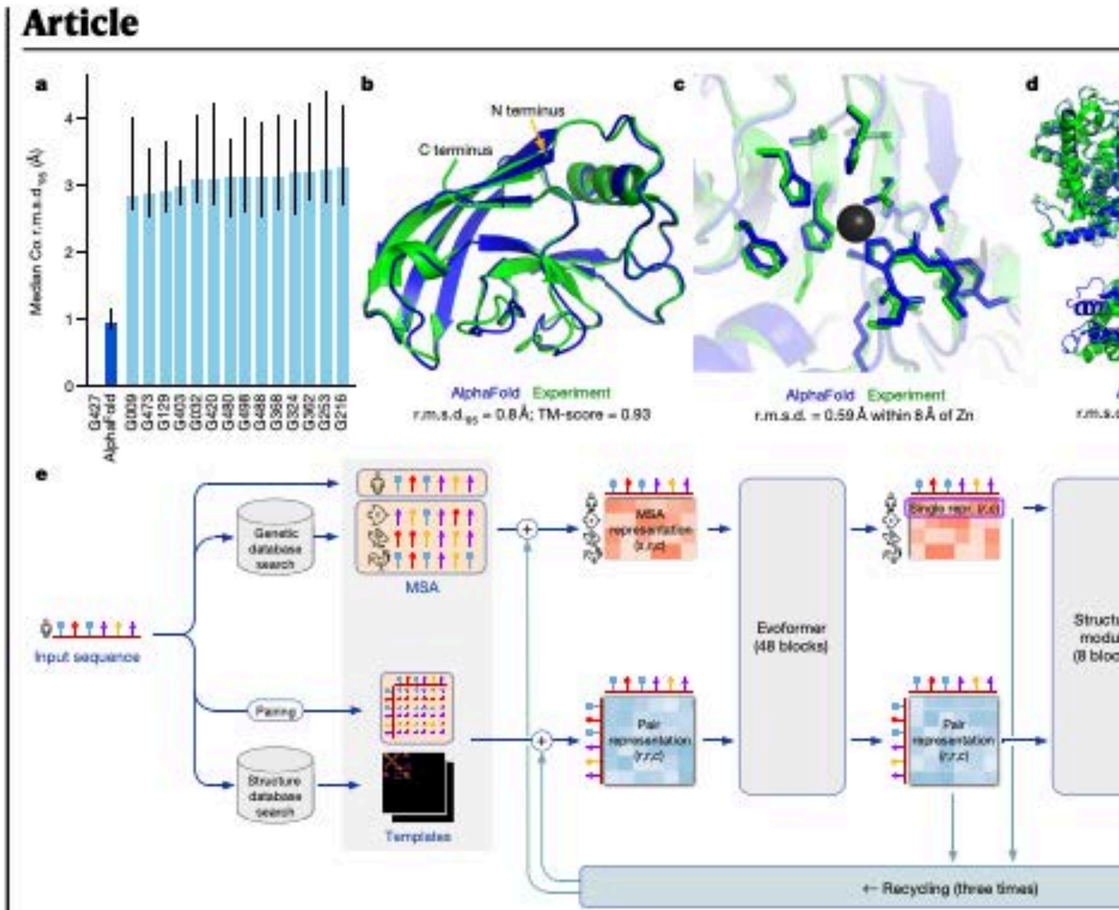


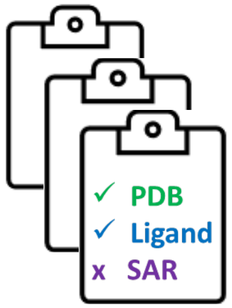
Fig. 1 | AlphaFold produces highly accurate structures. a, The performance of AlphaFold on the CASP14 dataset ($n = 87$ protein domains) relative to the top-15 entries (out of 146 entries), group numbers correspond to the numbers assigned to entrants by CASP. Data are median and the 95% confidence interval

An example of a well-predicted zinc-binding site (A
chains even though it does not explicitly predict th
T1044 (PDB 6VR4)—a 2,180-residue single chain—w
domain packing (the prediction was made after CA

CACHE:

Critical Assessment of Computational Hit-Finding Experiments

1. Challenge Issued



CACHE presents a variety of 'hit finding' challenges to the community, as well as assessment criteria

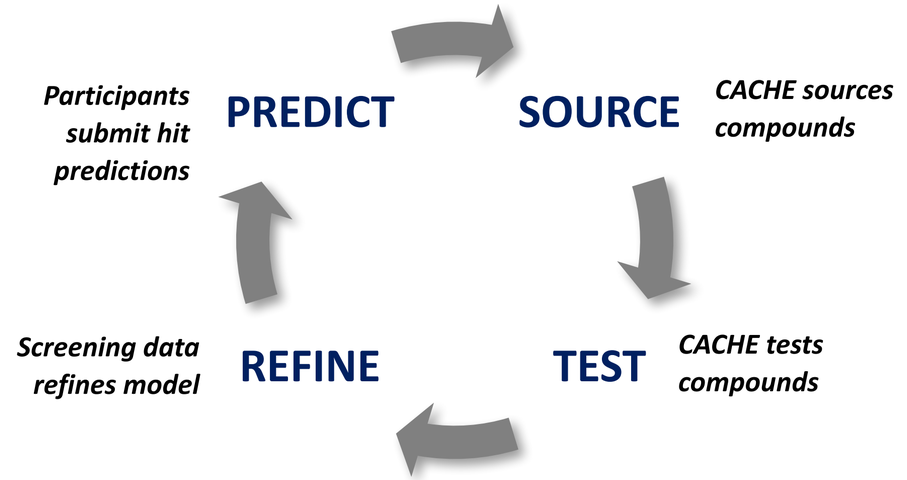
2. Virtual Library

Make-on-Demand
REAL, ZINC15

Crowd-Sourced
Bespoke Chemistry

CACHE establishes and hosts two virtual libraries: 'make-on-demand and 'crowd-sourced' (synthetically accessible by chemists)

3. Prediction and Validation



- **Participants** predict chemical matter
- Each participant has the opportunity to make two cycles of predictions per round
- **CACHE** sources compounds and experimentally tests predictions
- Outcomes help inform models and improve predictions

4. Outcomes Release

Open Chemistry

Open Data

Data package released to the public, including:

- # of compounds tested
- # of confirmed hits
- Potency of hits
- Chemical structures
- Physical properties
- Similarity to known actives

PDB = Protein data bank
SAR = structure-activity relationship

https://chemrxiv.org/engage/chemrxiv/article-details/6168ba62f718dfc39bdee0db

OPEN SCIENCE HAS THE POTENTIAL TO TRANSFORM DRUG DISCOVERY

THE COVID MOONSHOT 🌙

An open science antiviral discovery effort

The COVID Moonshot Consortium
http://postera.ai/covid

10 Feb 2020 - Frank von Delft

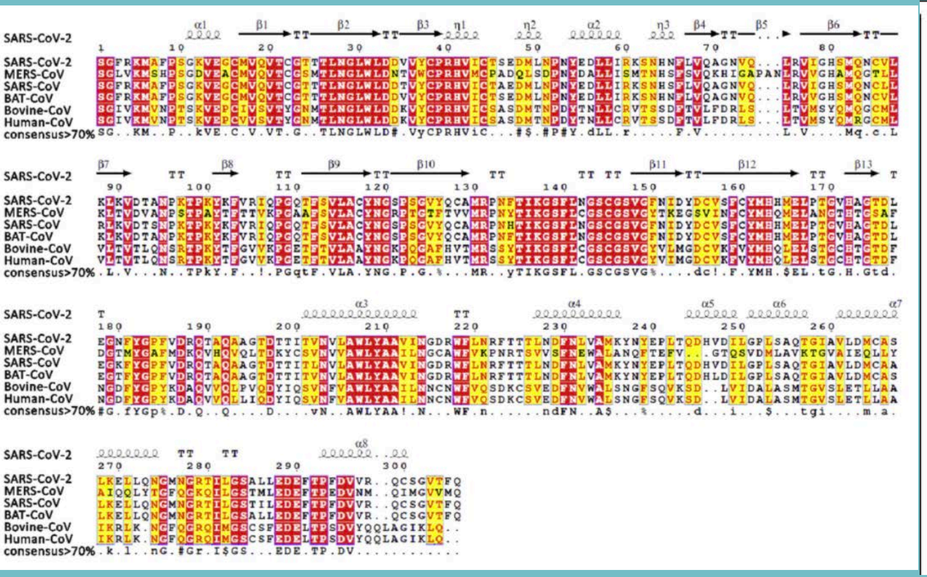
THE PROTEASE

THE PROTEASE

THE PROTEASE

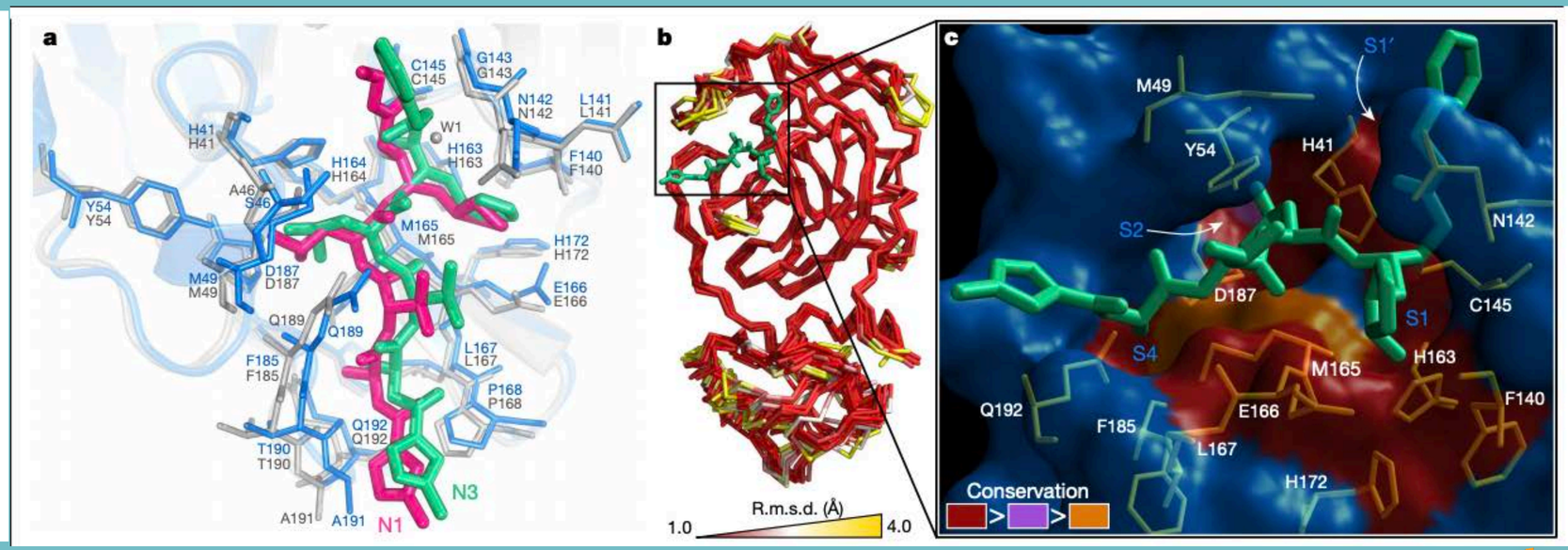
Both sequence and structure of Mpro are highly conserved among beta-coronaviruses

sequence (24 Jan 2020)



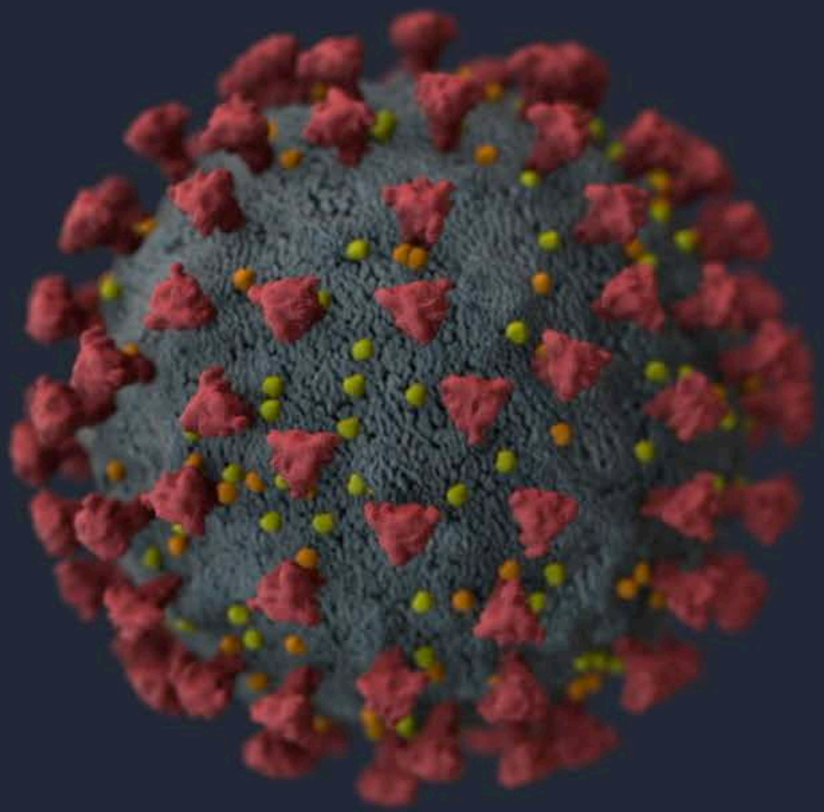
Tahir ul Qamal et al. J Pharm Anal, in press
doi:10.1016/j.jpha.2020.03.009

structure (PDB structure released 5 Feb 2020)



Jin et al. Nature 582:289, 2020
doi:10.1038/s41586-020-2223-y

Mpro appears to be a viable target for developing a SARS-CoV-2 antiviral as well as pan-coronavirus antivirals



OUR GOAL

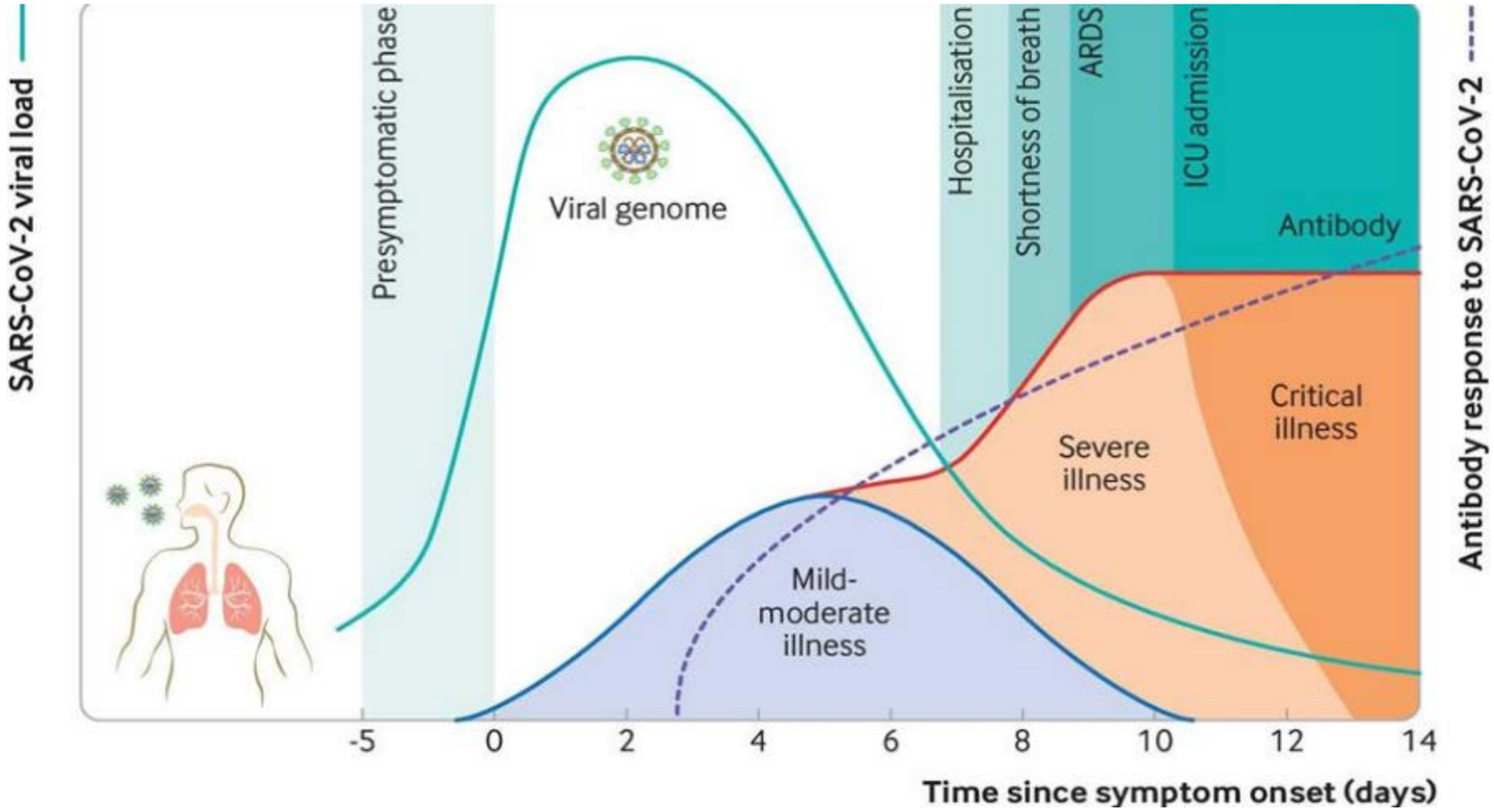
A VARIETY OF THERAPEUTICS WILL BE NECESSARY TO END THE PANDEMIC

Vaccines effective only if administered weeks prior to exposure

Oral antivirals window of opportunity (prophylaxis or early treatment) Requires coupling to **frequent testing**

Virus no longer drives disease; primarily inflammation-mediated injury
Antivirals less effective
Hospitalized: **access to IV drugs**

Antibodies useful if delivered sufficiently early, but challenges for IV administration



Muge Cevik et al. BMJ 2020;371:bmj.m3862
https://doi.org/10.1136/bmj.m3862

Why else might we want an oral antiviral?

A drug taken when needed doesn’t require 100% compliance by public

Inhibitors to essential, mutation-resistant targets could remain effective against Spike variants that may reduce vaccine effectiveness

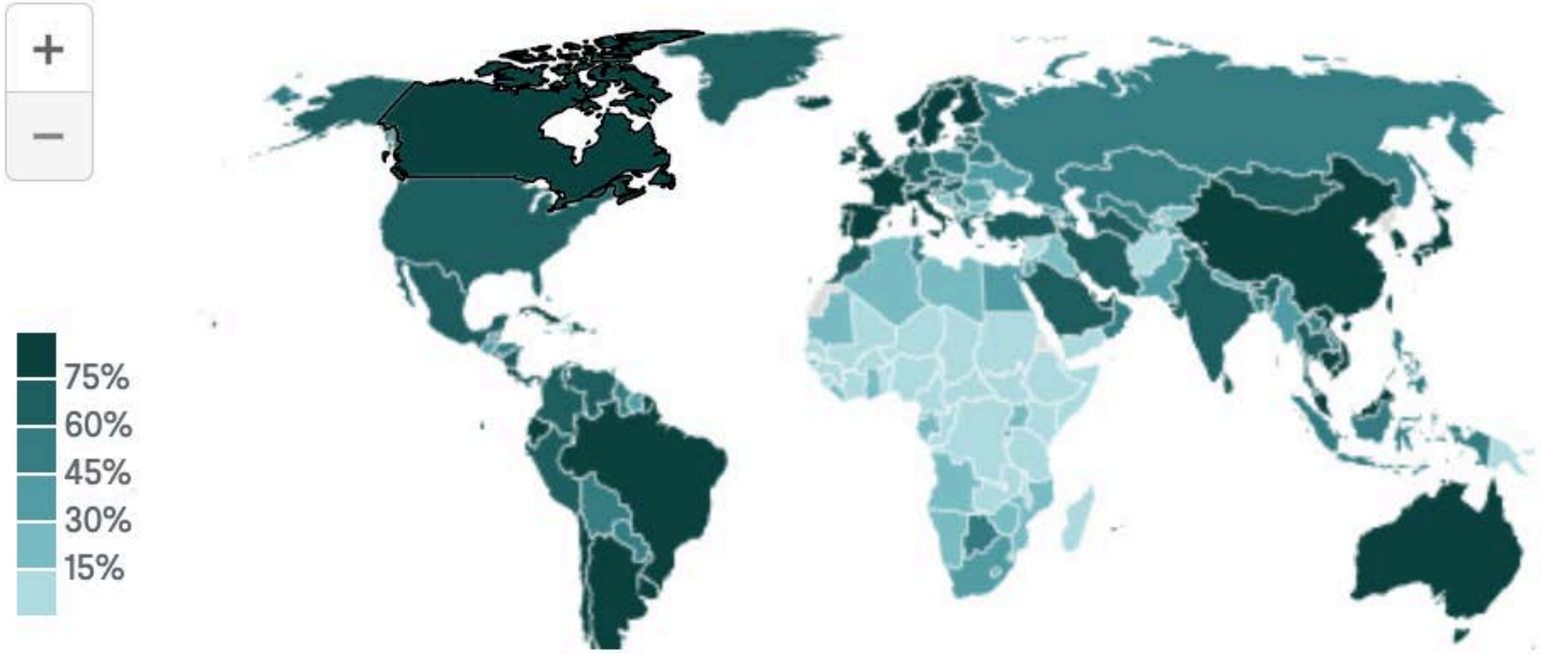
Shelf-stable oral inhibitor would enable practical global deployment without complications of cold chain storage

A simple synthetic route could enable rapid production at low cost

Much of the world is still waiting on vaccines

Global COVID-19 Vaccination Divide Widens

Share of people who have received at least one dose as of December 26 or most recent date available



Source: Our World in Data.

COUNCIL on FOREIGN RELATIONS

https://www.cfr.org/backgrounder/guide-global-covid-19-vaccine-efforts

GLOBAL, EQUITABLE ACCESS IS A HUGE PROBLEM

America And The TRIPS Waiver: You Can Talk The Talk, But Will You Walk The Walk?

Vineeta Gupta, Sreenath Namboodiri

JULY 13, 2021 10.1377/hblog20210712.248782

ADD TO FAVORITES SHARE

As nations grapple with the issues surrounding global COVID-19 vaccine manufacturing and distribution, the Trade-Related Aspects of Intellectual Property Rights (TRIPS) Agreement has found itself in mainstream conversation in the US more than ever before. A difficult concept to fully grasp, TRIPS refers to a World Trade Organization (WTO)-led international agreement about the protection of intellectual property rights and trade.

In October 2020, the governments of India and South Africa, with the support of 62 WTO member states, proposed a TRIPS Agreement waiver proposal that would temporarily waive intellectual property rights protections for technologies needed to prevent, contain, or treat COVID-19, including vaccines and vaccine-related technologies. More than 100 low-income countries support this proposal, but it is receiving much opposition from many high-income countries, including some European Union (EU) member states, the UK, Japan, Canada, and Australia. On May 5, 2021, the Biden administration announced support for negotiating this waiver, intensifying debate in the US and the EU—but so far the US has not gone further than its announcement of support.

The TRIPS waiver is critical to combating the COVID-19 pandemic around the world. Demand for the vaccine has already surpassed supply, with high-income countries taking a large share of reserved doses. Given that no single vaccine manufacturer could produce enough vaccines to meet the demand of the entire globe, supporters of the waiver ponder the ethics of multinational manufacturers holding exclusive rights to information and technology, preventing other companies from entering the markets that are not being served—primarily in low- and middle-income countries. Sharing vaccine-related information will not only help get the pandemic in check now, but it could also encourage firms to develop the next round of vaccines that will be necessary to address new variants.

The TRIPS waiver is critical to ensuring an equitable distribution of vaccines around the globe.

TRIPS patent waiver requests from India and 100 low-income countries to expand vaccine production have been pending since October 2020**, and nothing has happened**

Meanwhile….

Forbes

EDITORS' PICK | Jul 28, 2021, 01:48pm EDT | 40,696 views

Pfizer Expects $33.5 Billion In Vaccine Revenue In 2021

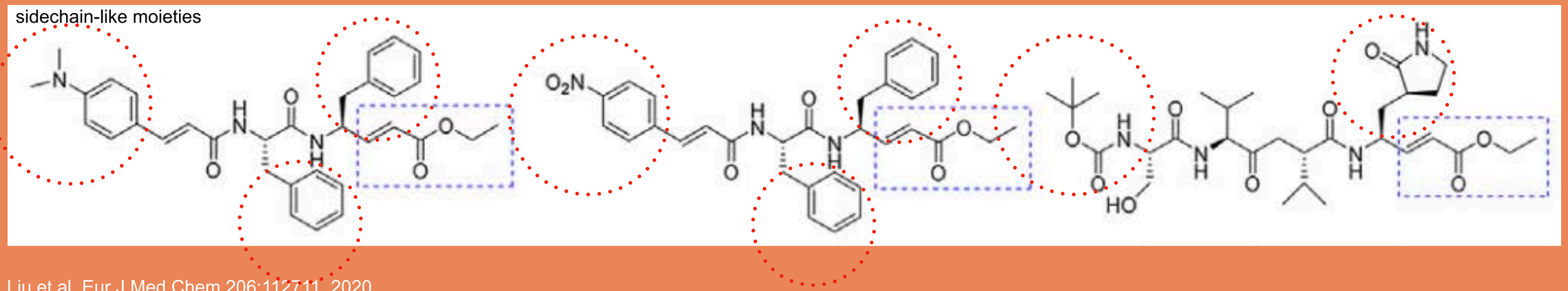
Moderna, Racing for Profits, Keeps Covid Vaccine Out of Reach of Poor

Some poorer countries are paying more and waiting longer for the company's vaccine than the wealthy — if they have access at all.

Moderna and U.S. at Odds Over Vaccine Patent Rights

https://www.healthaffairs.org/do/10.1377/hblog20210712.248782/full/
https://www.forbes.com/sites/aayushipratap/2021/07/28/pfizer-expects-335-billion-in-vaccine-revenue-in-2021/?sh=f49a83c217d4
https://www.nytimes.com/2021/11/09/us/moderna-vaccine-patent.html
https://www.nytimes.com/2021/10/09/business/moderna-covid-vaccine.html

Previously known Mpro inhibitors were peptidomimetics, which are difficult to develop into useful oral drugs

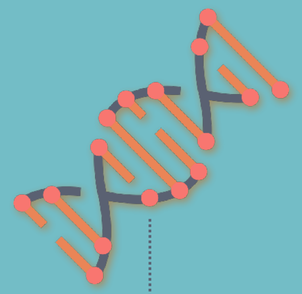


Liu et al. Eur J Med Chem 206:112711, 2020

Known inhibitors were also covalent inhibitors, which can be difficult to optimize to prevent off-target issues

Diamond Light Source prosecuted a high-throughput X-ray fragment screen in a matter of weeks

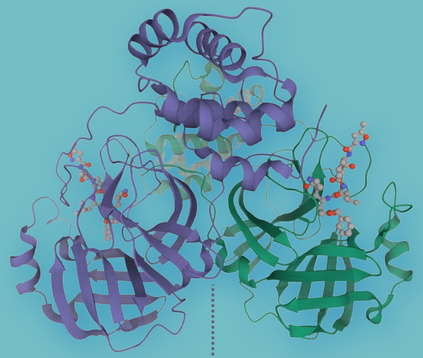
Frank von Delft
Diamond Light Source / XChem / SGC



February 14

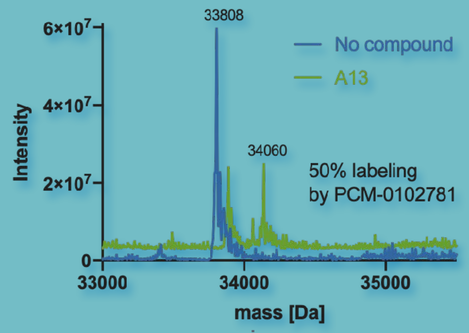
Main protease cloned and produced at Diamond after COVID shutdown of Haitao Yang lab in Shanghai

Martin Walsh



February 20

Atomic resolution structure of the protease determined

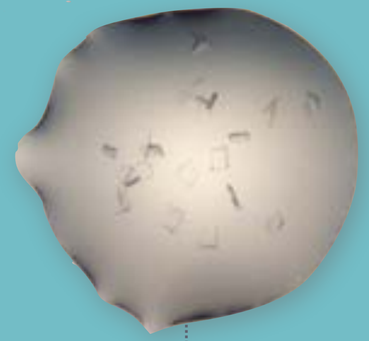


February 25

Covalent screen finds 150 active site hits

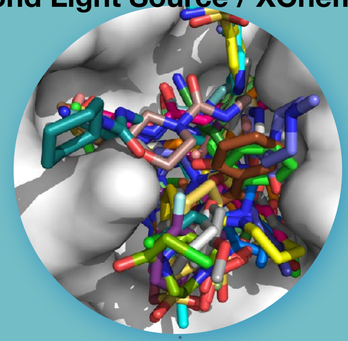
>40 hits validated

Nir London



March 5

1,500 crystals collected in one day (!)

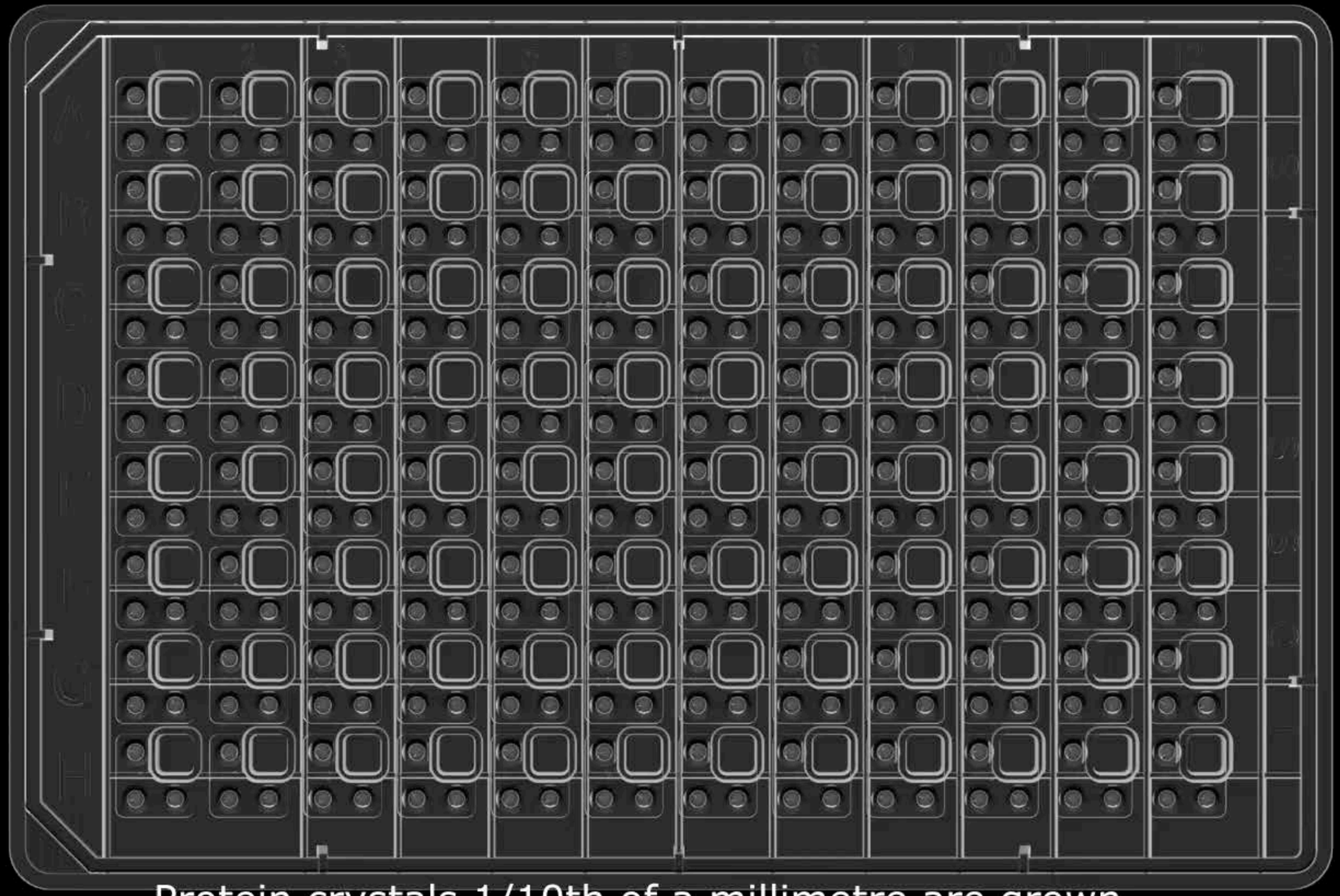


March 18

78 fragment-bound structures solved and released to the web

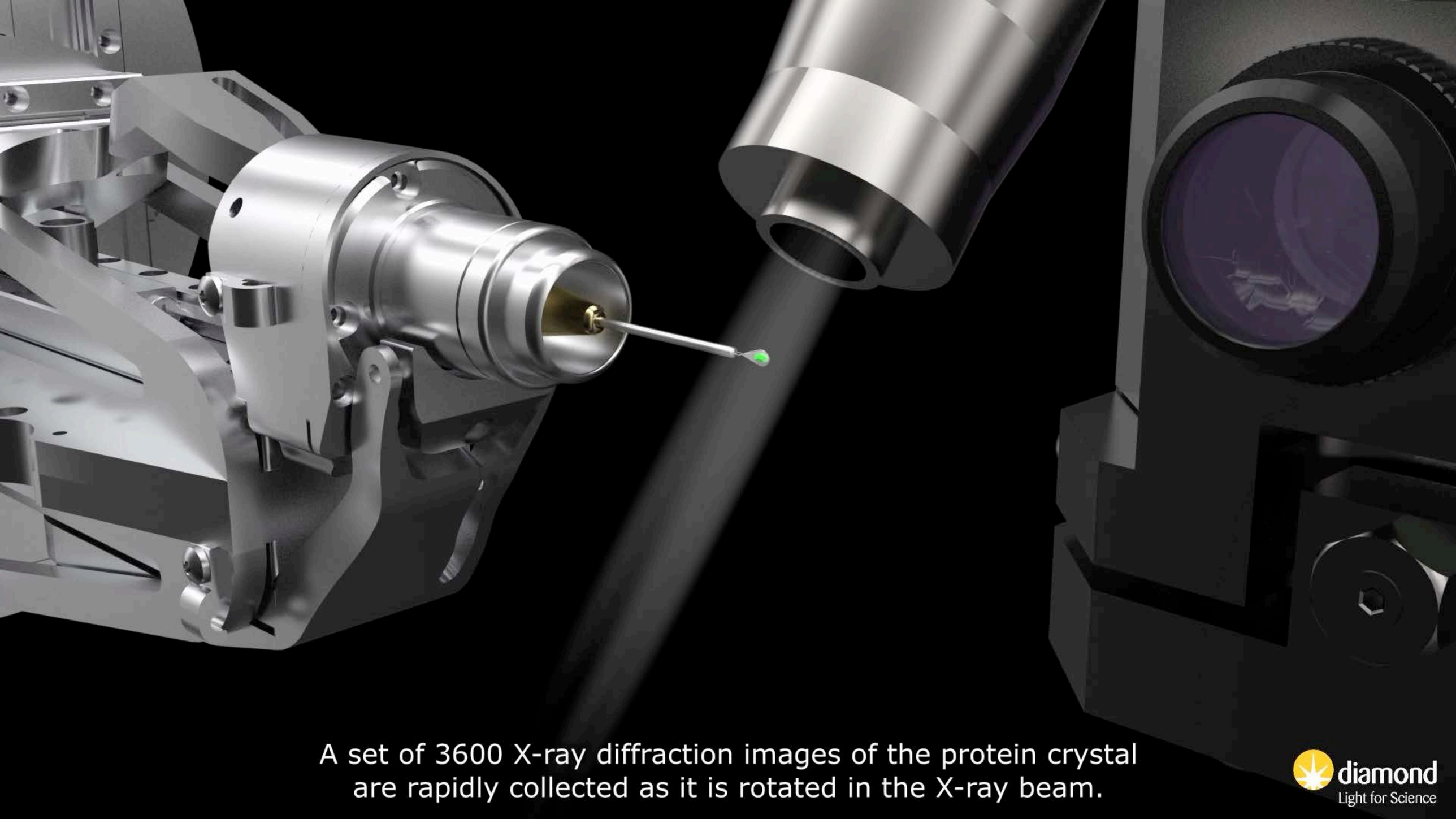
48 covalent fragments
71 active site fragments

https://www.diamond.ac.uk/covid-19/for-scientists/Main-protease-structure-and-XChem.html



Protein crystals 1/10th of a millimetre are grown
in microscopic drops no larger than 1 mm.

diamond
Light for Science



A set of 3600 X-ray diffraction images of the protein crystal are rapidly collected as it is rotated in the X-ray beam.

diamond
Light for Science

All data was immediately released online

diamond | Coronavirus Science

Home | For Scientists | For Journalists | For the Public | For Staff | Diamond Website

In This Section

- Main protease structure and XChem fragment screen
  - COVID MoonShot - Taking fragments to impact
  - Electron density evidence
  - Downloads
  - Highlights on progress
  - Credits
  - FAQ
- Nsp3 macrodomain ADP-ribosyl hydrolase and XChem fragment screen
- New scientific animations
- Rapid Access
- Research Areas
- Our collaborators

Main protease structure and XChem fragment screen

Summary

To contribute to the global effort to combat COVID-19, Diamond has been able to solve a new structure of the SARS-CoV-2 main protease (M^{Pro}) at high resolution (PDB ID: 6YB7), and complete a large XChem crystallographic fragment screen against it (detailed below). Data have been deposited with the PDB, but we are making the results available immediately to the world on this page; additional work is ongoing, and updates will be continually posted here in coming days and weeks.

This work builds on the sensationally fast crystal structure of M^{Pro} at 2.16 Å in complex with a covalent inhibitor, released in January this year by Prof Zihe Rao (6LU7, published here, described here). We thus ordered the synthetic gene and cloned the full length protein as previously described for the SARS main protease (Xue et al 2007). This yielded crystals of the unliganded enzyme that diffracted to high resolution (1.25 Å) on beamline I04-1, in a different space group to the inhibitor complex, and the structure was determined and refined rapidly. ***Critically, this showed it had the active site empty and solvent accessible - perfect for fragment screening.***

So it proved: the first 600-crystal experiment could be completed in 72 hours, through growing large numbers of crystals, optimising the soaking conditions, soaking and harvesting all 600 crystals and completing the data collection run on beamline I04-1. The hits from this initial run and other details were pre-released on March 6th.

By the 24th of March, the initial 1500-crystal experiment was complete, and the results made publicly available. Screening additional libraries throughout April brought the total number of active site fragments to 71, with 48 fragments binding covalently (full timeline here and download page here). This was an exceptionally large screen which yielded a remarkably rich readout, with vast opportunities for fragment growing and merging.

We have already triggered computationally-driven follow-up work internally, and externally joined forces to launch a fully-open crowdsourcing and crowdfunding initiative – the COVID Moonshot - to establish urgently the shortest route possible to clinical impact by maximally exploiting the readout - you can help, read more here.

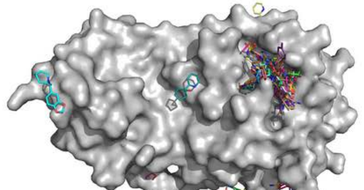
On the 11th of May, the first biochemical and structural data from Moonshot compounds was released and by the 12th of June over 500 compounds had been tested, demonstrating that the design-make-test process is fully in place.

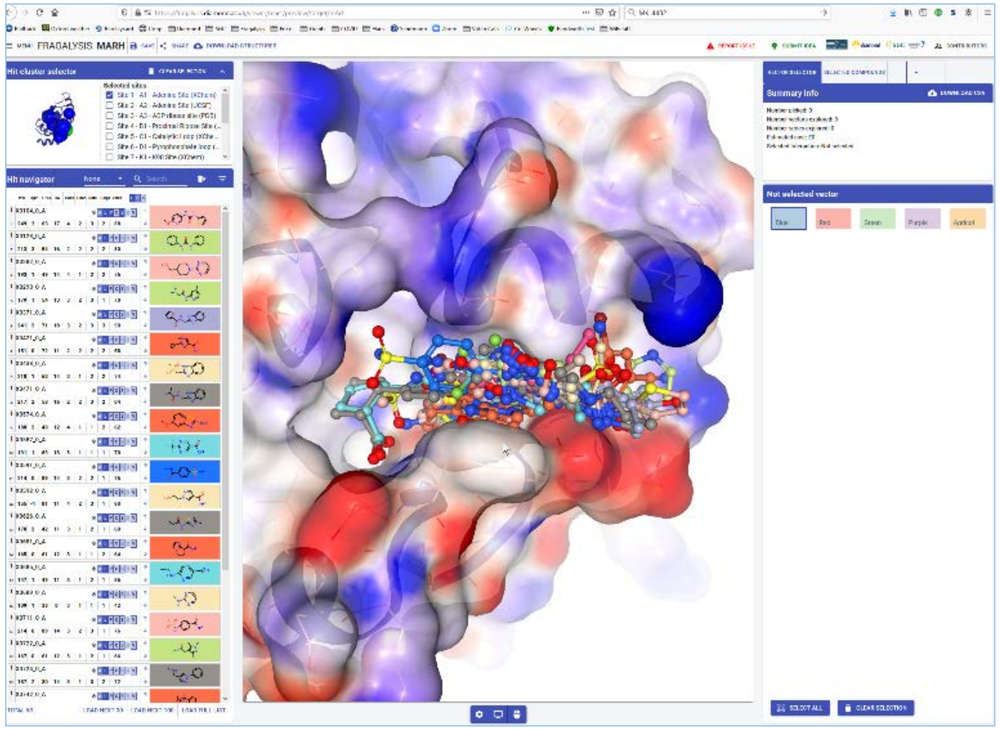
XChem fragment screen

The initial screen encompassed multiple fragment libraries: the DSI-poised library, MiniFrags (Astex) FragLites & Peplites (CRUK Newcastle Drug Discovery Unit (Newcastle University)), York3D (University of York), SpotFinder and heterocyclic electrophilic fragment library (Hungarian Academy of Sciences) and an electrophilic fragment library designed and pre-screened by mass spec at the Weizmann Institute (see below).

There were 74 hits of high interest - data and extensive details are here, and some interactive views here:

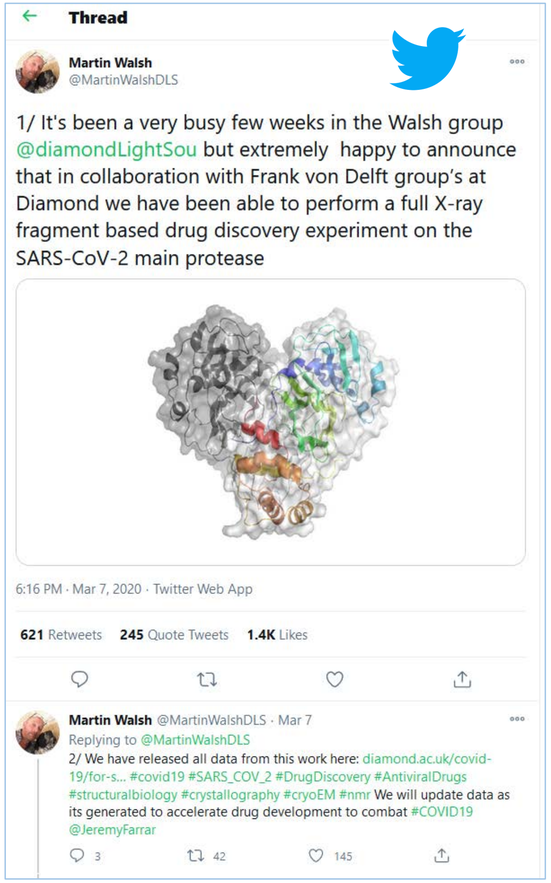
- 23 non-covalent hits in the active site
- 48 covalent hits in the active site
- 3 hits in the dimer interface, one in a calculated hotspot





https://fragalysis.diamond.ac.uk

https://www.diamond.ac.uk/covid-19/for-scientists/Main-protease-structure-and-XChem.html

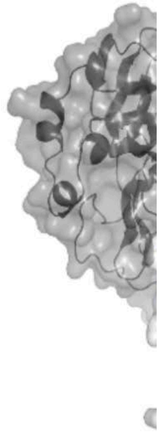


COVID Moonshot

Centre for Medicines Discovery
University of Oxford
XCHEM Diamond Fragment Screening
diamond

Martin Walsh
@MartinWalshDLS

1/ It's been a very busy few weeks in the Walsh group @diamondLightSou but extremely happy to announce that in collaboration with Frank von Delft group's at Diamond we have been able to perform a full X-ray fragment based drug discovery experiment on the SARS-CoV-2 main protease

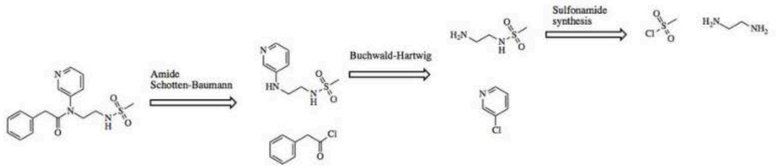


2:16 AM · Mar 8, 2020 · Twitter We

enthusiasm curbed @enthusiamcurbed · Mar 9, 2020

Replying to @MartinWalshDLS and @DiamondLightSou

Hey, @MartinWalshDLS, amazing work! Would love to help if possible - my old group (Lee Group @Cambridge_Uni) works on synthetically tractable hit expansion. E.g. can solve synthesis for this naive fragment merge practically instantly



1 2

Alpha Lee @iamalphalee · Mar 9, 2020

Hey @MartinWalshDLS this is awesome work! Would love to explore with @cuteargument whether our ML/synthesis-aware generative models can help. Happy to work on it if you need some extra hands!

Which strategies would most quickly carry fragment structures all the way to a useful antiviral drug?

Nir London
Weizmann Institute

What if we tried ALL OF THEM?

Alpha Lee (PostEra/Cambridge) quickly set up the COVID Moonshot website

Alpha Lee
Cambridge/PostEra

COVID Moonshot

Design a Compound, We Will Make It

After drawing the molecule, you will be asked for details on your design. After results are collected, we will prioritize compounds and send them out for synthesis and testing [see details]. There will be several rounds of design; the second round closed Thursday, April 2, 11:59 PM PST. Results will be posted live as we receive them so stay tuned!

View already submitted molecules here. Join the discussion with scientists around the world on our forum.

Draw or enter SMILES (add multiple by pressing "Add" after each entry)

SMILES | Add

100%

H C N O S P F Cl Br I

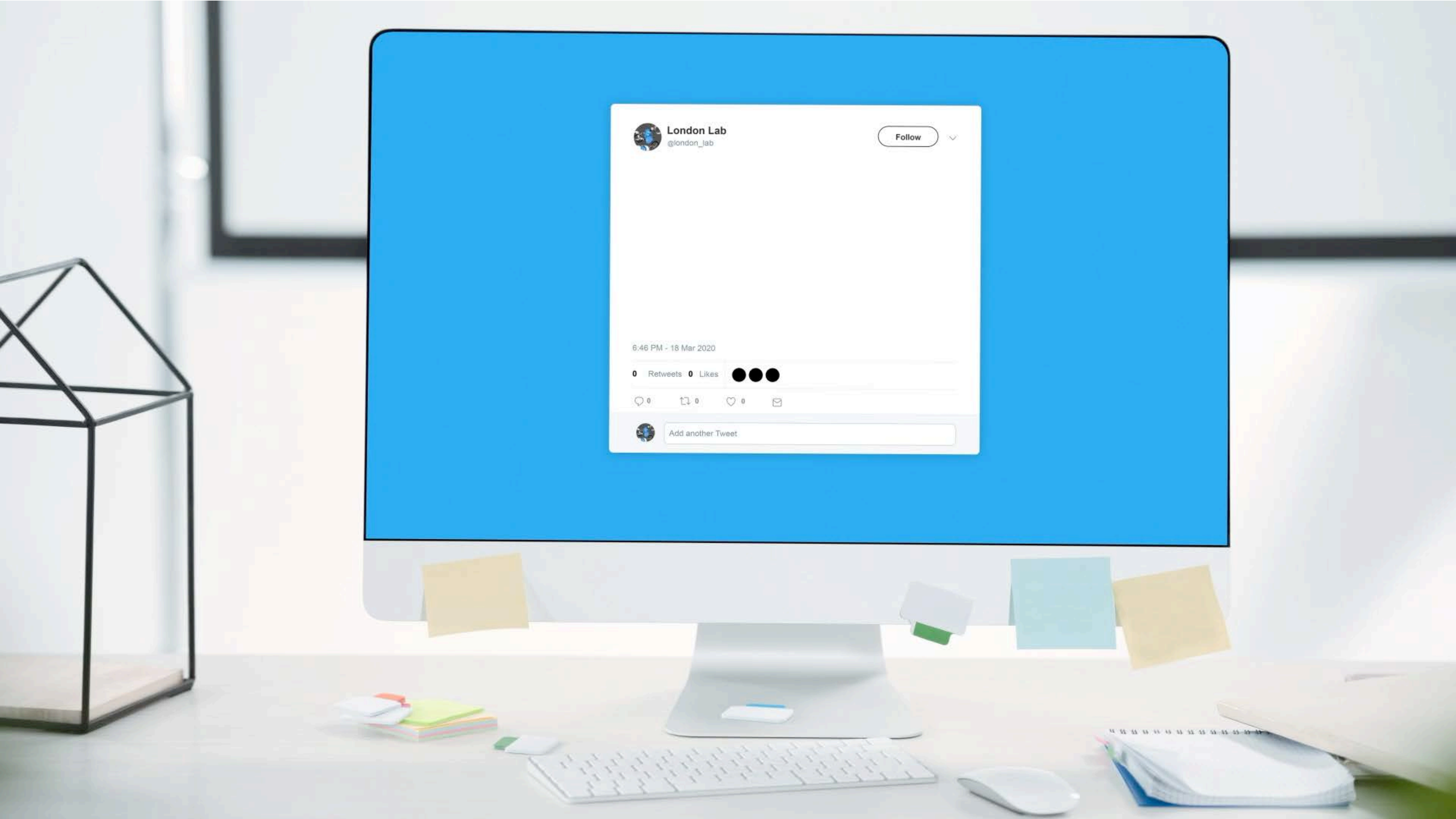
Contact Information

Name* | Email* | Affiliation

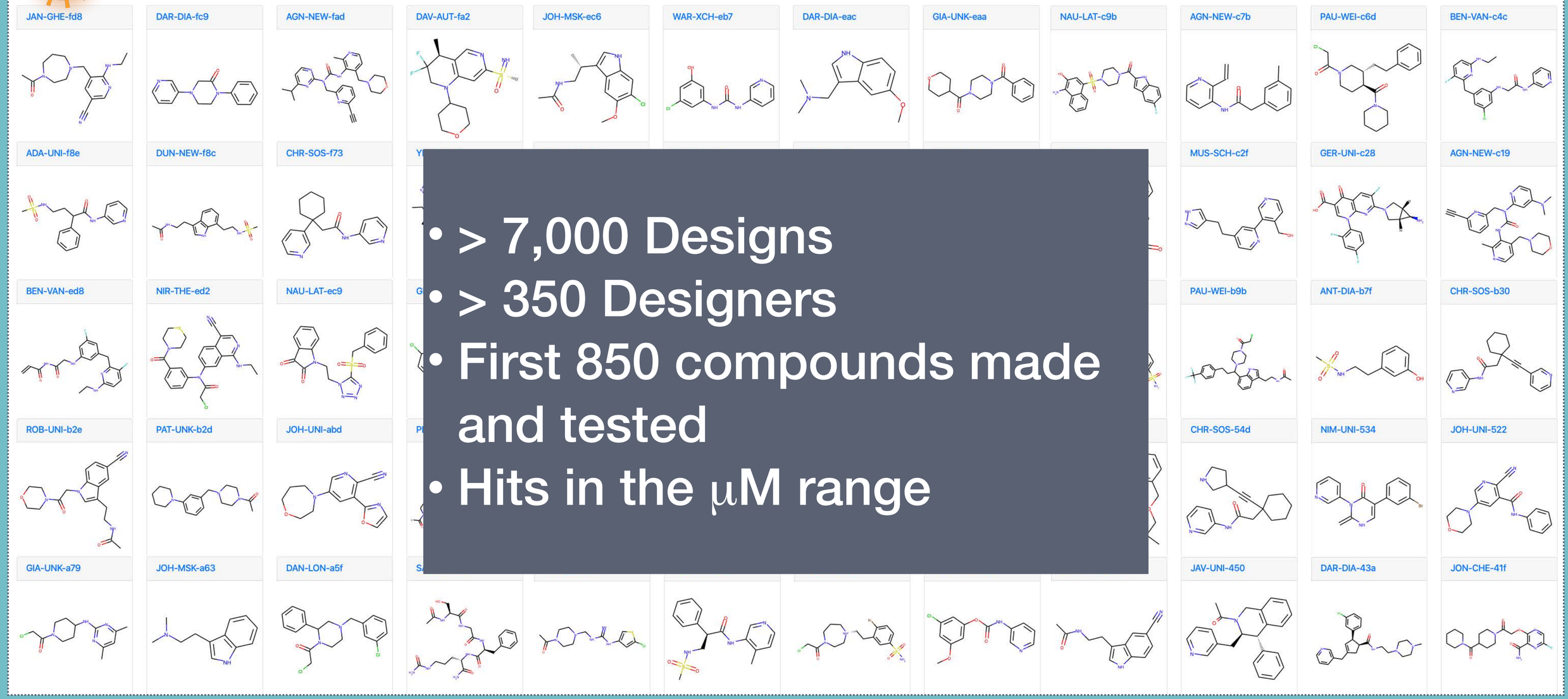
Background

- Please specify the rationale in some detail (by eye, docking, FEP, ...)
- Add any notes or special considerations regarding your compound (complex sythesis required, past experience, ...)
- If there are other compounds related to your main structure, submit them as a comma separated list of SMILES
- Please specify which fragments were used as inspiration (e.g. X_0072, X_0161)
- A PDB of the bound structure from simulations is optional

http://postera.ai/covid


London Lab
@london_lab
Follow
6:46 PM - 18 Mar 2020
0 Retweets 0 Likes
Add another Tweet

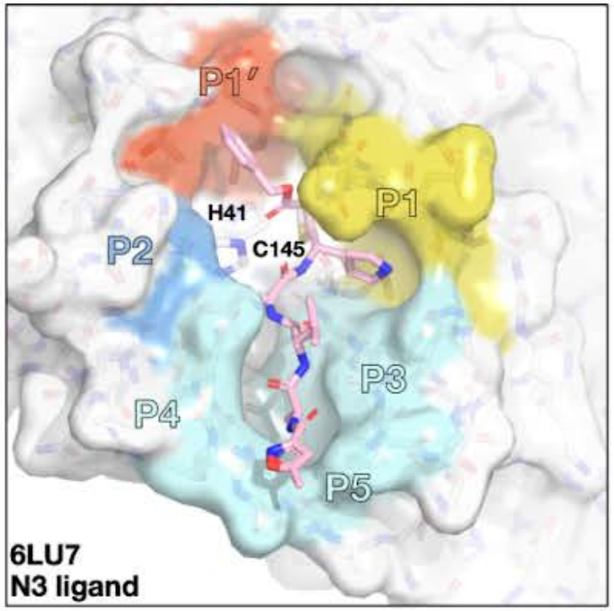
There was overwhelming response

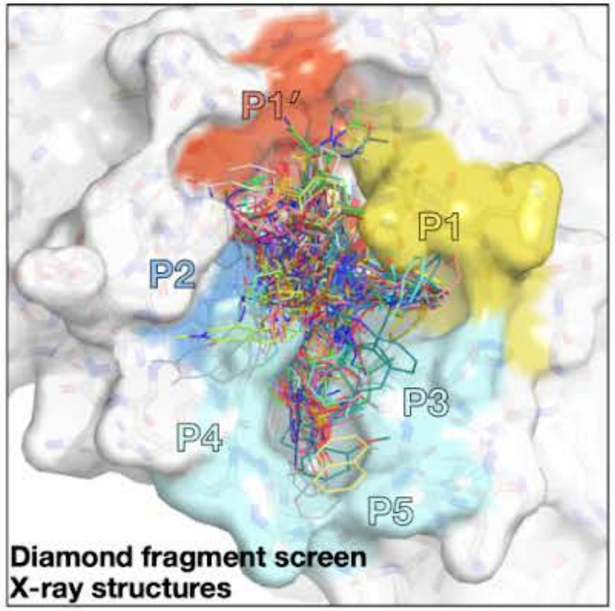


- > 7,000 Designs
- > 350 Designers
- First 850 compounds made and tested
- Hits in the μM range

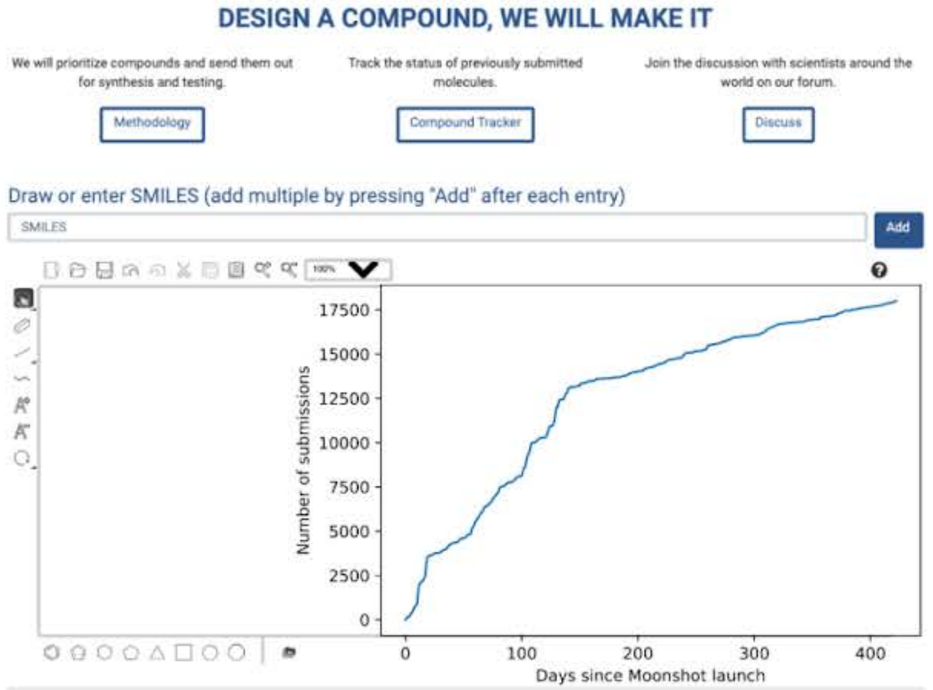
Crowdsourcing generated multiple novel hit chemotypes via fragment mergers

A

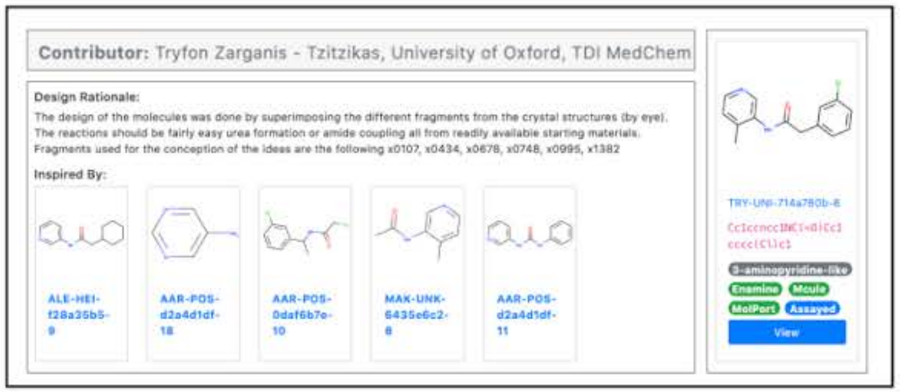




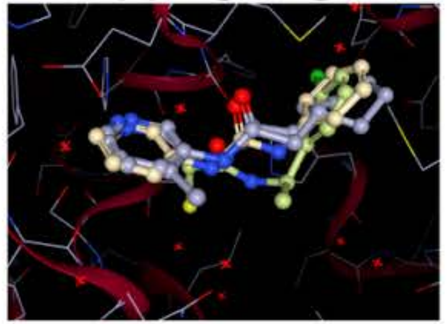
B



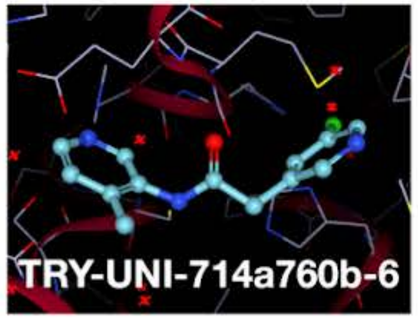
C

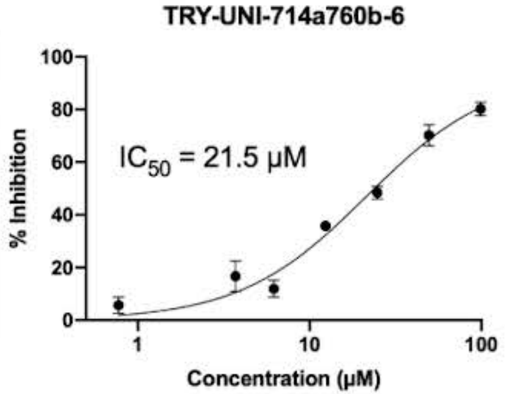


Overlapping fragments

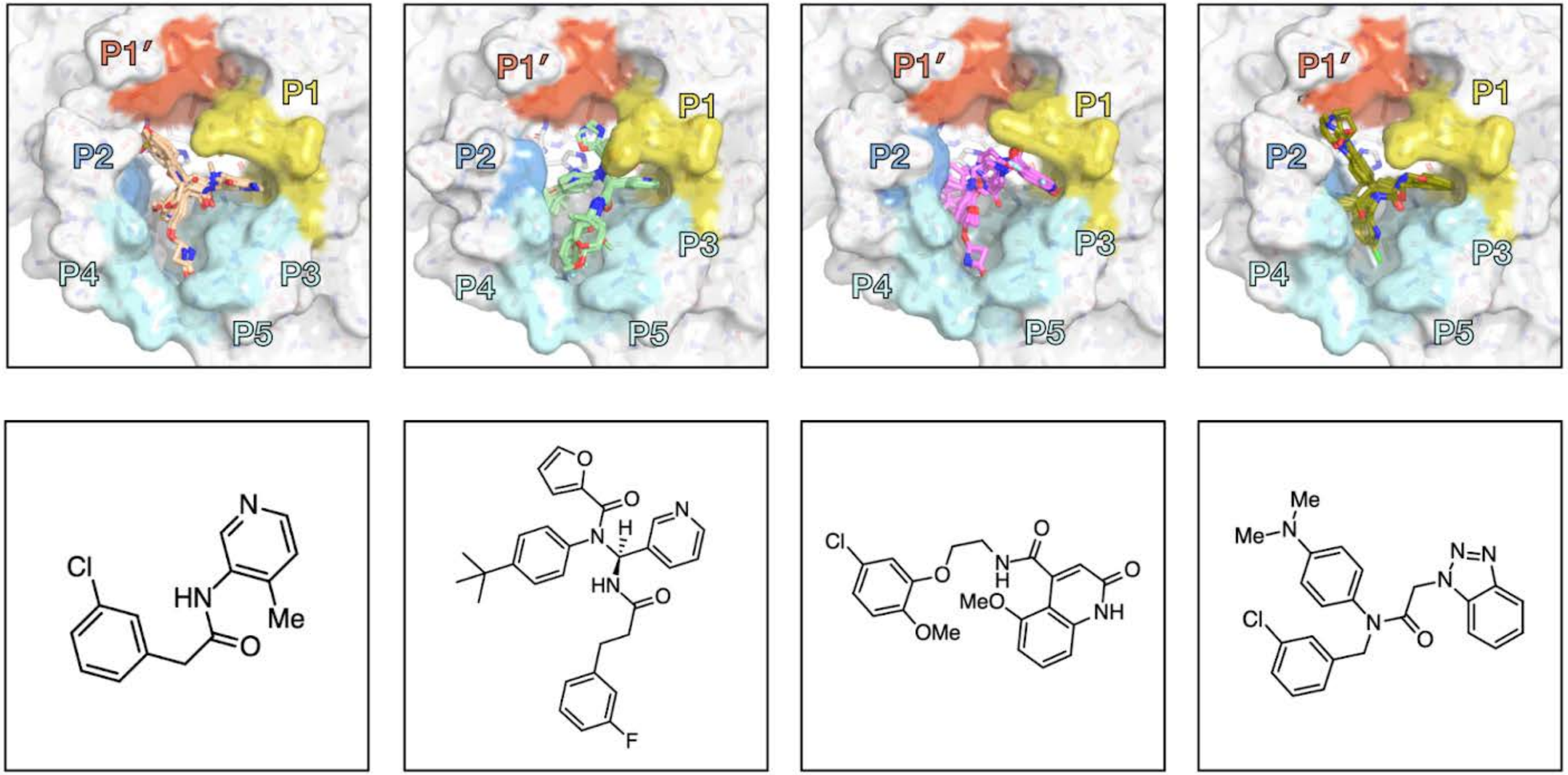


Merged fragments





Crowdsourcing generated multiple novel hit chemotypes via fragment mergers



The medicinal chemistry target product profile (TPP) defines the goal for producing a preclinical candidate

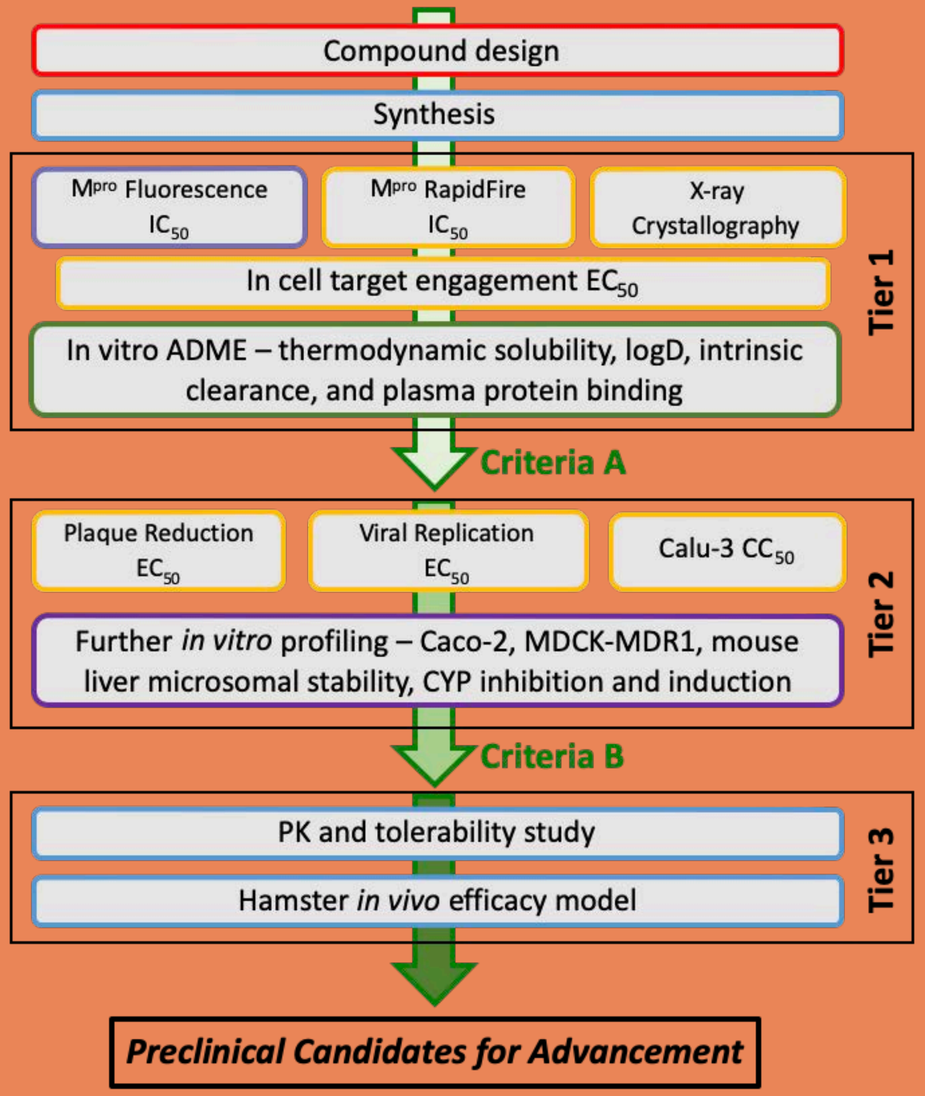
Ed Griffen (Medchemica) leads med chem design team of multiple industry veterans

TPP for 5-day oral antiviral course following exposure, SARS-CoV-2 PCR+, or onset of symptoms

Property	Target range	Rationale
protease assay	IC_{50} < 50 nM	Extrapolation from other anti-viral programs
viral replication	EC_{50} < 0.2µM	Suppression of virus at achievable blood levels
plaque reduction	EC_{50} < 0.2µM	Suppression of virus at achievable blood levels
Coronavirus spectrum	SARS-CoV2 B1.1.7 , 501.V2, B.1.1.248 variants essential, SARS-CoV1 & MERS desirable	Treat vaccine resistant variants and future pandemic preparation.
route of administration	oral	bid/tid(qid)- compromise PK for potency if pharmacodynamic effect achieved
solubility	> 5 mg/mL, >100µM tolerable	Aim for biopharmaceutical class 1 assuming <= 750 mg dose
half-life	Ideally>= 8 h (human) est from rat and dog	Assume PK/PD requires continuous cover over plaque inhibition for 24 h
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 / combination therapy, 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

We built an assay cascade to help us achieve the TPP in a rapid but cost-effective manner

Ed Griffen
Medchemica



Does it inhibit Mpro? How does it bind?

Does it enter cells and inhibit Mpro?

Does it have a chance of working in humans?

Does it kill virus in infected cells, sparing healthy cells?

Does it have a favorable safety profile?

Is it orally bioavailable at required concentrations?

Our lab had started to use Folding@home to aid experimental collaborators in pursuing COVID-19 drug discovery projects

FOLDING @HOME

Client statistics by OS

OS Type	Native TFLOPS*	x86 TFLOPS*	Active CPUs	Active Cores	Total CPUs
Windows	857	857	67,467	187,104	5,857,235
Mac OS X	91	91	8,083	85,382	217,033
Linux	87	87	6,383	26,457	882,200
NVIDIA GPU	1	2	4	4	348,371
ATI GPU	10,243	21,613	7,178	7,178	426,335
NVIDAI Fermi GPU	36,065	76,097	21,570	21,587	624,822
Total	47,344	98,747	110,685	327,712	8,355,996

1924085 people have non-anonymously contributed to Folding@home.

Table last updated at Sat, 19 Oct 2019 18:23:11

~100 pflop/s!

WE MOBILIZED THE FOLDING@HOME CONSORTIUM TO FOCUS ON COVID-19

About

Pande Lab

The Folding@home Consortium (FAHC)

Community volunteers

Partners

Donate

How does donor funding compare with federal grant funding?

Links

Donation FAQ

Stanford Donation Site

Highlight from the 2016 Stanford Chemistry Department Graduation

THE FOLDING@HOME CONSORTIUM (FAHC)

A number of research labs are involved in running and enhancing FAH.

BOWMAN LAB, WASHINGTON UNIVERSITY IN ST. LOUIS

The Bowman lab combines computer simulations and experiments to understand the mechanisms of allostery (i.e. long-range communication between different parts of a protein) and to exploit this insight to control proteins' functions with drugs and mutations. Examples of ongoing projects include (1) understanding how mutations give rise to antibiotic resistance, (2) designing allosteric drugs to combat antibiotic resistant infections, (3) understanding allosteric networks in G proteins and designing allosteric anti-cancer drugs, and (4) understanding and interfering with the mechanisms of Ebola infection. To rapidly converge on predictive models, we iterate between using simulations to gain mechanistic insight, conducting our own experimental tests of our models, and refining our simulations/analysis based on feedback from experiments. We also develop enhanced sampling algorithms for modeling rare events that are beyond the reach of existing simulation methodologies.

CHODERA LAB, MEMORIAL SLOAN-KETTERING CANCER CENTER

The Chodera lab at the Sloan-Kettering Institute uses Folding@home to better understand how we can design more effective therapies for cancer and other diseases.

Their mission is to completely redesign the way that therapeutics—especially anticancer drugs—are designed using computers, graphics processors (GPUs), distributed computing, robots, and whatever technology we can get our hands on. They are striving to make the design of new cancer drugs much more of an engineering science, where state-of-the-art computer models quantitatively and accurately predict many aspects of drug behavior before they are synthesized. Chodera Lab certainly won't get there overnight—lots of hard work is needed to improve algorithms, force fields, and theory. But by tapping into the enormous computing resources of F@h, they can more rapidly make predictions and then test them in the laboratory (with robots!) to quickly make improvements through learning from each cycle of prediction and validation.

VOELZ LAB, TEMPLE UNIVERSITY

Vincent Voelz lab at Temple University's Chemistry Department focuses on using transferrable, all-atom simulations for prediction and design of biomolecular dynamics and function. In particular, their interests include in silico prediction and design of proteins, peptide mimetics (e.g. peptoids), and binding sequences for cell signaling peptides.

HUANG LAB, HKUST

Xuhui Huang's lab at HKUST is interested in conformational change, which is crucial for a wide range of biological processes including biomolecular folding and the operation of key cellular machinery.

FOLDING@HOME TAKES UP THE FIGHT AGAINST COVID-19 / 2019-NCOV

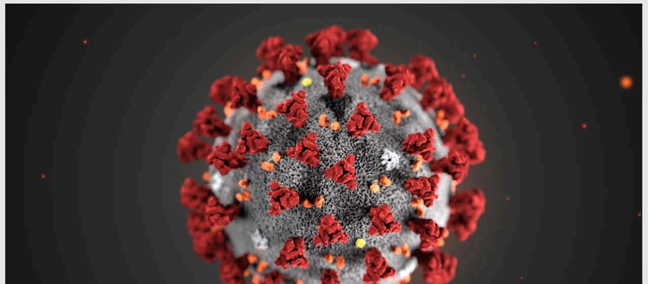
February 27, 2020
by Greg Bowman

We need your help! Folding@home is joining researchers around the world working to better understand the 2019 Coronavirus (2019-nCoV) to accelerate the open science effort to develop new life-saving therapies. By downloading Folding@Home, you can donate your unused computational resources to the Folding@home Consortium, where researchers working to advance our understanding of the structures of potential drug targets for 2019-nCoV that could aid in the design of new therapies. The data you help us generate will be quickly and openly disseminated as part of an open science collaboration of multiple laboratories around the world, giving researchers new tools that may unlock new opportunities for developing lifesaving drugs.

2019-nCoV is a close cousin to SARS coronavirus (SARS-CoV), and acts in a similar way. For both coronaviruses, the first step of infection occurs in the lungs, when a protein on the surface of the virus binds to a receptor protein on a lung cell. This viral protein is called the spike protein, depicted in red in the image below, and the receptor is known as ACE2. A therapeutic antibody is a type of protein that can block the viral protein from binding to its receptor, therefore preventing the virus from infecting the lung cell. A therapeutic antibody has already been developed for SARS-CoV, but to develop therapeutic antibodies or small molecules for 2019-nCoV, scientists need to better understand the structure of the viral spike protein and how it binds to the human ACE2 receptor required for viral entry into human cells.

Proteins are not stagnant—they wiggle and fold and unfold to take on numerous shapes. We need to study not only one shape of the viral spike protein, but all the ways the protein wiggles and folds into alternative shapes in order to best understand how it interacts with the ACE2 receptor, so that an antibody can be designed. Low-resolution structures of the SARS-CoV spike protein exist and we know the mutations that differ between SARS-CoV and 2019-nCoV. Given this information, we are uniquely positioned to help model the structure of the 2019-nCoV spike protein and identify sites that can be targeted by a therapeutic antibody. We can build computational models that accomplish this goal, but it takes a lot of computing power.

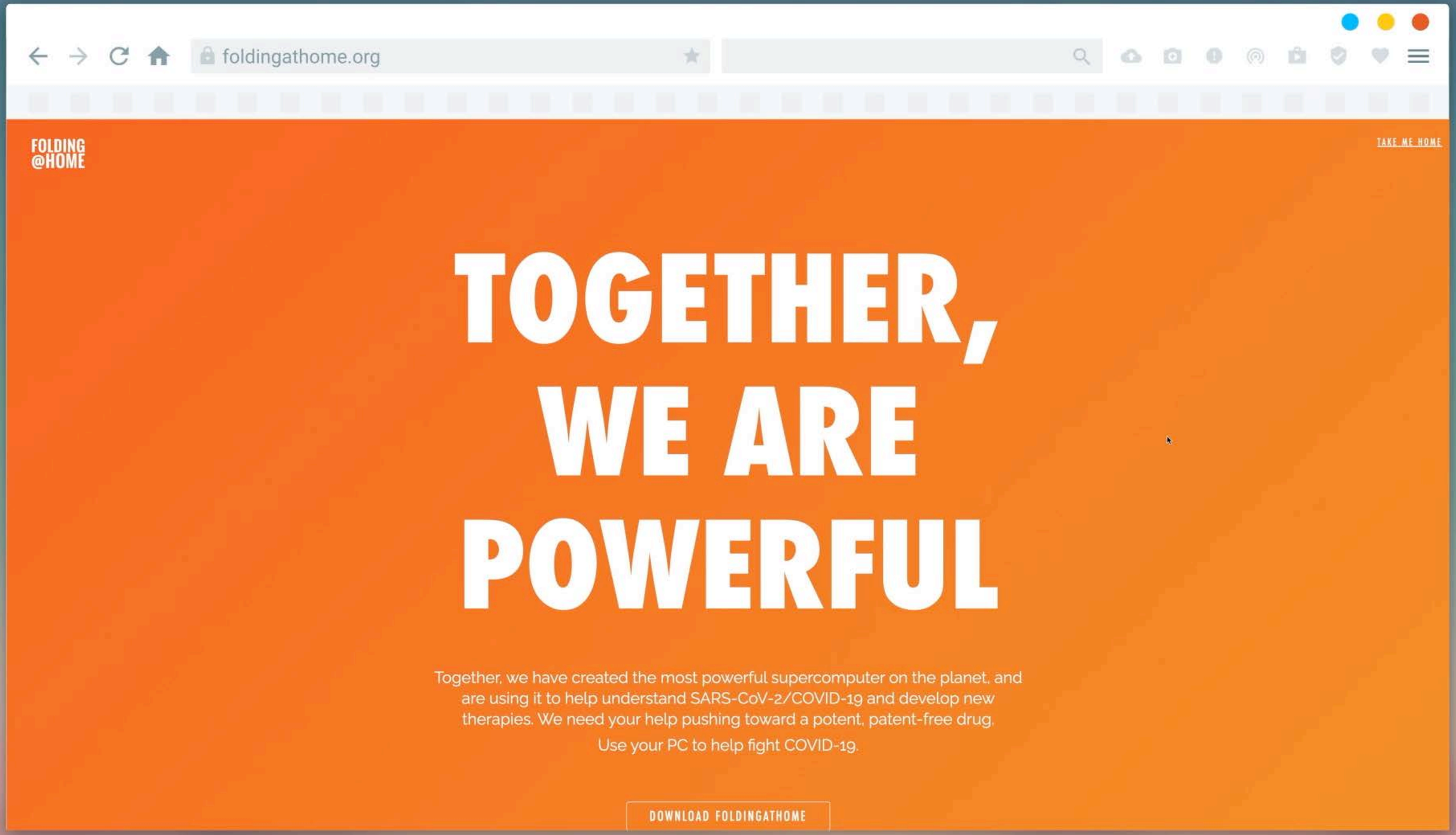
This is where you come in! With many computers working towards the same goal, we aim to help develop a therapeutic remedy as quickly as possible. By downloading Folding@home here [LINK] and selecting to contribute to "Any Disease", you can help provide us with the computational power required to tackle this problem. One protein from 2019-nCoV, a protease encoded by the viral RNA, has already been crystallized. Although the 2019-nCoV spike protein of interest has not yet been resolved bound to ACE2, our objective is to use the homologous structure of the SARS-CoV spike protein to identify therapeutic antibody targets.



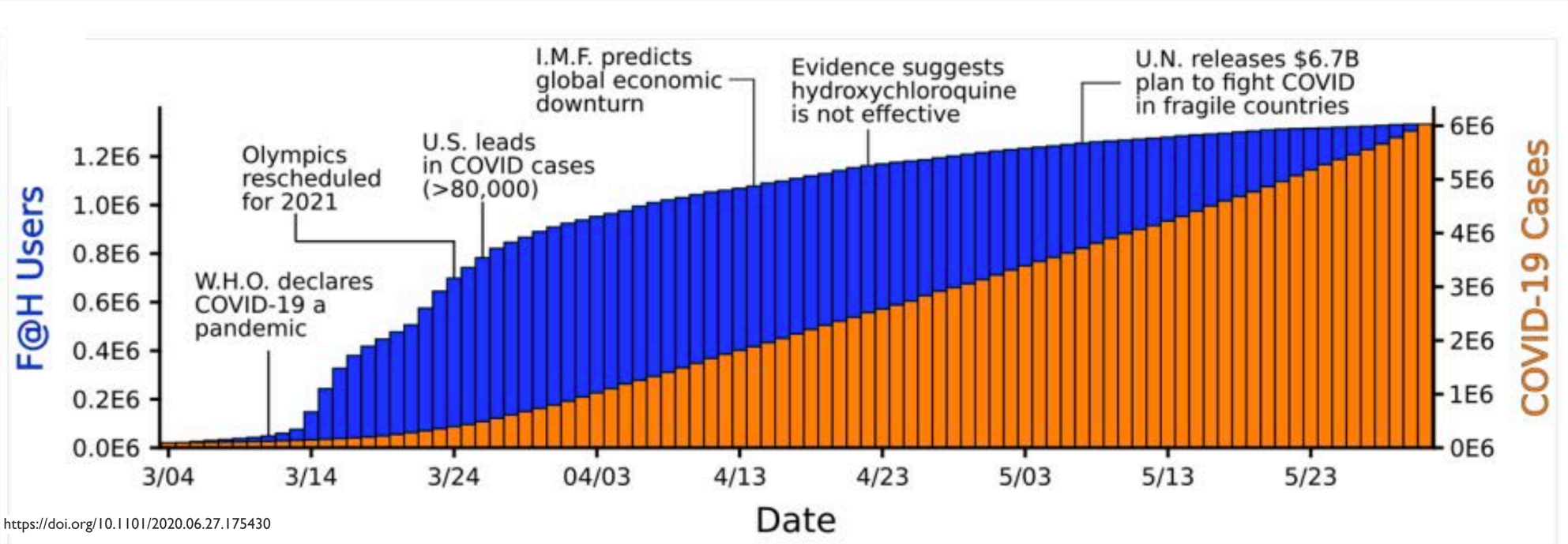
Ariana Brenner (CBM)

Rafal Wiewiora (TPCB)

Ivy Zhang (CBM)


foldingathome.org
FOLDING @HOME
TAKE ME HOME
TOGETHER, WE ARE POWERFUL
Together, we have created the most powerful supercomputer on the planet, and are using it to help understand SARS-CoV-2/COVID-19 and develop new therapies. We need your help pushing toward a potent, patent-free drug.
Use your PC to help fight COVID-19.
DOWNLOAD FOLDINGATHOME

We built the first exaFLOP/s computing platform as the public joined in our effort



SARS-CoV-2 simulations go exascale to predict dramatic spike opening and cryptic pockets across the proteome
https://www.nature.com/articles/s41557-021-00707-0

This honestly came as a bit of a surprise

Folding@home

Team Monthly | Team | Donor | OS Stats

Active CPUs & GPUs by OS

OS	AMD GPUs	NVidia GPUs	CPUs	CPU cores	TFLOPS	x86 TFLOPS
Windows	75,823	314,952	474,277	3,588,315	680,371	1,384,998
Linux	3,675	41,113	78,124	811,997	85,028	167,152
macOSX	0	0	41,582	230,198	2,578	2,578
Totals	79,498	356,065	593,983	4,630,510	767,977	1,554,728

CPUs and GPUs which have returned Work Units within the last 50 days are listed by OS. FLOPS per core is estimated.

TFLOPS is Tera Floating-point OPerations per Second or trillions of math operations per second. Please see our FLOPS FAQ for more information.

Reported on Wed, 25 Mar 2020 23:42:36 GMT

~1.5 exaflops

> sum of top-10 supercomputers

Use Your Computer To Help Folding@Home Solve The COVID-19 Virus Pandemic

Longmont Observer · Yesterday

- 400,000 new people have joined Folding@Home's fight against COVID-19

  Engadget · 2 days ago

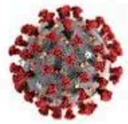
View Full Coverage

Folding@home software diverts users' excess processing power to finding coronavirus cure

Dezeen · 22 hours ago

Folding@Home Network Breaks the ExaFLOP Barrier In Fight Against Coronavirus

Tom's Hardware · 5 hours ago



How to Fight Coronavirus With Folding@home and a Gaming PC

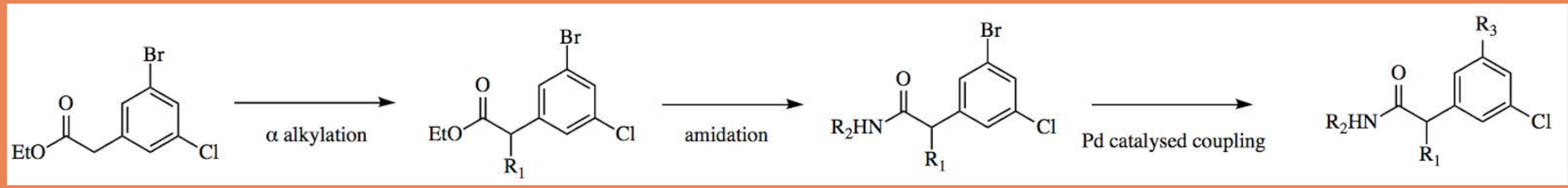
How-To Geek · 5 days ago

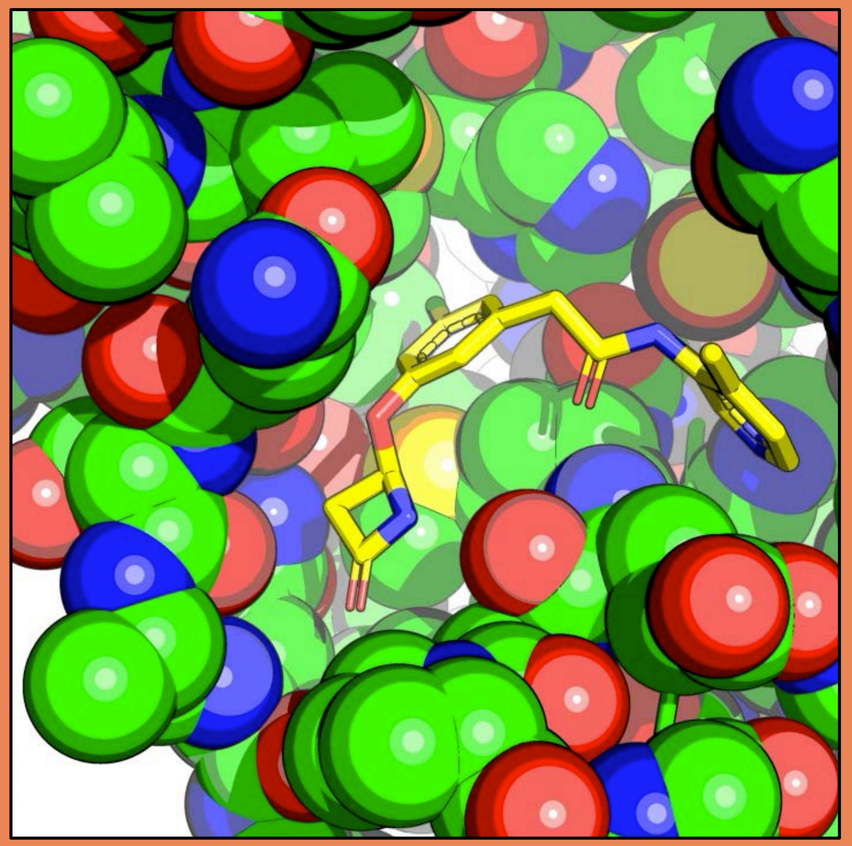
Join Team Hackaday To Crunch COVID-19 Through Folding@Home

Hackaday · 7 days ago

Coronavirus And Folding@Home; More On How Your Computer Helps Medical Research

To prioritize compounds for further potency optimization, we enumerated a huge variety of molecules that can be quickly synthesized by Enamine

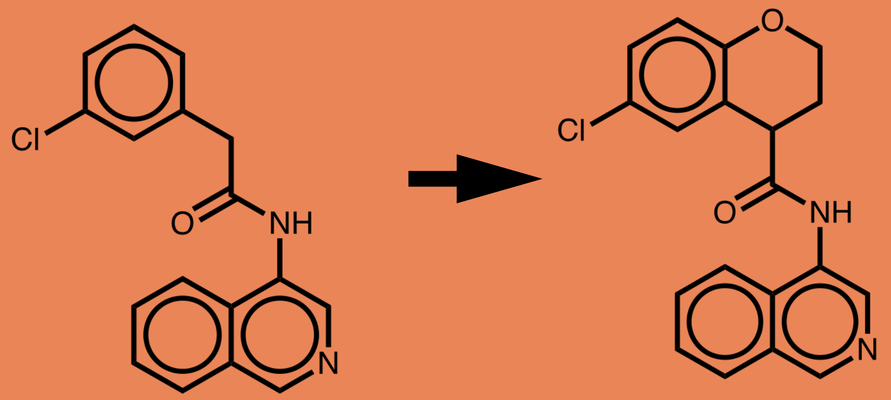




We can use Folding@home to run alchemical free energy calculations to evaluate which designs should bind better

Instead of transmuting **lead** into **gold…**

…we change **one molecule** into **another!**

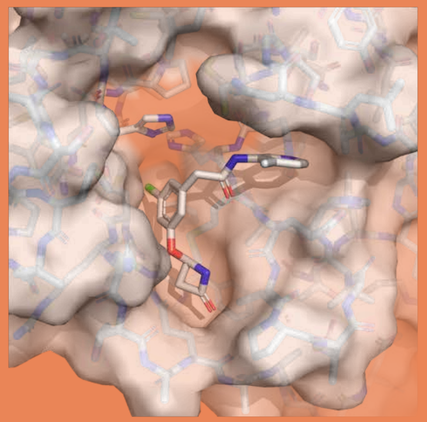


Neither process can be done with chemistry, but we can do it in a computer!

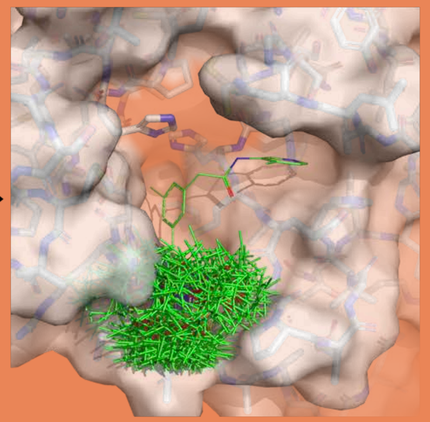
We used Folding@home to run relative alchemical free energy calculations at planetary scale, performing tens of thousands of transformations/week

Dominic Rufa
Tri-I TPCB PhD student

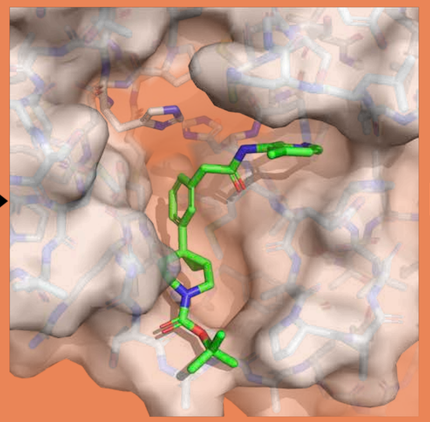
X-ray structure as reference



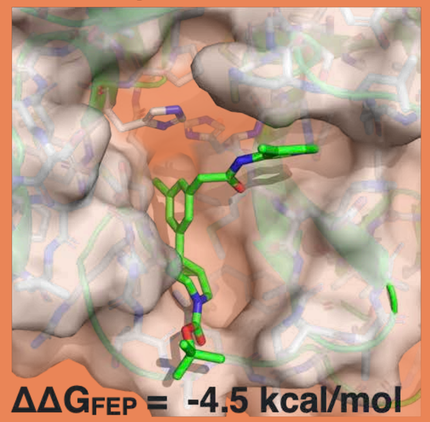
constrained enumeration of poses for proposed molecule



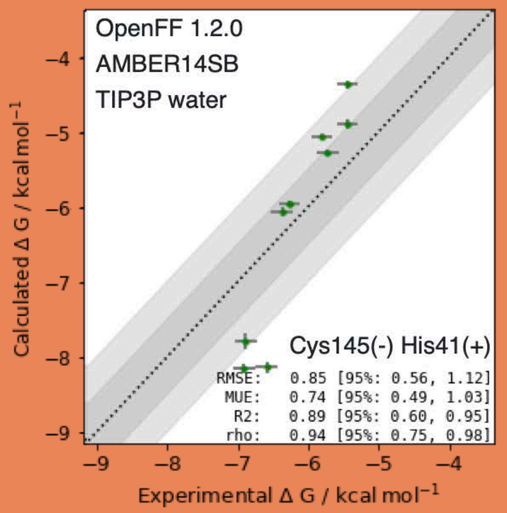
selection of pose with best docking score



nonequilibrium alchemical free energy calculation final posed structure



retrospective performance on 3-aminopyridine lead series



perses: open source relative alchemical free energy calculations
http://github.com/choderalab/perses
Open Force Field Initiative OpenFF (“Parsley”) small molecule force field
http://openforcefield.org

+ Hannah Bruce Macdonald
William Glass
Matt Wittman
David Dotson

The Open Force Field 1.0 small molecule force field, our first optimized force field (codename "Parsley")

At the end of our first year, the Open Force Field Consortium releases its first optimized force field: the Open Force Field 1.0 (codename "Parsley") small molecule force field

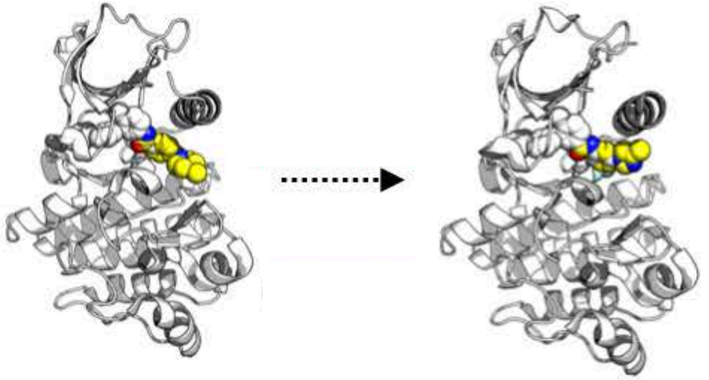
35 minute read, Published: 10 Oct, 2019

We're delighted to announce the release of "Parsley", the Open Force Field 1.0 small molecule force field---the first in a series of iteratively-improved small molecule force fields for biomolecular simulation funded in part by the Open Force Field Consortium. This is the first optimized force field to use the SMIRNOFF force field specification for atom type-free direct chemical perception, and provides substantially improved valence (bond, angle, and torsion) parameters relative to its predecessor, the AMBER-lineage SMIRNOFF99Frosst. This force field was optimized to improve agreement with quantum chemical geometries, energetics, and vibrational frequencies, and will likely provide improved accuracy (relative to its predecessor) for a wide variety of properties, especially energetics and geometries relative to gas phase quantum chemical calculations

https://openforcefield.org/news/introducing-openforcefield-1.0/

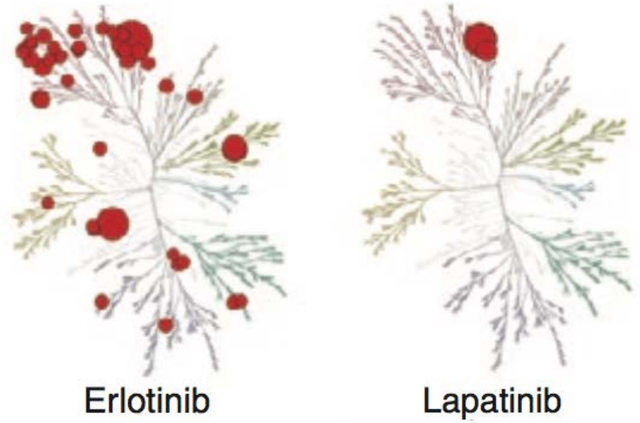
ALCHEMICAL FREE ENERGY CALCULATIONS HAVE THE POTENTIAL TO COMPUTE MULTIPLE PROPERTIES OF INTEREST

driving affinity / potency



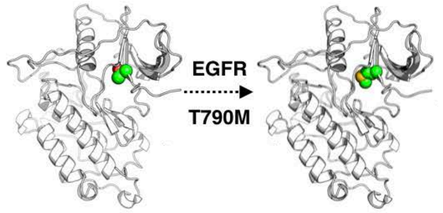
driving selectivity

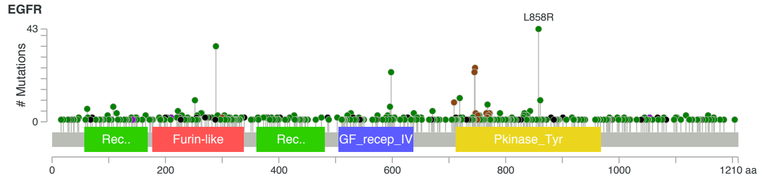
Moraca, Negri, de Olivera, Abel JCIM 2019
https://doi.org/10.1021/acs.jcim.9b00106
Aldeghi et al. JACS 139:946, 2017.
https://doi.org/10.1021/jacs.6b11467



predicting clinical drug resistance/sensitivity

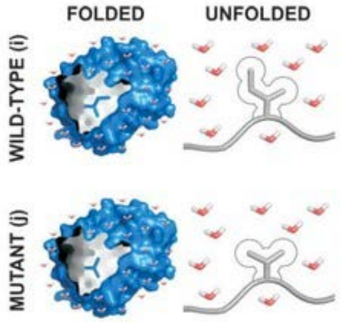
Hauser, Negron, Albanese, Ray, Steinbrecher, Abel, Chodera, Wang.
Communications Biology 1:70, 2018
https://doi.org/10.1038/s42003-018-0075-x
Aldeghi, Gapsys, de Groot. ACS Central Science 4:1708, 2018
https://doi.org/10.1021/acscentsci.8b00717

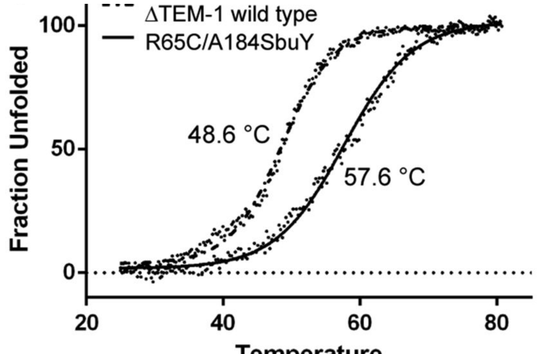




optimizing thermostability

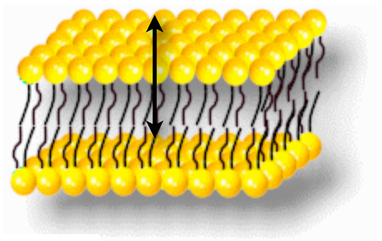
Gapsys, Michielssens, Seeliger, and de Groot. Angew Chem 55:7364, 2016
https://doi.org/10.1002/anie.201510054



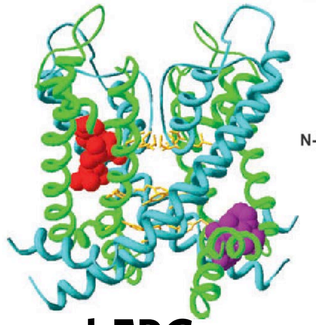


ALCHEMICAL FREE ENERGY CALCULATIONS HAVE THE POTENTIAL TO COMPUTE MULTIPLE PROPERTIES OF INTEREST

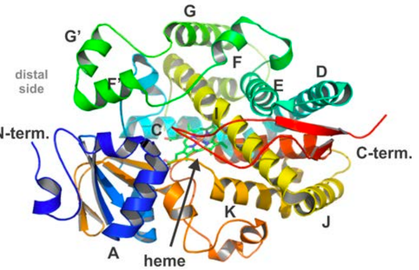
partition coefficients (logP, logD) and permeabilities



structure-enabled ADME/Tox targets

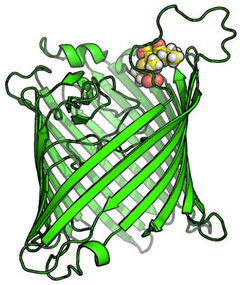


hERG

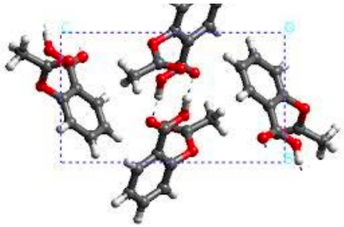


CYP3A4

porin permeation



crystal polymorphs, etc.



TOGETHER, WE ARE POWERFUL

Together, we have created the most powerful supercomputer on the planet, and are using it to help understand SARS-CoV-2/COVID-19 and develop new therapies. We need your help pushing toward a potent, patent-free drug.

Use your PC to help fight COVID-19.

DOWNLOAD FOLDINGATHOME

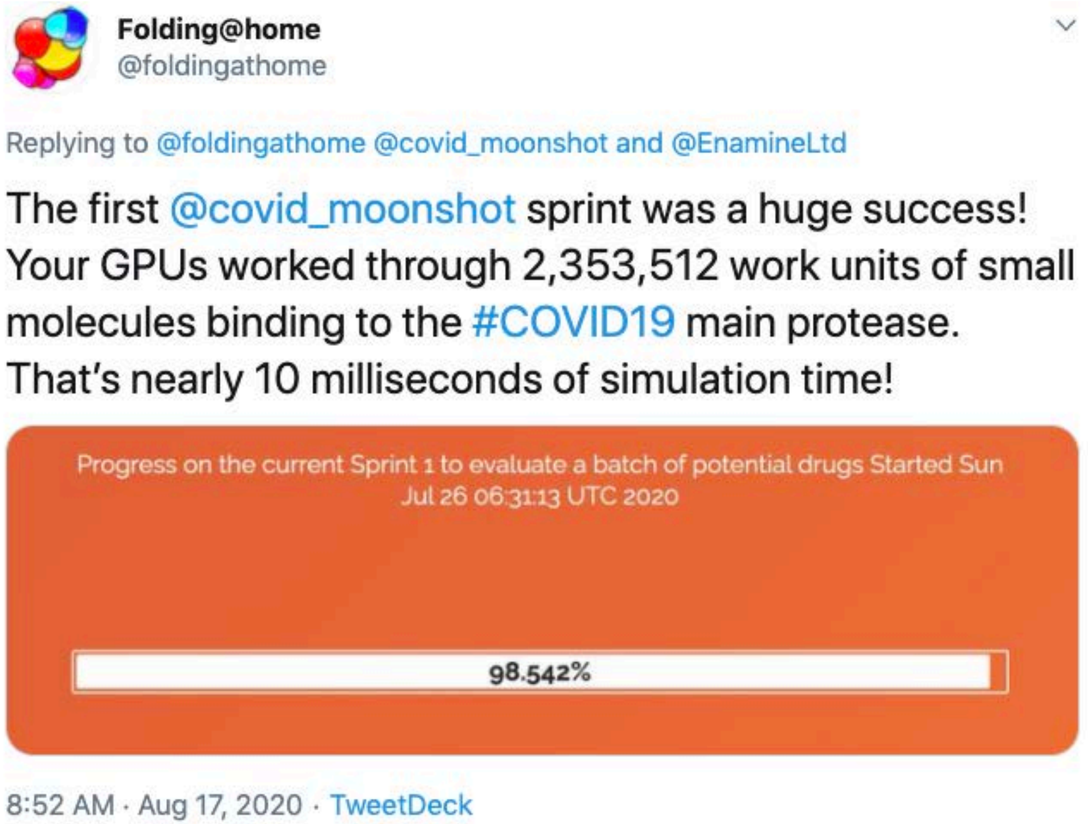
[Available for Windows, Mac, Linux]

Progress on the current Sprint 2 to evaluate a batch of potential drugs Started Sun Aug 16 01:00:00 UTC 2020

25.996%

The **progress bar** measures the fraction of compounds we could synthesize that we've evaluated for each sprint

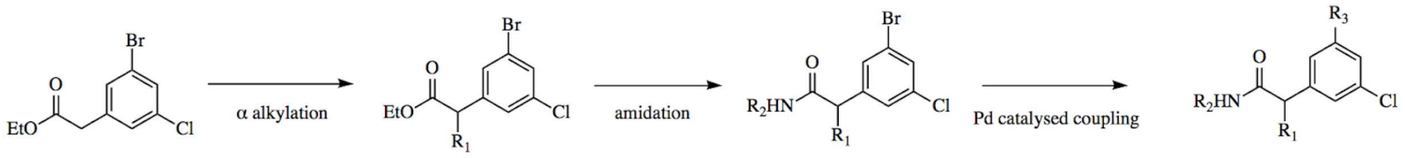
We generated a *lot* of data, which we have shared online via AWS



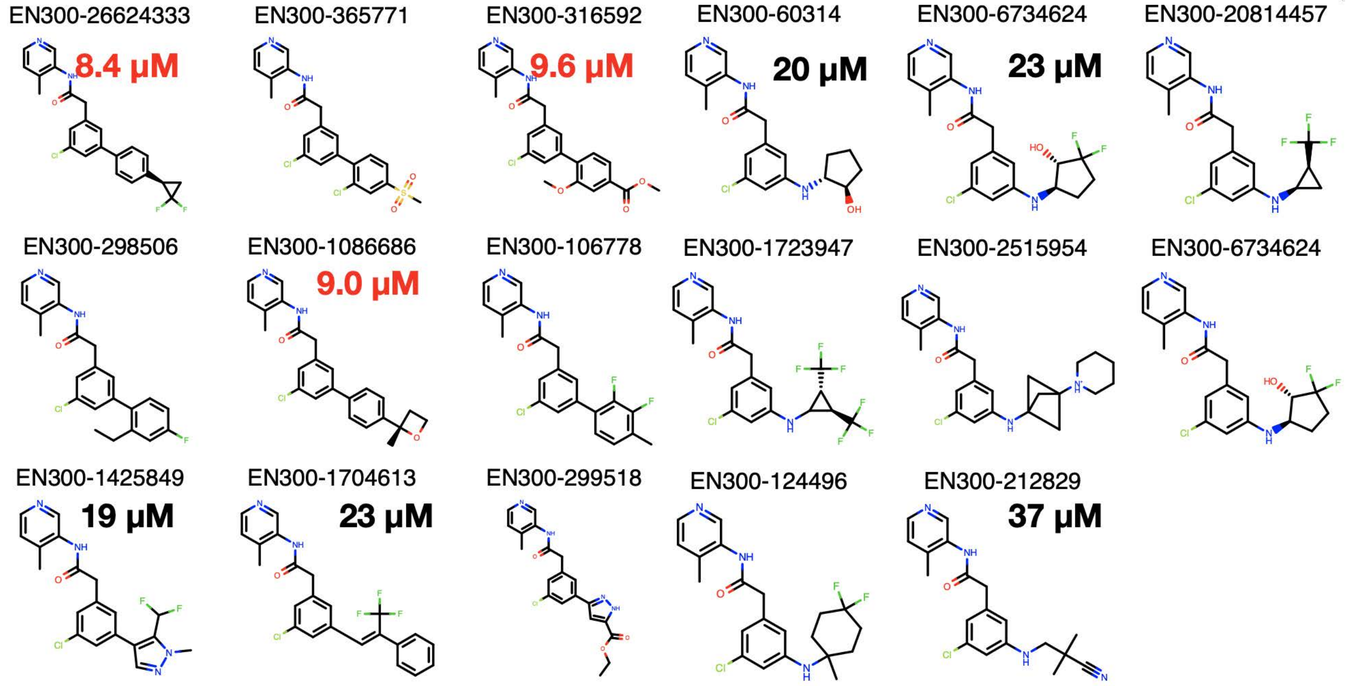
https://registry.opendata.aws/foldingathome-covid19/
https://covid.molssi.org//org-contributions/#folding--home

FREE ENERGY CALCULATIONS CAN RAPIDLY PRIORITIZE COMPOUNDS FROM LARGE VIRTUAL SYNTHETIC LIBRARIES

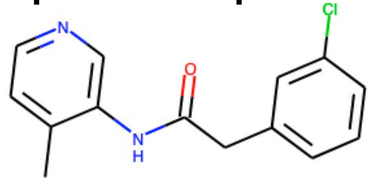
Can we engage S4 from this 5,000-compound virtual synthetic library varying R3



Top free energy calculation compounds and experimental affinity measurements:



parent compound



IC_{50} = 25 μM

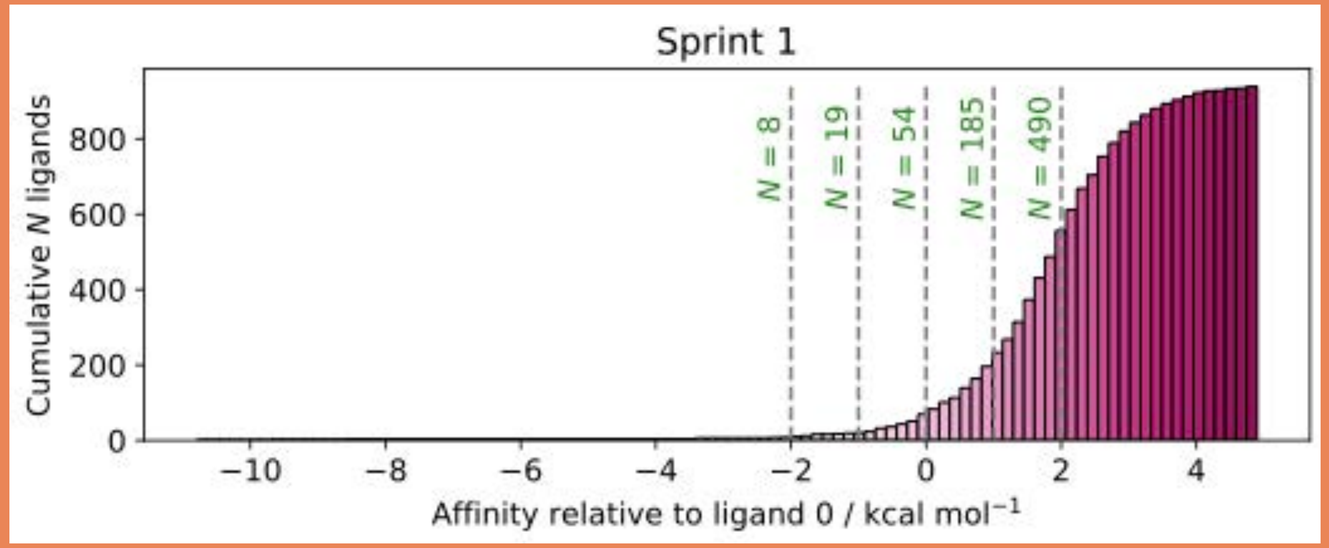
TRY-UNI-714a760b-6

top compounds from free energy calculations

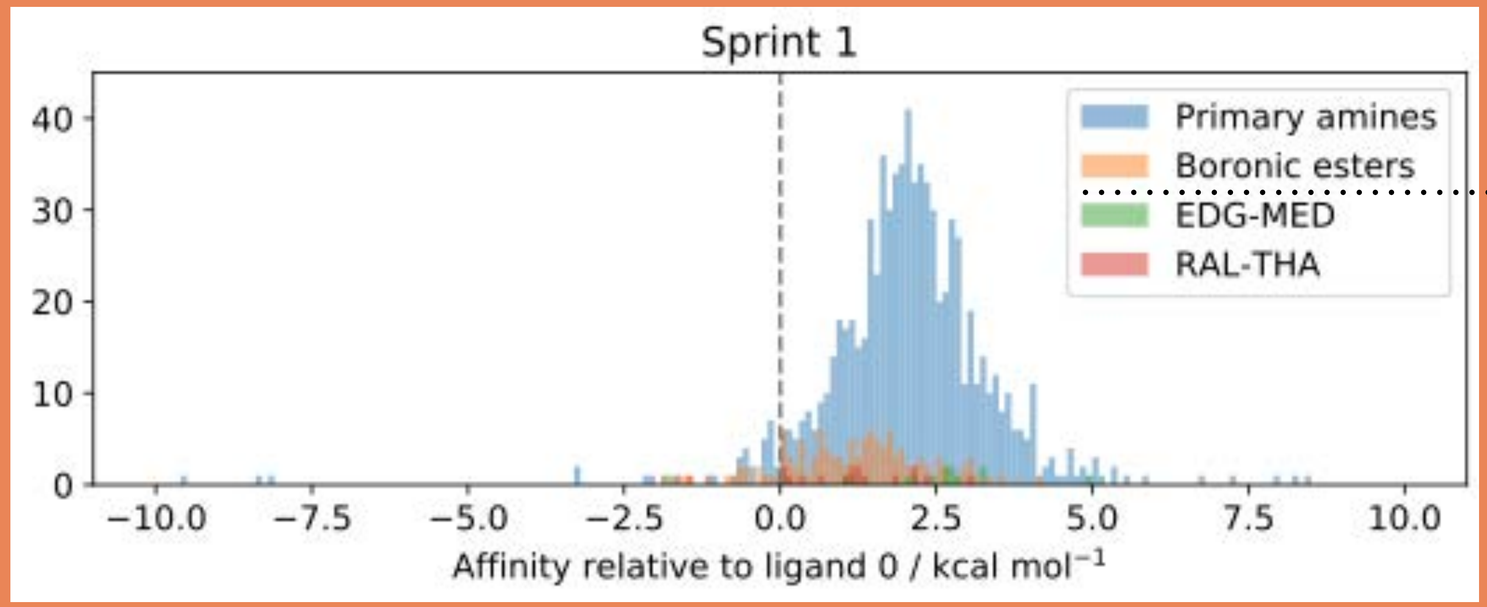
COVID Moonshot: [Moonshot] [Fragalysis] [Dashboard]

Most ideas were bad ideas

better **worse**



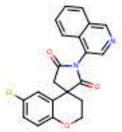
Human chemists seem better than random, but it's hard to get them to generate enough ideas



computer

humans

Sprint 5
Science Dashboard

7	BEN-BAS-c2bc0d80-6		c1ccc2c(c1)cncc2N3C(=0)CC4(C3=0)CC0c5c4cc(cc5)Cl

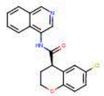
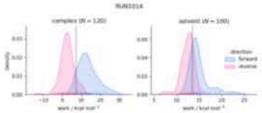
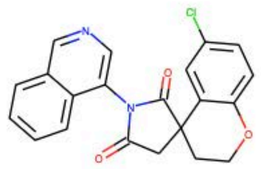
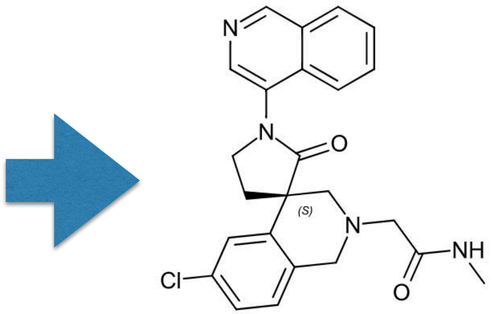
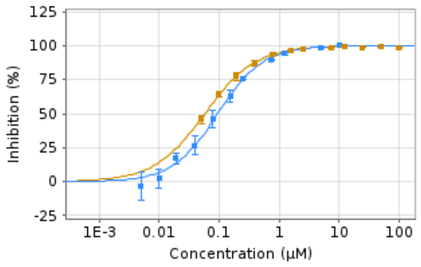
RUN1014	MAT-POS-b3e365b9-1_1		sdf	pdb	BEN-BAS-c2bc0d80-6_1		sdf	pdb	−6.2 ± 0.2	 

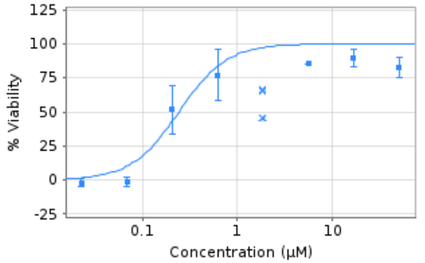
Image	Molecule	IC50 Curves	IC50 (μM) - Fluorescence
	BEN-BAS-c2bc0d80-6 O=C1CC2(CCOc3ccc(Cl)cc32)C(=O)N1c1cncc2ccccc12 3-aminopyridine-like Assayed Check Availability on Manifold	Fluorescence RapidFire	0.49



CVD-0019465
COVID Moonshot



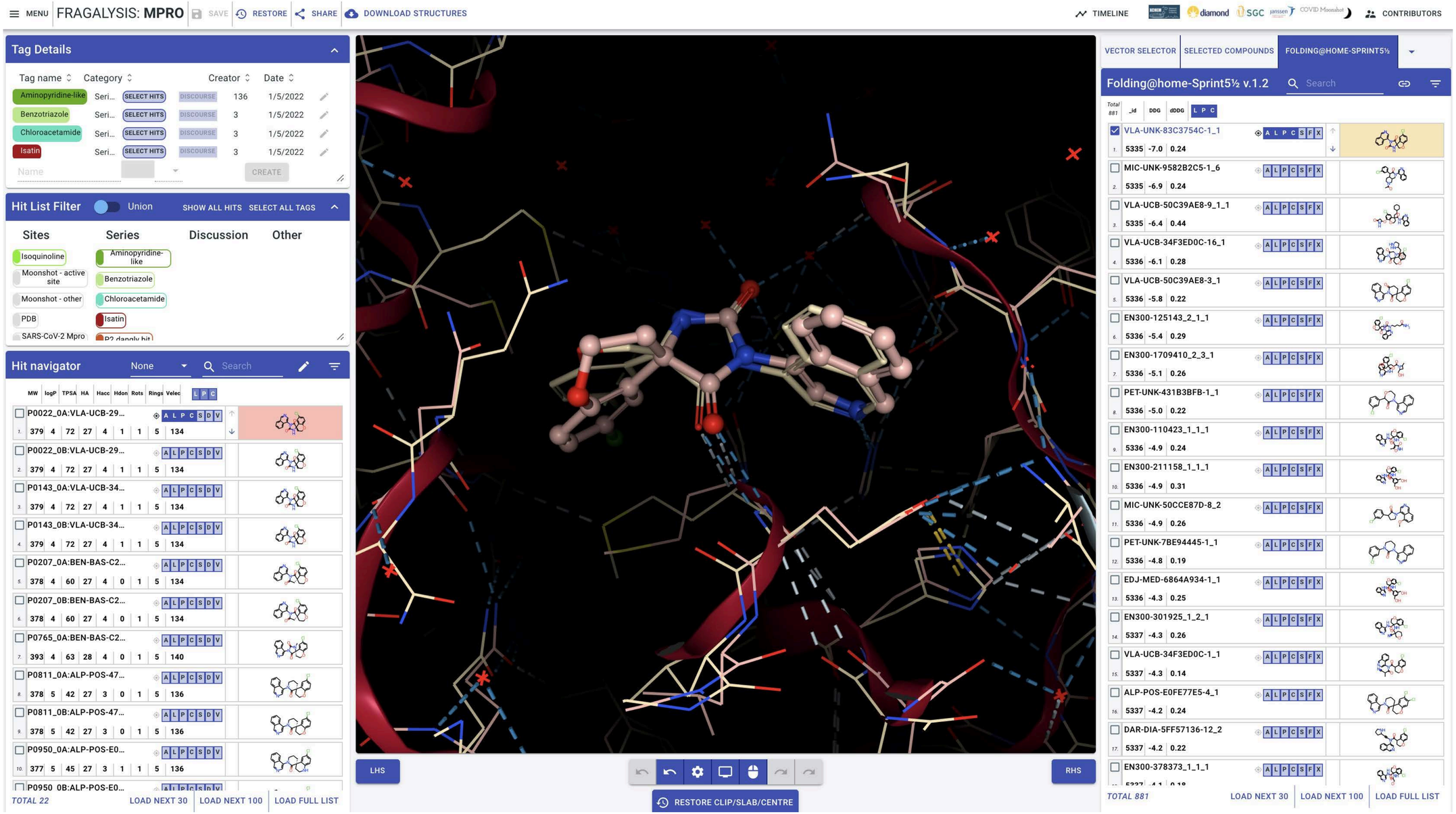
biochemical
IC50 : **74 nM**



antiviral
EC50 : **241 nM**

dashboard: https://tinyurl.com/fan-sprint-5-dimer
Fragalysis viewer: https://fragalysis.diamond.ac.uk/viewer/react/preview/target/Mpro

IT'S SURPRISING HOW WELL BINDING POSES CAN BE PREDICTED

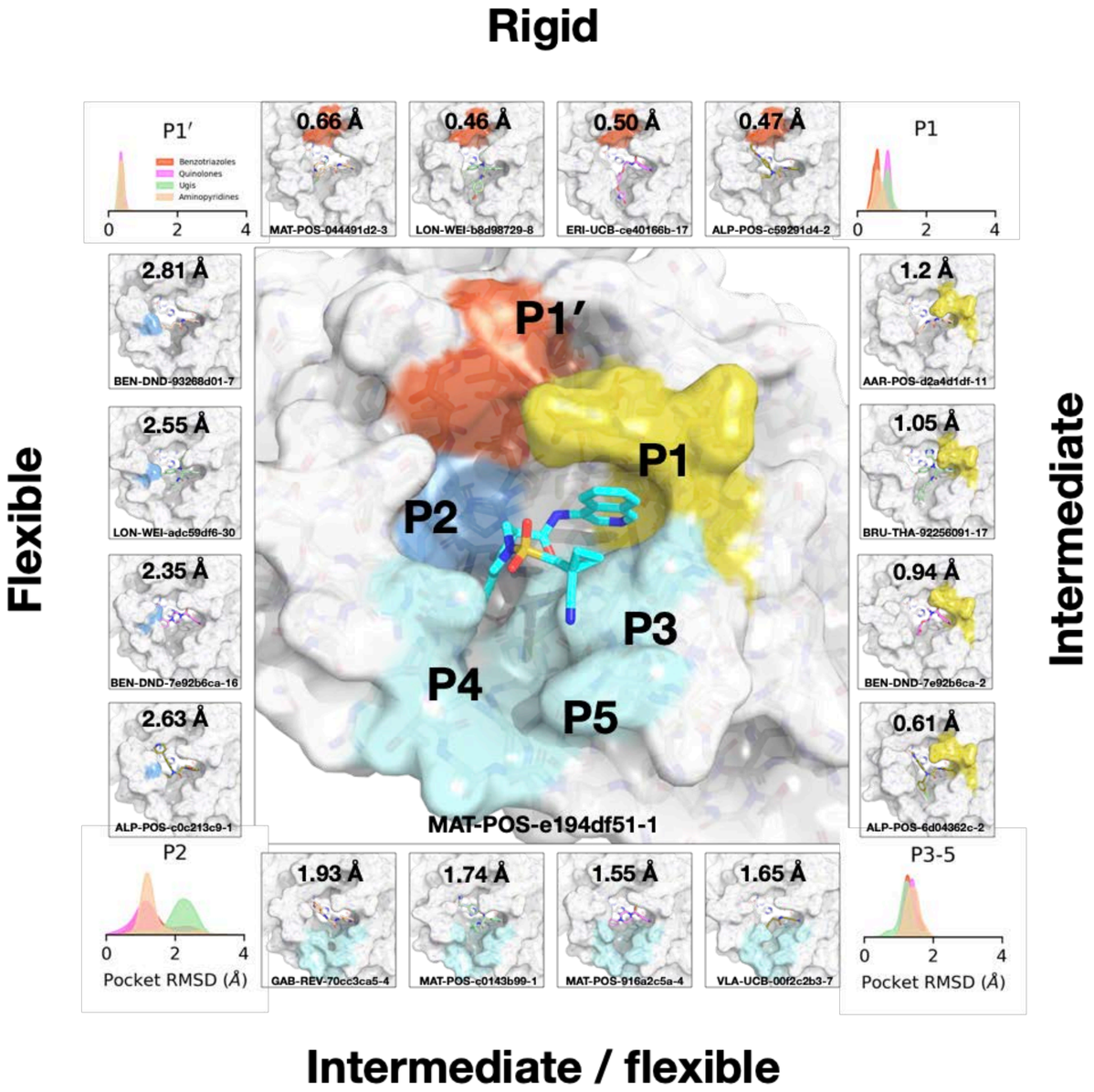


dashboard: https://tinyurl.com/fah-sprint-5-dimer

Fragalysis viewer: https://fragalysis.diamond.ac.uk/viewer/react/preview/target/Mpro

https://fragalysis.diamond.ac.uk/viewer/react/projects/1264/924

WE GENERATED HUNDREDS OF X-RAY STRUCTURES THAT MAP THE PLASTICITY OF THE BINDING SITE



534 X-ray structures posted
+ 357 more in refinement

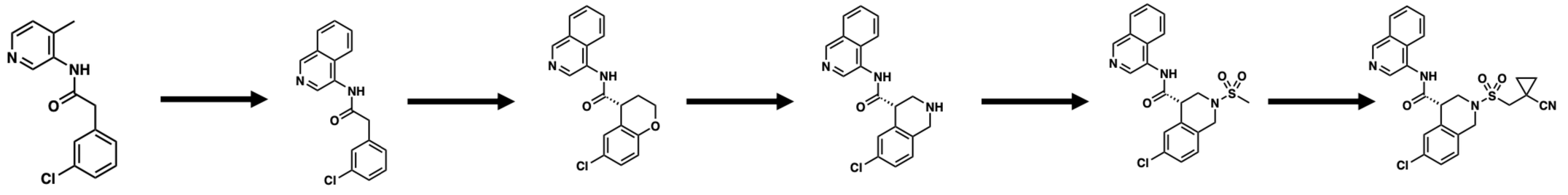
There are only 1188 X-ray structures of *all* SARS-CoV-2 proteins in the PDB!

https://fragalysis.diamond.ac.uk/viewer/react/preview/target/Mpro

SUCCESSIVE ROUNDS OF MEDICINAL CHEMISTRY PRODUCED POTENT MPRO INHIBITORS WITH ANTIVIRAL ACTIVITY

P1

P2

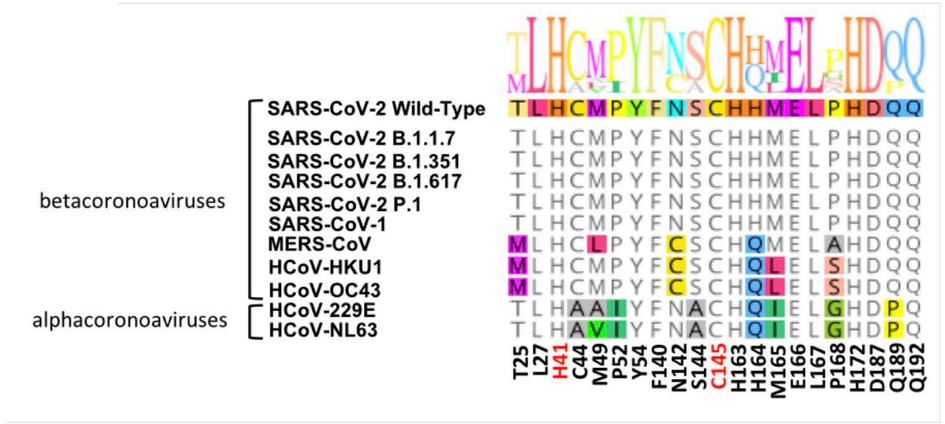


	TRY-UNI-714a760b-6	ADA-UCB-6c2cb422-1	MAT-POS-b3e365b9-1	MAT-POS-3ccb8ef6-1	MAT-POS-e194df51-1	MAT-POS-e194df51-1
IC_{50}(Mpro)/uM	25	0.73	0.21	0.28	0.141	0.037
EC_{50}(SARS-CoV-2, A549)/uM	n.d.	4.5	7.0	1.9	1.65	0.064

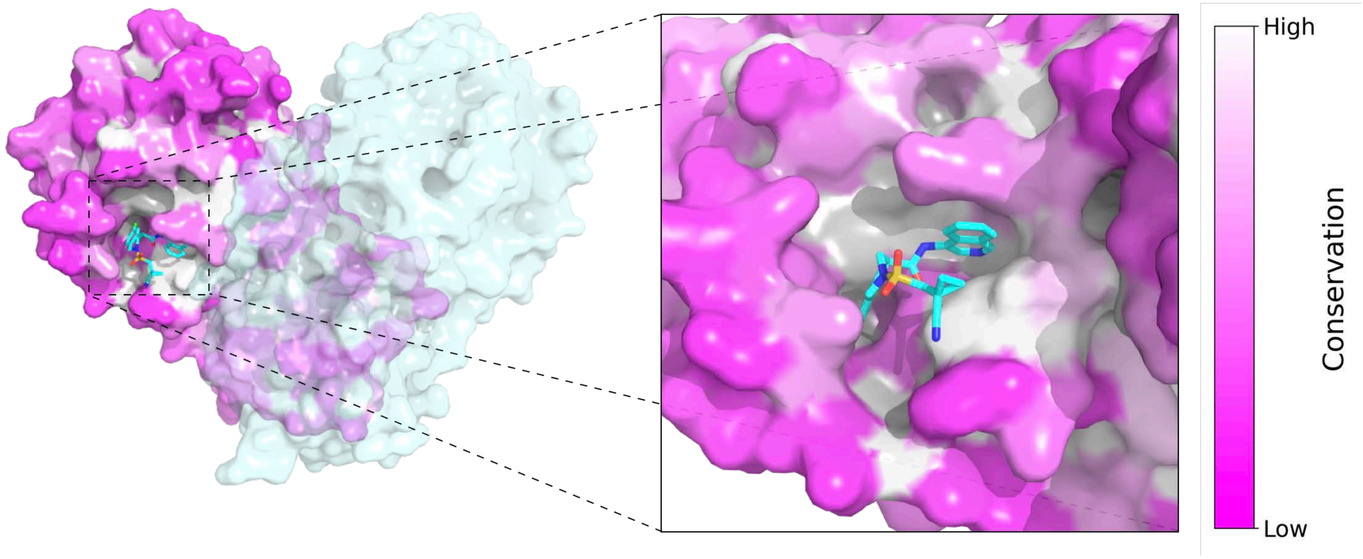
crowdsourced
merged fragment hit

OUR INHIBITORS ARE SMALL, NONCOVALENT, AND ENGAGE HIGHLY CONSERVED RESIDUES

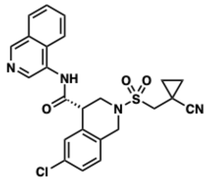
active-site residue conservation
of pathogenic coronaviruses



residue conservation
mapped onto Mpro structure



THE FIRST COMPOUND TO MEET OUR MEDICINAL CHEMISTRY TARGET PRODUCT PROFILE HAS ACHIEVABLE HUMAN DOSE PREDICTIONS



MAT-POS-e194df51-1

Antiviral efficacy				
Mpro IC50 /uM	0.037			
A549 IC50 /uM	0.064			
In vitro ADME				
LogD [measured]	2.5			
MDCK-LE FA (%)	92.9			
	Rat	**Dog**	**Minipig**	**Human**
Liver microsomes Cl ul/min/kg	604	164	542	152
Liver microsomes t ½ (min)	2.4	8.5	2.6	9.1
Heps Cl ul/min/kg	67.6	61.4	65.9	10.3
Heps t ½ (min)	10.3	11.3	10.5	67.5
PPB free fraction (%)	5.4			10.1
Safety / Drug-drug interactions				
Cyp450 (uM) 2C9/2D6/3A4	25/9.4/10.3			
PXR risk	Low			
Herg (uM)	>30			
In vivo pharmacokinetics				
Rat IV Vd (l/kg)	1.05			
Rat IV CL	34.8			
Rat t ½ IV/PO (h)	0.448 / 1.4			
Rat Bioavailability (%)	18			

human dose projections of 100-350 mg t.i.d.

CSH Cold Spring Harbor Laboratory

bioRxiv

THE PREPRINT SERVER FOR BIOLOGY

bioRxiv posts many COVID19-related papers. A reminder: they have not been formally peer-reviewed and should not guide health-related behavior or be reported in the press as conclusive.

New Results

Follow this preprint

Open Science Discovery of Oral Non-Covalent SARS-CoV-2 Main Protease Inhibitor Therapeutics

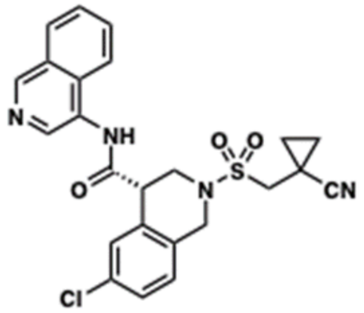
https://doi.org/10.1101/2020.10.29.339317

(updated Mon 31 Jan)

Over 180 contributors/authors:
https://tinyurl.com/covid-moonshot-authors

We're still actively pursuing multiple backups to enter an accelerated preclinical program

THIS COMPOUND HAS EXCELLENT ANTIVIRAL ACTIVITY AGAINST ALL VARIANTS



MAT-POS-e194df51-1

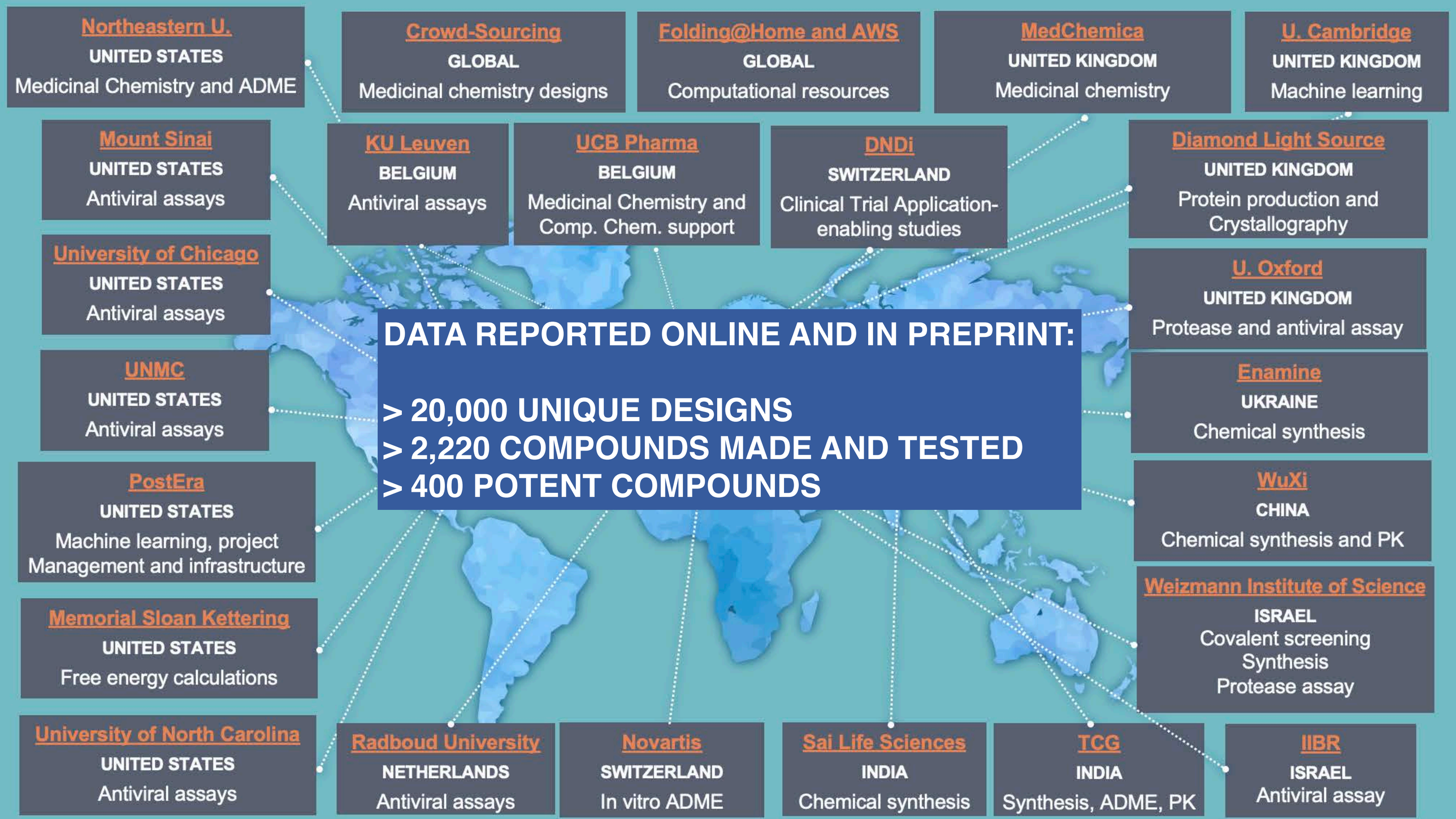
37 nM SARS-CoV-2 Mpro IC_{50} (enzymatic)
64 nM SARS-CoV-2 antiviral EC_{50} (A549 cells)

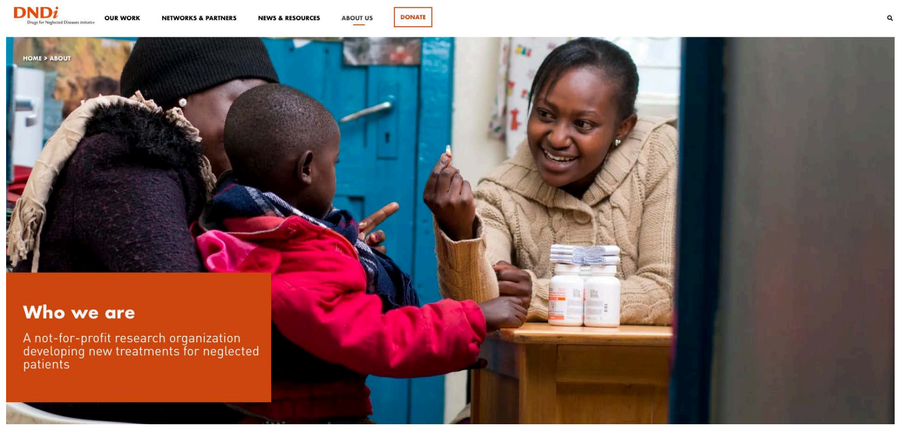
	MAT-POS-e194df51-1		Nirmatrelvir	
	IC50	CC50	IC50	CC50
Alpha variant (B.1.1.7.	0.38	>20	0.12	>10
Beta variant (B.1.351)	1.48	>20	0.21	>10
Delta variant (B.1.617.2)	1.52	>20	0.21	>10
Omicron variant (B.1.529)	0.29	>20	0.07	>10
MA-SARS-CoV-2/WA1	0.43	>20	0.14	>10

(micromolar)

CPE assay in HelaACE2 cells

https://doi.org/10.1101/2020.10.29.339317


Northeastern U.
UNITED STATES
Medicinal Chemistry and ADME
Crowd-Sourcing
GLOBAL
Medicinal chemistry designs
Folding@Home and AWS
GLOBAL
Computational resources
MedChemica
UNITED KINGDOM
Medicinal chemistry
U. Cambridge
UNITED KINGDOM
Machine learning
Mount Sinai
UNITED STATES
Antiviral assays
KU Leuven
BELGIUM
Antiviral assays
UCB Pharma
BELGIUM
Medicinal Chemistry and Comp. Chem. support
DNDi
SWITZERLAND
Clinical Trial Application-enabling studies
Diamond Light Source
UNITED KINGDOM
Protein production and Crystallography
University of Chicago
UNITED STATES
Antiviral assays
U. Oxford
UNITED KINGDOM
Protease and antiviral assay
DATA REPORTED ONLINE AND IN PREPRINT:
> 20,000 UNIQUE DESIGNS
> 2,220 COMPOUNDS MADE AND TESTED
> 400 POTENT COMPOUNDS
UNMC
UNITED STATES
Antiviral assays
Enamine
UKRAINE
Chemical synthesis
PostEra
UNITED STATES
Machine learning, project Management and infrastructure
WuXi
CHINA
Chemical synthesis and PK
Weizmann Institute of Science
ISRAEL
Covalent screening
Synthesis
Protease assay
Memorial Sloan Kettering
UNITED STATES
Free energy calculations
University of North Carolina
UNITED STATES
Antiviral assays
Radboud University
NETHERLANDS
Antiviral assays
Novartis
SWITZERLAND
In vitro ADME
Sai Life Sciences
INDIA
Chemical synthesis
TCG
INDIA
Synthesis, ADME, PK
IIBR
ISRAEL
Antiviral assay



COVID Moonshot funded by COVID-19 Therapeutics Accelerator to rapidly develop a safe, globally accessible and affordable antiviral pill

GENEVA / OXFORD / NEW YORK / TEL AVIV — 27 SEP 2021

The COVID Moonshot, a non-profit, open-science consortium of scientists from around the world dedicated to the discovery of globally affordable and easily-manufactured antiviral drugs against COVID-19 and future viral pandemics has received key funding of £8 million from Wellcome, on behalf of the Covid-19 Therapeutics Accelerator.

'*Faced with global vaccine inequality and the rapid spread of variants of concern, the need for easily-accessible antiviral therapeutics to treat people with COVID-19 is as pressing as ever, especially in low- and middle-income countries*,' said Annette von Delft, Translational Scientist at the University of Oxford and NIHR Oxford Biomedical Research Centre.

'*Most of the research and funding efforts early in the pandemic focused predominantly on repurposing of existing small molecule drugs and the more rapid development of novel monoclonal antibodies. Now, with the realization that COVID-19 will be a global issue for the foreseeable future we urgently need to develop novel antiviral therapeutics. We are therefore thrilled to receive this critical funding from Wellcome and hope it can lead to more support*,' said Alpha Lee, Chief Scientific Officer at PostEra and Faculty Member at the University of Cambridge.

The Moonshot started as a spontaneous virtual collaboration in March 2020. As countries locked down, a group of scientists, academics, pharmaceutical research teams and students began a worldwide, twitter-fuelled race against the clock to identify new molecules that could block SARS-CoV-2 infection and develop pills that would be readily available to the most vulnerable communities.

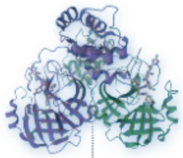
Ultimately more than 150 scientists – including dozens of students who put their own projects on hold – joined Moonshot to crowdsource ideas for molecular compounds, model them and evaluate them in-vitro against the virus. Their goal: a safe, globally affordable, not-for-profit oral treatment for COVID-19 and related viral pandemics.

https://dndi.org/press-releases/2021/covid-moonshot-funded-by-wellcome-to-rapidly-develop-safe-globally-accessible-affordable-antiviral-pill/

WE WENT FROM FRAGMENT SCREEN TO PRECLINICAL PHASE IN JUST 18 MONTHS, SPENDING LESS THAN $1M

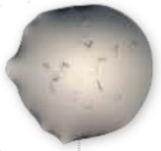
Feb 14 2020

Main protease cloned and produced at Diamond



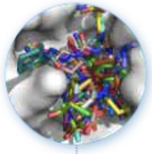
Feb 20 2020

Atomic resolution structure of the protease determined



Mar 5 2020

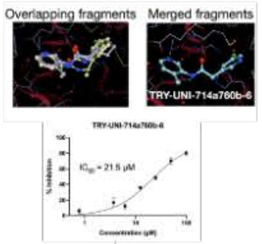
1,500 crystals collected in one day



Mar 7 2020

78 fragment-bound structures solved and released to the web

48 covalent fragments
71 active site fragments

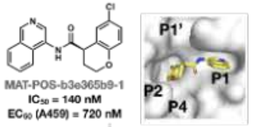


Mar 17 2020

Registration of the hit designed using a fragment merge strategy

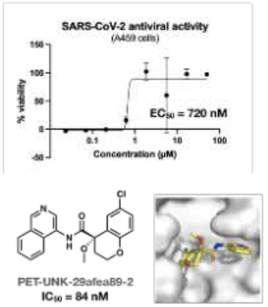
Jul 19 2020

First nM inhibitor in the series



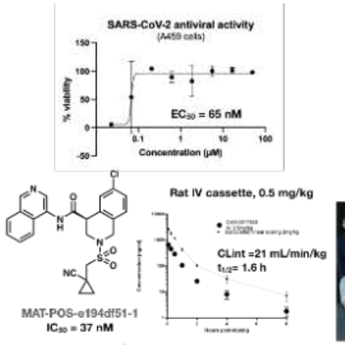
Aug 08 2020

First compound with measurable antiviral activity



Feb 02 2021

First compound with sub nM antiviral activity



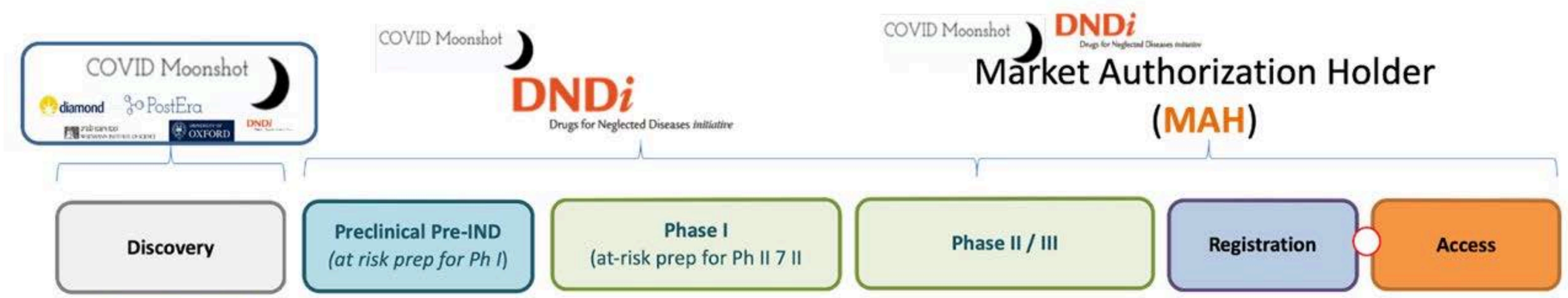
Jun 9 2021

First compound with sub nM antiviral activity and acceptable rodent PK

Jul 1 2021

Wellcome Trust/ ACTA announces $11m funding to The COVID Moonshot for IND-enabling studies

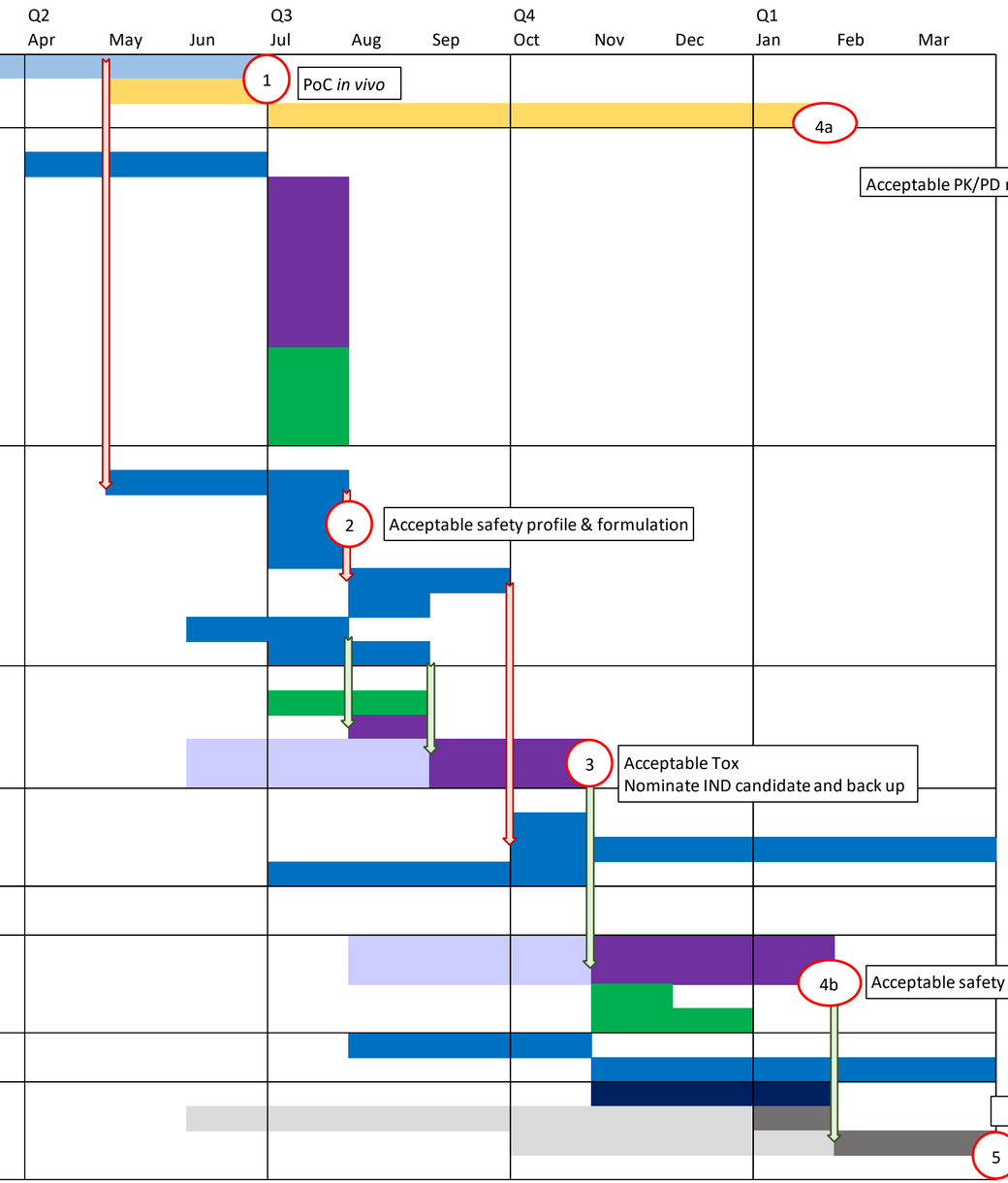
WE'RE AIMING TO BRING AN ANTIVIRAL TO MANUFACTURE WITH MINIMAL OR NO IP



We have a path to go "straight to generics" (potentially entirely free of patents)
to enable true, low-cost global access to meet the needs of underserved LMICs

Getting to Investigational New Drug (IND) approval in <1 year is complex and expensive

			No/cmpds
Preparation of regulatory phase			
WP 1	**Identify up to 5 optimized leads**	Medicinal Chemistry	
		Efficacy	
		Efficacy vs PK/PD	
WP2	**Build Optimized Lead profile package**	**5 compounds**	
		Compound synthesis at risk for WP2 (5-10g)	at risk
		In Silico DEREK / UCB D2P2	5
	pure sample compound	Safety pharmacology (Ames, Genotox)	5
		Hepatotoxicity cellular assay	5
		Patch clamp cardiomyocytes	5
		Safety 44 panel receptor binding panel	5
		Check CNS exposure - initial assessment (Irwin test)	5
		Acceptable drug interaction profile (CYP, TDI)	5
		ADME for human dose prediction	5
		Multi-dosing study (rat)	5
		Multi-dosing study (dog)	5
		Ascending dose study (rat)	5
WP3	**Optimized lead scale-up**	**5 compounds**	
		Process chemistry assessment of scale up feasibility	at risk 5
		Formulation assessment	5
		Formulation for PK	5
		2-3 compounds	
		Formulation Phase 1	3
		Forced degradation study	3
		Compound synthesis for 7 day tox	at risk 5
		Compound synthesis for DRF pilot toxicology	at risk 5
WP4	**Exploratory toxicology**	**2-3 compounds**	
		Bioanalysis validated (ICH), rat, dog, human	3
		7 d ascending dose (non GLP) tox rat	3
		Dose Range Finding (DRF) pilot toxicology - rat	3
		Dose Range Finding (DRF) pilot toxicology - dog	3
WP5	**Large-scale synthesis/stability**	**2-3 compounds**	
		Stability in capsule	at risk 3
		Stability (3 - 6 months)	at risk 3
		Large scale synthesis for GLP tox	at risk 5
Regulatory phase		**1 compound (and back-up)**	
WP6	GLP toxicology and safety	GLP tox - 1 month dog (assume 5 days dosing)	1
		GLP tox - 1 month rat (assume 5 days dosing)	1
		Acceptable PK (with a validated bioanalytical method) rat/dog GLP	1
		PK scaling and Dose to Human prediction	1
WP7	GMP manufacture	GMP manufacture feasibility	at risk 3
		GMP manufacture	1
WP8	Regulatory assessments	Develop clinical endpoints	1
		Regulatory assessments	1
		Clear IND regulatory path	1
		HPOC/CPOC plan is acceptable to regulatory agency	1



$/milestone	
1	570000
2	607500
3	995000
4	1085000
5	1710000
	4967500 Total

Efficacy
CMC
Pharmacokinetics
Regulatory
Clinical
Safety

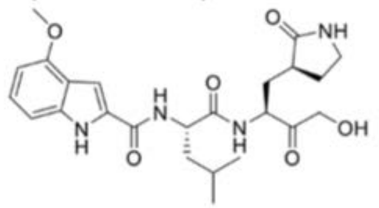
COVID Moonshot

PFIZER DEVELOPED THEIR IV MPRO INHIBITOR INTO AN ORAL ANTIVIRAL IN RECORD TIME

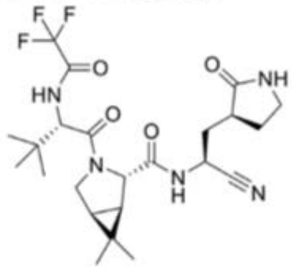
intravenous antiviral
(clinical trials paused)

paxlovid oral antiviral
(co-dosed with ritonavir as bait for CYPs)

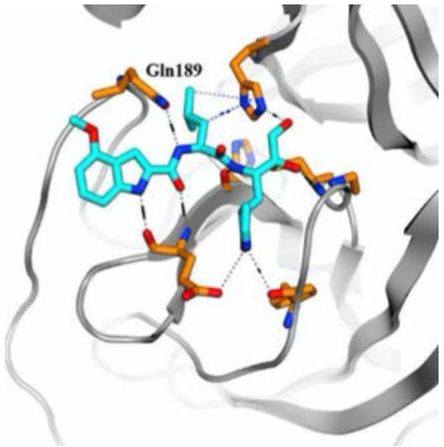
1 (PF-00835231)



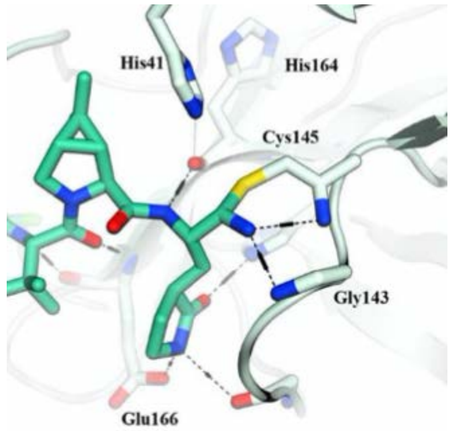
6 (PF-07321332)



350 people
roughly $1B
11 months from start to clinic
clinical trials Mar-Nov 2021



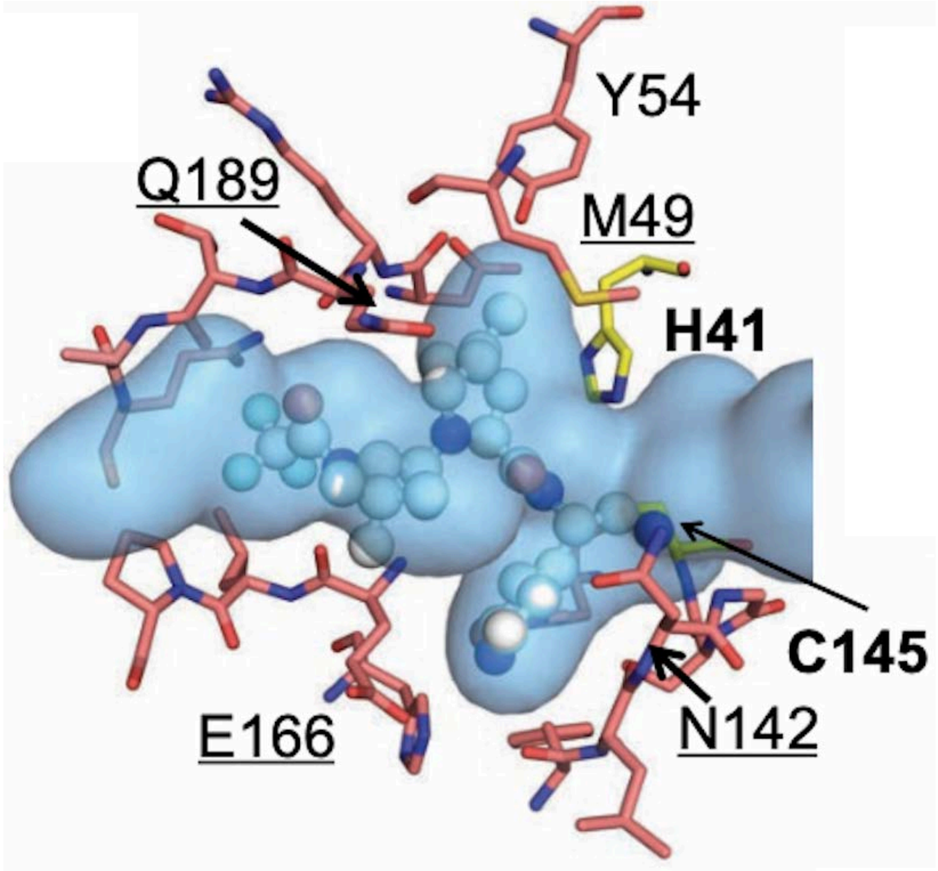
Ki 0.3 [0.2, 0.5] nM
EC50 230 [160, 340] nM (VeroE6)
1.4% oral bioavailability



Ki 3 [1,7] nM
EC50 75 [66,83] nM (VeroE6)
50% oral bioavailbility

https://www.science.org/doi/10.1126/science.abl4784

WE STILL NEED MORE THAN ONE ORAL ANTIVIRAL



Defining the Substrate Envelope of SARS-CoV-2 Main Protease to Predict and Avoid Drug Resistance

Ala M. Shaqra, Sarah Zvornicanin, Qiu Yu Huang, Gordon J. Lockbaum, Mark Knapp, Laura Tandeske, David T. Barkan, Julia Flynn, Daniel N.A. Bolon, Stephanie Moquin, Dustin Dovala, Nese Kurt Yilmaz, Celia A. Schiffer

doi: https://doi.org/10.1101/2022.01.25.477757

https://www.biorxiv.org/content/10.1101/2022.01.25.477757v1

4 CONTRAINDICATIONS

PAXLOVID is contraindicated in patients with a history of clinically significant hypersensitivity reactions [e.g., toxic epidermal necrolysis (TEN) or Stevens-Johnson syndrome] to its active ingredients (nirmatrelvir or ritonavir) or any other components of the product.

PAXLOVID is contraindicated with drugs that are highly dependent on CYP3A for clearance and for which elevated concentrations are associated with serious and/or life-threatening reactions *[see Drug Interactions (7.3)]*:

- Alpha$_1$-adrenoreceptor antagonist: alfuzosin
- Analgesics: pethidine, piroxicam, propoxyphene
- Antianginal: ranolazine
- Antiarrhythmic: amiodarone, dronedarone, flecainide, propafenone, quinidine
- Anti-gout: colchicine
- Antipsychotics: lurasidone, pimozide, clozapine
- Ergot derivatives: dihydroergotamine, ergotamine, methylergonovine
- HMG-CoA reductase inhibitors: lovastatin, simvastatin
- PDE5 inhibitor: sildenafil (Revatio®) when used for pulmonary arterial hypertension (PAH)
- Sedative/hypnotics: triazolam, oral midazolam

PAXLOVID is contraindicated with drugs that are potent CYP3A inducers where significantly reduced nirmatrelvir or ritonavir plasma concentrations may be associated with the potential for loss of virologic response and possible resistance. PAXLOVID cannot be started immediately after discontinuation of any of the following medications due to the delayed offset of the recently discontinued CYP3A inducer *[see Drug Interactions (7.3)]*:

- Anticancer drugs: apalutamide
- Anticonvulsant: carbamazepine, phenobarbital, phenytoin
- Antimycobacterials: rifampin
- Herbal products: St. John's Wort (*hypericum perforatum*)

EUA contains **seven pages** of drug-drug interactions, leaving a significant vulnerable untreated population

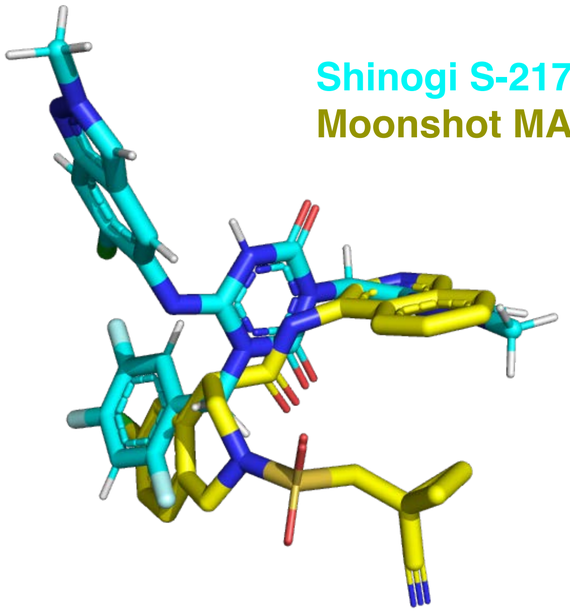
FDA Paxlovid Emergency Use Authorization
https://www.fda.gov/media/155050/download

SHINOGI RECENTLY REPORTED THE DISCOVERY OF S-217622, DISCOVERED WITH THE HELP OF MOONSHOT DATA

COVID Moonshot molecules and X-ray structures informed pharmacophore used to identify compound in internal collection for pain program

Rapidly developed into potent antiviral with extraordinary PK (one pill/day!)

Currently in Phase 3 trials with readout expected soon



Shinogi S-217622
Moonshot MAT-POS-e194df51-1

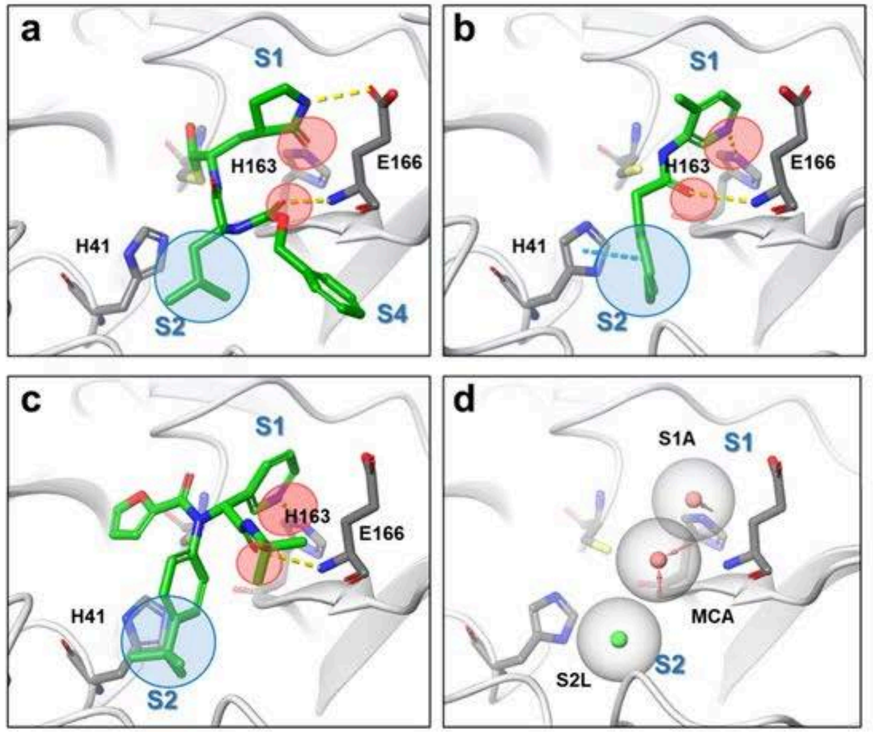
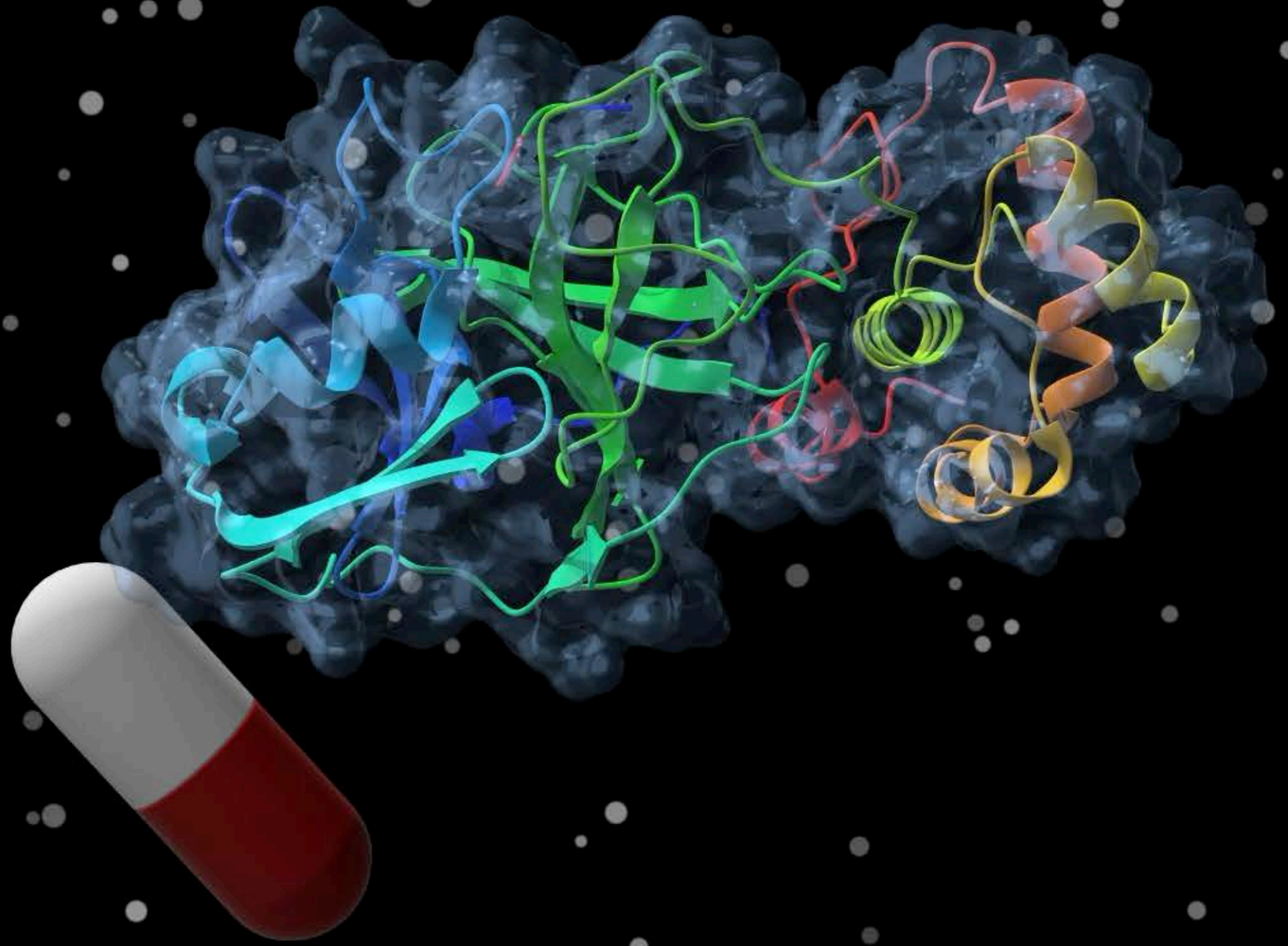


Figure 2. Binding modes of 3CL^{pro} inhibitors, their pharmacophores, and defined pharmacophore filters for virtual screening. (a) Crystal structures of GC376 (PDB: 6WTT), (b) 3-aminopyridine-like compound of the Postera COVID moonshot project (PDB: 5RH2) and (c) ML188 (PDB: 7L0D). The common H-bond acceptors are circled in red; the common hydrophobic pharmacophores are circled in blue. (d) Common pharmacophores shared with inhibitors A–C. Red and green spheres represent H-bond acceptors and lipophilic features, respectively.

https://www.biorxiv.org/content/10.1101/2022.01.26.477782v1

A GLIMMER OF HOPE



THE ONLY REASON WE DIDN'T HAVE ANTIVIRALS FOR SARS-COV-2 WAS DUE TO MARKET FAILURE

Comment

A white-knuckle ride of open COVID drug discovery

Frank von Delft, John Chodera, Ed Griffen, Alpha Lee, Nir London, Tatiana Matviuk, Ben Perry, Matt Robinson, Mark Calmiano & Annette von Delft

In early 2020, a spontaneous global collaboration came together to design a new, urgent antiviral treatment. There are lessons in what happened next.

Nearly 15 months ago, a large, fast-moving and unscheduled experiment began: probing a key protein of the coronavirus SARS-CoV-2 to find chemical starting points for drug discovery. The end point was to develop pills that people could take to treat COVID-19 and related diseases.

This experiment pulled together a spontaneous, open, global, Twitter-fuelled collaboration called the COVID Moonshot. Urgency and a commitment to working openly recruited more than 150 active participants, spanning a huge range of expertise and technology across academia, biotechnology, pharmaceuticals and more, all working without claiming intellectual property. Open drug-discovery efforts are invariably super slow — ours has been an express train on tracks we have laid down as we go. It is a way of working that none of us realized was possible.

The intention for the original experiment was simply to help jump-start large drug-discovery initiatives that could draw directly on our data. In those first weeks, before the pandemic had taken hold in the United Kingdom or Israel (where the experiment started), we expected that some international effort was already in the works for countries and companies to collaborate on finding COVID-19 treatments, as was happening with vaccines.

Disappointingly, from the start of the COVID-19 fight, international funders decided to support only the development of repurposed small-molecule drugs and monoclonal antibodies to deliver treatments quickly, neglecting other approaches. The world seemed to give up on new antivirals before they even started, agreeing on a self-fulfilling prophesy that such drugs would take years to develop. Few seemed willing to contemplate such a timescale for this pandemic. Our first grant proposal was rejected, so we had to find a different way to press on.

Amazing virtual collaborations sprang up around the pandemic in many fields: bioinformaticians and phylogeneticists worked out ways to track new variants. Epidemiologists and computer modellers ran simulations. The World Health Organization activated a network of experts to vet new publications and preprints. Military personnel transported medical equipment and vaccines, and set up community testing centres.

Our COVID Moonshot is different. Rather than engaging with patients while using personal protective equipment, we work in chemistry hoods and with spectrometers, X-rays, computer models and courier companies. It's driven by a conviction that conventional wisdom is wrong about *de novo* drug discovery being a job only for big pharma and peripheral to a fast-moving global outbreak: the pandemic is still here, and antiviral drugs against COVID-19 are not.

The screens

Drug-discovery efforts generally require a target, such as a protein that has an important role in disease. Promising drug compounds bind to the protein, affect its function and act safely in the body. Diamond Light Source near Oxford is the UK national synchrotron — a particle accelerator essential for modern X-ray crystallography, the go-to technique for determining 3D structures of proteins. There, one of us (F.v.D.) leads the XChem facility that uses the technique to screen for very small compounds called fragments that bind to drug targets. Although these 'fragment hits' bind weakly and the throughput is low compared with other techniques (screening fewer than 1,000 compounds per experiment), the 3D structures show exactly how each fragment binds. This provides powerful clues about how to create bigger, more potent molecules.

By late January 2020, scientists in China had solved the first 3D crystal structures of the SARS-CoV-2 main protease (M^{pro}), an essential viral enzyme, and made them public. With their guidance, a group at Diamond led by Martin Walsh generated new, high-quality crystals by mid-February — lightning fast for such work. The group also shipped M^{pro} protein to the Weizmann Institute of Science in Rehovot, Israel, where N.L.'s group uses mass spectrometry to quickly identify covalent fragments that attach to proteins irreversibly. This is another way to find useful starting points for drugs.

Racing to exploit the two weeks before scheduled shutdown of the synchrotron on 6 March last year, more than a dozen scientists from the Walsh, F.v.D. and N.L. groups dropped everything to complete an XChem experiment four times the normal size[1]. All the data were analysed within one month, and as soon as we had the first batch of results, we posted downloadable data and a short write-up on the Diamond web page, then tweeted the link on 7 March (see go.nature.com/3vju8vb).

The tweets

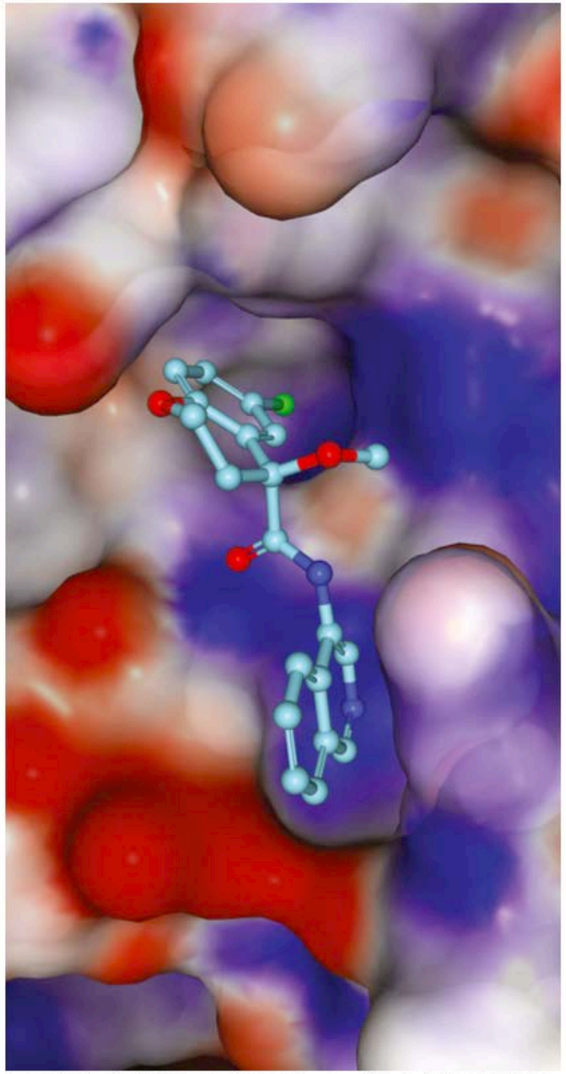
The response surprised us: almost 1,000 retweets in a week, and diverse offers for help. A.L. and M.R., two co-founders of the US-based technology firm PostEra, got in touch to suggest that their machine-learning technology could propose synthetic routes to make new molecules inspired by the fragment hits. But first we needed drug-like molecules to be designed, and N.L. realized whom we could ask: medicinal chemists newly under lockdown restrictions but full of expertise and desperate to help.

The next step was a tweet to crowdsource ideas for such molecules, declaring that we would make and test the best ones. A web page built by M.R. and his team in 48 hours enabled participants to submit machine-readable suggestions for compounds. The site made clear that contributions would have no strings attached, no intellectual property and no remuneration. We expected a few hundred submissions at most — in two weeks, we had more than 4,000, and had to work out how to test them.

The experiments

From March to May last year, we were on Zoom calls almost daily, lining up collaborators, logistics, expertise, funding, institutional support and permissions. All around us, the world was shutting down. We were trying to work out how to keep ourselves, our colleagues and our families sane, and our laboratories open.

We tapped an inexhaustible wellspring of goodwill. At the Ukrainian company Enamine, T.M. convinced management to commit to doing synthesis at cost, and to handle compound logistics. Its 650 chemists made molecules to order and have a renowned collection of building blocks for quick synthesis. In early May, new compounds were being shipped weekly from Enamine to organizations in four countries, and that work continues. Two other contract research organizations, WuXi in China and Sai Life Sciences in India, pitched in with offers of chemists and discounts.



Crystal structure of a COVID Moonshot advanced compound (turquoise) in the active site of the SARS-CoV-2 main protease. The molecular surface colours show electrostatic charge.

Chris Schofield and his team at the University of Oxford, UK, together with Haim Barr and his colleagues at the Weizmann Institute, developed distinct biochemical assays that were key to cross-validating how well molecules inhibited the working M^{pro} enzyme. At the same time, for all compounds, the 3D mode of binding was assessed at Diamond in crystal structures. Half a dozen graduate students and postdocs suspended their own projects to coordinate, run and evaluate these assays, week after week. The work hasn't stopped since.

By mid-April 2020, a volunteer troop of industry-based medicinal chemists, chaired by E.G., were holding weekly meetings to scrutinize submissions, review results, discuss strategies, design molecules and coordinate with synthetic chemists at Enamine. This work continues, too.

Computational chemists assembled their own team through their own network, then met weekly to work out algorithms to rank submissions. J.C. developed new ways to use Folding@home, the world's largest crowdsourced supercomputer, which was already being used to generate models of viral proteins. It crunched 'free energy' calculations to predict the best binders for up to 10,000 compounds a week: 100 times more than had been attempted before.

Pharmaceutical companies develop elaborate information systems to track, store and analyse compounds and their associated data; our global effort urgently needed this, too. The informatics web platform CDD Vault donated us cloud space in its infrastructure just hours after a phone call, also arranging training and support. Many other vendors provided licences for free, and XChem's platform for sharing 3D data, the Fragalysis cloud, had fortunately just been released. M.R. built a back-end system that sent all data live on GitHub, which is more often used as a repository for programming code.

As the pandemic unfolded, on some calls, you could hear the ambulance sirens from half a world away. The first agenda item of every meeting was a list of participants' latest constraints — lockdowns, lab closures and home-schooling. Children made regular Zoom appearances, and at least two of us came down with COVID-19 ourselves. People pulled their weight not for glory or reward, but because there was a job that needed doing, and it was one that they could do.

To cells and live virus

By June 2020, the Zoom-based collaboration had identified sets of molecules that clearly inhibited a crucial viral protein. The next step was to test antiviral activity in living cells. These are complex experiments, requiring level-three biosafety labs certified for airborne pathogens. A.v.D., a translational clinician, coordinated a shifting coalition of groups. One virologist friend and colleague lived a 10-minute walk away, and they planned experiments on lockdown evening strolls. Other virology groups responded to our tweet for help, and offered a variety of assays. Compounds were shipped, early results trickled in and some compounds unambiguously stalled the virus. These initial successes were crucial, both scientifically and for morale.

Researchers at the Israel Institute for Biological Research near Rehovot agreed to run a single test plate once we had molecules that were sufficiently potent. When that test showed signs of drug-like activity, they worked out how to conduct regular measurements, filling a crucial gap in our testing cascade.

By September, we had reached a milestone with a chemical series that instilled confidence: the compounds inhibited enzymes at submicromolar concentrations, and blocked viral activity at single-digit micromolar concentrations.

The slog

Since then, for the past nine months, the project has entered familiar territory in medicinal chemistry: we have been tweaking and testing compound designs, and optimizing early lead molecules so that they behave like drugs — entering the blood and staying there without being toxic. Potency against the M^{pro} enzyme has improved 100-fold, as has antiviral activity, and we are honing compounds' solubility and rate of metabolism by the liver.

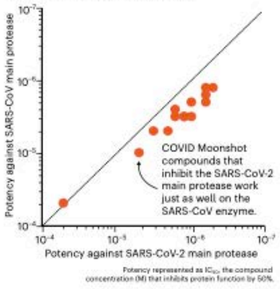
Above all, we can start predicting that these molecules will be straightforward to synthesize and will work as pills that are suitable for vaccine-hesitant or immunocompromised individuals, health-care workers and others in risky situations who could take them prophylactically. Furthermore, we expect them to work against vaccine-resistant variants: whereas vaccines target the spike protein on the virus capsule, our compounds target a conserved part of the virus machinery that works inside cells.

We've also had to deal with rejected grant proposals to advance antiviral drugs. Still, as vaccines have showed their dramatic successes, further variants have arrived and funders have begun calling urgently for antivirals and looking at how projects might be accelerated. In April this year, 16 months after the outbreak of SARS-CoV-2 in Wuhan, China, the United Kingdom finally launched a task force focusing on antivirals[2].

Pfizer's March announcement of early clinical trials for its antiviral pill is confirmation that an accelerated approach can work, and that we should persevere. Our molecules also inhibit proteins of the coronavirus that causes severe acute respiratory syndrome (SARS; see 'Missed opportunity'): had drug discovery persevered during the SARS epidemic in 2003, antiviral drugs would have been available when this pandemic hit. Above all, it has become much clearer how an antiviral would be most effective: the treatment must be readily available to everybody, long before they are hospitalized. Accordingly, we have been able to develop a clear plan for how to proceed, and the resources required.

MISSED OPPORTUNITY

Had direct-acting antivirals been developed for SARS, they would have worked for COVID-19.



Potency represented as IC_{50}, the compound concentration (M) that inhibits protein function by 50%.

We are approaching the capital-intensive, highly regulated phases of animal studies, producing kilograms of substance for clinical trials and, beyond that, worldwide manufacture and distribution of billions of pills. Our initial goal of delivering a drug straight from the discovery pipeline, free from patents and available for anyone to manufacture, cannot offer investors any conventional return on investment. Yet COVID-19 is not conventional, and vaccines have elevated the normally arcane question of intellectual property into a major political concern. Perhaps the COVID Moonshot can also shape how open drug discovery reaches patients.

> **"People pulled their weight not for glory or reward, but because there was a job that needed doing."**

The moral

So, what has made our approach work? Presumably, the fact that the mission was clear, even if distant, and the ethos was unambiguous and clearly signposted[3,4]. Initially, a few of us, fuelled by the urgency of the moment, acted on a conviction that our various combined technologies would accelerate drug discovery. We were soon joined by many people who did the hard work because they felt it was the right thing to do.

Also crucial was the existing large ecosystem of expertise and biopharma supply chains, coupled with new capabilities driven by long-term strategic investments in national infrastructure and research institutes. Tools for online collaboration have reached a critical mass, both general ones (such as Zoom or Google Docs) and those specific to drug discovery (in our case, CDD Vault). Serendipitously, for the segments of our project that had the most collaborators — such as submitting ideas for molecules — the requested contributions broke into discrete, doable tasks that easily accommodated each contributor's availability and know-how.

The project self-selected a team of reflexively collaborative people, with no big egos. So far, we have avoided bureaucracy — no one claims to be the head of the COVID Moonshot. We retained momentum with collective trust, combined with sufficiently diverse expertise and perspectives, which allowed us to rapidly reach and implement strategic decisions. Reassuringly, people seemed to leave the collaboration only once their part of the project had been completed.

Perhaps the most surprising asset was that we did not have time to plan much at all — if we had, we'd have been paralysed. It seems you just have to get started and set deadlines for when to move on. Even now, we are astonished at how quickly this infrastructure self-assembled, just by scientists unabashedly asking for help from colleagues, distant connections or vendors. With so clear a goal, so obvious a need and the complete absence of contracts, people across the world stepped up.

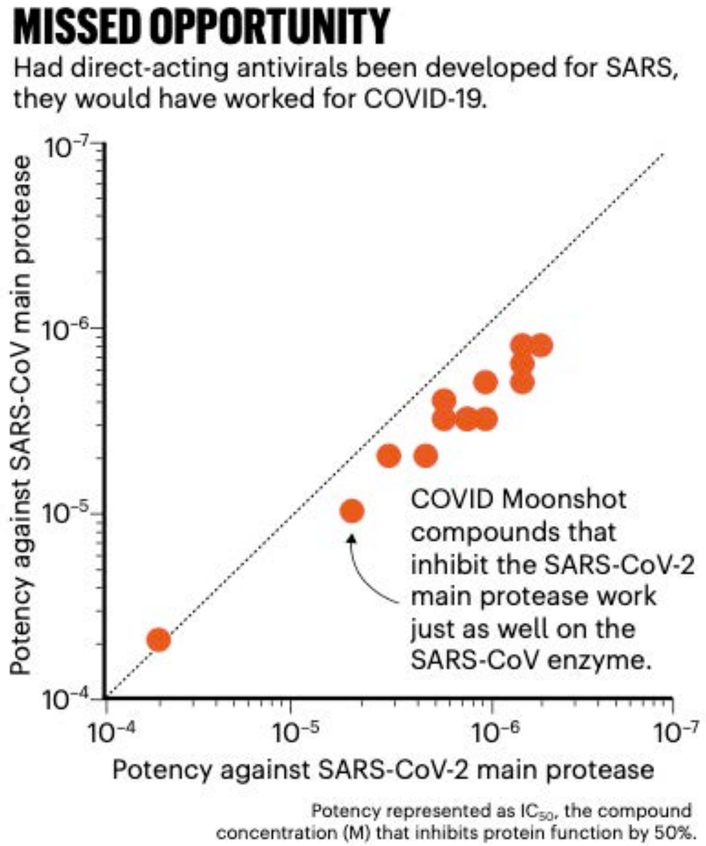
The authors

Frank von Delft is professor of structural chemical biology at the University of Oxford, UK, and principal beamline scientist at Diamond Light Source, Didcot, UK. **John Chodera** is associate member at the Memorial Sloan Kettering Cancer Center, New York, USA. **Ed Griffen** is technical director and co-founder of MedChemica, Ryecraft, UK. **Alpha Lee** is group leader in the Department of Physics, University of Cambridge, UK, and chief scientific officer at PostEra, Boston, Massachusetts, USA. **Nir London** is assistant professor in the Department of Organic Chemistry at the Weizmann Institute of Science, Rehovot, Israel. **Tatiana Matviuk** is principal scientist at Enamine, Kiev, Ukraine. **Ben Perry** is discovery open innovation leader at the Drugs for Neglected Diseases initiative, Geneva, Switzerland. **Matt Robinson** is chief technology officer of PostEra, Boston, Massachusetts, USA. **Mark Calmiano** is a computational chemist at UCB Biopharma, Brussels, Belgium. **Annette von Delft** is a translational scientist at the University of Oxford, UK
e-mail: frank.vondelft@cmd.ox.ac.uk

1. Douangamath, A. *et al. Nature Commun.* **11**, 5047 (2020).
2. Mahase, E. *Br. Med. J.* **373**, n1077 (2021).
3. The COVID Moonshot Consortium *et al.* Preprint at bioRxiv https://doi.org/10.1101/2020.10.29.339317 (2020).
4. Chodera, J., Lee, A. A., London, N. & von Delft, F. *Nature Chem.* **12**, 581 (2020).

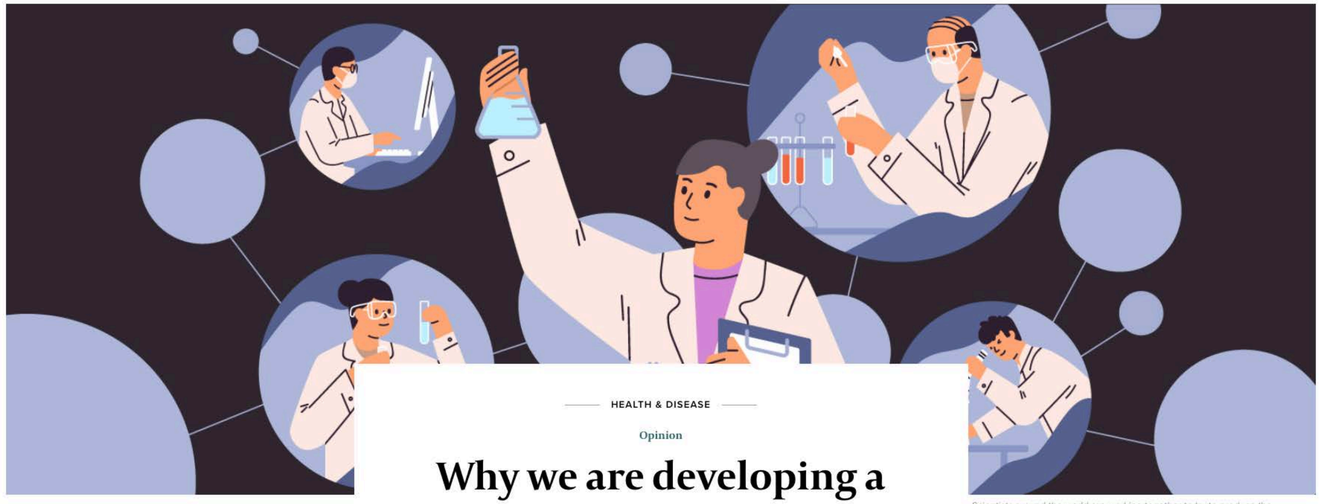
J.C., E.G., A.L., N.L. & M.R. declare competing interests.

THE ONLY REASON WE DIDN'T HAVE ANTIVIRALS FOR SARS-COV-2 WAS DUE TO MARKET FAILURE



Our compounds are **equipotent** against SARS-CoV-1.

There's **no reason we couldn't have done this in 2004** after the first SARS pandemic.



Scientists around the world are working together to try to produce the world's first patent-free antiviral therapy aimed at Covid-19. During a deadly pandemic, this is how drug development should proceed, the researchers argue.

CREDIT: GOODSTUDIO / SHUTTERSTOCK

HEALTH & DISEASE

Opinion

Why we are developing a patent-free Covid antiviral therapy

OPINION: During global health crises such as pandemics, drug discovery should be publicly funded and open, with no research secrets locked away

By Alpha Lee and John Chodera | By Frank von Delft | 09.27.2021

The rapid development of vaccines against Covid-19 is a scientific triumph. But the recipes for making these vaccines are the exclusive intellectual property of pharmaceutical companies, which means countries cannot manufacture an approved vaccine themselves, thus limiting distribution worldwide. For this and other reasons — such as problems with medical infrastructure and a lack of trained workers to administer the vaccine — most poor countries won't be widely vaccinated until at least 2024.

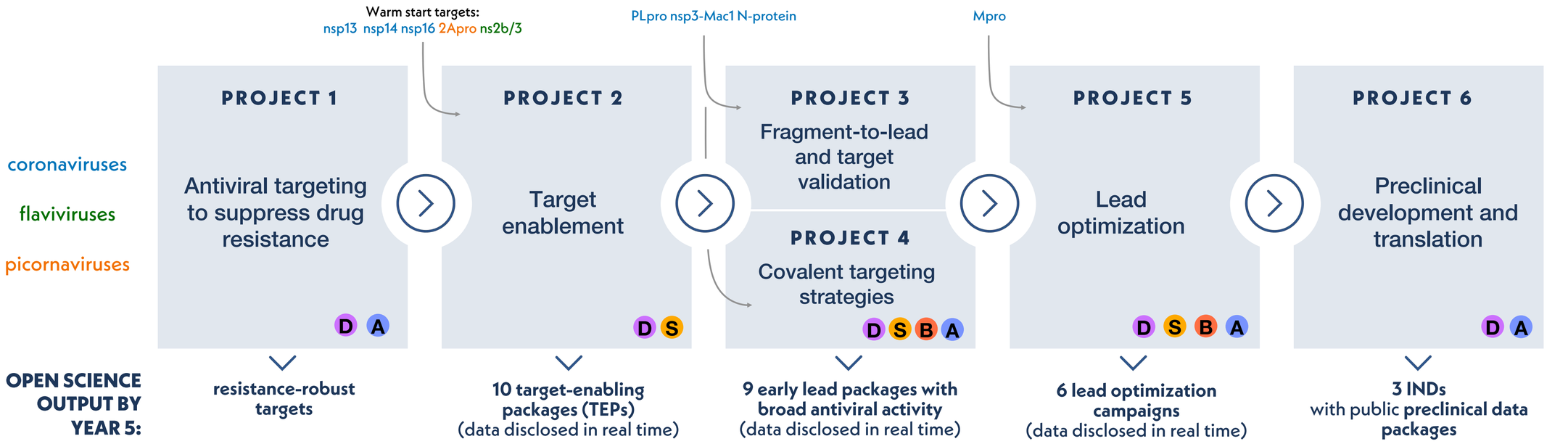
Much of the process of discovering a new drug or vaccine — as researchers hunt for new candidates, and companies develop those into safe, effective products — is typically conducted behind closed doors. Even once a product is approved, patent protections prevent other manufacturers from making and selling it. Eventually, patents expire; but some aspects of the lifesaving science behind the development of those patented products — such as which candidates don't work — often remain forever locked up in corporate silos, hindering research that may prevent future pandemics.

More from *Reset* — An ongoing series exploring how the world is navigating the coronavirus pandemic, its consequences and the way forward.

OPEN SCIENCE CAN TRANSFORM ANTIVIRAL DISCOVERY

- **Nobody knows how IP should work for diseases that don’t yet exist.**
  Openly disclosing all stages of discovery is only way to ensure investment in antiviral research is available when we need it.
- **We need an open science nexus for antiviral discovery.**
  We must share data, coordinate resources, and enable seamless collaboration to avoid wasted and duplicated effort.
- **We can evaluate and apply new technologies to accelerate discovery.**
  We can leverage the latest advances in the open source drug discovery ecosystemto increase success rates, benchmark the utility of methods, and disseminate learnings.
- **We can exercise this platform to produce clinic-ready oral drug candidates** to stop new outbreaks before they become pandemics.
- **We must develop therapeutics with global access in mind from day one.**
  No one is safe unless we’re all safe.

AI-DRIVEN STRUCTURE-ENABLED ANTIVIRAL PLATFORM (ASAP)



P1: Karla Kirkegaard (Stanford)
Matt Bogyo (Stanford)
Jesse Bloom (Fred Hutch)

P2: Frank von Delft (Diamond Light Source)
Martin Walsh (Diamond Light Source)
Oxford CMD SRF [service facility]

P3: Alpha Lee (PostEra)
John Chodera (MSKCC)
Frank von Delft (Diamond)
Ed Griffen (Medchemica)
Nir London (Weizmann)
Karla Kirkegaard (Stanford)
Martin Walsh (Diamond)

P4: Nir London (Weizmann)
Matt Bogyo (Stanford)

P5: Ed Griffen (Medchemica)
Ben Perry (DNDi)
Alpha Lee (PostEra)
John Chodera (MSKCC)

P6: Ben Perry (DNDi)
Laurent Fraisse (DNDi)
Annette von Delft (Medchemica)

SUPPORTING LETTERS

ADMINISTRATIVE CORE

John Chodera (MSKCC)
Ben Perry (DNDi)
Alpha Lee (PostEra)
Administrative Director
Project Coordinator

PIs

D DATA INFRASTRUCTURE CORE
Alpha Lee (PostEra)
Matthew Robinson (PostEra)
Frank von Delft (Diamond)
John Chodera (MSKCC)

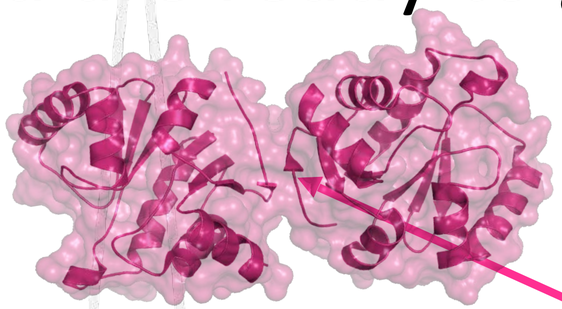
S STRUCTURAL BIOLOGY CORE
Frank von Delft (Diamond Light Source)
Daren Fearon (Diamond Light Source)
Martin Walsh (Diamond Light Source)

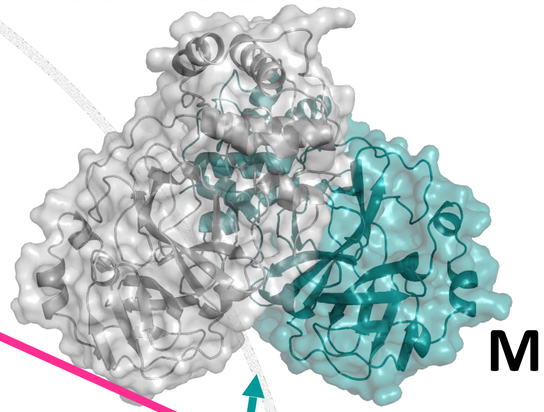
B BIOCHEMICAL ASSAY CORE
Nir London (Weizmann)
Haim Barr (Weizmann)

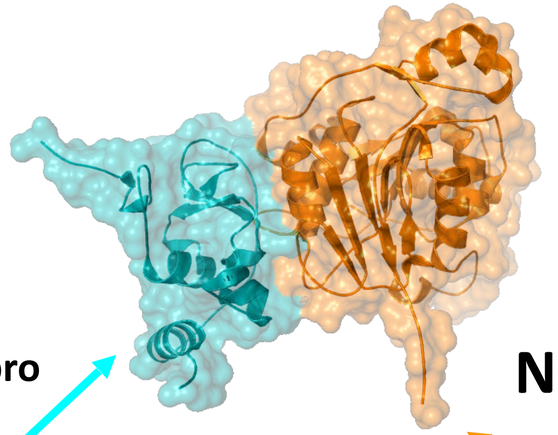
A ANTIVIRAL EFFICACY AND RESISTANCE CORE
Kris White (Mount Sinai)
Adolfo García-Sastre (Mount Sinai)
Randy Albrecht (Mount Sinai)
Johan Neyts (Leuven) [service facility]

https://www.choderalab.org/news/2021/10/26/asap-avidd-proposal

We already have fragment screens for multiple targets, and are ready to get to work







PLpro cleavage sites

Mpro cleavage sites

PLpro

Mpro

Nsp16/10

1 2 3 4 5 6 8 10 11 12 13 14 15 16

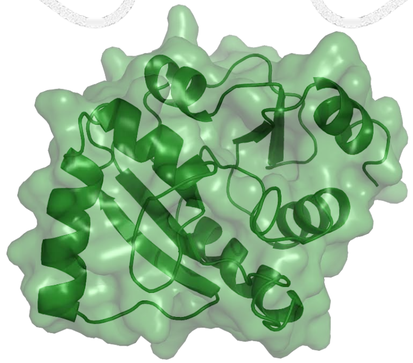
NSP3 TM TM Mpro TM NiRAN RdRp ZBD HEL1 ExoN N7-MT EndoU 2’-O-MT RBD HR2 S

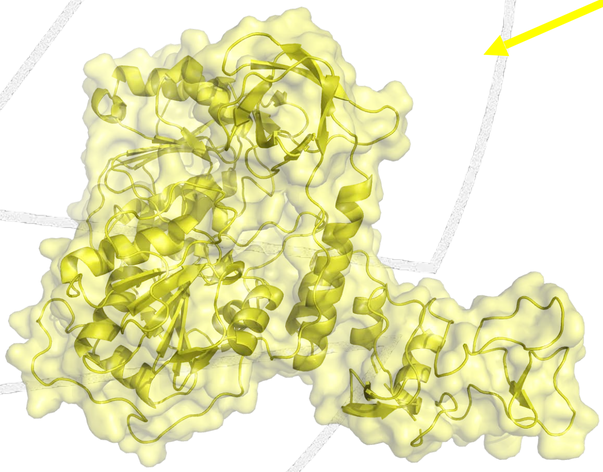
Mac1

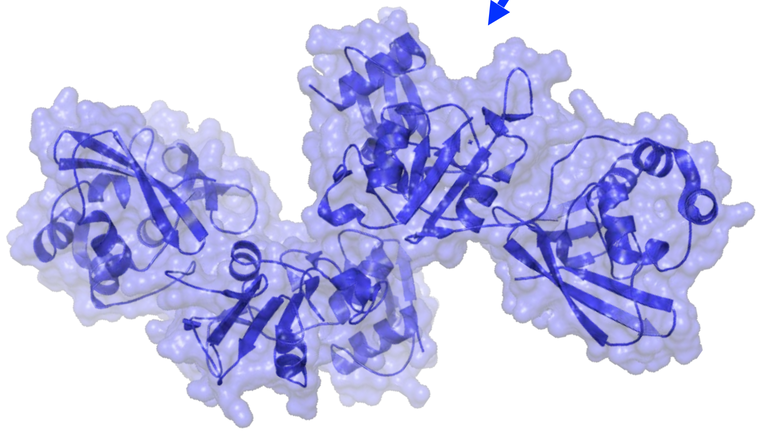
Nsp13

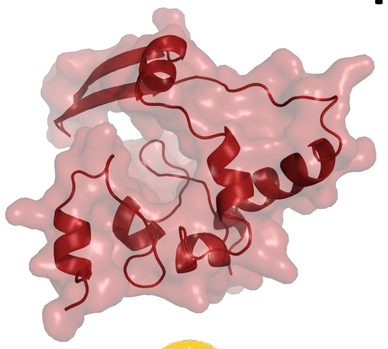
Nsp15

Nprot







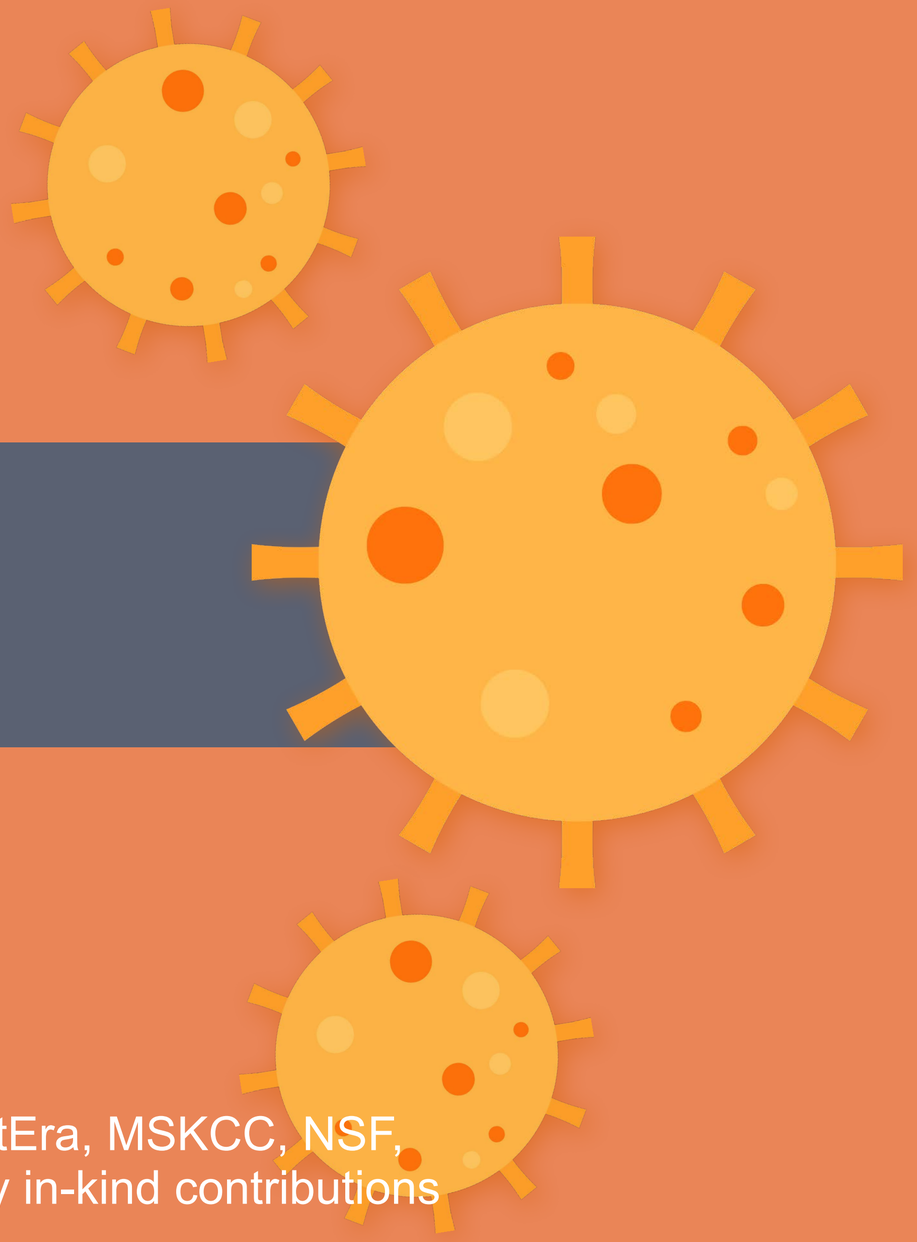


diamond

CHODERA LAB

- All funding: http://choderalab.org/funding

THANK YOU!



preprint: https://doi.org/10.1101/2020.10.29.339317
contributors: https://tinyurl.com/covid-moonshot-authors
twitter: https://twitter.com/covid_moonshot
slides: http://choderalab.org/news
Moonshot data: http://postera.ai/covid
Folding@home data: http://covid.molssi.org
funding: Diamond, Oxford COVID Response Fund, Weizmann, PostEra, MSKCC, NSF, DNDi, LifeArc, Wellcome Trust TEP Strategic Award, and so many in-kind contributions