REDESIGNING DRUG DESIGN

John D. Chodera
MSKCC Computational and Systems Biology Program
http://www.choderalab.org

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

24 Oct 2022 - SKI Talk - NYC

CANCER IS A LEADING CAUSE OF DEATH

In 2021, in the United States alone:

1,900,000 new cancer cases will be diagnosed
600,000 are expected to die of cancer
More than **16,900,000** are living with a history of cancer

Many cancers have very **poor five-year survival rates** and are in need of significantly better therapies

percent surviving after five years

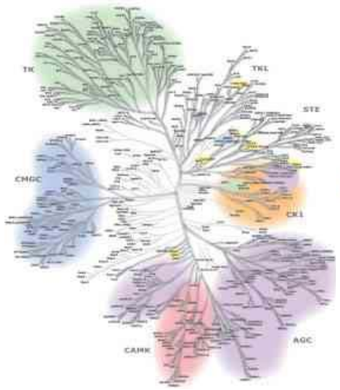
	1975-77	1987-89	2010-16
All sites	49	55	67
Brain & other nervous system	23	29	33
Breast (female)	75	84	90
Colon & rectum	50	60	65
Colon	51	60	63
Rectum	48	58	67
Esophagus	5	9	20
Hodgkin lymphoma	72	79	87
Kidney & renal pelvis	50	57	75
Larynx	66	66	61
Leukemia	34	43	64
Liver & intrahepatic bile duct	3	5	20
Lung & bronchus	12	13	21
Melanoma of the skin	82	88	93
Myeloma	25	27	54
Non-Hodgkin lymphoma	47	51	73
Oral cavity & pharynx	53	54	66
Ovary	36	38	49
Pancreas	3	4	10
Prostate	68	83	98
Stomach	15	20	32
Testis	83	95	95
Thyroid	92	94	98
Urinary bladder	72	79	77
Uterine cervix	69	70	66
Uterine corpus	87	82	81

Cancer Facts & Figures 2021 - The American Cancer Society

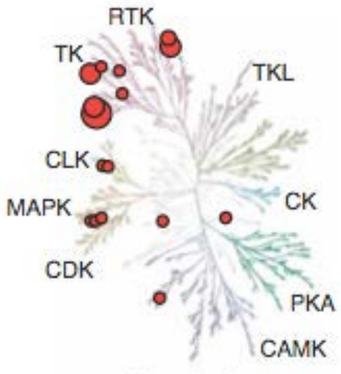
SMALL MOLECULE KINASE INHIBITORS CAN HAVE SIGNIFICANT THERAPEUTIC BENEFITS ON CANCERS INVOLVING KINASE DYSREGULATION

molecular target 🎯: Bcr-Abl fusion constitutively activates ABL in CML patients, resulting in unchecked white blood cell proliferation

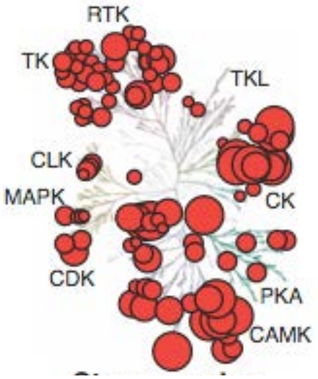
imatinib bound to Abl
[PDB:1IEP]



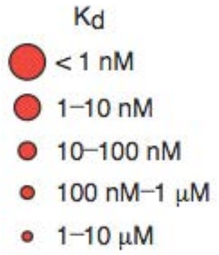
human kinome
[~500 kinases]



imatinib
[blockbuster drug]



staurosporine
[toxic natural product]



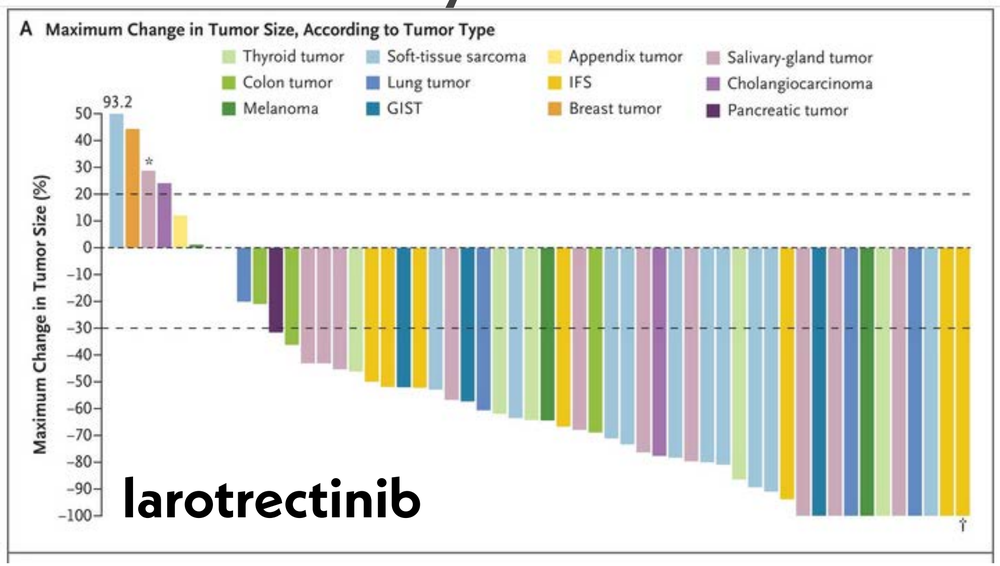
[1]

71 small molecule kinase inhibitors have been approved by the FDA [2]
Global annual kinase inhibitor market $40B in 2020; expect $65B in 2027 [3]

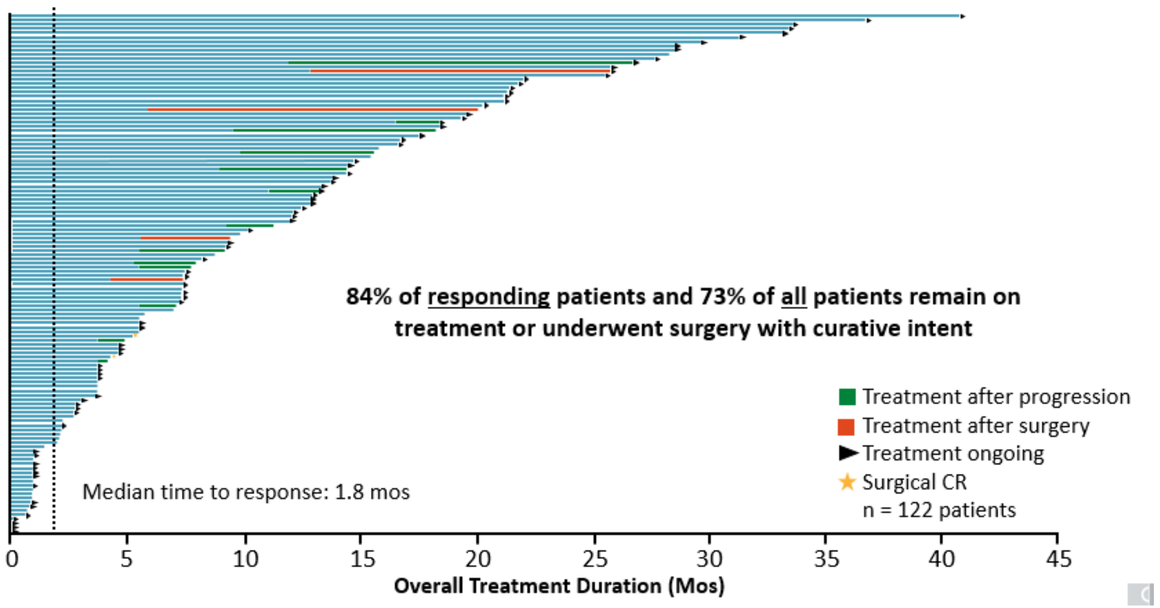
[1] Nature Biotech 23:329, 2005
[2] as of 29 Mar 2022: http://www.brimr.org/PKI/PKIs.htm
[3] Global Kinase Inhibitor Markets 2019-2020 and 2027 [URL]

KINASE INHIBITORS CAN HAVE SIGNIFICANT THERAPEUTIC BENEFITS ON CANCERS INVOLVING KINASE DYSREGULATION

anti-tumor activity



treatment duration

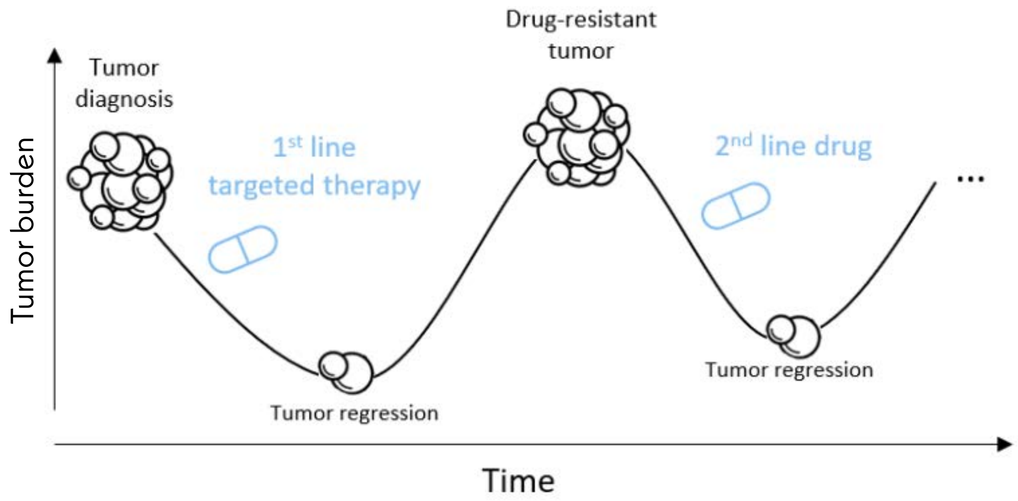


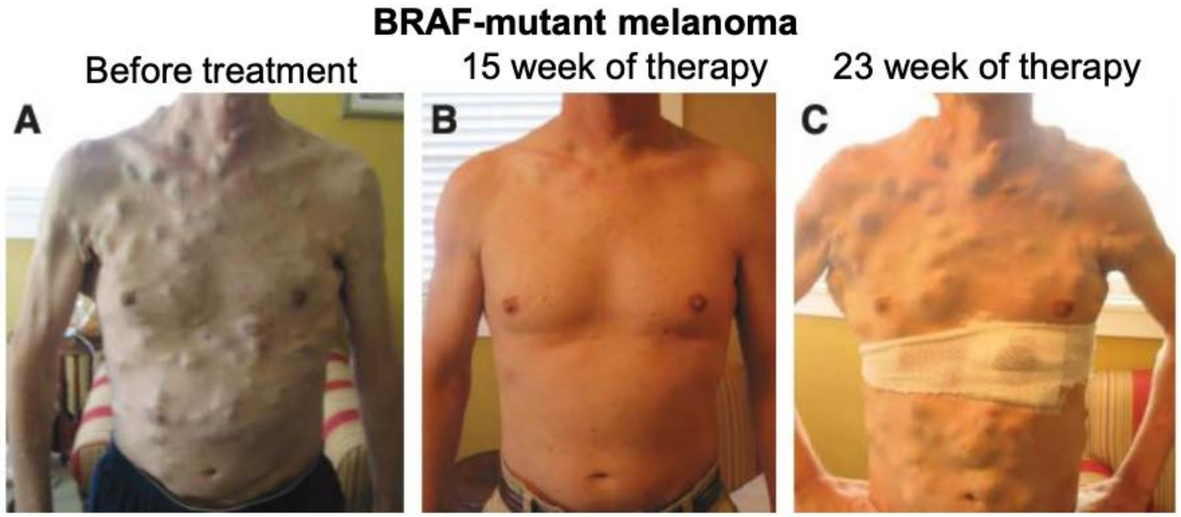
Analysis of three open-label trials (phase I, adults; phase I/II, children; phase II, adolescents/adults) assessing larotrectinib for treating advanced solid tumors with *NTRK* gene fusion (N = 55). Trial IDs NCT02122913, NCT02637687, and NCT02576431.

How can we develop more therapeutics that yield durable response?

Drilon et al., NEJM 378:731, 2018 [DOI]

DRUG RESISTANCE IS A MAJOR CHALLENGE FOR TARGETED KINASE INHIBITOR THERAPY



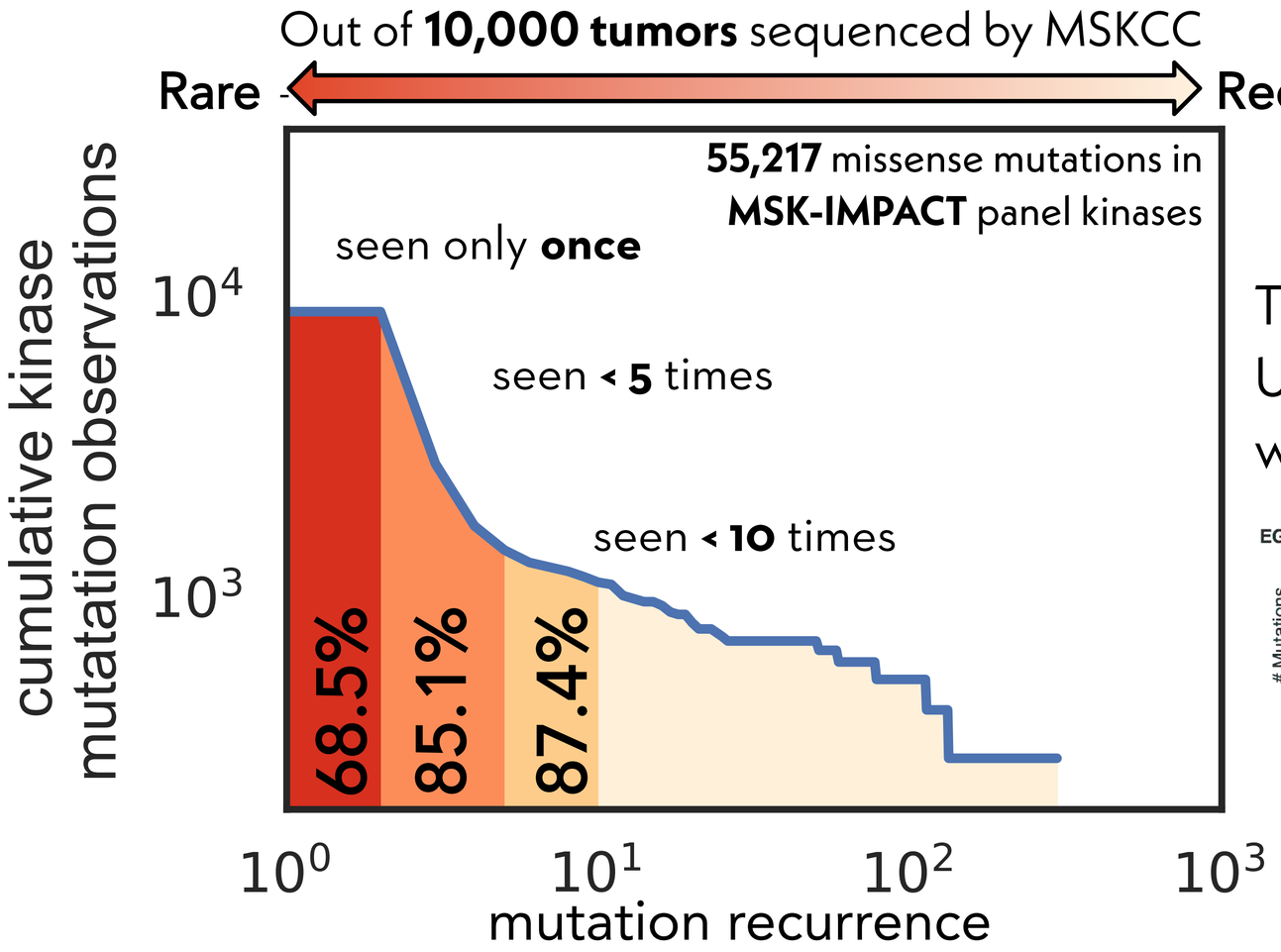


Mutations in the target of therapy can reduce the durability of response

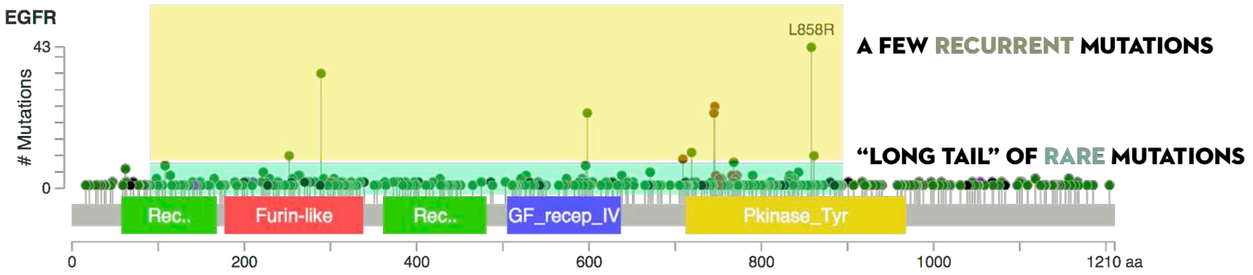
Need an armamentarium of second-line therapeutics or **broad activity against mutants**

Wagle et al. J Clin Oncol 2011 [DOI]

THE LONG TAIL OF CANCER MUTATIONS FRUSTRATES THE PREDICTION OF RESISTANCE



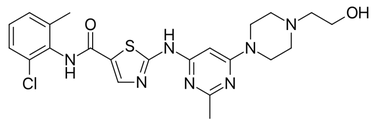
The **vast majority of mutations** are rare:
Unlikely to be clinical or biochemical evidence of whether they confer **drug resistance** or **susceptibility**



Zehir et al. Nature Medicine 23:703, 2017 [DOI]
Hauser et al. Communications Biology 1:70, 2018 [DOI]

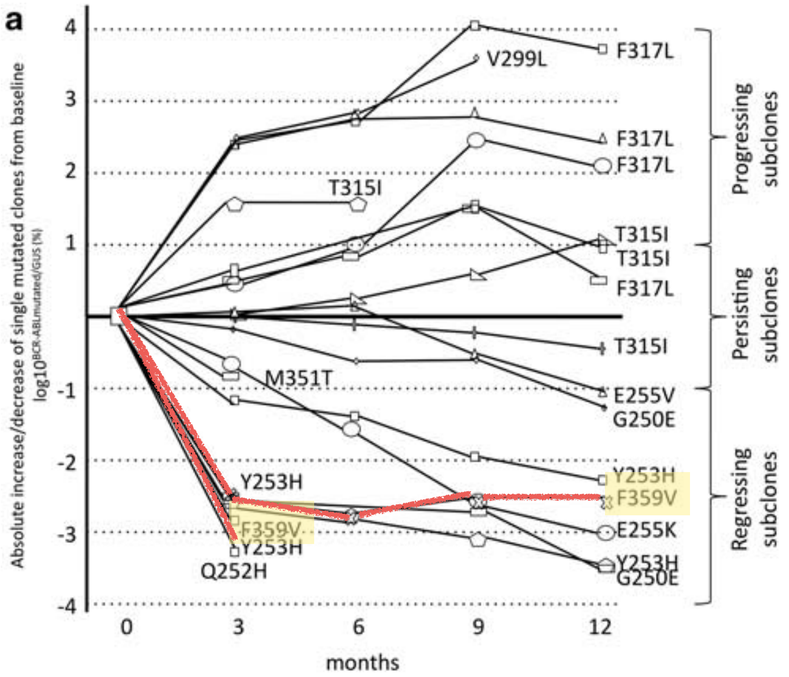
DIFFERENT DRUGS APPEAR TO EXERT DISTINCT SELECTIVE EVOLUTIONARY PRESSURES

CML patients failing out of **imatinib** therapy often different kinds of resistance depending on the choice of second-line therapy:



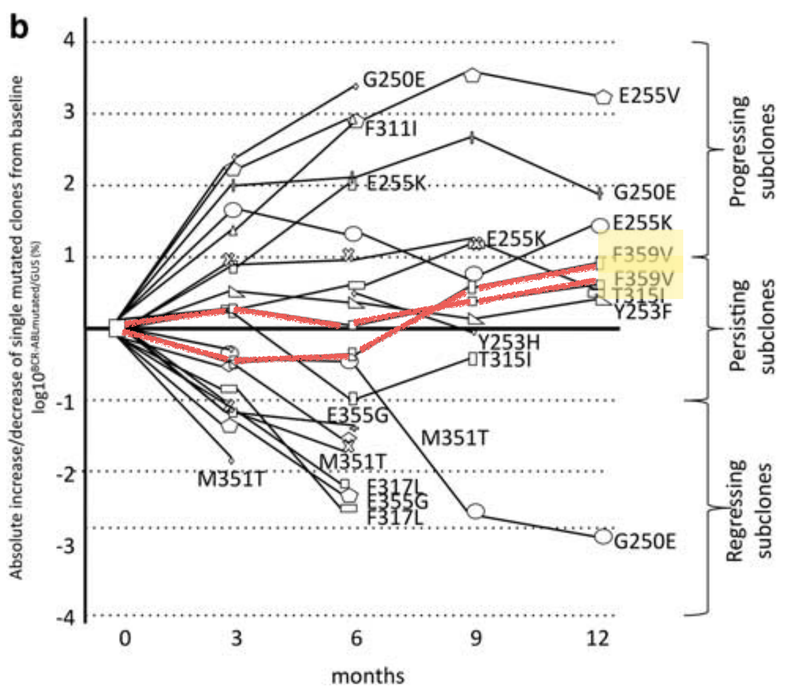
dasatinib

dasatinib treatment

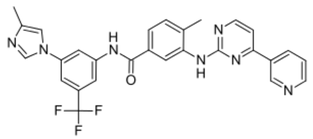


F359V DEPLETED

nilotinib treatment



F359V ENRICHED



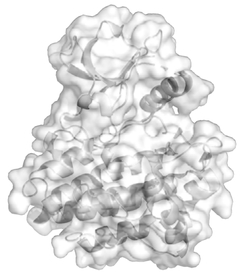
nilotinib

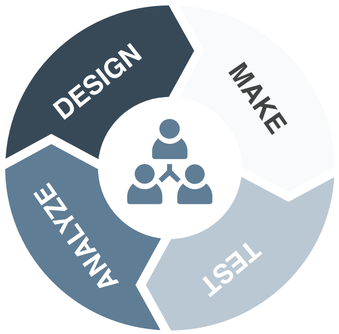
Data suggests **we should be able to predict** how different mutants confer resistance/susceptibility

Gruber et al. Leukemia 26:172, 2012.

DRUG DISCOVERY AND DEVELOPMENT IS COSTLY, TIME-CONSUMING, AND INEFFICIENT

	Target Candidate Profile (TCP)						
value*		$8M	$24M	$58M	$321M		$1.66B
stage cost*		$13.5M	$3M	$3M	$8M	$20M	$3M
% survival		48%	38%	25%	12%	7%	6%
duration*		4.5 y	2.3 y	2.1 y	3.1 y	2.5 y	9 mo
	Target ID	Discovery	Preclinical	Phase 1 safety	Phase 2 efficacy	Phase 3 effectiveness	NDA new drug approval





Human-driven design iterations

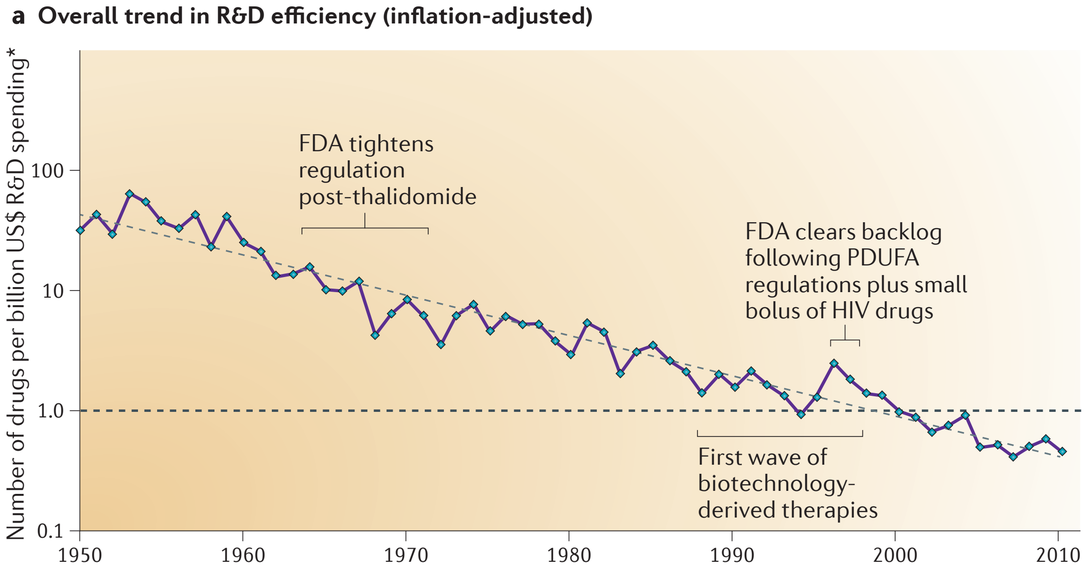
* denotes mean

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

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

DRUG DISCOVERY USUALLY ENDS IN FAILURE

Drugs are getting more expensive to develop due to **low success rates (~2%)**

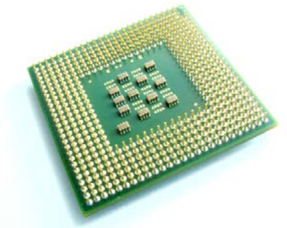


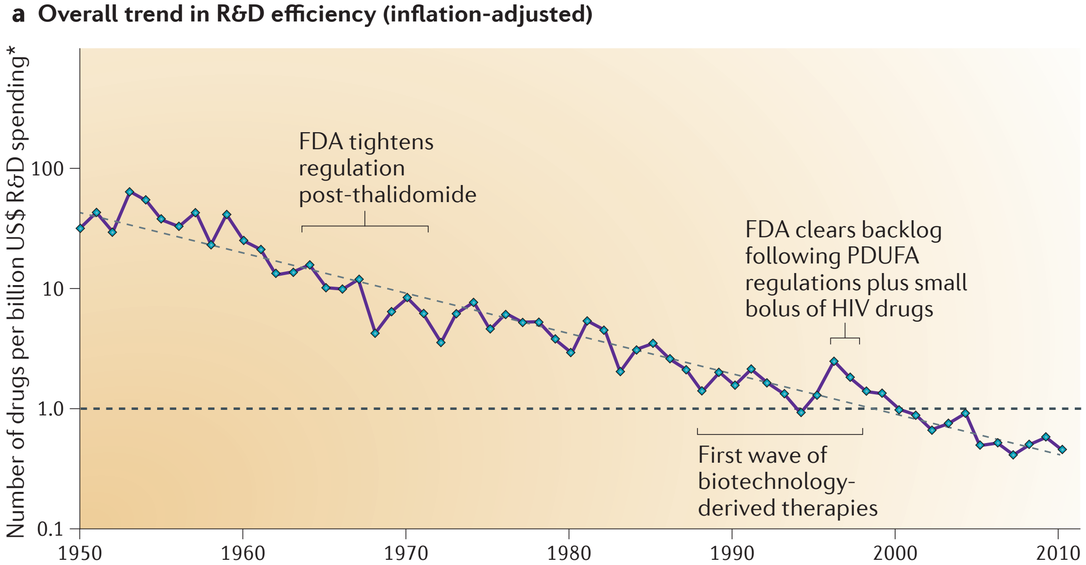
now: $2.6/drug*

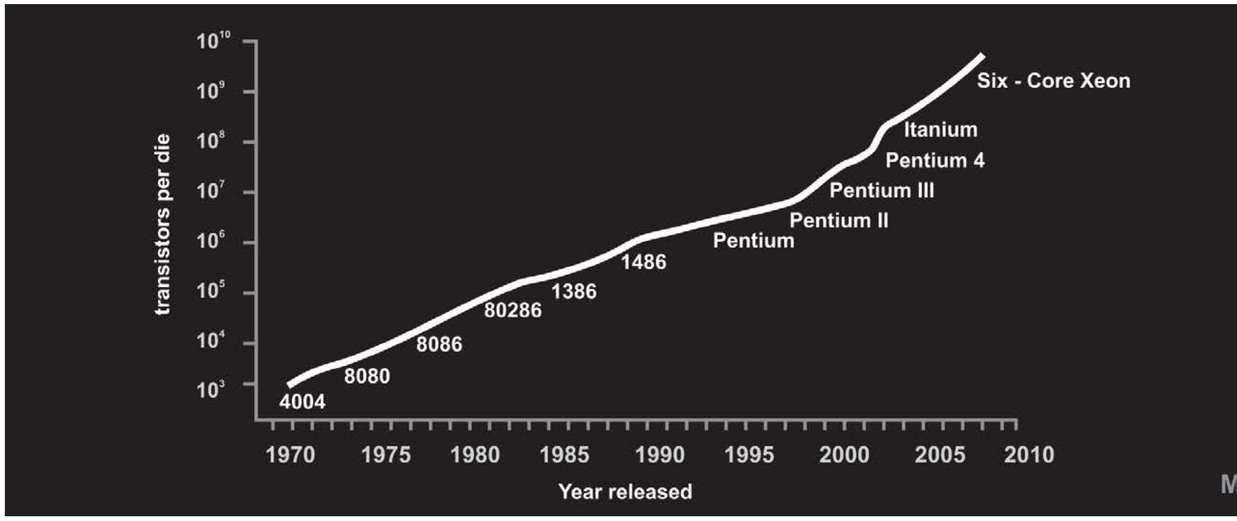
EROOM'S LAW

https://www.nature.com/articles/nrd3681 * https://www.nature.com/articles/nrd4507

DRUG DISCOVERY USUALLY ENDS IN FAILURE





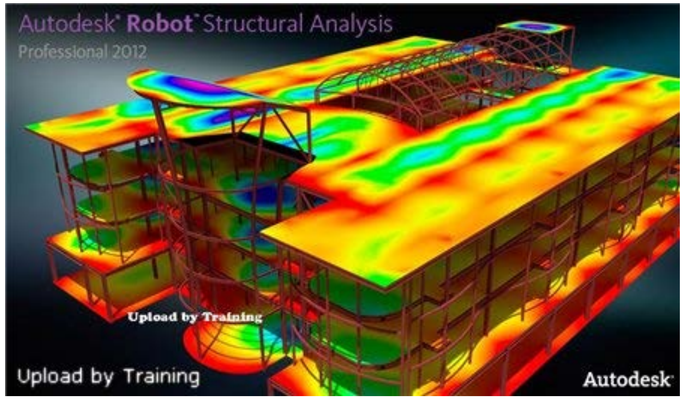


EROOM'S LAW

MOORE'S LAW

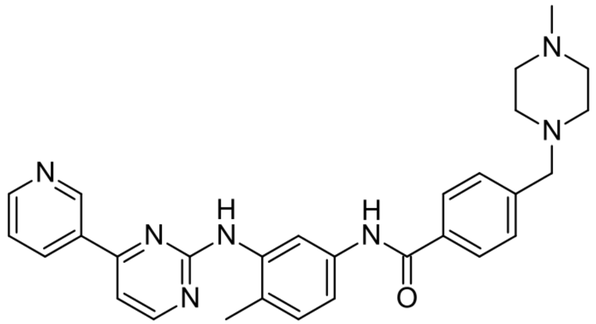
https://www.nature.com/articles/nrd3681 * https://www.nature.com/articles/nrd4507

DRUG DISCOVERY IS A COMPLEX MULTI-OBJECTIVE DESIGN PROBLEM



10^3 - 10^6 parts

WHY NOT SMALL MOLECULE DRUGS?

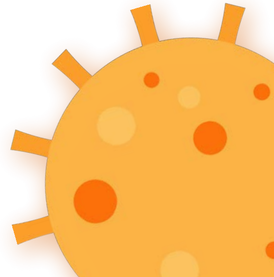


< 10^2 atoms

DRUG DISCOVERY IS A COMPLEX MULTI-OBJECTIVE DESIGN PROBLEM

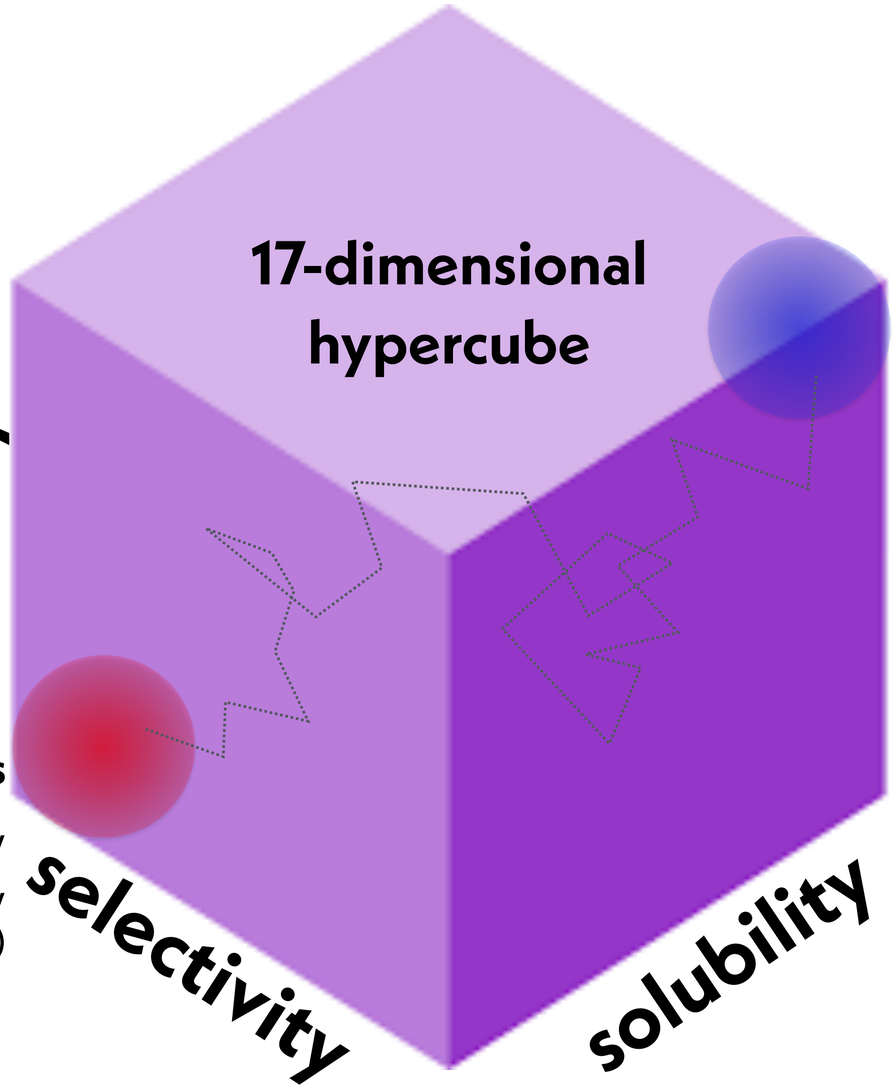
Target Candidate Profile (TCP) for oral SARS-CoV-2 main viral protease (Mpro) inhibitor

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



https://doi.org/10.1101/2020.10.29.339317
https://covid.postera.ai/covid

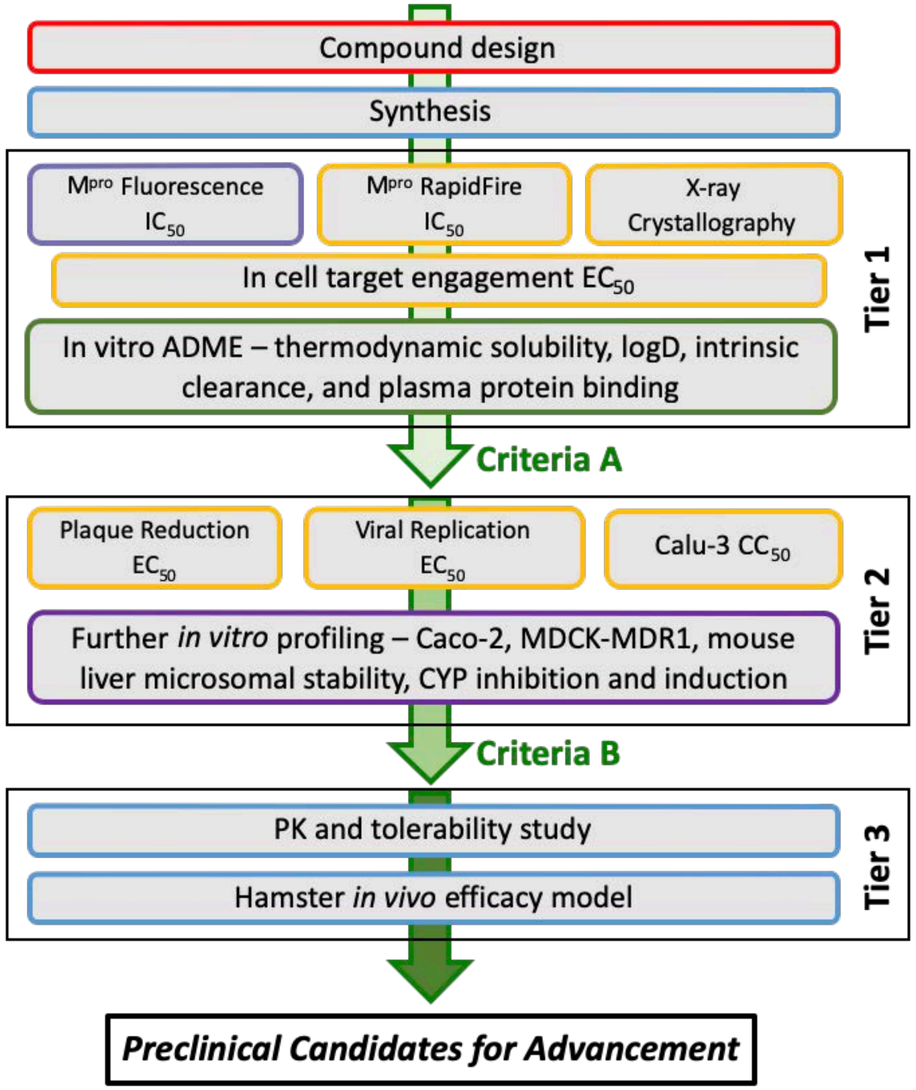
DRUG DISCOVERY IS A COMPLEX MULTI-OBJECTIVE DESIGN PROBLEM



target goals for druglike molecule

initial hits
(from high-throughput screening, DNA-encoded libraries, virtual screening, etc.)

TO MEET THESE OBJECTIVES, TYPICAL DISCOVERY PROJECTS GROUP ASSAYS INTO SEQUENTIAL TIERS



assay purpose

$ — Does it inhibit the target? How does it bind?

Does it bind the target in cells?

Does it have a chance of working in humans?

$$ — Does it actually work in cells?

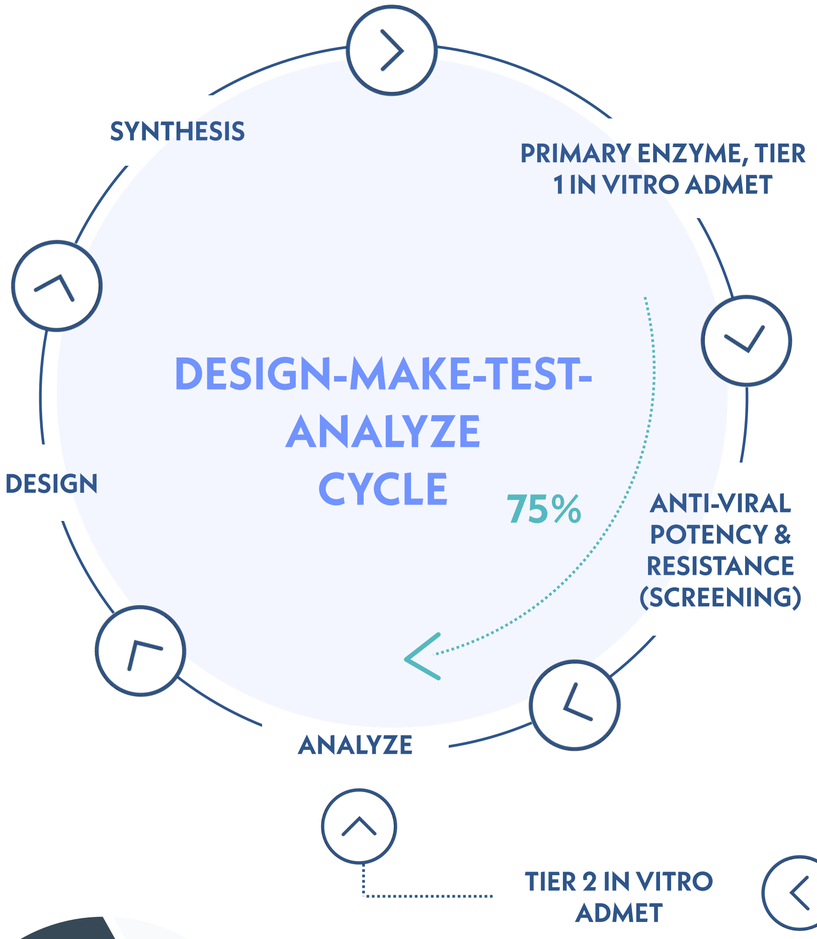
Could it cause bad side effects?

$$$ — Can oral dosing deliver sufficient drug?

Does it actually work against the disease?

$$$$

DRUG DISCOVERY PROGRESSES THROUGH MANY DESIGN-MAKE-TEST-ANALYZE CYCLES



20% (~ 500 compounds)

PRIMARY IN VIVO ADMET

6% of total (120 compounds)

TIER 2 IN VITRO ADMET

Anti-viral Animal (profiling)

Anti-viral Cellular (profiling)

45 compounds

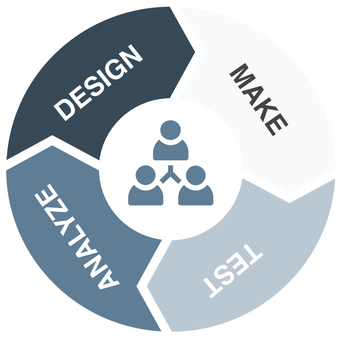
TIER 3 IN VITRO ADMET

10 compounds

SECONDARY IN VIVO ADMET

In vivo tool compounds

DOSE TO HUMAN PREDICTION



THE CHODERA LAB AIMS TO DEVELOP PREDICTIVE MODELS WITH REAL IMPACT ON HUMAN HEALTH

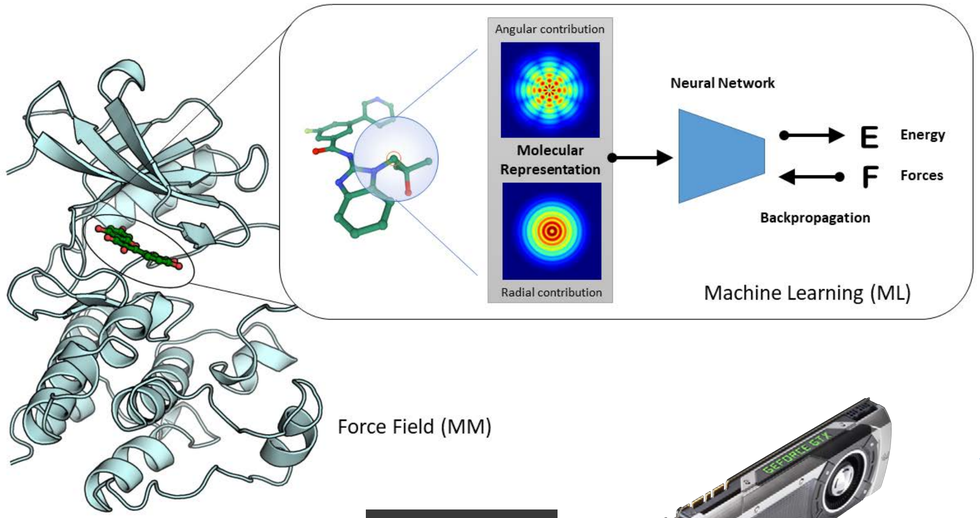
Develop predictive models useful for guiding drug discovery

Make predictions that enable statistically sound decisionmaking

Impact both drug discovery and clinical applications

CHODERA LAB

COMPUTATION



Force Field (MM)

amazon web services EC2

FOLDING @HOME

informs model improvement

EXPERIMENT

CHODERA LAB, Z17

CLOUD LABORATORIES

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

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

WE COLLABORATE BROADLY TO IMPACT DRUG DISCOVERY

open forcefield consortium

OpenMM

Open Molecular Software Foundation

Open Free Energy

Molecular Sciences Software Institute

Folding@home

INTERLINE therapeutics

BAYER

Ventus THERAPEUTICS

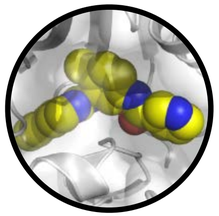
SILICON Therapeutics

OpenEye SCIENTIFIC

open source software development initiatives

NIH National Center for Advancing Translational Sciences

industry collaborations



choderalab
(algorithms and open source software)

data generators, community challenges, and resources

ASAP ACCESSIBLE ANTIVIRALS TO PREVENT PANDEMICS

The SAMPL challenges
Challenge information and announcements

Drug Design Data Resource

academia

VIR

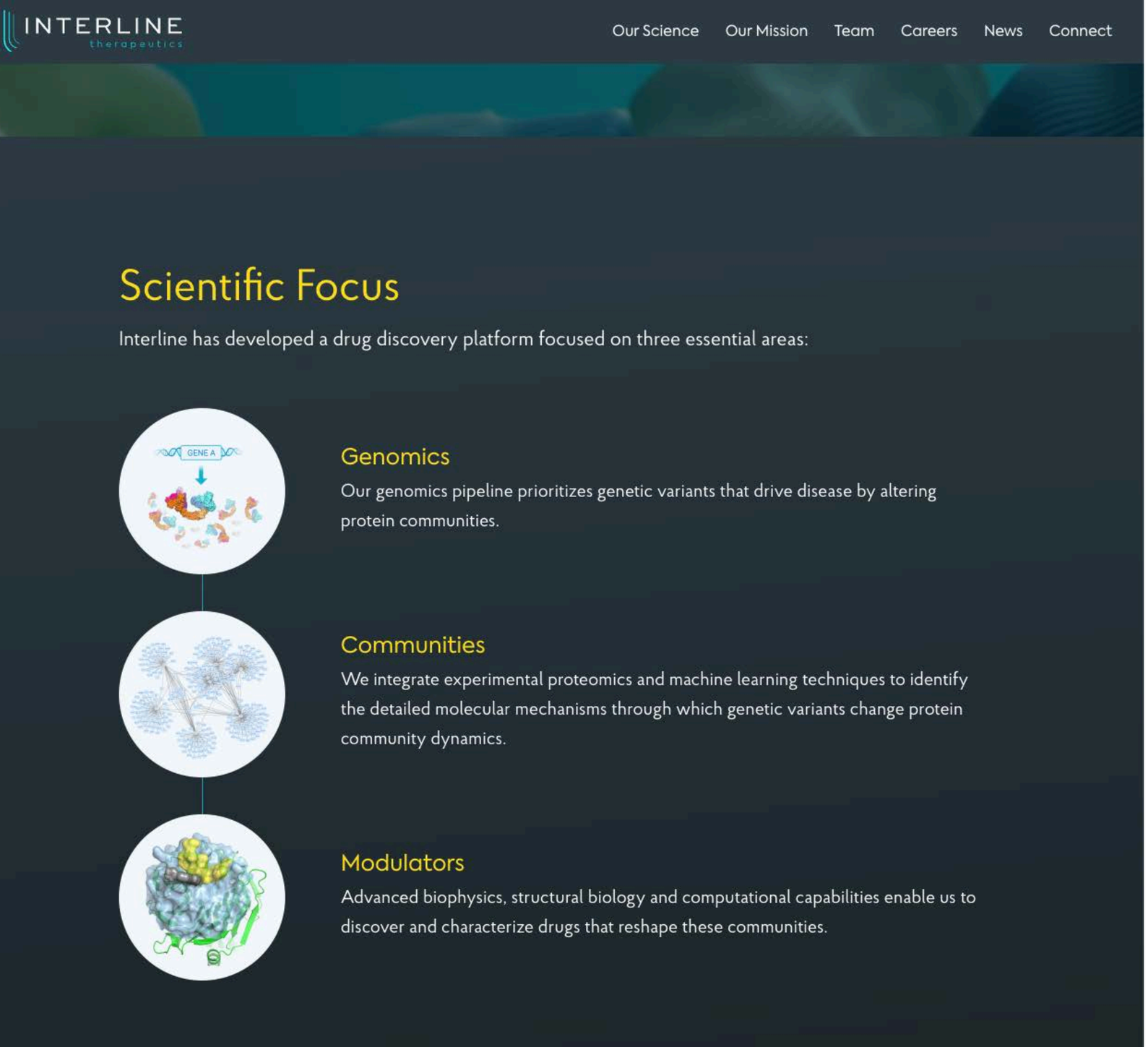
1884

XCHEM Diamond Fragment Screening

Diamond Light Source / XChem

IP-generating collaborations

open science / open source software



We also found **new companies** to deploy our technologies to maximize **impact**

Interline Therapeutics

Licensed technologies from MSK
Launched in May 2021
With **$92M Series A**

aims to use our technologies to **design selective modulators of protein communities**

JDC is a Founding SAB member
MSKCC has equity in Interline

DRUG DISCOVERY IS NOT A BIG DATA PROBLEM

DALL-E 2 was trained on a dataset of **650 million** images

Q: Who is the president during WWII?
A: Franklin D. Roosevelt was the president during WWII.

GPT-3 was trained on a corpus of **22.5 billion pages of text** (45 TB)

vs

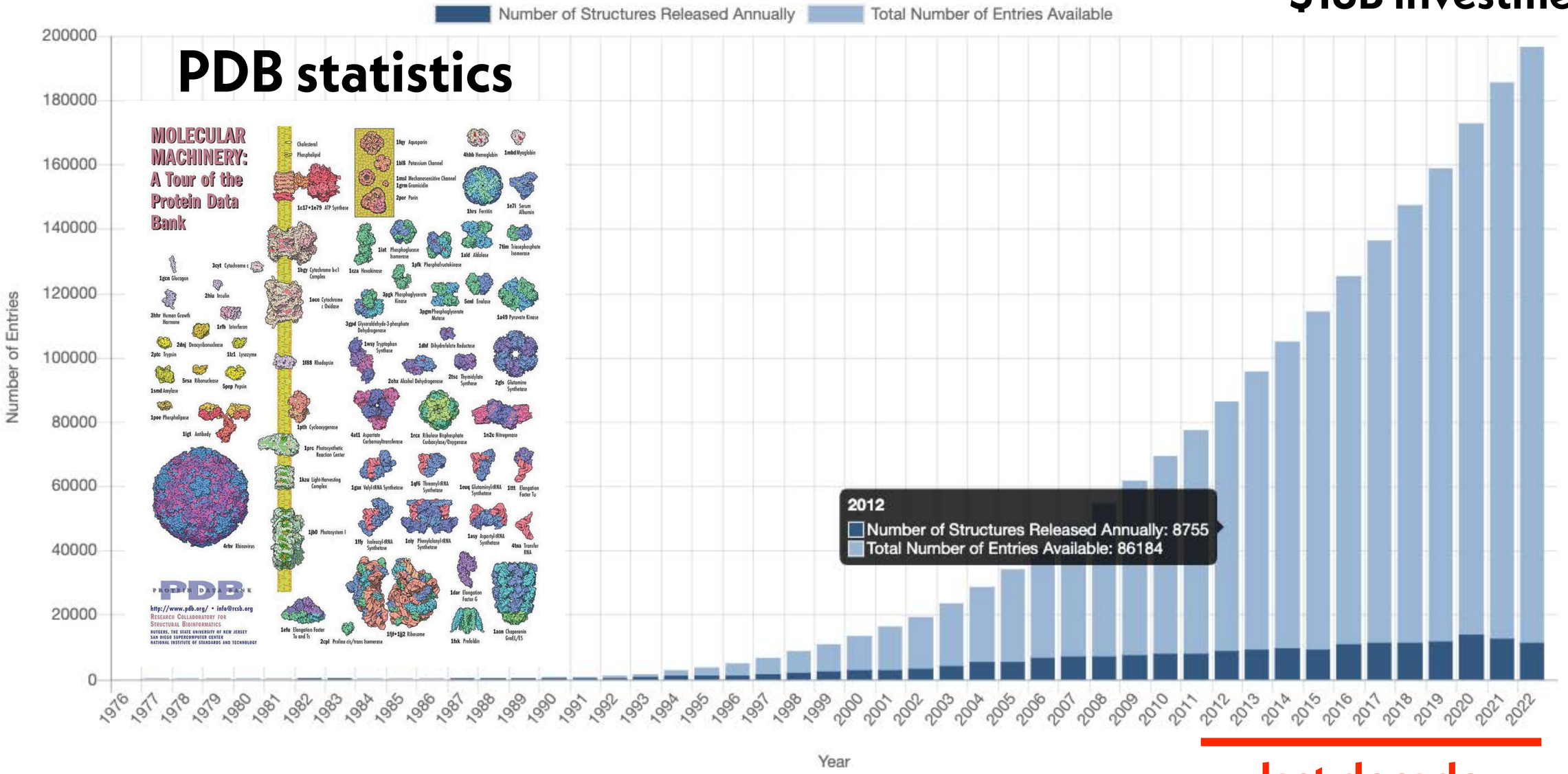
Typical drug discovery programs make and test **~2000 compounds** and largest opportunity for impact is **early on in the program**

We need methods that can **extrapolate from little or no data**

STRUCTURAL DATA IS NOW AN ABUNDANT RESOURCE FOR DRUG DISCOVERY

$18B investment

PDB statistics



http://www.rcsb.org/stats

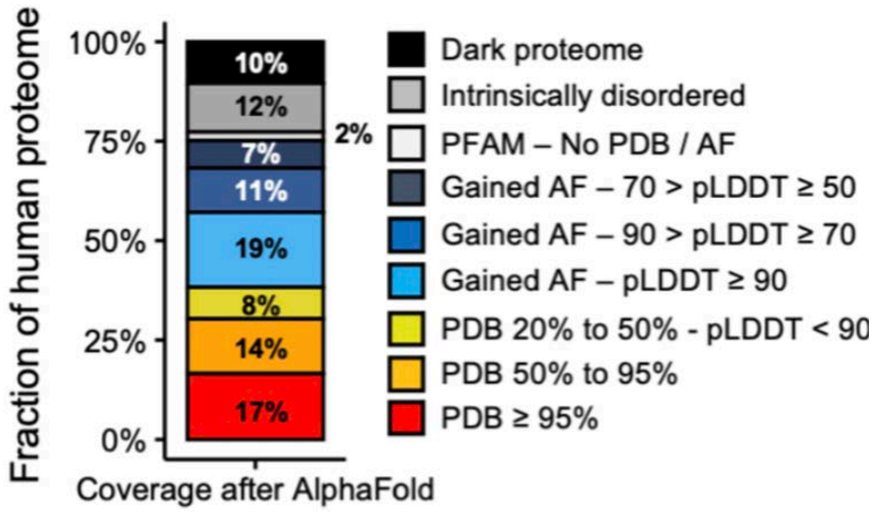
ALPHAFOLD-LIKE METHODS HAVE DRAMATICALLY EXTENDED THE REACH OF STRUCTURAL MODELS

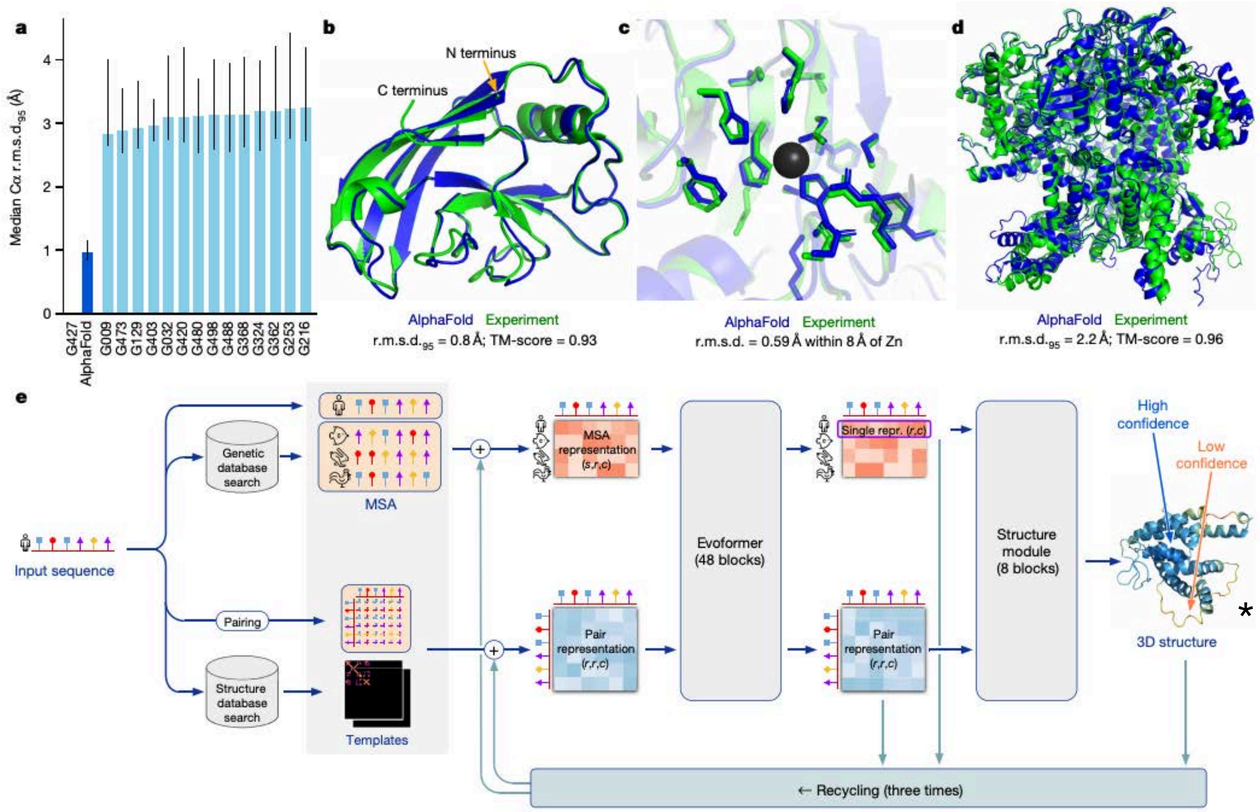
Article | Open Access | Published: 15 July 2021

Highly accurate protein structure prediction with AlphaFold

John Jumper, Richard Evans, [...]Demis Hassabis

Nature (2021) | Cite this article

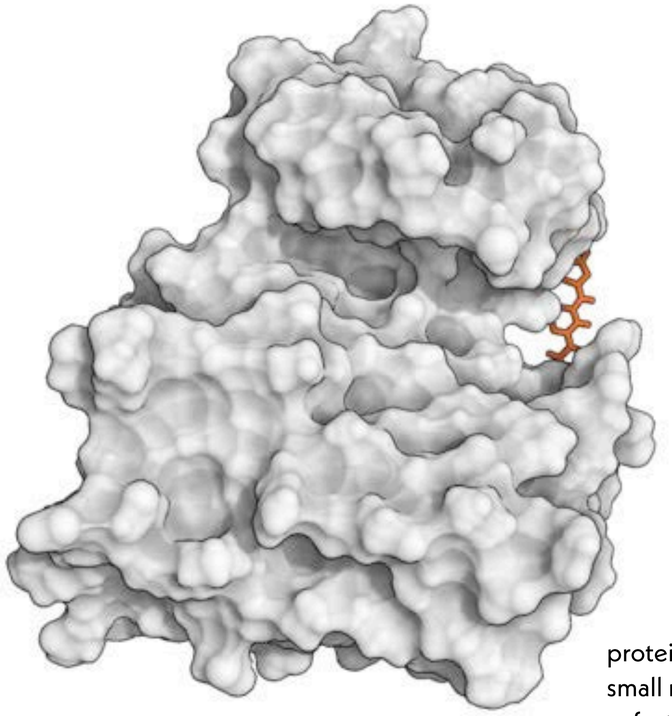




*Used our **OpenMM** molecular modeling framework!
http://openmm.org/ · over 1.2 million downloads

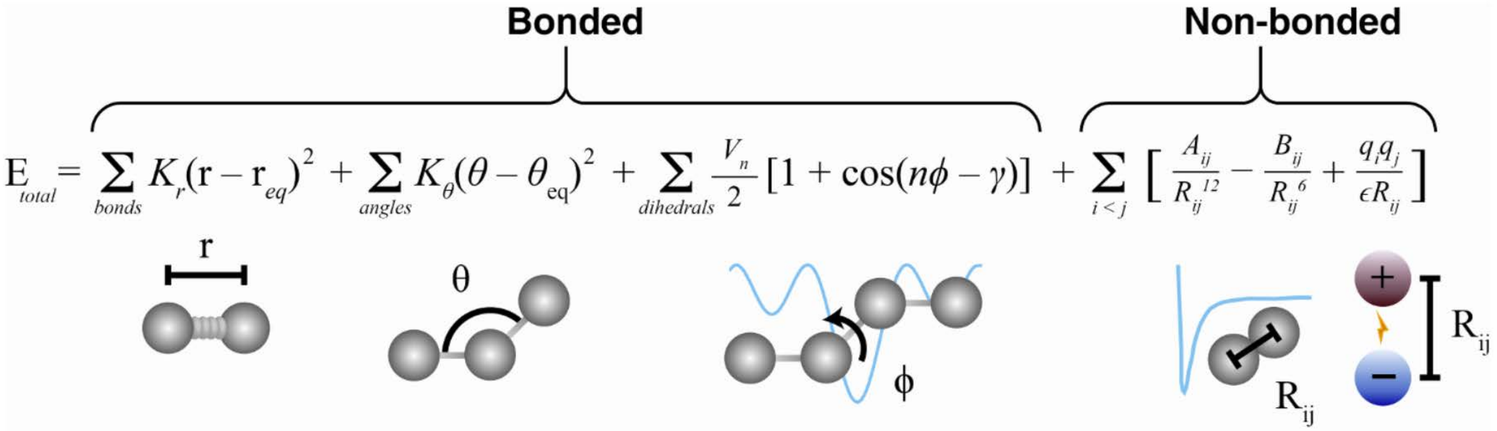
AlphaFold2: https://www.nature.com/articles/s41586-021-03819-2
OpenFold: https://github.com/aqlaboratory/openfold
Structural coverage: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1009818

STRUCTURAL DATA ENABLES PHYSICAL MODELS TO PROVIDE A DATA EFFICIENT WAY TO GENERALIZE FROM SPARSE DATA



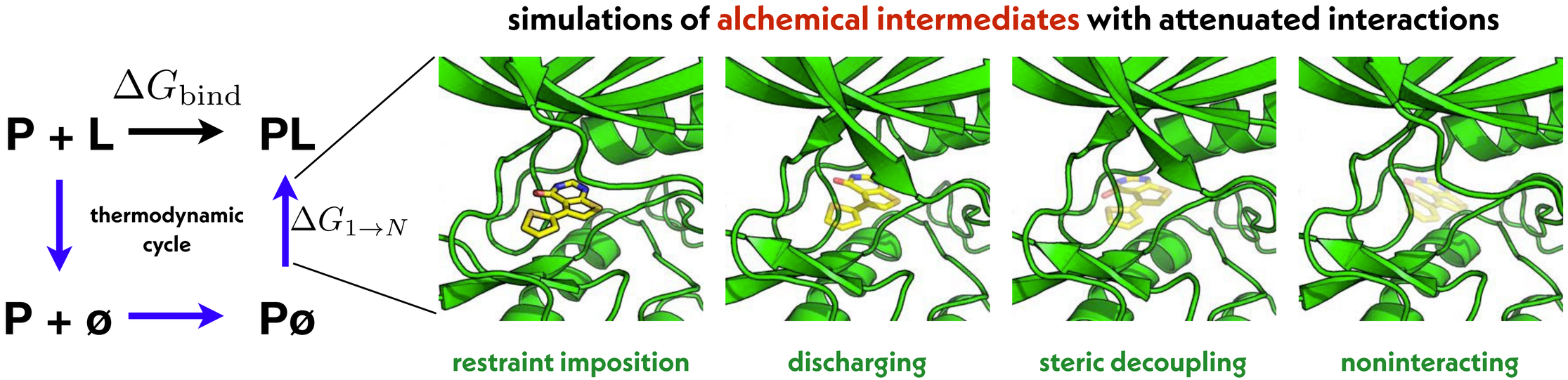
protein target
small molecule(s)
cofactors
ions
waters

typical class I molecular mechanics force field



Shan, Kim, Eastwood, Dror, Seeliger, Shaw. JACS 133:9181, 2011
Durrant, McCammon. Molecular dynamics simulations and drug discovery. BMC Biology, 2011

WE DEVELOPED ALCHEMICAL FREE ENERGY CALCULATIONS INTO A USEFUL TECHNOLOGY TO EXPLOIT STRUCTURAL DATA TO PREDICT AFFINITIES



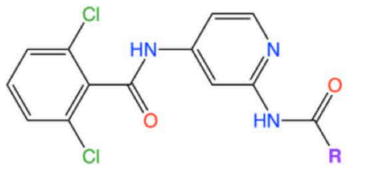
Includes all contributions from enthalpy and entropy of binding to a flexible receptor

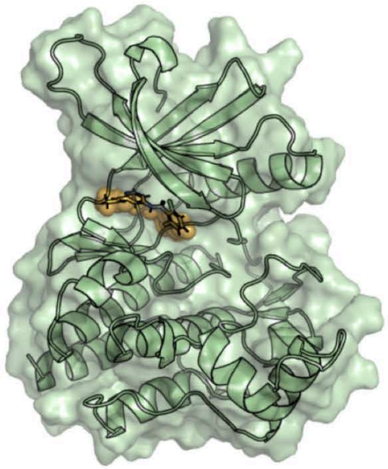
$$\Delta G_{1\to N} = -\beta^{-1} \ln \frac{Z_N}{Z_1} = -\beta^{-1} \ln \frac{Z_2}{Z_1} \cdot \frac{Z_3}{Z_2} \cdots \frac{Z_N}{Z_{N-1}}$$

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

JDC contributions to theory and algorithms: J Chem Phys 125:084902, 2006; J Chem Theor Comput 3:1231, 2007; J Phys CHem B 111:2242, 2007; J Phys Chem B 111:13052, 2007; J Chem Phys 126:155101, 2007; J Chem Phys 129:124105, 2008; J Med Chem 51:769, 2008; J Chem Phys 134:174105, 2011; PNAS 108:E1009, 2011; J Computer Aided Molecular Design 27:989, 2013; J Phys Chem B 119:12912, 2015; J Phys Chem B 122:5579, 2018; J Phys Chem B 122:5466, 2018; Entropy 20:318, 2018; Living Journal of Computational Molecular Sciences 2:1, 2020

WE CAN PREDICT SMALL MOLECULE AFFINITIES WITHIN A LEAD SERIES TO USEFUL ACCURACY



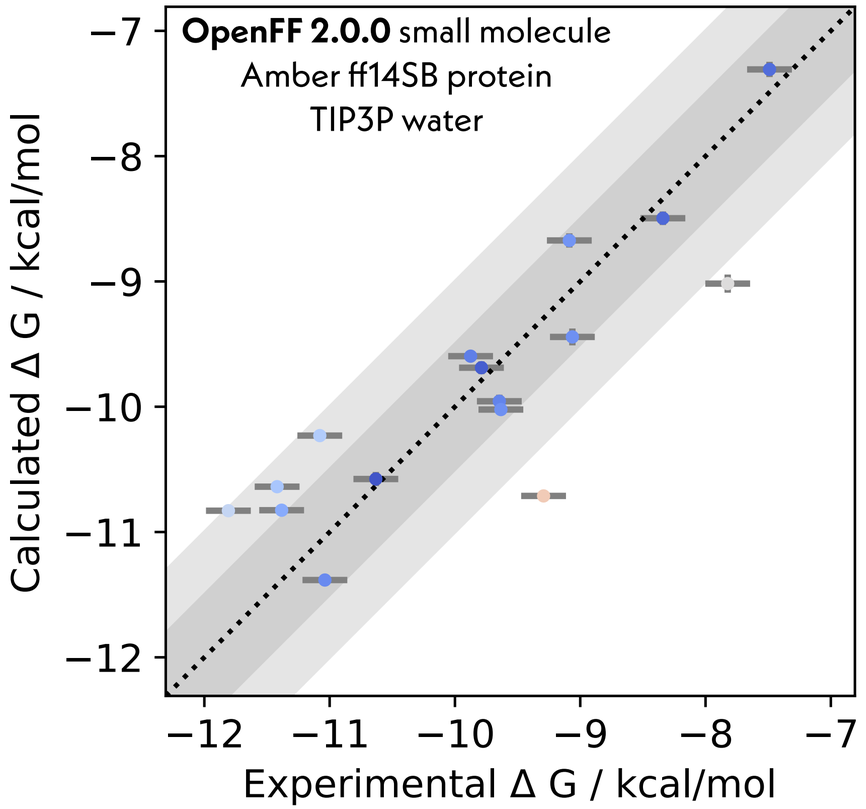


Tyk2
PDBID: 4GIH

	ΔG_{exp} / kcal mol^{-1}		ΔG_{exp} / kcal mol^{-1}
1	-9.54	9	-9.56
2	-10.94	10	-7.42
3	-8.98	11	-11.28
4	-11.31	12	-9.00
5	-9.21	13	-9.70
6	-8.26	14	-11.70
7	-10.91	15	-9.78
8	-7.75	16	-10.53

Absolute binding energies - benchmark benchmark (N = 16)

RMSE:	0.66 [95%: 0.45, 0.89]
MUE:	0.52 [95%: 0.35, 0.75]
R2:	0.74 [95%: 0.43, 0.90]
rho:	0.86 [95%: 0.64, 0.95]




MIKE HENRY


IVÁN PULIDO


IVY ZHANG


DOMINIC RUFA

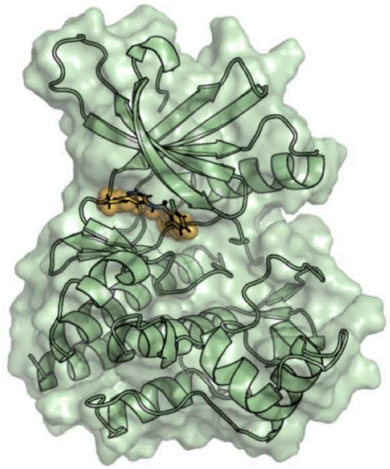

HANNAH BRUCE MACDONALD


MELISSA BOBY

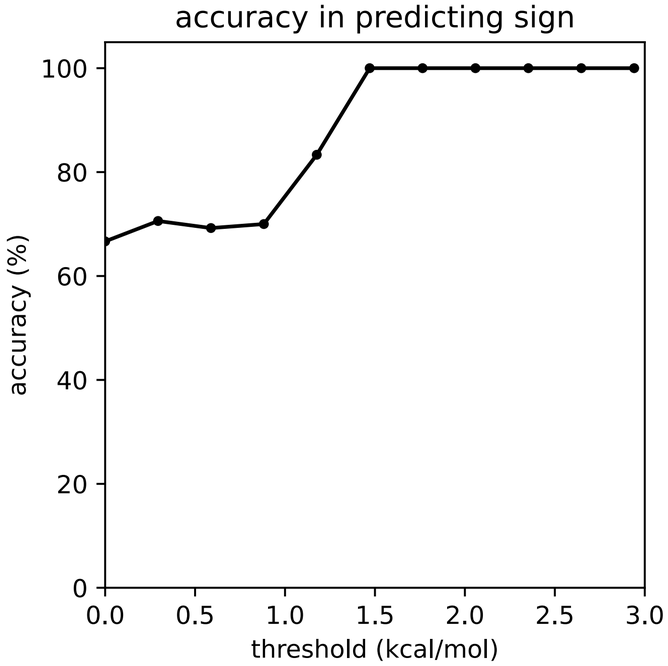
free energy calculations with http://github.com/choderalab/perses

WE CAN PREDICT SMALL MOLECULE AFFINITIES WITHIN A LEAD SERIES TO USEFUL ACCURACY

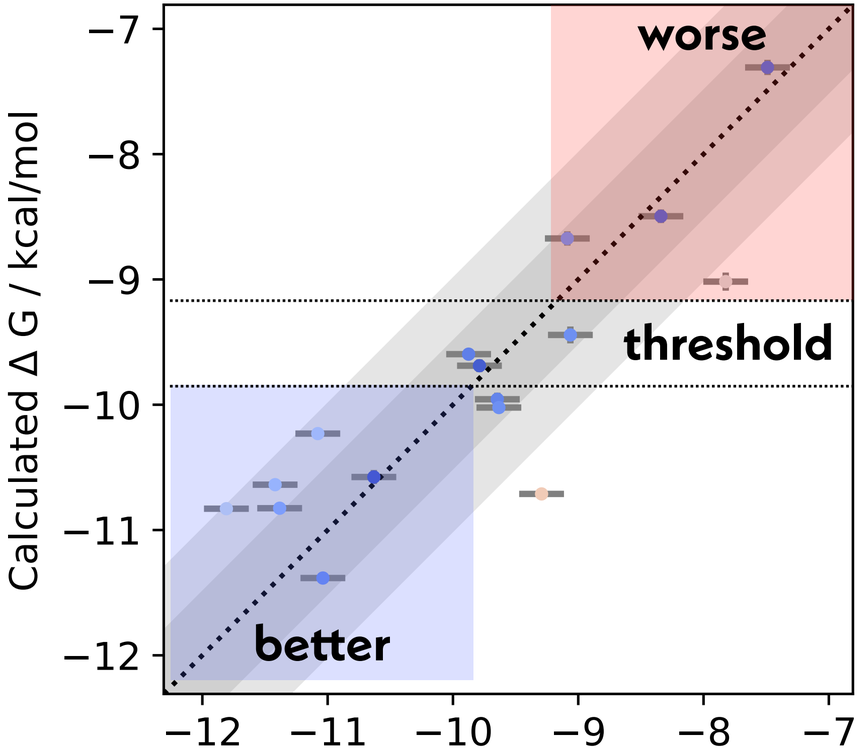
How often can this help us make the right decision about which molecules to synthesize?



Tyk2
PDBID: 4GIH




```
Absolute binding energies - benchmark
                    benchmark (N = 16)
RMSE:    0.66 [95%: 0.45, 0.89]
MUE:     0.52 [95%: 0.35, 0.75]
R2:      0.74 [95%: 0.43, 0.90]
rho:     0.86 [95%: 0.64, 0.95]
```



MIKE HENRY

IVÁN PULIDO

IVY ZHANG

DOMINIC RUFA

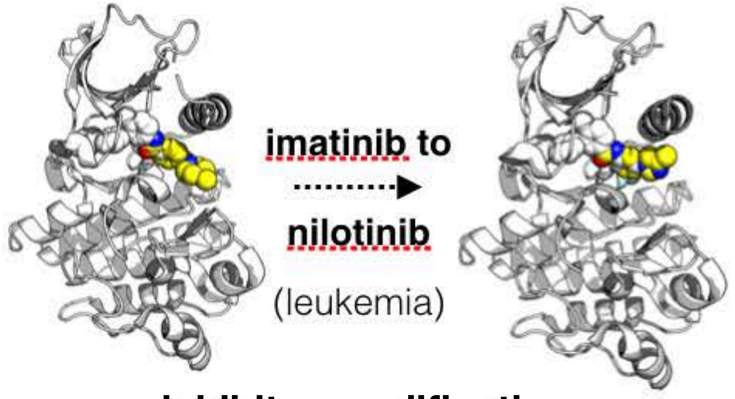
HANNAH BRUCE MACDONALD

MELISSA BOBY

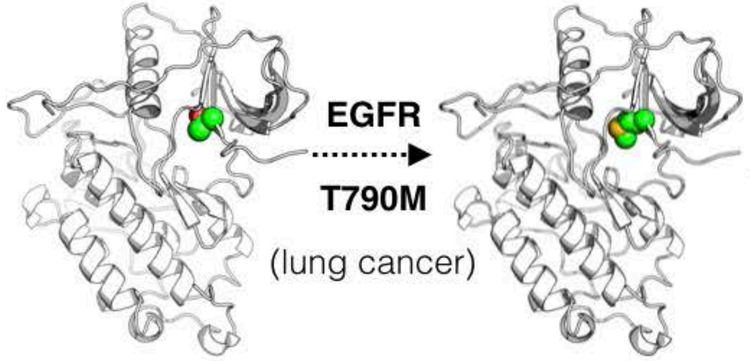
free energy calculations with http://github.com/choderalab/perses

WE PREVIOUSLY SHOWED WE CAN USE FREE ENERGY CALCULATIONS TO ADDRESS MAJOR QUESTIONS IN CANCER DRUG DISCOVERY AND THERAPY

CHANGES OF A FEW ATOMS



inhibitor modification
for drug design



tumor-specific mutation
for therapeutic biomarkers

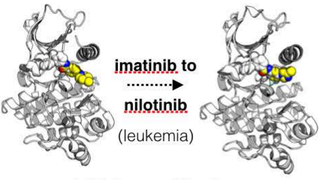
HOW CAN WE DESIGN SPECIFICALLY TARGETED CANCER DRUGS?

▲ Albanese, Chodera, Volkamer, Keng, Abel, Wang.
J Chem Inf Model 60:6211, 2020
https://doi.org/10.1021/acs.jcim.0c00815

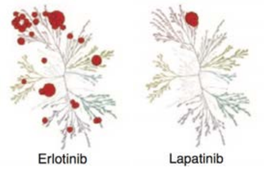
HOW CAN WE PREDICT DRUG RESISTANCE AND SUSCEPTIBILITY?

▲ Hauser, Negron, Albanese, Ray, Steinbrecher, Abel, Chodera, Wang.
Communications Biology 1:70, 2018
https://doi.org/10.1038/s42003-018-0075-x

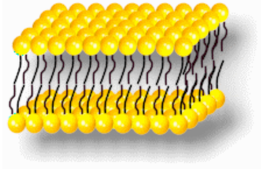
ALCHEMICAL FREE ENERGY CALCULATIONS CAN BE USED TO COMPUTE MANY DRUG PROPERTIES



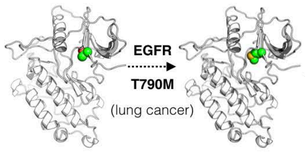
driving potency



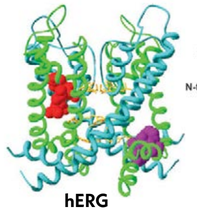
driving selectivity



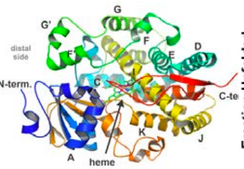
partition coefficients and permeabilities



predicting drug resistance/sensitivity

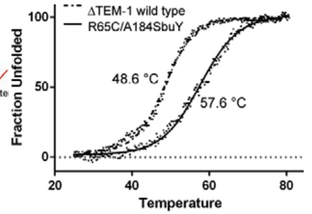


hERG



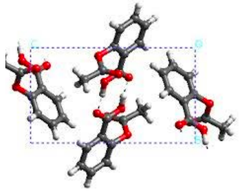
CYP3A4

structure-enabled ADME/Tox targets



optimizing thermostability

bacterial porin permeation



crystal polymorphs, solubility

Target Candidate Profile (TCP)

HOW WELL DO THESE METHODS WORK IN A REAL DRUG DISCOVERY PROGRAM?

"All the News That's Fit to Print"

The New York Times

Late Edition

Today, clouds and sunshine, breezy, mild for late January, high 46. Tonight, partly cloudy, low 33. Tomorrow, clouds, sunshine, high 44. Weather map appears on Page B6.

VOL. CLXIX ... No. 58,585 © 2020 The New York Times Company NEW YORK, MONDAY, JANUARY 27, 2020 $3.00

Jan 27, 2020 (4 days after SARS-CoV-2 genome was published)

PHILIP CHEUNG FOR THE NEW YORK TIMES

A vigil Sunday in Los Angeles for Kobe Bryant, who spent his entire 20-year career with the Lakers and won five N.B.A. titles.

Helicopter Crash Kills N.B.A. Star Known to All as Kobe

By SCOTT CACCIOLA

Kobe Bryant, the retired Los Angeles Lakers basketball star who was one of the greatest to play the game, and his 13-year-old daughter were among nine people killed in a helicopter crash on Sunday outside Los Angeles, rocking the sports world and generating an outpouring of grief and shock across the country.

The helicopter went down near Calabasas, Calif., about 30 miles northwest of downtown Los Angeles, in foggy conditions, though the authorities were investigating the cause. The helicopter was on its way from Orange County, Calif., where the Bryant family lives, to Mr. Bryant's youth basketball academy northwest of Los Angeles, where he coaches his daughter Gianna, who died in the crash.

It was a moment of national mourning, coast to coast. Thousands of people converged at Staples Center, the Lakers' home arena in downtown Los Angeles; condolences poured in from presidents, celebrities and sports luminaries; and several entertainers paid tribute to Mr. Bryant at the Grammy Awards, which took place at the arena hours later. A shrine emerged at Mr. Bryant's high school alma mater in suburban Philadelphia.

A Giant of the Sport, Recognized in Beijing or in Beverly Hills

Mr. Bryant, 41, a quiet force of nature on the court who gave himself the nickname Black Mamba, retired in 2016 with five N.B.A. championship rings and a long list of N.B.A. records — he was surpassed by LeBron James on Saturday night for third on the N.B.A. career scoring list. Signing with the league right out of high school in 1996, he changed the way the N.B.A. identified, groomed and developed its youngest stars.

Yet he was far more than a basketball giant. He was among the world's best-known athletes, a star on the order of Tiger Woods and Michael Jordan, swarmed by fans whether he was in Beijing or Beverly Hills. It is not uncommon to hear young people in some quarters shout, "Kobe!" when

Continued on Page A11

MONEY TO UKRAINE TIED TO INQUIRIES, BOLTON BOOK SAYS

President Rebuffed Top Cabinet Officials Who Urged Him to Release Aid

By MAGGIE HABERMAN and MICHAEL S. SCHMIDT

WASHINGTON — President Trump told his national security adviser in August that he wanted to continue freezing $391 million in security assistance to Ukraine until officials there helped with investigations into Democrats including the Bidens, according to an unpublished manuscript by the former adviser, John R. Bolton.

The president's statement as described by Mr. Bolton could undercut a key element of his impeachment defense: that the hold-up in aid was separate from Mr. Trump's requests that Ukraine announce investigations into his perceived enemies, including former Vice President Joseph R. Biden Jr. and his son Hunter Biden, who had worked for a Ukrainian energy firm while his father was in office.

Mr. Bolton's explosive account of the matter at the center of Mr. Trump's impeachment trial, the third in American history, was included in drafts of a manuscript he has circulated in recent weeks to close associates. He also sent a draft to the White House for a standard review process for some current and former administration officials who write books.

Multiple people described Mr. Bolton's account of the Ukraine affair.

The book presents an outline of what Mr. Bolton might testify to if he is called as a witness in the Senate impeachment trial, the people said. The White House could use the prepublication review process, which has no set time frame, to delay or even kill the book's publication or omit key passages.

Over dozens of pages, Mr. Bolton described how the Ukraine affair unfolded over several months until he departed the White House in September. He described not only the president's private disparagement of Ukraine but also new details about senior cabinet officials who have publicly tried to sidestep involvement.

For example, Secretary of State Mike Pompeo acknowledged privately that there was no basis to claims by the president's lawyer Rudolph W. Giuliani that the ambassador to Ukraine was corrupt and believed Mr. Giuliani may have been acting on behalf of other clients, Mr. Bolton wrote.

Mr. Bolton also said that after

Continued on Page A12

DOUG MILLS/THE NEW YORK TIMES

The former national security adviser John R. Bolton.

Disease Surges, And Lockdown May Not Halt It

This article is by ***Chris Buckley, Raymond Zhong, Denise Grady and Roni Caryn Rabin.***

WUHAN, China — A top Chinese health official warned on Sunday that the spread of the dangerous new coronavirus, already extraordinarily rapid, is accelerating further, deepening global fears about an illness that has sickened more than 2,700 people worldwide and killed at least 80 people in China.

The grim diagnosis came amid concerns that China's efforts to contain the spread of the disease, despite a lockdown of unprecedented scope affecting 56 million people, may not only have come too late but could even make the situation worse, including by exacerbating shortages of medical supplies.

Adding to the growing global alarm, people who are carrying the virus but not showing symptoms may still be able to infect others, according to the Chinese official, Ma Xiaowei, the director of China's National Health Commission. Such asymptomatic transmissions would make the disease much more difficult to control, as seemingly healthy people travel and interact with others.

"The epidemic is now entering a more serious and complex period," Mr. Ma said during a Sunday news conference in Beijing. "It looks like it will continue for some time, and the number of cases may increase."

China's attempts to curb the disease's spread — essentially cordoning off the major cities in the province of Hubei, including its

Continued on Page A8

HECTOR RETAMAL/AGENCE FRANCE-PRESSE — GETTY IMAGES

Hospitals in Wuhan, China, the epicenter of the coronavirus outbreak, remain intensely crowded.

Novel Virus Tests China's Authoritarian Bargain

By STEVEN LEE MYERS and CHRIS BUCKLEY

Xi Is Under Pressure to Tame Growing Crisis

BEIJING — It took thousands of infections and scores of deaths from a mysterious virus for China's authoritarian leader to publicly say what had become glaringly obvious to many in recent weeks: The country is facing a grave public health crisis.

After his declaration, the leader, Xi Jinping, put China on a virtual war footing to cope with the unfolding epidemic of the coronavirus. He convened an extraordinary session of the Communist Party's top political body, issuing orders for handling the crisis with the crisp, somber stoicism of a field marshal.

"We're sure to be able to win in this battle," he proclaimed on Saturday before his six grim-faced colleagues on the party's Politburo Standing Committee.

Compared to the very low bar set by the Chinese leadership's secrecy and inaction during the SARS epidemic in 2002 and 2003, Mr. Xi has responded with speed and alacrity to the latest health emergency, a pneumonialike virus that at last official count has killed at least 80, sickened thousands in China and spread around the world.

But there are also signs that the government, especially at the regional level in Hubei Province, the source of the outbreak, was slow to recognize the danger and is continuing to mishandle the crisis. Some public health experts have asked whether the sweeping travel restrictions that have been imposed are leaving people without access to medical care, while many Chinese remain unconvinced the government is being completely forthcoming about the toll of the disease.

"Substantively, the response

Continued on Page A8

NEWS ANALYSIS

G.O.P. Sees a Kyiv Sideshow; Democrats See Russia's Hand

By DAVID E. SANGER

WASHINGTON — When Secretary of State Mike Pompeo, in a curse-laden tirade to a reporter on Friday, asked, "Do you think Americans care about Ukraine?" he was getting at an essential element of President Trump's defense in the impeachment trial. White House officials are convinced that Americans are indifferent to what happens in the struggling former Soviet republic, and they may well be right.

But the impeachment trial is about more than the fate of Ukraine — and whether Mr. Trump sold it out for a "domestic political errand," as his former adviser, Fiona Hill, put it so bitingly. To Democrats, it's about a president who undercut his own administration's stated goal of pushing back hard against Vladimir V. Putin's Russia — the geopolitical challenge of a new, very different Cold War.

It is one of those cases where the geography of the debate shapes the politics of the argument.

As long as the president's lawyers can focus the debate on the narrower question of Ukraine, they can argue that the charges against the president focus on a foreign policy side-

Continued on Page A14

For Sanders, an Internet Army That Hits Critics Close to Home

This article is by ***Matt Flegenheimer, Rebecca R. Ruiz and Nellie Bowles.***

The defense from Bernie Sanders was straightforward: It wasn't me.

He had been milling about on the Senate floor one day in the summer of 2017 when a colleague, Kamala Harris, stepped toward him. "Do we have a problem?" Ms. Harris asked, according to Democrats familiar with the exchange.

Some prominent Sanders supporters had been flaming Ms. Harris publicly as the preferred choice of the corporate Democratic establishment against which Mr. Sanders had long railed, a view amplified among Sanders-boosting accounts across social media. "Pre-emptive strike," one wrote on the popular SandersForPresident Reddit group, where Sanders fans were sharing details of Ms. Harris's recent fund-raising swing in the Hamptons with former Hillary Clinton donors. "Start the conversation now, end it before 2020."

Mr. Sanders assured Ms. Harris

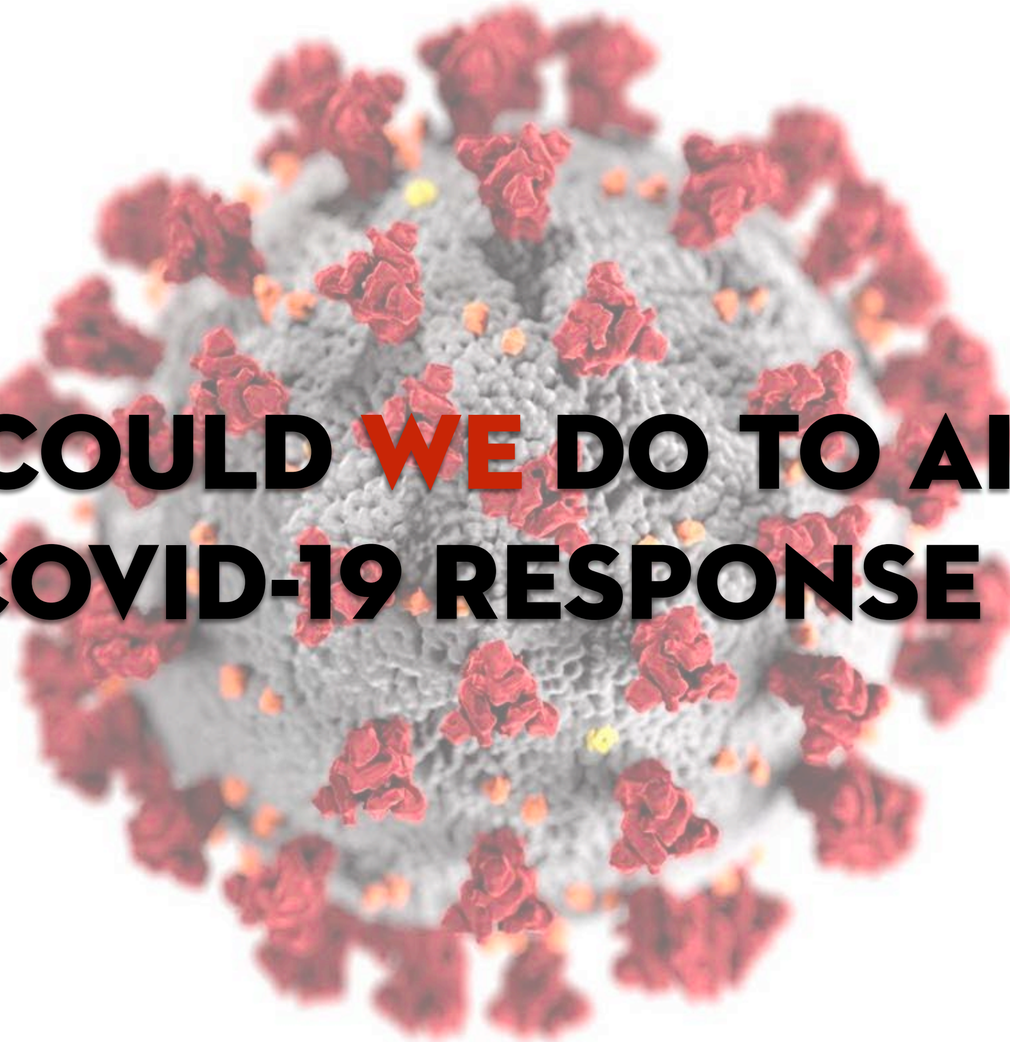
Continued on Page A17

doi:10.3389/fcimb.2021.697876

Will antibodies still bind against escape mutants?
How can we optimize antibodies to ensure broad specificity against mutations?

VIR

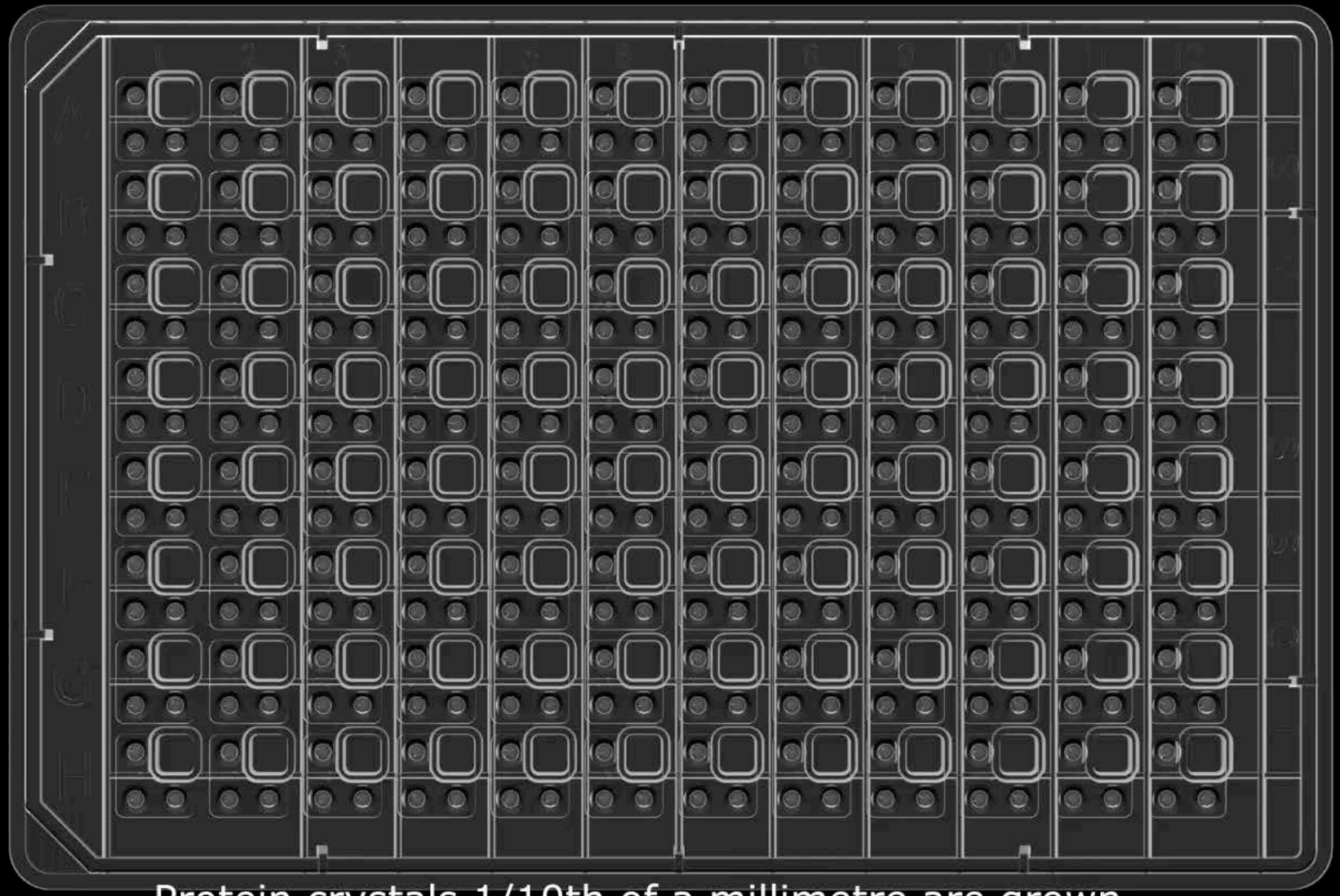
IVY ZHANG
CBM student



WHAT COULD WE DO TO AID THE GLOBAL COVID-19 RESPONSE EFFORT?

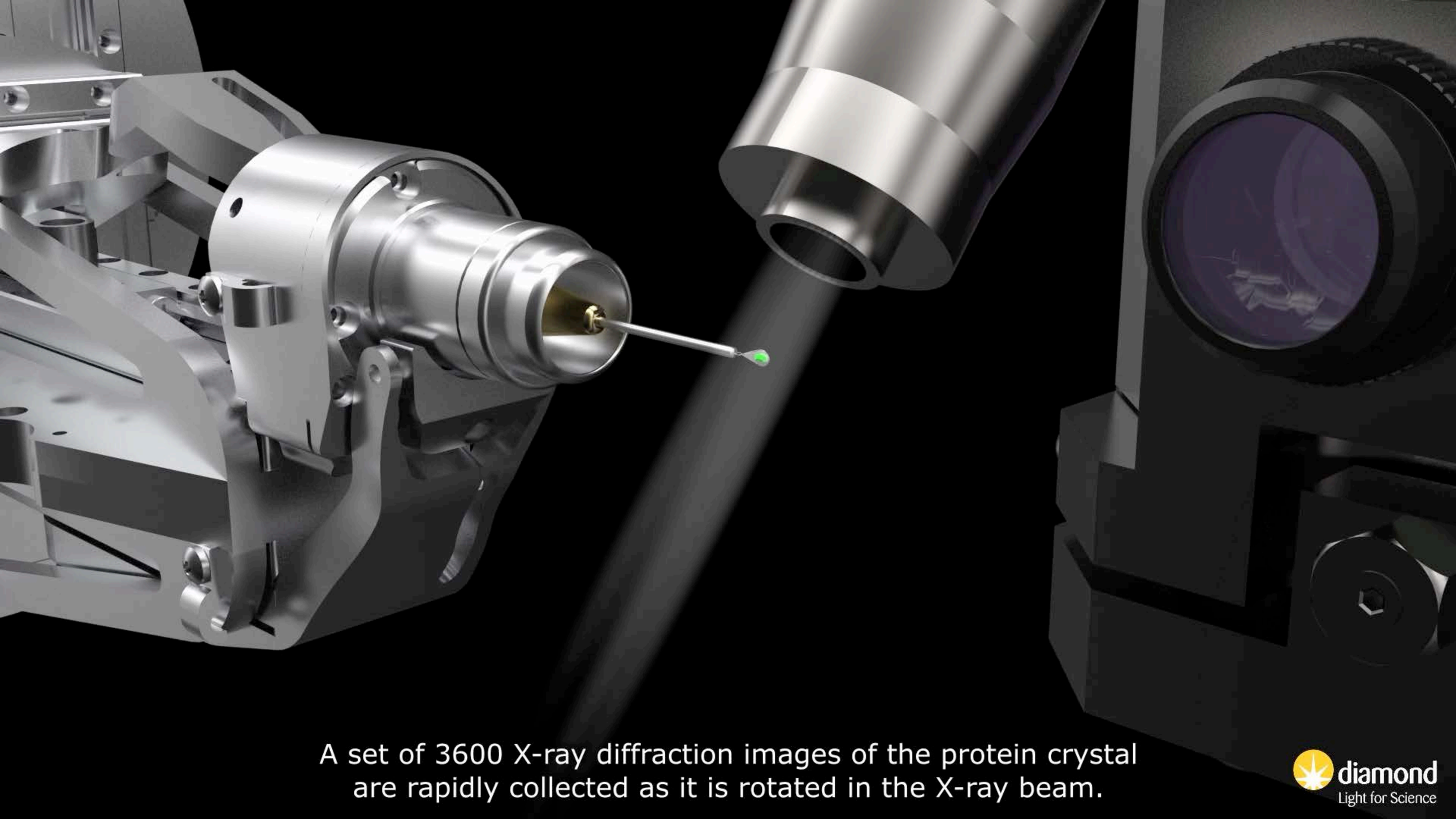
WE HAD ALSO BEEN COLLABORATING WITH DIAMOND LIGHT SOURCE IN THE UK

Diamond Light Source, UK



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

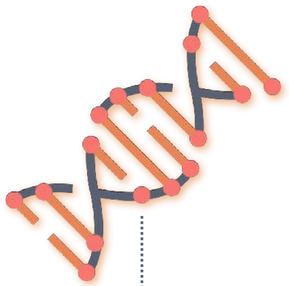
diamond
Light for Science



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

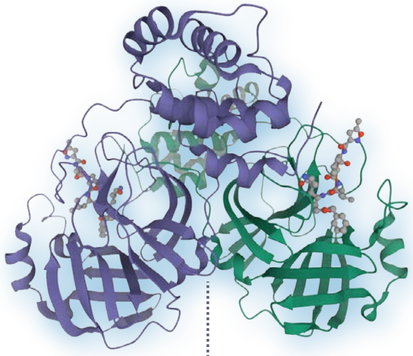
DIAMOND LIGHT SOURCE PROSECUTED A HIGH-THROUGHPUT X-RAY FRAGMENT SCREEN IN JUST WEEKS



February 14

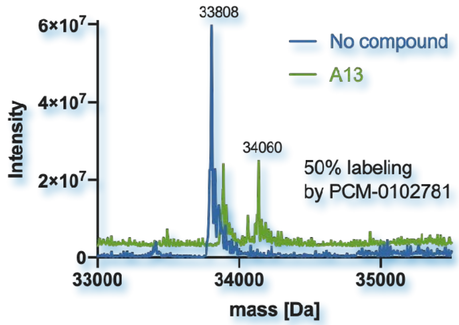
Main protease cloned and produced at Diamond after COVID shutdown of Haitao Yang lab in Shanghai

Martin Walsh



February 20

Atomic resolution structure of the protease determined

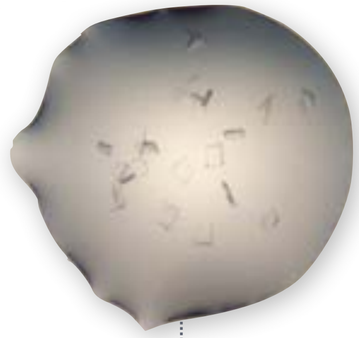


February 25

Covalent screen finds 150 active site hits

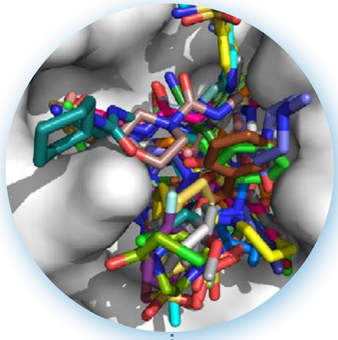
>40 hits validated

Nir London



March 5

1,500 crystals collected in one day (!)



March 18

78 fragment-bound structures solved and released to the web

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

Frank von Delft
Diamond Light Source / XChem / SGC

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.

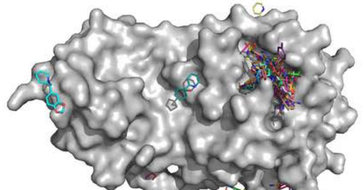
On the 11th of May, the first biochemical and structural data from Moonshot compounds was released and by the 12th of June over 500 compounds had been tested, demonstrating that the design-make-test process is fully in place.

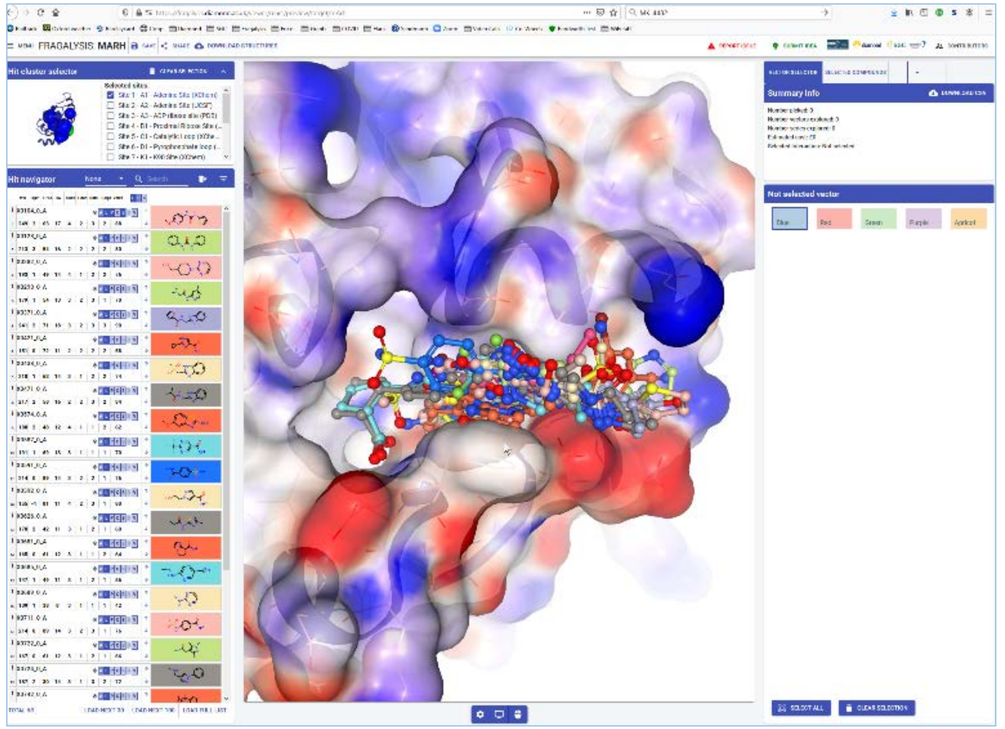
XChem fragment screen

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

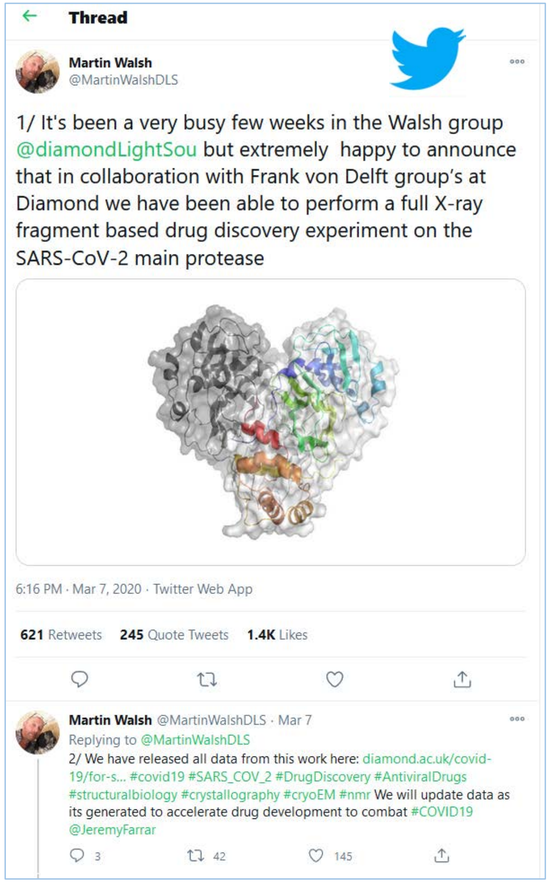




https://fragalysis.diamond.ac.uk

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

(pre-preprinting!)



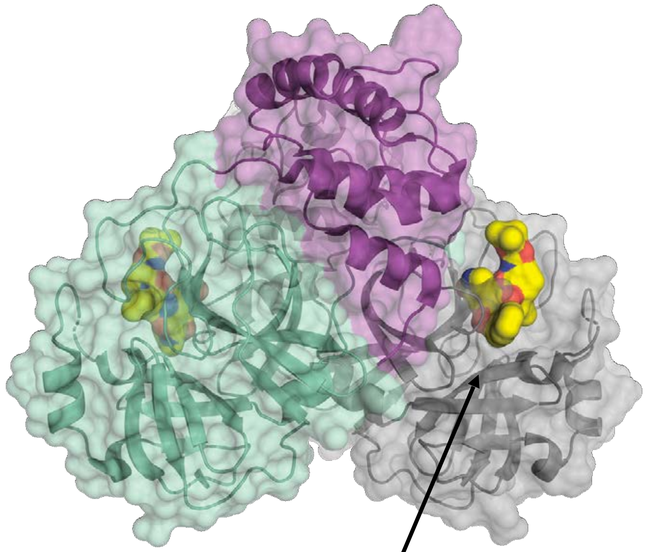
Centre for Medicines Discovery | University of Oxford

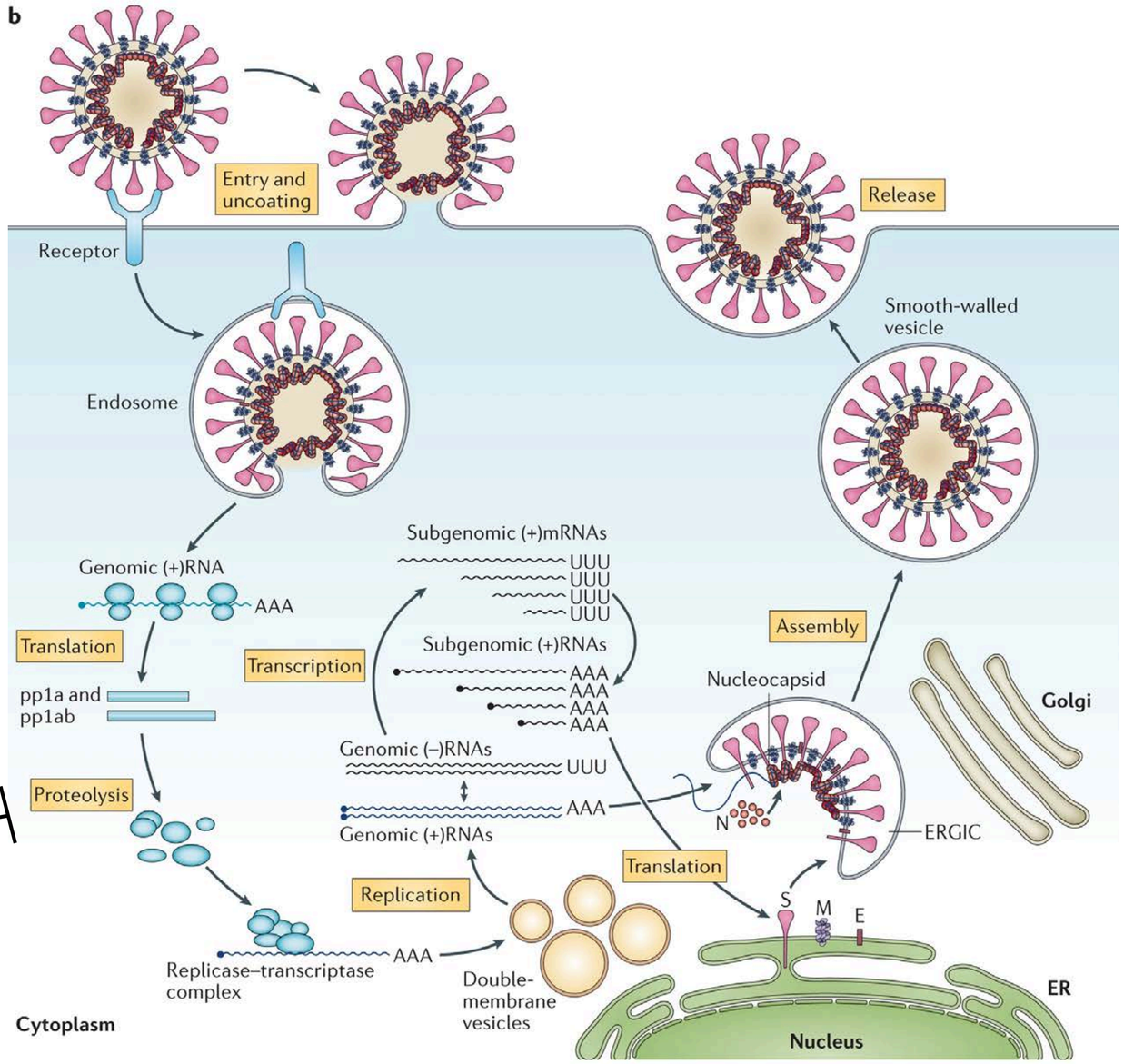
XCHEM Diamond Fragment Screening | diamond

WE KNOW FROM SARS-COV THAT THE MAIN VIRAL PROTEASE (MPRO) IS ESSENTIAL FOR VIRAL REPLICATION

3CL^{Pro}

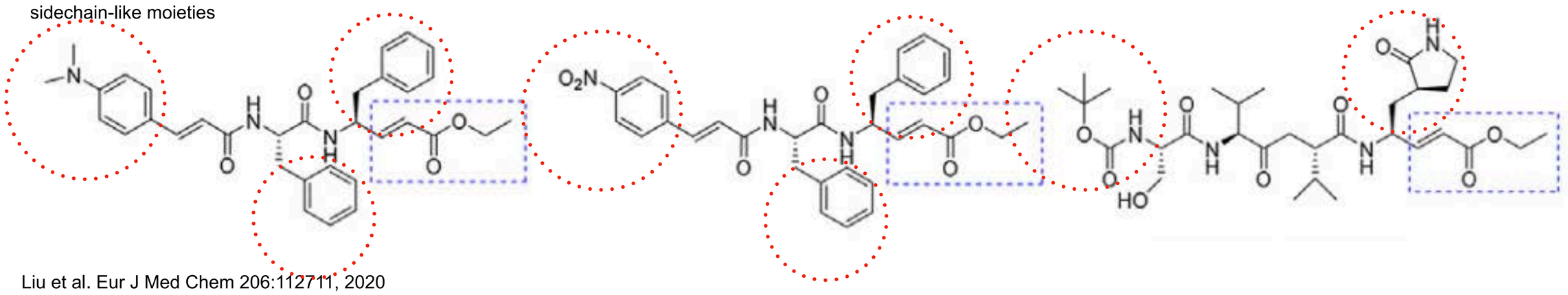
or **Mpro**





de Wit et al. Nat. Rev. Microbiology (2016)

PREVIOUSLY KNOWN SARS-COV MPRO INHIBITORS WERE PEPTIDOMIMETICS, WHICH ARE DIFFICULT TO DEVELOP INTO ORAL DRUGS



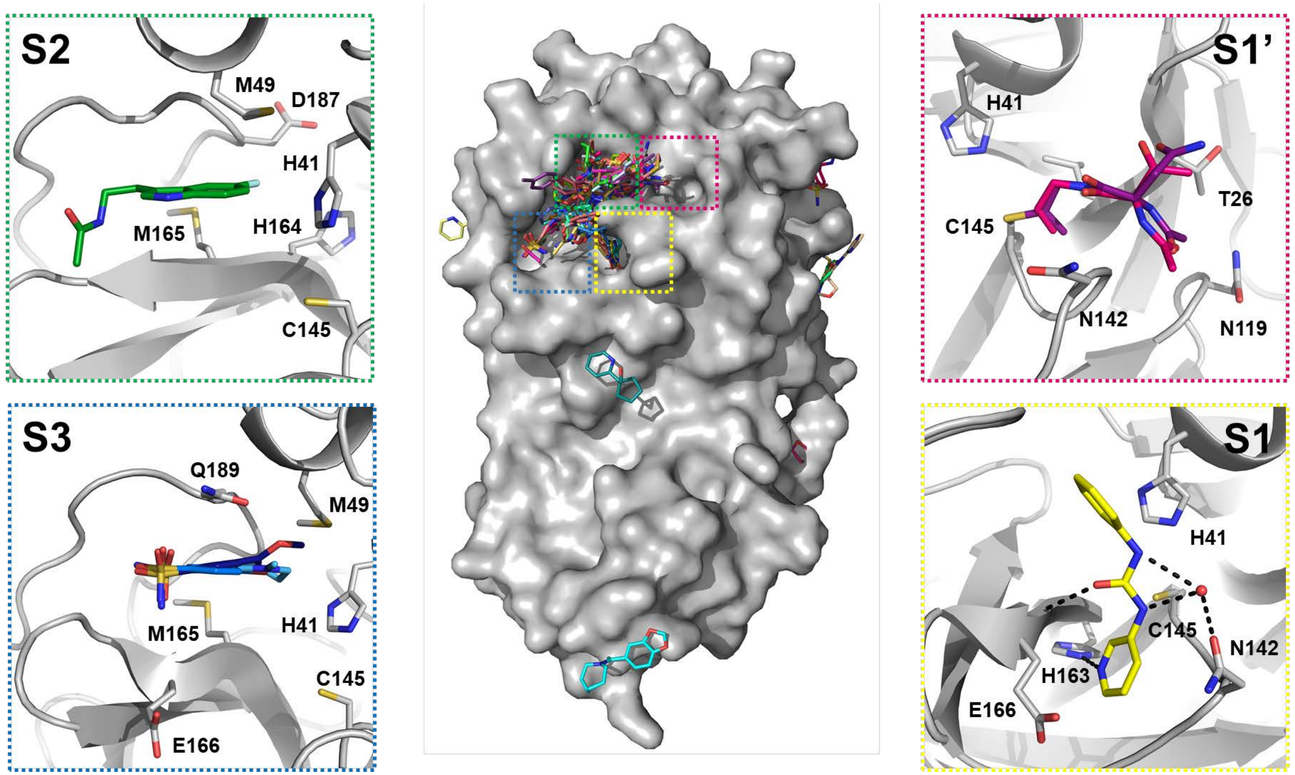
Liu et al. Eur J Med Chem 206:112711, 2020

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

Could X-ray fragment hits be a route to an *oral* SARS-CoV-2 antiviral?

FRAGMENT HITS COMPLETELY COVER THE ACTIVE SITE, SUGGEST FRAGMENT MERGES COULD IMPROVE POTENCY

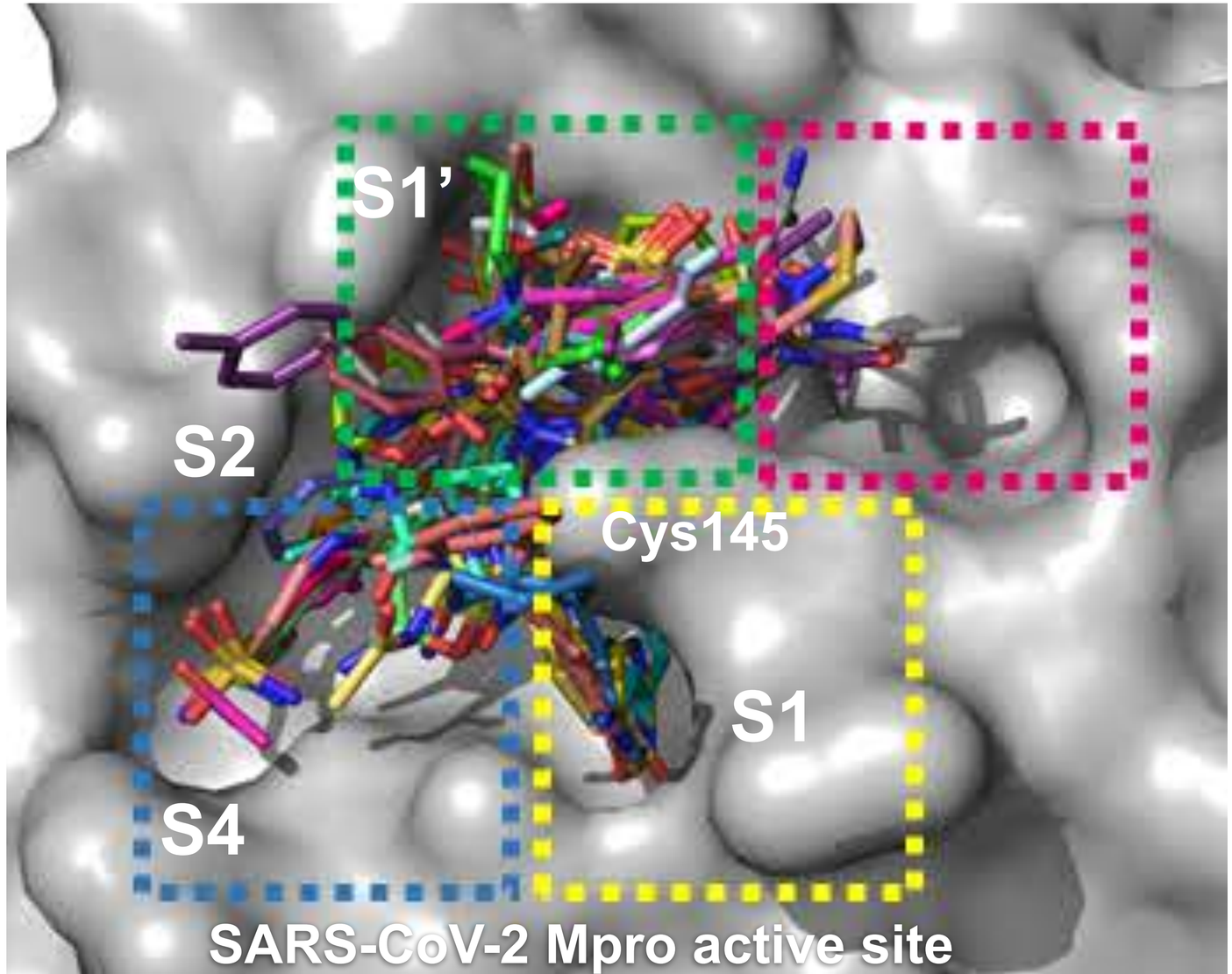
interactive view: https://fragalysis.diamond.ac.uk/viewer/react/preview/target/Mpro

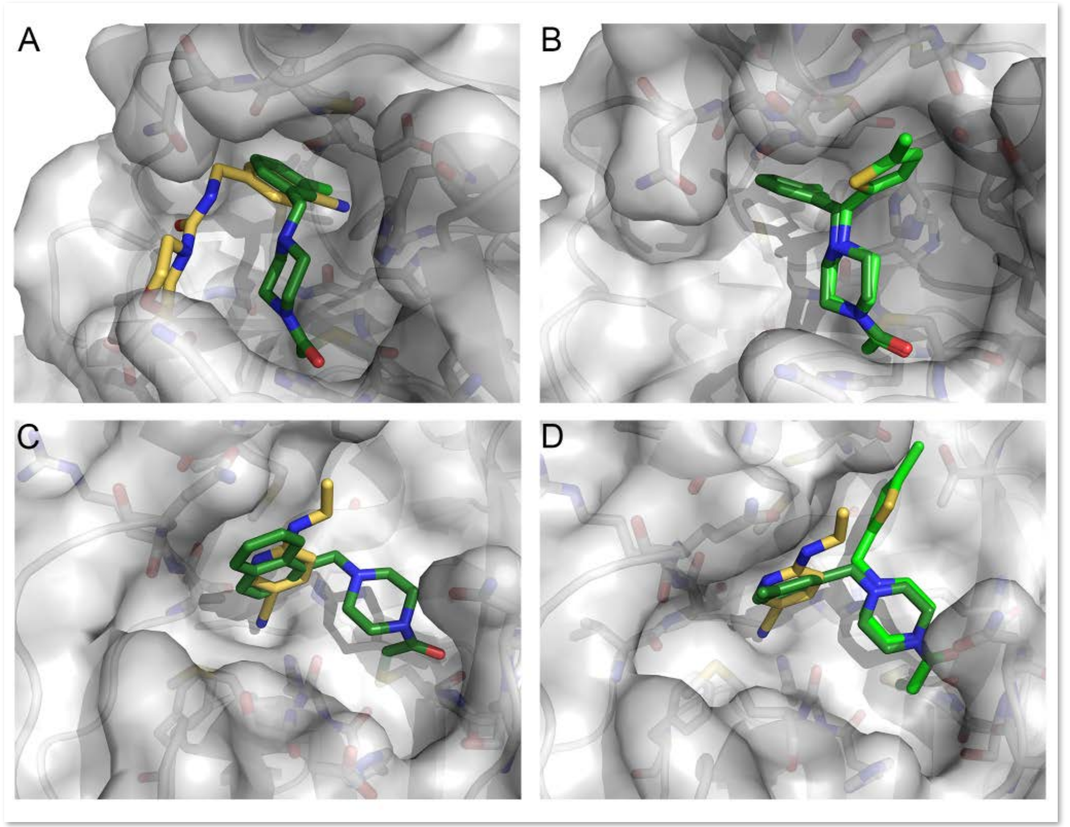


Douangamath et al., Nature Communications 11:5047, 2020
https://www.nature.com/articles/s41467-020-18709-w

FRAGMENT HITS COMPLETELY COVER THE ACTIVE SITE, SUGGEST FRAGMENT MERGES COULD IMPROVE POTENCY

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

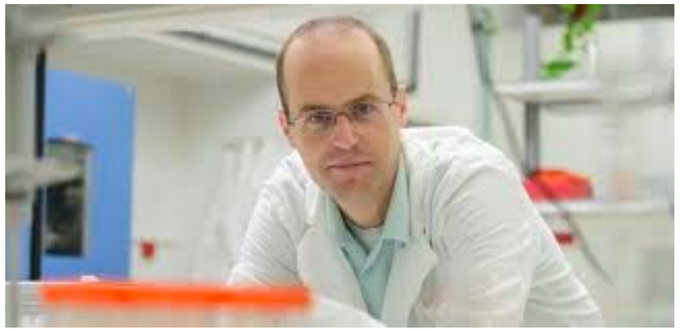




Could we merge our way to potent lead compounds directly?

Douangamath et al., Nature Communications 11:5047, 2020
https://www.nature.com/articles/s41467-020-18709-w

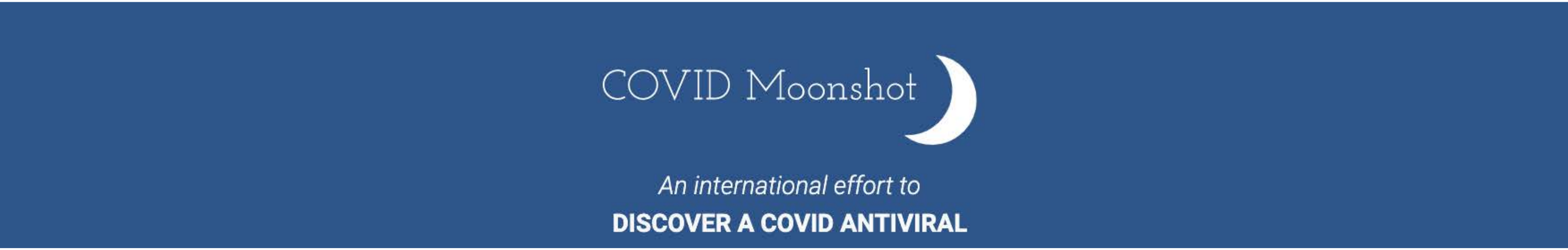
WHICH COMPUTATIONAL STRATEGIES WOULD MOST RAPIDLY PROGRESS FRAGMENTS TO EARLY LEADS WITH MEASURABLE POTENCY?



Nir London
Weizmann Institute

“…what if we tried **all of them?**”

FIRST, WE NEEDED A COOL NAME TO MOTIVATE PEOPLE



THE COVID MOONSHOT ADOPTED A GLOBAL OPEN SCIENCE, PATENT-FREE, COLLABORATIVE APPROACH TO DRUG DISCOVERY

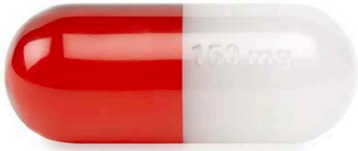
Open science

COVID Moonshot

Open data

http://postera.ai/covid

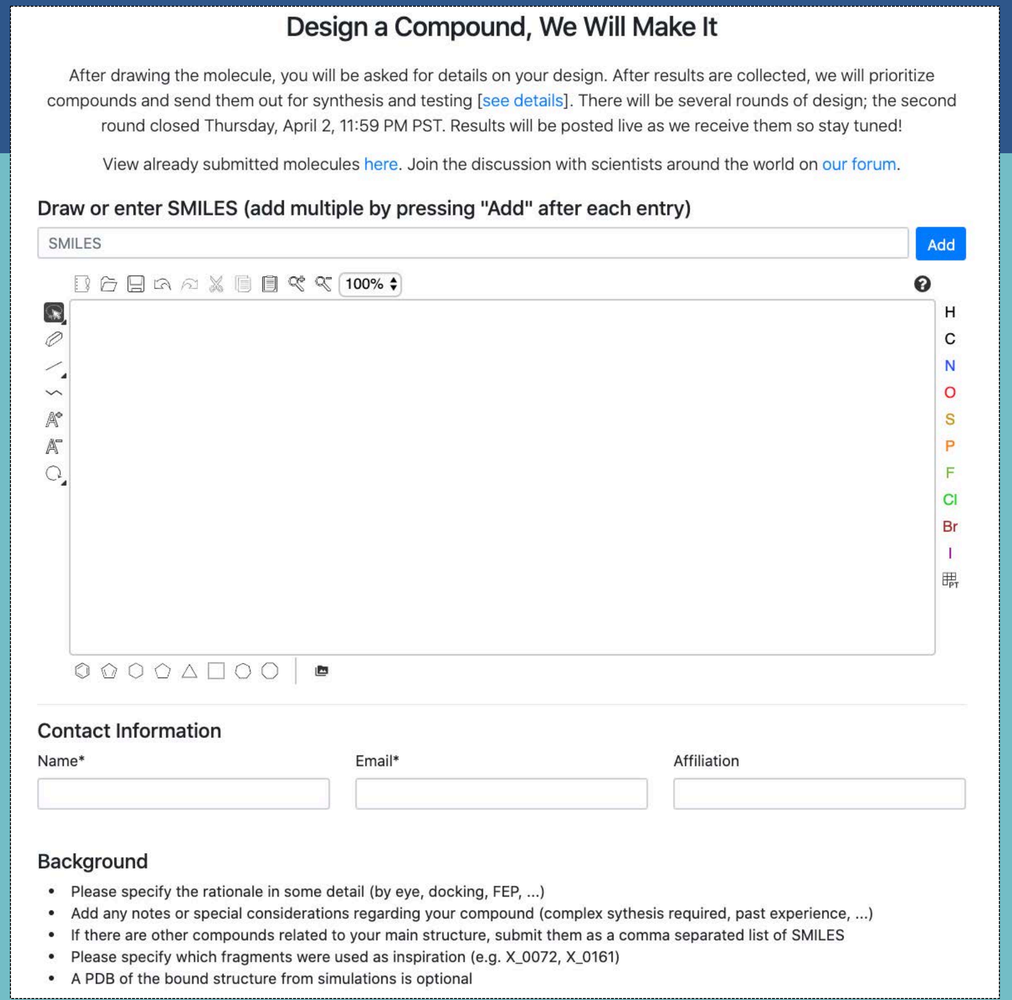
Patent-free



Alpha Lee (Cambridge) tapped his startup company (PostEra) to create an open drug discovery commons website

Alpha Lee
Cambridge/PostEra

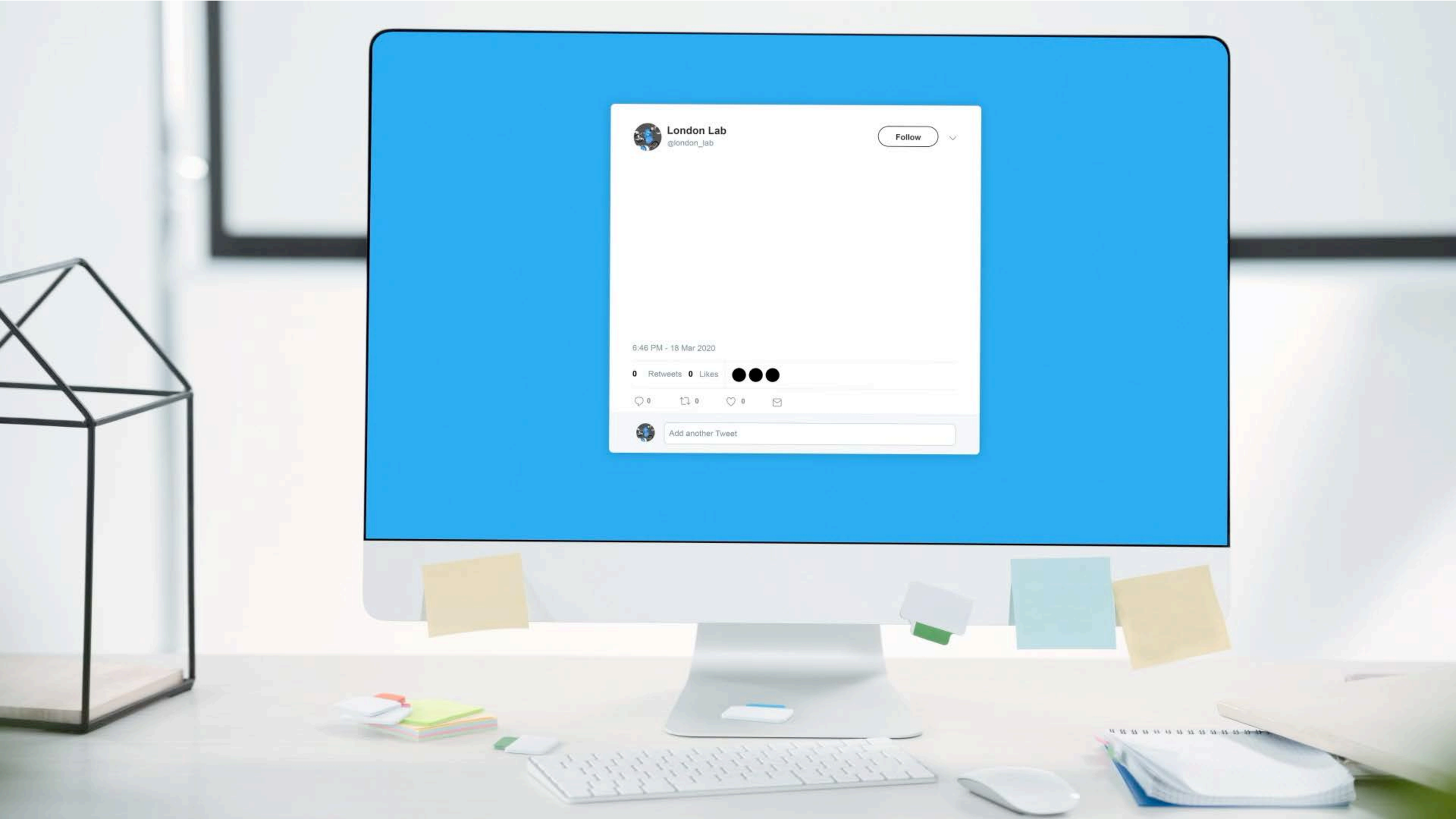
COVID Moonshot



Molecule sketcher!
2D compound design viewer!
Discussion boards!

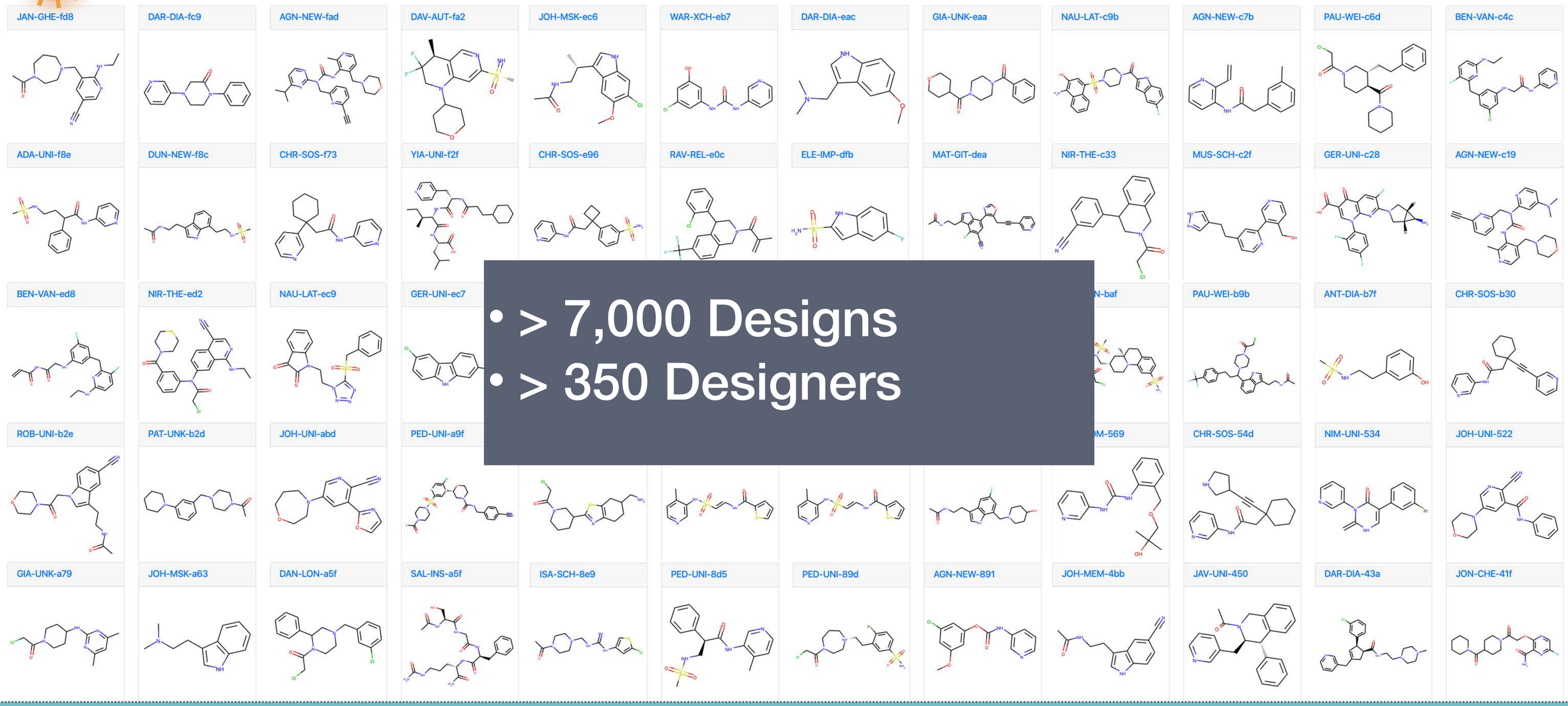
http://postera.ai/covid

+ Matthew Robinson (PostEra)


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

…and there was overwhelming response

- > 7,000 Designs
- > 350 Designers



THERE WERE SOME EXCELLENT IDEAS

MAK-UNK-e05327b2-1

MAK-UNK-e05327b2-2

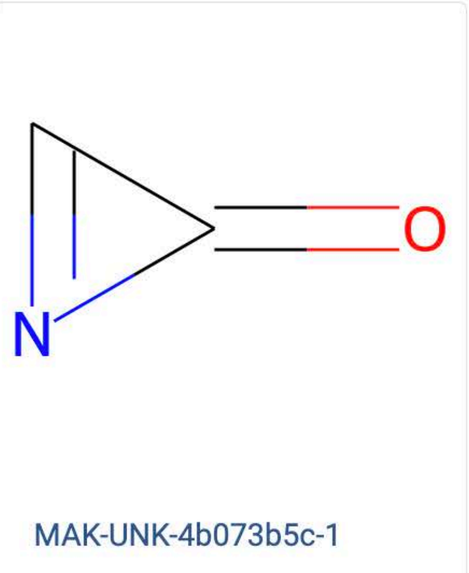
MAK-UNK-e05327b2-3

MAK-UNK-e05327b2-5

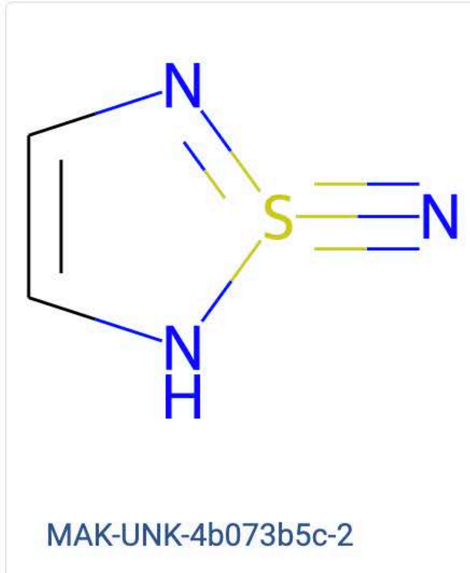
Design Rationale:

using https://molmatinf.com/covid19/ as a score reference

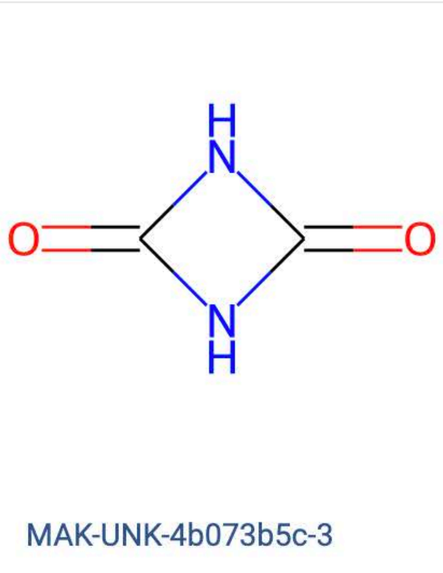
THERE WERE SOME ... INTERESTING ... IDEAS TOO



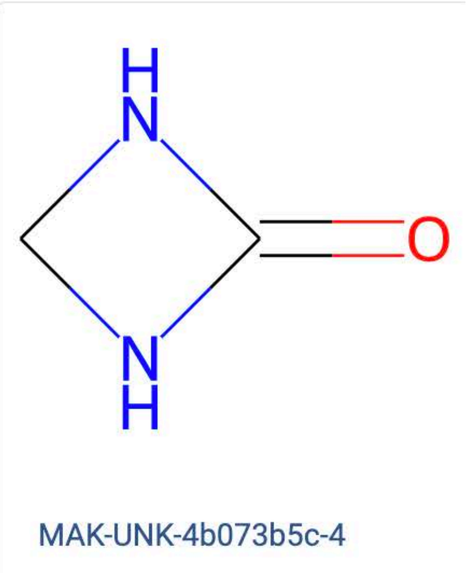
MAK-UNK-4b073b5c-1



MAK-UNK-4b073b5c-2



MAK-UNK-4b073b5c-3

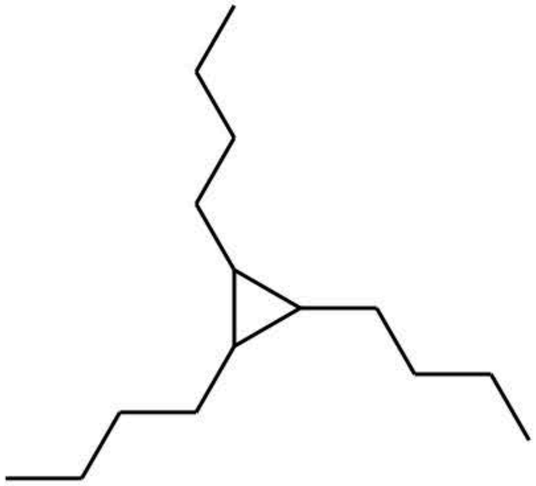


MAK-UNK-4b073b5c-4

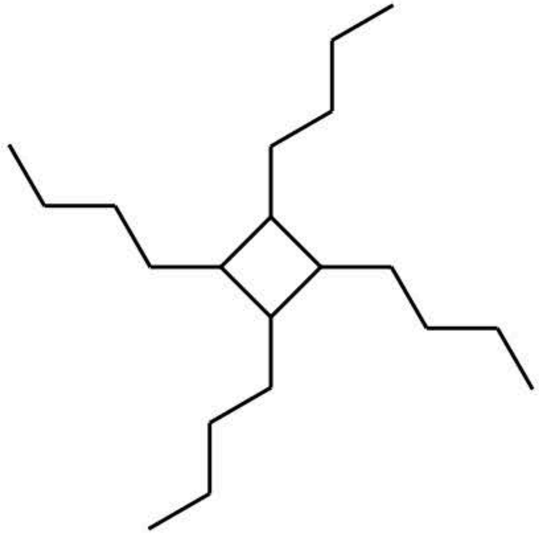
Design Rationale:

by eye, tiny molecules

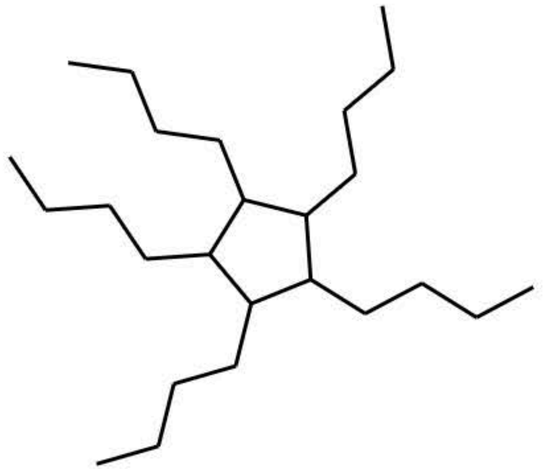
THERE WERE SOME ... INTERESTING ... IDEAS TOO



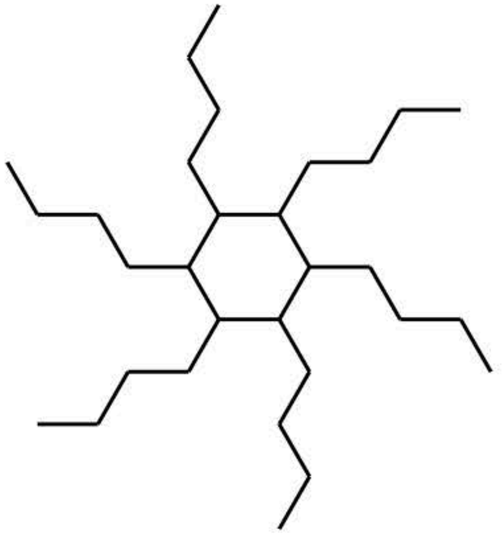
THE-UNK-833274f3-1



THE-UNK-833274f3-2



THE-UNK-833274f3-3



THE-UNK-833274f3-4

Design Rationale:

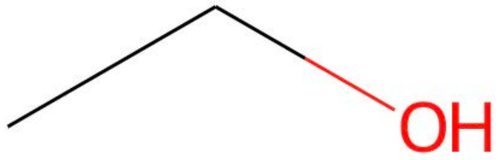
These substances are only carbon, and they have no alarm.

THERE WERE SOME ... INTERESTING ... IDEAS TOO

Design Rationale:

I used random numbers to find this compound.

THERE WERE SOME ... INTERESTING ... IDEAS TOO



JAM-UNK-fcc74568-1

Design Rationale:

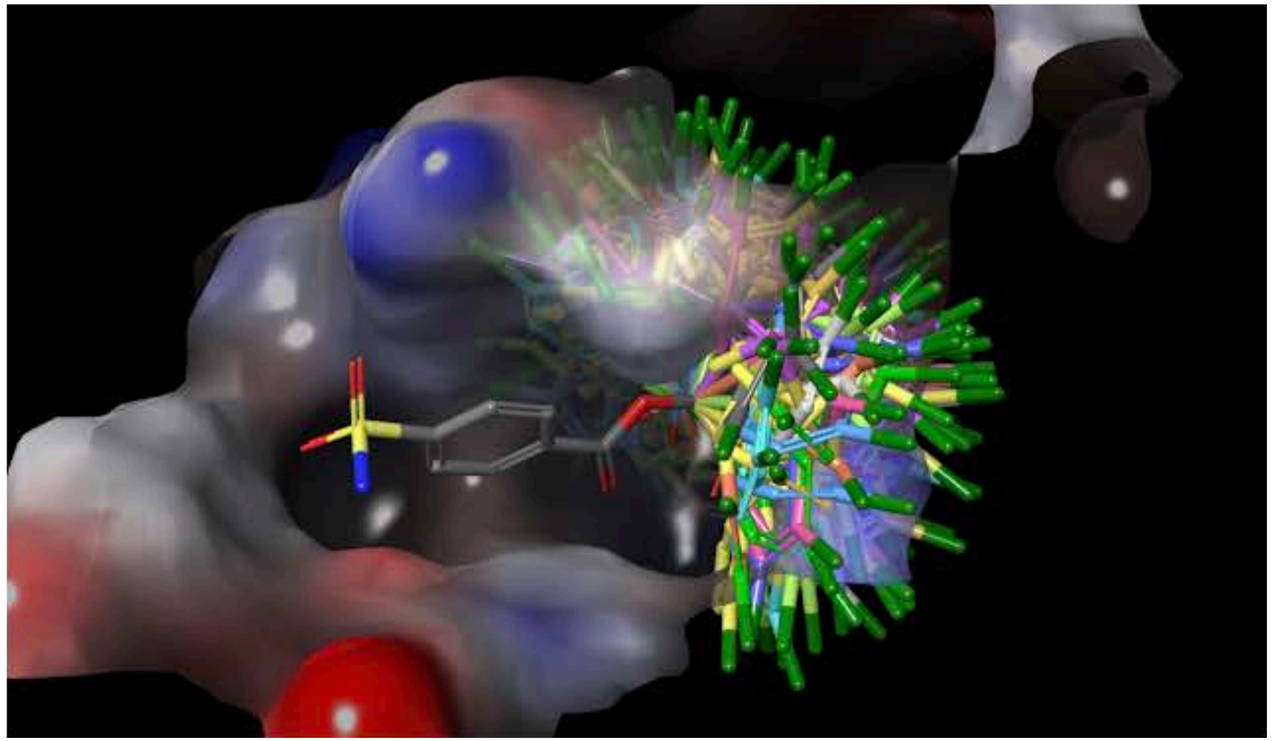
Common sense

Other Notes:

I'm sure it works, on a dish at least.

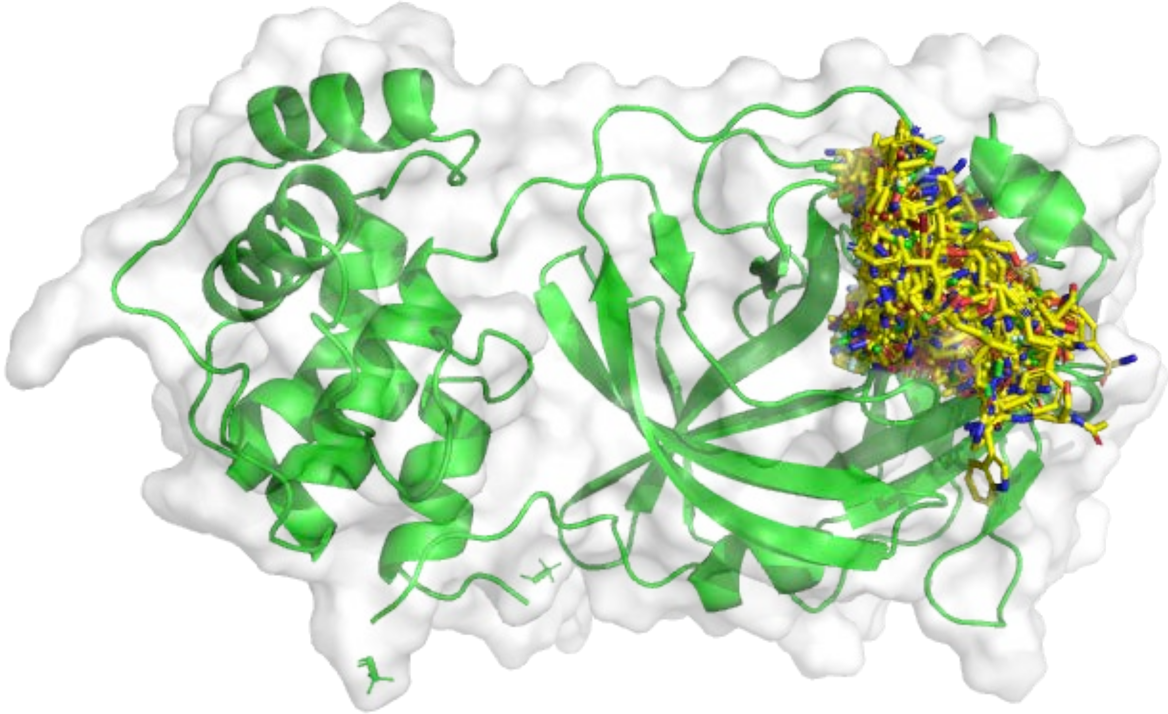
WE USED FAST PHYSICAL MODELS TO WEED OUT BAD IDEAS

docking of a single compound, showing all possible conformers



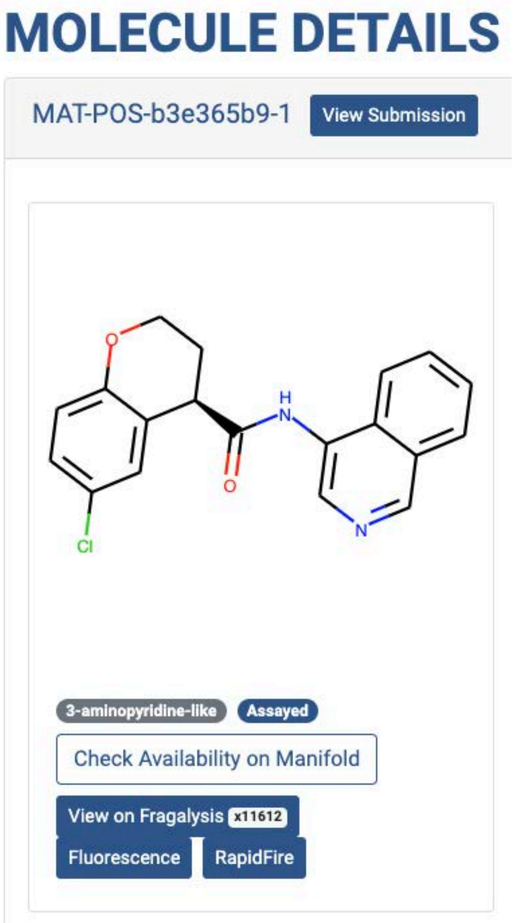
Pat Walters blog: http://practicalcheminformatics.blogspot.com

all final docked ligand structures



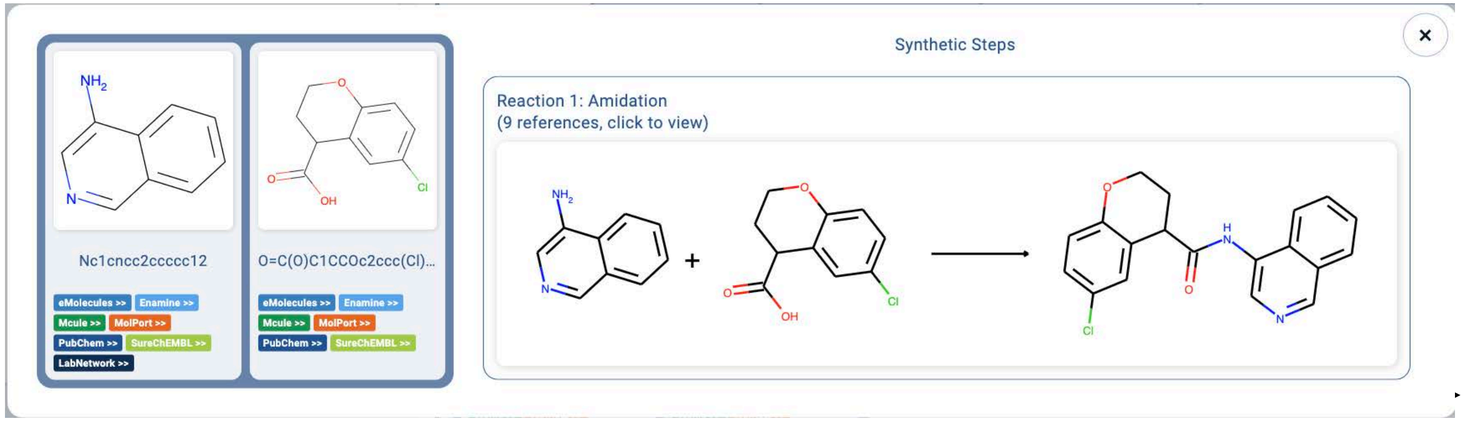
code and docking results: https://github.com/FoldingAtHome/covid-moonshot/tree/master/moonshot-submissions

MACHINE LEARNING BASED SYNTHETIC ROUTE PREDICTION MODELS IDENTIFIED DESIGNS THAT COULD BE EASILY SYNTHESIZED



http://postera.ai/covid

CRO catalogue-aware optimal synthetic route



Input: reactants-reagents (atom-wise tokenization)

Br c 1 c c c 2 ...c(c1)c1cc3c4ccccc4c4ccccc4c3cc1n2-c1ccc2c(c1)c1ccccc1n2-c1ccccc1.CCO.
Cc1ccccc1.OB(O)c1ccc2ccc3cccnc3c2n1.c1ccc([PH](c2ccccc2)(c2ccccc2)[Pd]([PH](c2ccccc2)
(c2ccccc2)c2ccccc2)([PH](c2ccccc2)(c2ccccc2)c2ccccc2)[PH](c2ccccc2)(c2ccccc2)c2ccccc2)cc1

c 1 c c c (– n
...2c3ccccc3c3cc(–n4c5ccc(–
c6ccc7ccc8cccnc8c7n6)cc5c5cc6c7ccccc7c7ccccc7c6cc54)ccc32)cc1

Target: most likely products

Molecular Transformer:
http://postera.ai/manifold

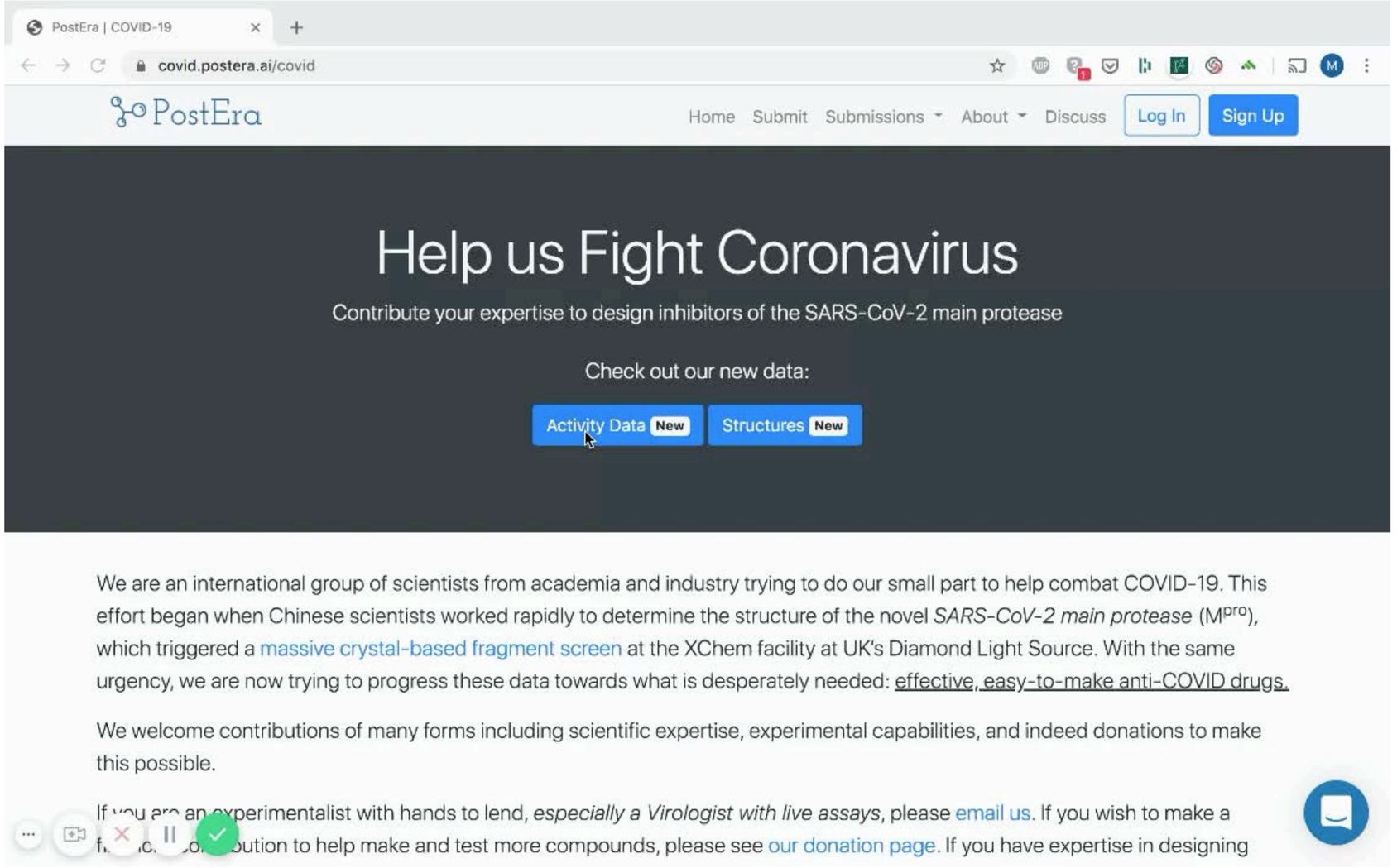
Contract Research Organizations (CROs)

Enamine
WuXi
Sai

Quickly made 850 compounds in a few weeks!

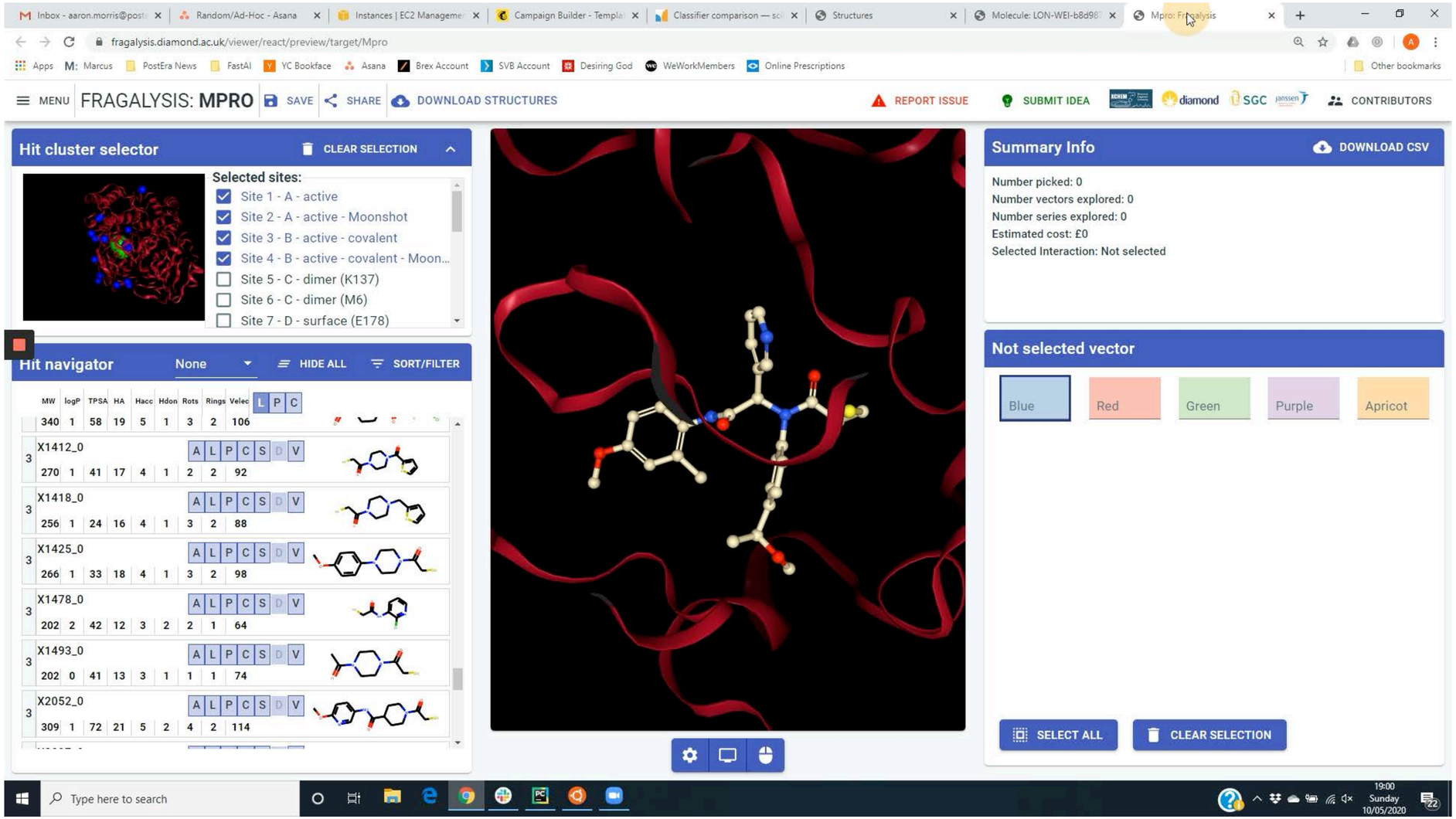
Schwaller et al. ACS Central Science 5:9, 2019
https://pubs.acs.org/doi/10.1021/acscentsci.9b00576

DATA WAS IMMEDIATELY REPORTED BACK TO THE COMMUNITY



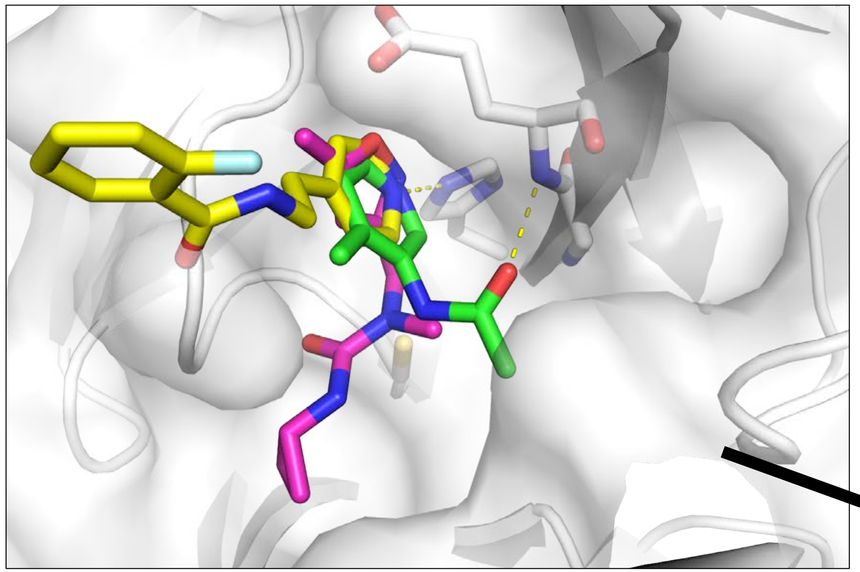
http://postera.ai/covid

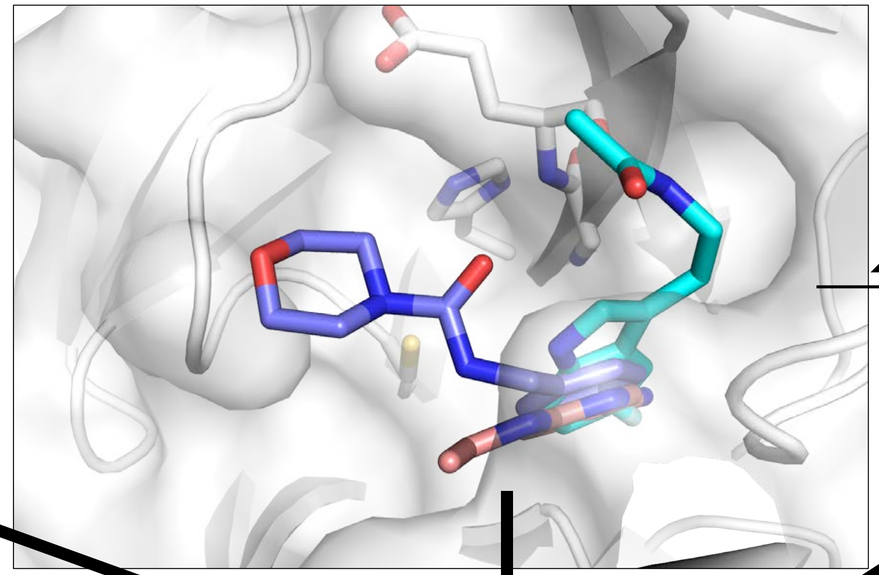
DIAMOND'S AUTOMATED BEAM LINE ENABLED US TO GENERATE STRUCTURAL DATA IN JUST DAYS

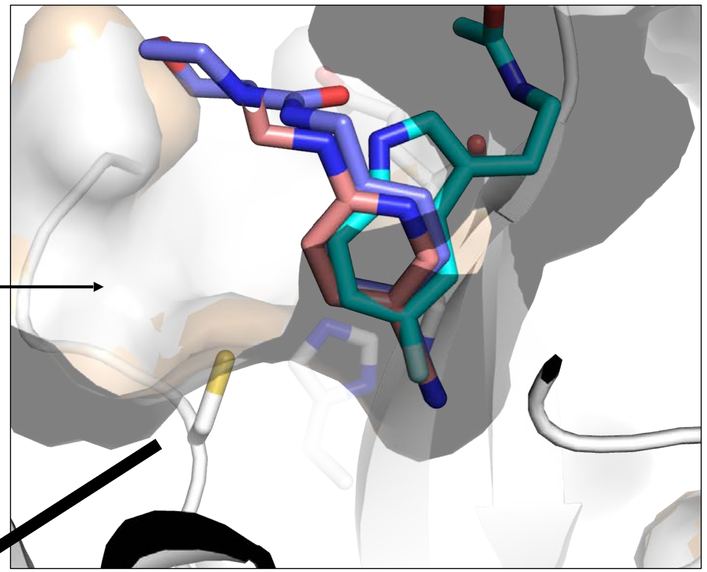


https://fragalysis.diamond.ac.uk/viewer/react/preview/target/Mpro

CROWDSOURCED DESIGNS GENERATED A NUMBER OF NOVEL CHEMICAL SERIES BY FRAGMENT MERGING





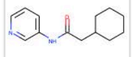


Contributor: Tryfon Zarganis - Tzitzikas, University of Oxford, TDI MedChem

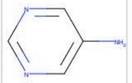
Design Rationale:

The design of the molecules was done by superimposing the different fragments from the crystal structures (by eye). The reactions should be fairly easy urea formation or amide coupling all from readily available starting materials. Fragments used for the conception of the ideas are the following x0107, x0434, x0678, x0748, x0995, x1382

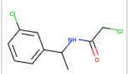
Inspired By:



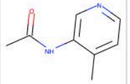
ALE-HEI-f28a35b5-9



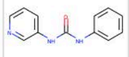
AAR-POS-d2a4d1df-18



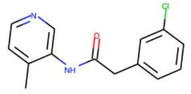
AAR-POS-0daf6b7e-10



MAK-UNK-6435e6c2-8



AAR-POS-d2a4d1df-11



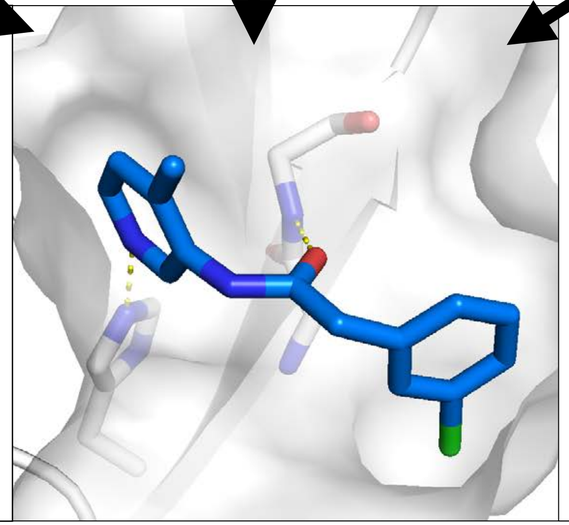
TRY-UNI-714a760b-6

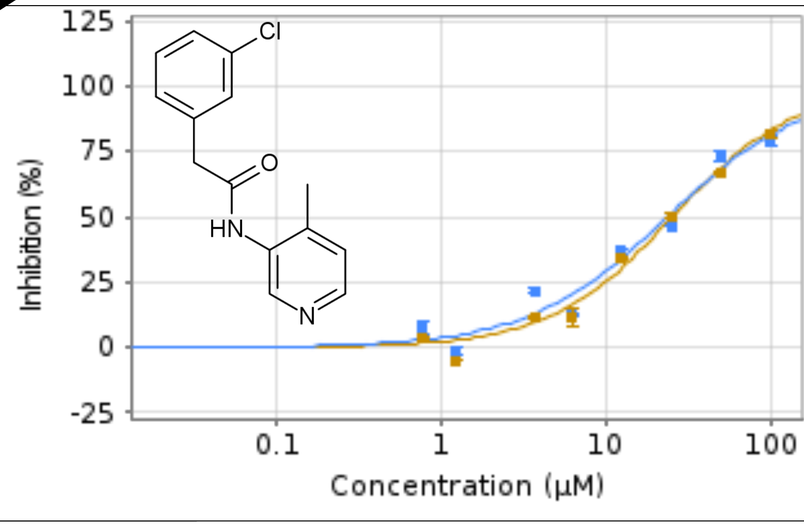
Cc1ccncc1NC(=O)Cc1cccc(Cl)c1

3-aminopyridine-like

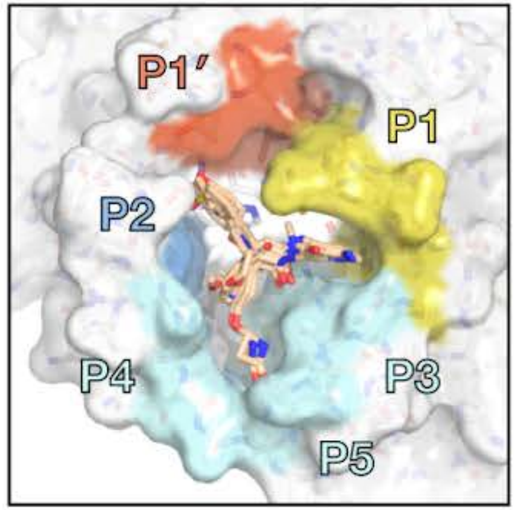
Enamine Mcule MolPort Assayed

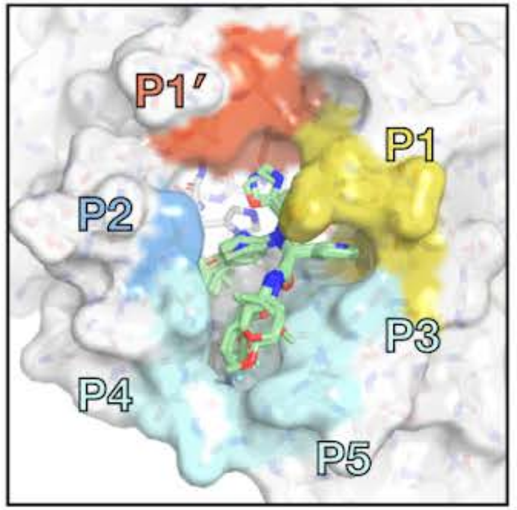
View

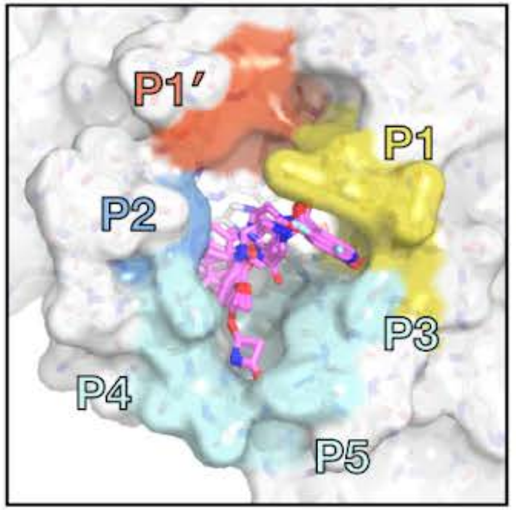


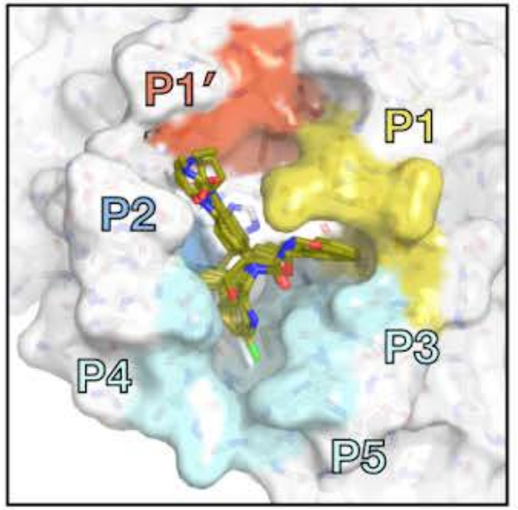


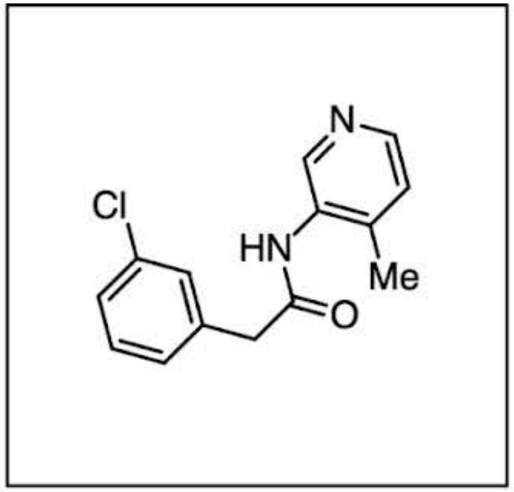
CROWDSOURCING GENERATED MULTIPLE LEADS WITH NOVEL NONCOVALENT CHEMOTYPES



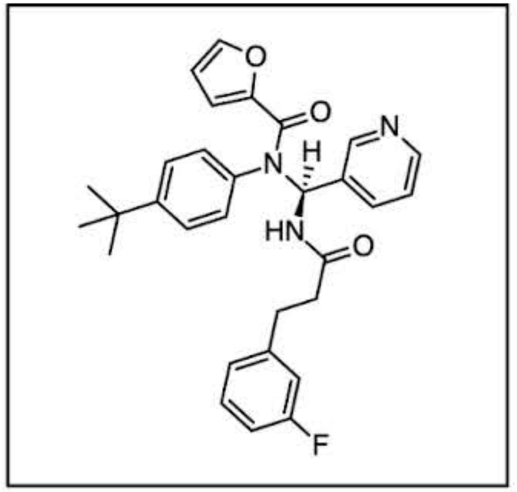




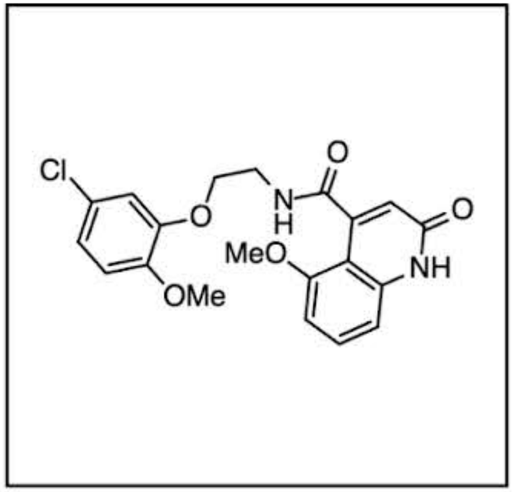




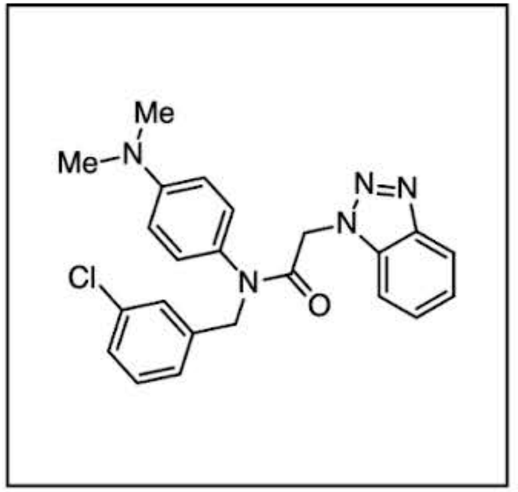
Aminopyridines



Ugis



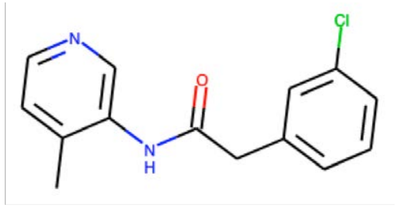
Quinolones



Benzotriazoles

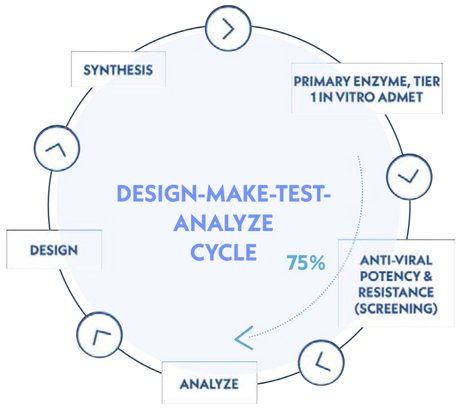
DESIGN-MAKE-TEST-ANALYZE CYCLES SHARE A COMMON OPERATION:

1. Select a current lead molecule

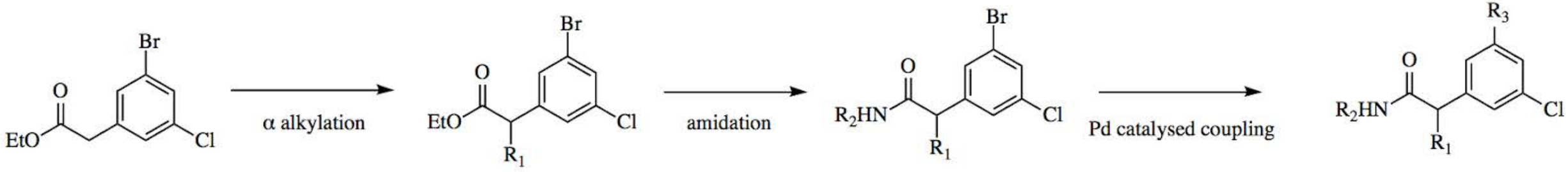


IC_{50} = 25 µM

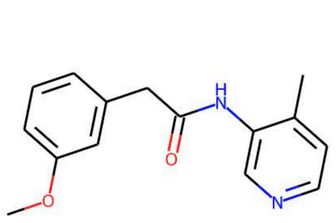
TRY-UNI-714a760b-6



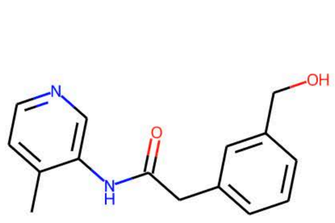
2. Select a retrosynthetic pathway capable of installing Enamine building blocks to replace part of the molecule



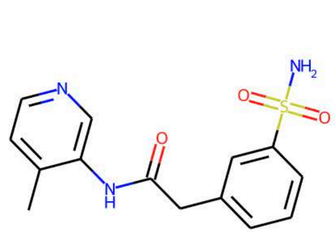
3. Chemists conservatively select analogues from the (often very) large enumerated synthetic space



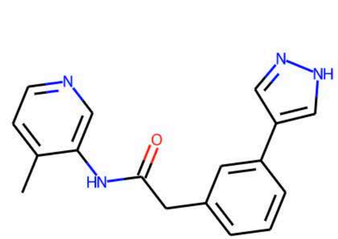
EDJ-MED-e58735b6-1



EDJ-MED-e58735b6-2

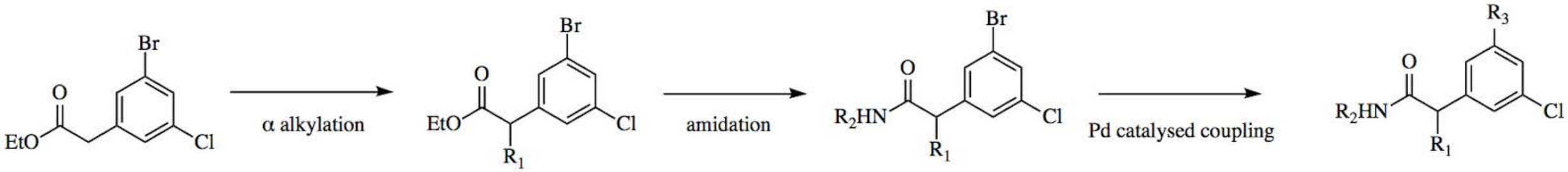


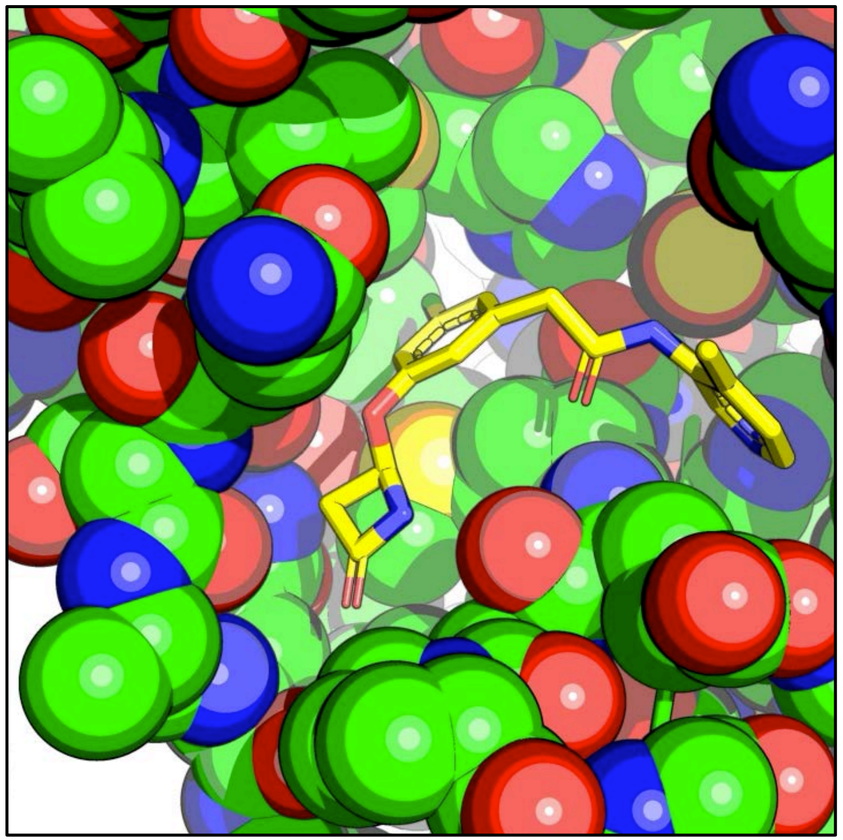
EDJ-MED-e58735b6-3



EDJ-MED-e58735b6-4

COULD WE USE FREE ENERGY CALCULATIONS TO ASSESS THESE DESIGNS AND FIND IDEAS THE CHEMISTS HAD OVERLOOKED?





~15,000
Potential
R3 groups

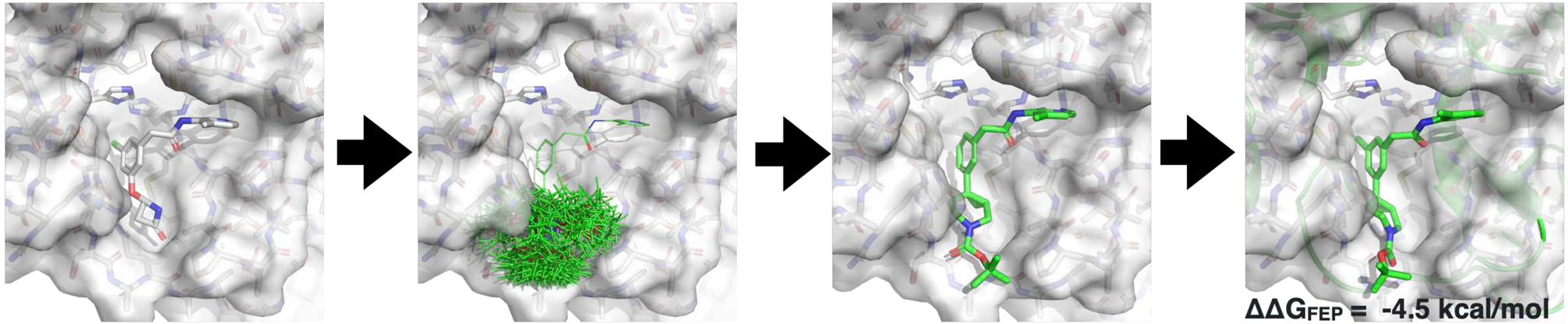
RETROSPECTIVE CALCULATIONS SUGGESTED OUR TOOLS DID A REASONABLY GOOD JOB OF PREDICTING WHICH COMPOUNDS WERE MORE POTENT

X-ray structure as reference

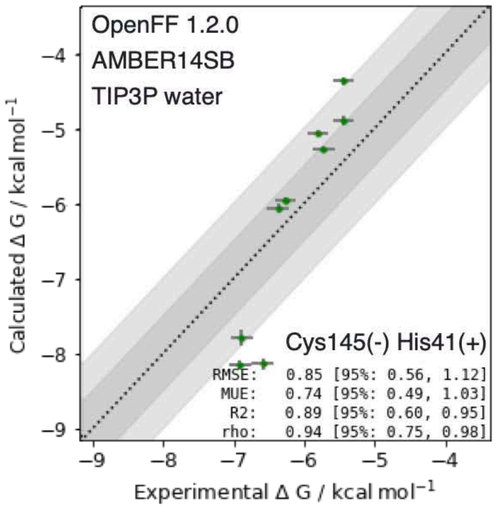
constrained enumeration of poses for proposed molecule

selection of pose with best docking score

free energy calculation final posed structure



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
Simple star maps

DOMINIC RUFA
TPCB student

HANNAH BRUCE MACDONALD
postdoc

WILLIAM GLASS
postdoc

OK, BUT WHERE DO WE GET ENOUGH GPUS?
OUR VIRTUAL LIBRARIES ARE > 15,000 COMPOUNDS!

OUR LAB HAD HAD STARTED TO USE FOLDING@HOME TO AID EXPERIMENTAL COLLABORATORS PURSUING COVID-19 DRUG DISCOVERY PROGRAMS

FOLDING @HOME

CHOOSE YOUR PLATFORM

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

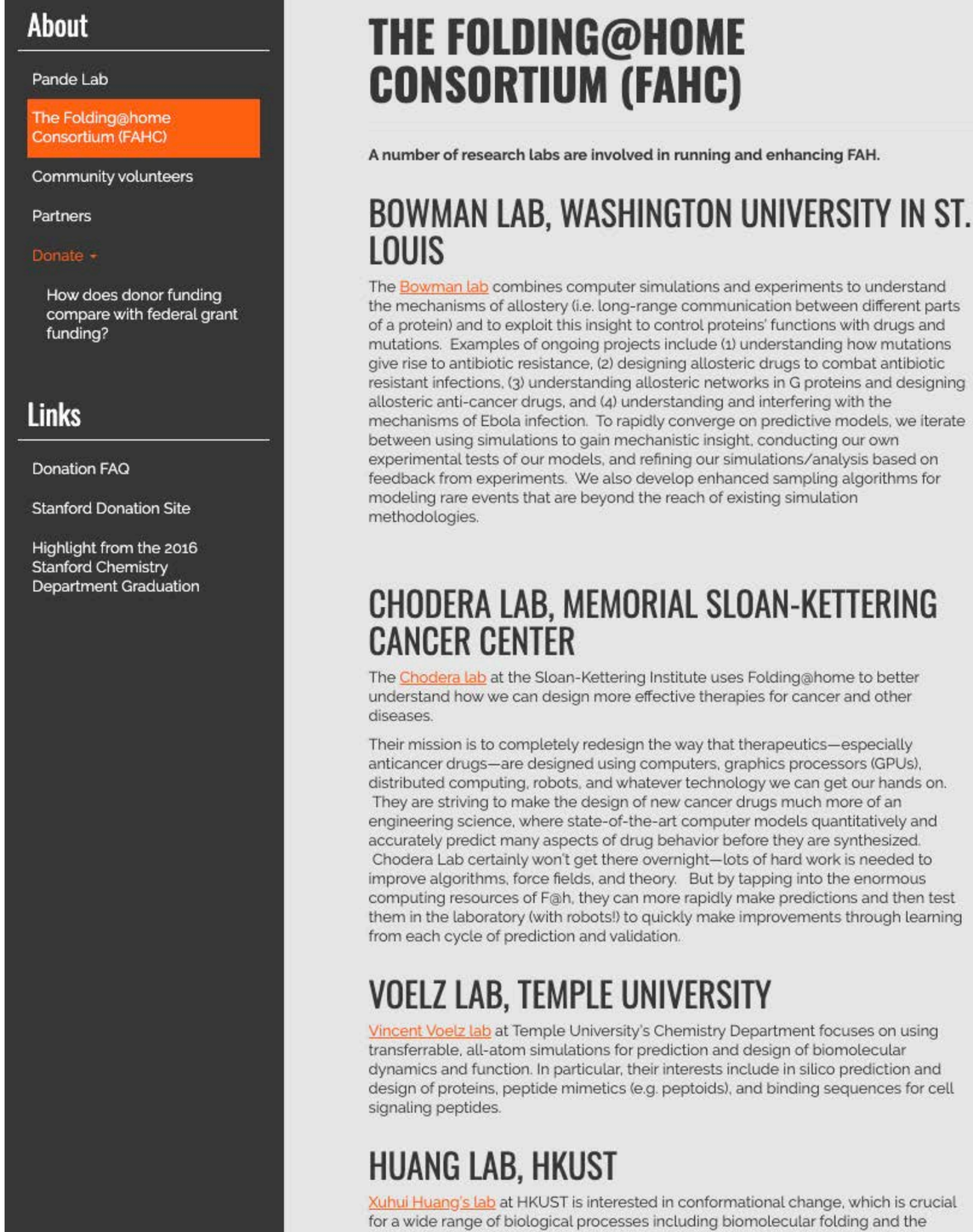
1924085 people have non-anonymously contributed to Folding@home.

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

* **probing mutations at the RBD:ACE2 interface** to optimize Ab therapeutics
* **free energy calculations** for prioritizing compounds tested by experimental collaborators
* **identifying cryptic pockets** for potential allosteric inhibition mechanisms
* **Simulating multiple targets** to understand their potential for drug discovery



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

AS PEOPLE FROM AROUND THE WORLD STARTED RUNNING FOLDING@HOME, WE QUICKLY CREATED THE WORLD'S FIRST EXASCALE COMPUTING RESOURCE

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

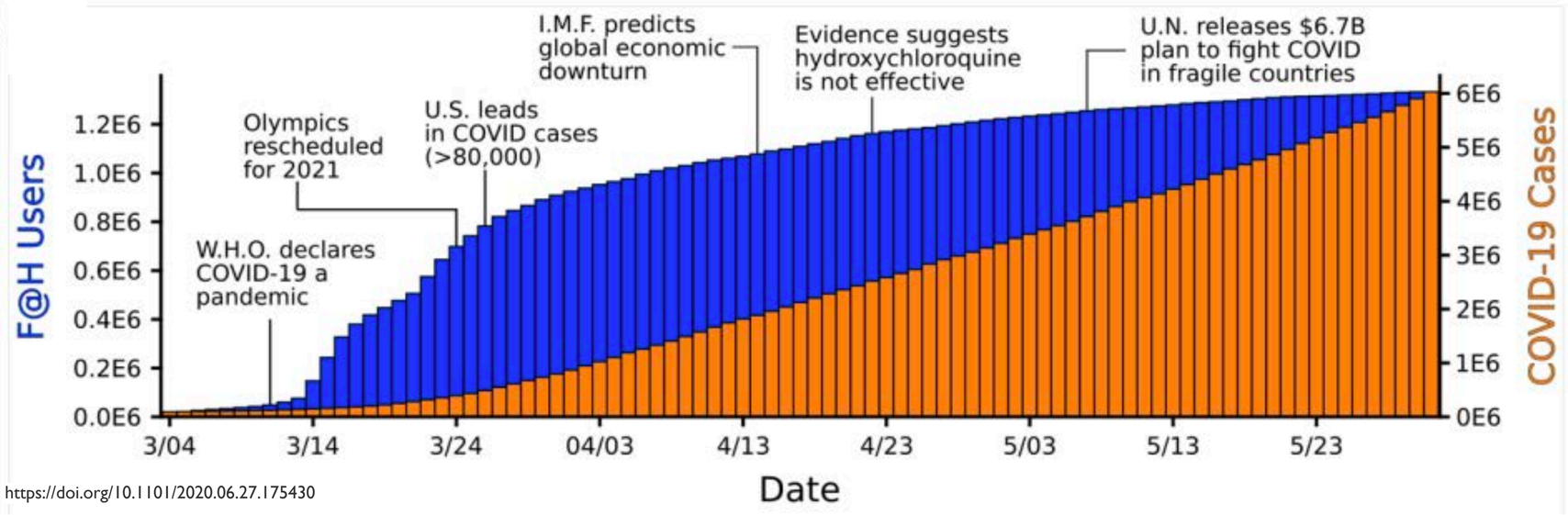
February 27, 2020
by Greg Bowman

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.



https://doi.org/10.1101/2020.06.27.175430

Ariana Brenner (CBM)

Rafal Wiewiora (TPCB)

Ivy Zhang (CBM)

THIS WAS AN ENORMOUS INCREASE IN COMPUTATIONAL POWER

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

DB date 2019-10-19 23:22:42

Active CPUS are defined as those which have returned WUs within 50 days. The FLOPS per core was last updated based on a FAH core performance report run on Wed May 11 11:56:35 PDT 2016.

*TFLOPS is the actual teraflops from the software cores, not the peak values from CPU/GPU specs. Please see our FAQ and FLOPS FAQ.

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

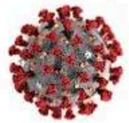
View Full Coverage

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

THIS WAS AN ENORMOUS INCREASE IN COMPUTATIONAL POWER

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

This would cost $6.8B/year on AWS.

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

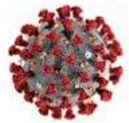
View Full Coverage

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



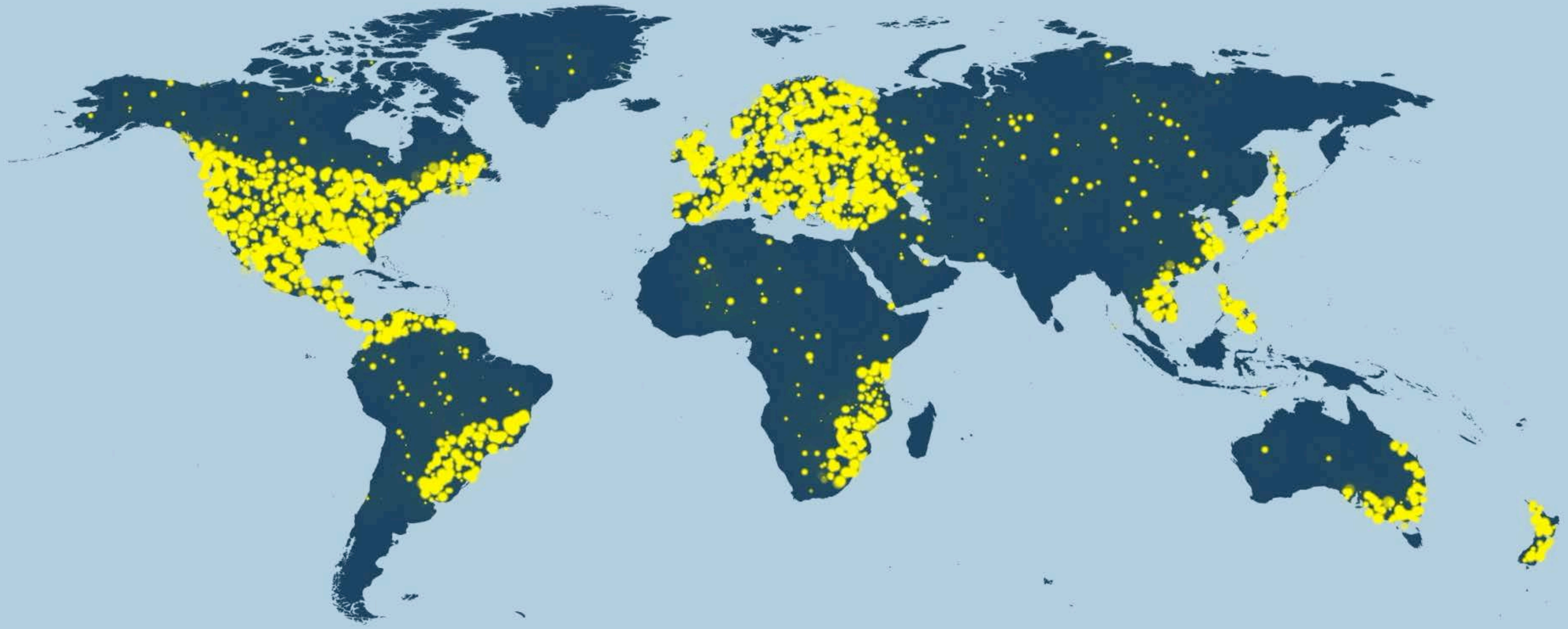
How to Fight Coronavirus With Folding@home and a Gaming PC

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



BOTH COMPUTING AND SCIENCE CONTRIBUTORS WERE TRULY GLOBAL

nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › articles › article

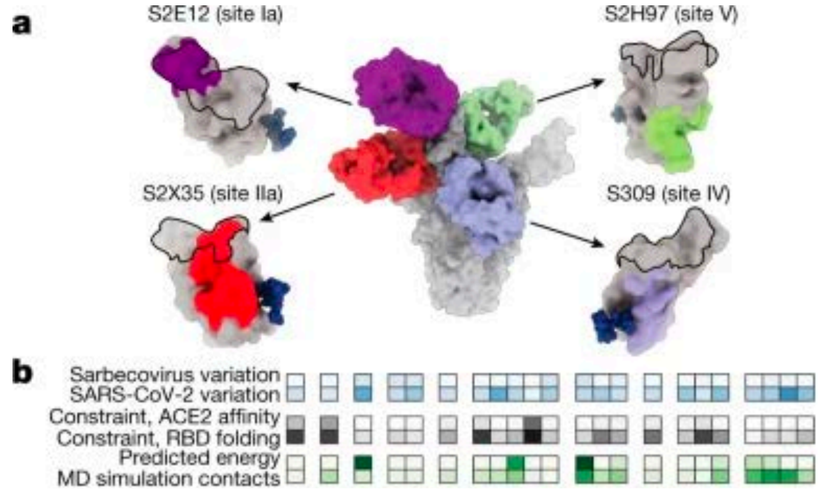
Article | Published: 14 July 2021

SARS-CoV-2 RBD antibodies that maximize breadth and resistance to escape

Tyler N. Starr, Nadine Czudnochowski, Zhuoming Liu, Fabrizia Zatta, Young-Jun Park, Amin Addetia, Dora Pinto, Martina Beltramello, Patrick Hernandez, Allison J. Greaney, Roberta Marzi, William G. Glass, Ivy Zhang, Adam S. Dingens, John E. Bowen, M. Alejandra Tortorici, Alexandra C. Walls, Jason A. Wojcechowskyj, Anna De Marco, Laura E. Rosen, Jiayi Zhou, Martin Montiel-Ruiz, Hannah Kaiser, Josh R. Dillen, Heather Tucker, Jessica Bassi, Chiara Silacci-Fregni, Michael P. Housley, Julia di Iulio, Gloria Lombardo, Maria Agostini, Nicole Sprugasci, Katja Culap, Stefano Jaconi, Marcel Meury, Exequiel Dellota Jr, Rana Abdelnabi, Shi-Yan Caroline Foo, Elisabetta Cameroni, Spencer Stumpf, Tristan I. Croll, Jay C. Nix, Colin Havenar-Daughton, Luca Piccoli, Fabio Benigni, Johan Neyts, Amalio Telenti, Florian A. Lempp, Matteo S. Pizzuto, John D. Chodera, Christy M. Hebner, Herbert W. Virgin, Sean P. J. Whelan, David Veesler, Davide Corti ✉, Jesse D. Bloom ✉ & Gyorgy Snell ✉ — Show fewer authors

Nature **597**, 97–102 (2021) | Cite this article

86k Accesses | **129** Citations | **526** Altmetric | Metrics



VIR

IVY ZHANG
CBM student

WILLIAM GLASS
postdoc

Cell

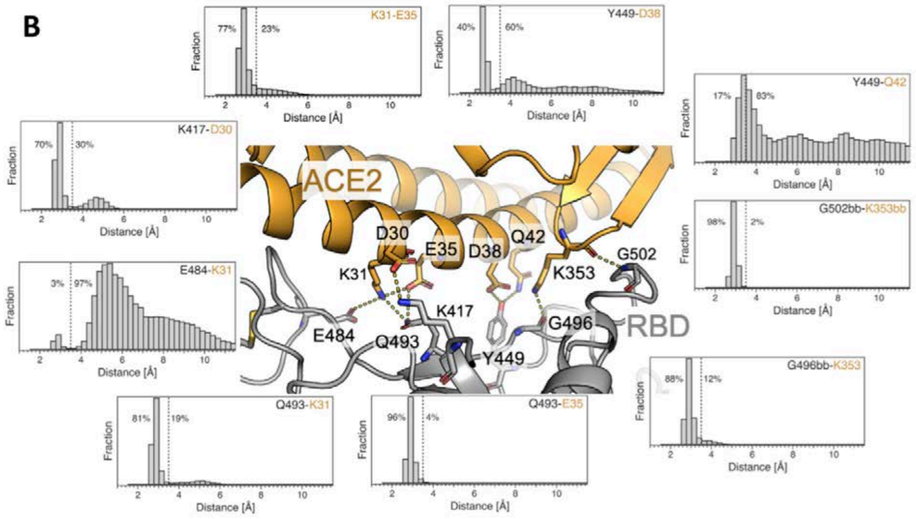
CellPress

Volume 184, Issue 5, 4 March 2021, Pages 1171-1187.e20

Article

Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity

Emma C. Thomson [1, 2, 29], Laura E. Rosen [3, 29], James G. Shepherd [1, 29], Roberto Spreafico [3, 29], Ana da Silva Filipe [1], Jason A. Wojcechowskyj [3], Chris Davis [1], Luca Piccoli [4], David J. Pascall [5], Josh Dillen [3], Spyros Lytras [1], Nadine Czudnochowski [3], Rajiv Shah [1], Marcel Meury [3], Natasha Jesudason [1], Anna De Marco [4], Kathy Li [1], Jessica Bassi [4], Aine O'Toole [6], Dora Pinto [4], Rachel M. Colquhoun [6], Katja Culap [4], Ben Jackson [6], Fabrizia Zatta [4], Andrew Rambaut [6], Stefano Jaconi [4], Vattipally B. Sreenu [1], Jay Nix [7], Ivy Zhang [8, 9], Ruth F. Jarrett [1], William G. Glass [8], Martina Beltramello [4], Kyriaki Nomikou [1], Matteo Pizzuto [4], Lily Tong [1], Elisabetta Cameroni [4], Tristan I. Croll [10], Natasha Johnson [1], Julia Di Iulio [3], Arthur Wickenhagen [1], Alessandro Ceschi [11, 12, 13], Aoife M. Harbison [14], Daniel Mair [1], Paolo Ferrari [15, 16], Katherine Smollett [1], Federica Sallusto [17, 18], Stephen Carmichael [1], Christian Garzoni [19], Jenna Nichols [1], Massimo Galli [20], Joseph Hughes [1], Agostino Riva [20], Antonia Ho [1], Marco Schiuma [20], Malcolm G. Semple [21, 22], Peter J.M. Openshaw [23], Elisa Fadda [14], J. Kenneth Baillie [24, 25], John D. Chodera [8], The ISARIC4C Investigators [26], the COVID-19 Genomics UK (COG-UK) Consortium [27], Suzannah J. Rihn [1], Samantha J. Lycett [24], Herbert W. Virgin [3, 28], Amalio Telenti [3], Davide Corti [4], David L. Robertson [1] ✉, Gyorgy Snell [3, 30] ✉



OK, WE HAVE COMPUTING RESOURCES NOW...

...BUT WE HAD NEVER RUN FREE ENERGY CALCULATIONS ON FOLDING@HOME

ALCHEMICAL FREE ENERGY CALCULATIONS GENERALLY USE CLEVER BUT COMPLEX MARKOV CHAIN MONTE CARLO ALGORITHMS TO SAMPLE ALCHEMICAL STATES

Independent simulations

Easy to parallelize, but sampling problems at any λ can make calculations unreliable

simple but too dangerous to use

Hamiltonian replica exchange ★

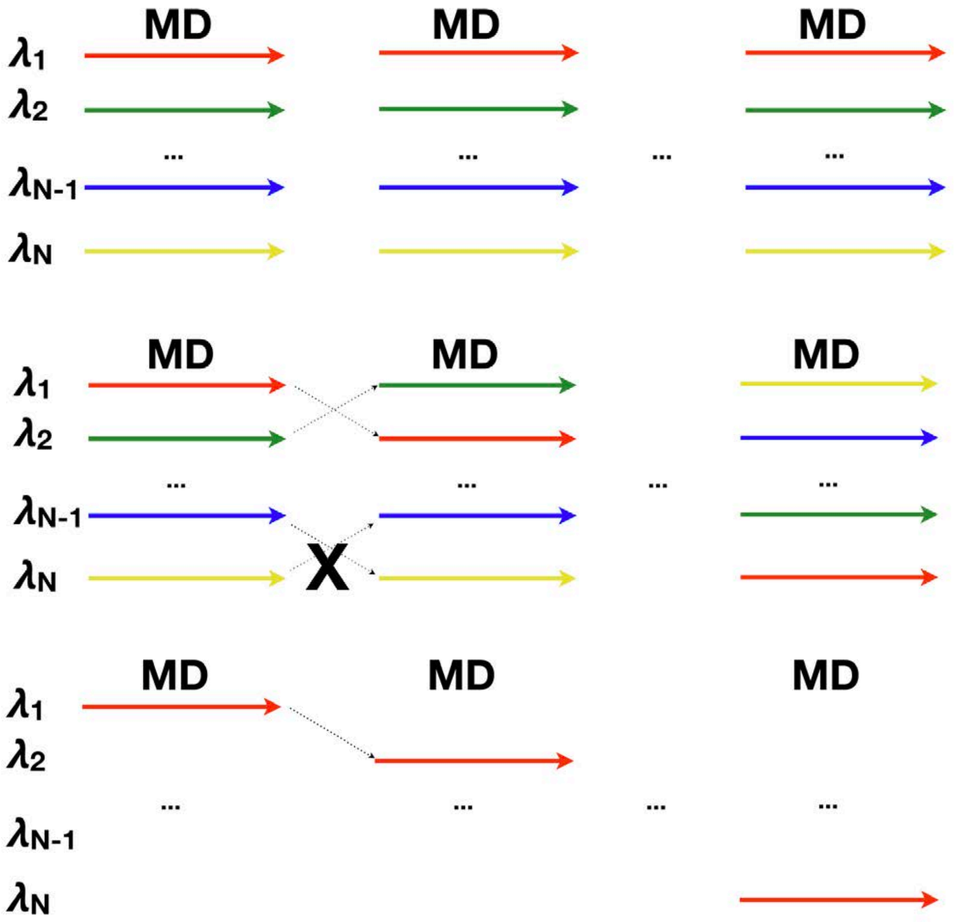
Good sampling at any λ can rescue problems at other λ if good λ overlap

reliable but complex for Folding@home

Single-replica methods

For certainly problems, can converge extremely quickly in a fraction of computer effort; tricky to make reliable

immature and tricky to implement



Replica exchange and expanded ensemble simulations as Gibbs sampling: Simple improvements for enhanced mixing

J. Chem. Phys. **135**, 194110 (2011); https://doi.org/10.1063/1.3660669

John D. Chodera[1, a)] *and* Michael R. Shirts[2, b)]

Excursions in Statistical Dynamics

by

Gavin Earl Crooks

B.Sc. (University of East Anglia) 1992
M.Sc. (University of East Anglia) 1993

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Chemistry

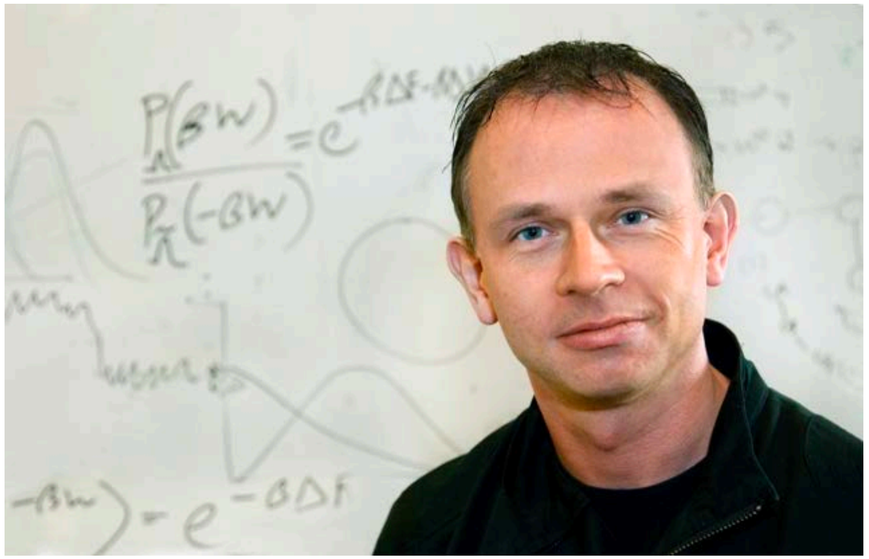
in the

GRADUATE DIVISION
of the
UNIVERSITY of CALIFORNIA at BERKELEY

Committee in charge:

Professor David Chandler, Chair
Professor Robert A. Harris
Professor Daniel S. Rokhsar

1999



JOURNAL OF COMPUTATIONAL PHYSICS 22, 245–268 (1976)

Efficient Estimation of Free Energy Differences from Monte Carlo Data

CHARLES H. BENNETT

IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598

Received February 13, 1976; accepted May 3, 1976

Near-optimal strategies are developed for estimating the free energy difference between two canonical ensembles, given a Metropolis-type Monte Carlo program for sampling each one. The estimation strategy depends on the extent of overlap between the two ensembles, on the smoothness of the density-of-states as a function of the difference potential, and on the relative Monte Carlo sampling costs, per statistically independent data point. The best estimate of the free energy difference is usually obtained by dividing the available computer time approximately equally between the two ensembles; its efficiency (variance × computer time)$^{-1}$ is never less, and may be several orders of magnitude greater, than that obtained by sampling only one ensemble, as is done in perturbation theory.

I. INTRODUCTION

A well-known deficiency of the Monte Carlo [1, 2] and molecular dynamics [3] methods, commonly used to study the thermodynamic properties of classical systems having 10^2 to 10^4 degrees of freedom, is their inability to calculate quantities such as the entropy or free energy, which cannot be expressed as canonical or microcanonical ensemble averages. In general, the free energy of a Monte Carlo (MC) or molecular dynamics (MD) system can be determined only by a procedure analogous to calorimetry, i.e., by establishing a reversible path between the system of interest and some reference system of known free energy. "Computer calorimetry" has a considerable advantage over laboratory calorimetry in that the reference system may differ from the system of interest not only in its thermodynamic state variables but also in its Hamiltonian, thereby making possible a much wider variety of reference systems and reversible paths. Often the path between an analytically tractable reference system and the system of ultimate physical interest will include one or more intermediate systems. These may be interesting in their own right (e.g., the hard sphere fluid), or they may be special systems, important only as calorimetric stepping stones, whose Hamiltonians contain artificial terms designed to stabilize the system against phase transitions [4, 5], induce favorable importance weighting [6, 7], or otherwise enhance the system's efficiency as a computational tool [8–10].

Excursions in Statistical Dynamics

by

Gavin Earl Crooks

B.Sc. (University of East Anglia) 1992
M.Sc. (University of East Anglia) 1993

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Chemistry

in the

GRADUATE DIVISION
of the
UNIVERSITY of CALIFORNIA at BERKELEY

Committee in charge:

Professor David Chandler, Chair
Professor Robert A. Harris
Professor Daniel S. Rokhsar

1999

Vol 437|8 September 2005|doi:10.1038/nature04061 nature

LETTERS

Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies

D. Collin[1]*, F. Ritort[2]*, C. Jarzynski[3], S. B. Smith[4], I. Tinoco Jr[5] & C. Bustamante[4,6]

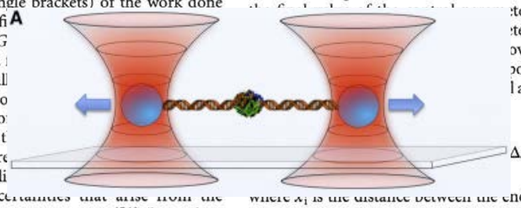
Atomic force microscopes and optical tweezers are widely used to probe the mechanical properties of individual molecules and molecular interactions, by exerting mechanical forces that induce transitions such as unfolding or dissociation. These transitions often occur under nonequilibrium conditions and are associated with hysteresis effects—features usually taken to preclude the extraction of equilibrium information from the experimental data. But fluctuation theorems[1–5] allow us to relate the work along nonequilibrium trajectories to thermodynamic free-energy differences. They have been shown to be applicable to single-molecule force measurements[6] and have already provided information on the folding free energy of a RNA hairpin[7,8]. Here we show that the Crooks fluctuation theorem[9] can be used to determine folding free energies for folding and unfolding processes occurring in weak as well as strong nonequilibrium regimes, thereby providing a test of its validity under such conditions. We use optical tweezers[10] to measure repeatedly the mechanical work associated with the unfolding and refolding of a small RNA hairpin[11] and an RNA three-helix junction[12]. The resultant work distributions are then analysed according to the theorem and allow us to determine the difference in folding free energy between an RNA molecule and a mutant differing only by one base pair, and the thermodynamic stabilizing effect of magnesium ions on the RNA structure.

The Crooks fluctuation theorem[9] (CFT) predicts a symmetry relation in the work fluctuations associated with the forward and reverse changes a system undergoes as it is driven away from thermal equilibrium by the action of an external perturbation. This theorem applies to processes that are microscopically reversible, and its experimental evaluation in small systems is crucial to understand better the foundations of nonequilibrium physics[13]. A consequence of the CFT is Jarzynski's equality[14], which relates the equilibrium free-energy difference ΔG between two equilibrium states to an exponential average (denoted by angle brackets) of the work done on the system, W, taken over an infi
quilibrium experiments, $\exp(-\Delta G$
equality has been developed[6] into a
nonequilibrium single-molecule pul
free-energy profiles or potentials o
coordinates. Experimental testing o
molecule force experiments[16] used t
can be folded and unfolded quasi-r
occur far from equilibrium, the appli
hampered by large statistical unc
sensitivity of the exponential average to rare events[17,18] (low values of W). Moreover, although the equality $\langle W \rangle = \Delta G$ holds for processes occurring near equilibrium, spatial drift in the experimental system usually makes it difficult in practice to extract unfolding free energies using small loading rates (below a few pN s^{-1}). Drift effects decrease noticeably for larger pulling speeds, making it possible to obtain more reliable experimental data (and also good statistics as a large number of pulls can be executed in a reasonable time), but at the expense of a more irreversible unfolding process. Here we show that significant improvements can be obtained by using the CFT, which provides a more robust and more rapidly converging method to extract equilibrium free energies from non-equilibrium processes.

The CFT allows us to quantify the amount of hysteresis observed in the values of the irreversible work done to unfold and refold a macromolecule. Let $P_U(W)$ denote the probability distribution of the values of the work performed on the molecule in an infinite number of pulling experiments along the unfolding (U) process, and define $P_R(W)$ analogously for the reverse (R) process. For the CFT to be applicable, the unfolding and refolding processes need to be related by time-reversal symmetry, that is, in our experiments, the optical trap used to manipulate the molecule must be moved at the same speeds during unfolding and refolding. Moreover, the molecular transition probed always has to start in an equilibrium state (folded in the unfolding process, and denatured or unfolded in the refolding process) and reach a well-defined final state. The CFT[9] then predicts that:

$$\frac{P_U(W)}{P_R(-W)} = \exp\left(\frac{W - \Delta G}{k_B T}\right) \quad (1)$$

where ΔG is the free-energy change between the final and the initial states, and thus equal to the reversible work associated with this process. Note that the CFT does not require that the system studied reaches its final equilibrium state immediately after the unfolding and refolding processes have been completed; it is only the control parameter that needs to attain its final value, whereas the system may continue to equilibrate to a well-defined state that is consistent with
ter. The equilibration occurs
ter, and therefore contributes
over the external variation of a
osition of the optical trap or
l approximated by the familiar



$$\Delta x_i \quad (2)$$

where x_i is the distance between the ends of the molecule and N_s is the number of intervals used in the sum (see ref. 6 for a thorough discussion of this issue). Relation (1) quantifies hysteretic effects in the pulling experiment: work values larger than ΔG occur most often

[1]Merck & Co. Inc., Automated Biotechnology Department, North Wales, Pennsylvania 19454, USA. [2]Departament de Física Fonamental, Facultat de Física, Universitat de Barcelona, Diagonal 647, 08028 Barcelona, Spain. [3]T-13 Complex Systems, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA. [4]Howard Hughes Medical Institute, [5]Department of Chemistry, [6]Departments of Physics and Molecular & Cell Biology, University of California, Berkeley, California 94720, USA.
*These authors contributed equally to this work.

F COMPUTATIONAL PHYSICS 22, 245–268 (1976)

Estimation of Free Energy Differences from Monte Carlo Data

CHARLES H. BENNETT

BM Thomas J. Watson Research Center, Yorktown Heights, New York 10598

Received February 13, 1976; accepted May 3, 1976

r-optimal strategies are developed for estimating the free energy difference between nonical ensembles, given a Metropolis-type Monte Carlo program for sampling ne. The estimation strategy depends on the extent of overlap between the two bles, on the smoothness of the density-of-states as a function of the difference ial, and on the relative Monte Carlo sampling costs, per statistically independent oint. The best estimate of the free energy difference is usually obtained by dividing vailable computer time approximately equally between the two ensembles; its ncy (variance × computer time)$^{-1}$ is never less, and may be several orders of tude greater, than that obtained by sampling only one ensemble, as is done in bation theory.

I. INTRODUCTION

l-known deficiency of the Monte Carlo [1, 2] and molecular dynamics ods, commonly used to study the thermodynamic properties of classical aving 10^2 to 10^4 degrees of freedom, is their inability to calculate quantities the entropy or free energy, which cannot be expressed as canonical or onical ensemble averages. In general, the free energy of a Monte Carlo molecular dynamics (MD) system can be determined only by a procedure s to calorimetry, i.e., by establishing a reversible path between the f interest and some reference system of known free energy. "Computer try" has a considerable advantage over laboratory calorimetry in that the system may differ from the system of interest not only in its thermo- state variables but also in its Hamiltonian, thereby making possible a ider variety of reference systems and reversible paths. Often the path an analytically tractable reference system and the system of ultimate interest will include one or more intermediate systems. These may be g in their own right (e.g., the hard sphere fluid), or they may be special important only as calorimetric stepping stones, whose Hamiltonians artificial terms designed to stabilize the system against phase transitions duce favorable importance weighting [6, 7], or otherwise enhance the efficiency as a computational tool [8–10].

WE KNOW HOW TO OPTIMIZE NONEQUILIBRIUM PROTOCOLS!

The **thermodynamic metric tensor** measures how rapidly the equilibrium distribution changes as control parameters are twiddled.

$$g_{ij}(\boldsymbol{\lambda}) \equiv \left\langle \frac{\partial \ln \pi(\boldsymbol{x};\boldsymbol{\lambda})}{\partial \lambda_i} \frac{\partial \ln \pi(\boldsymbol{x};\boldsymbol{\lambda})}{\partial \lambda_j} \right\rangle_{\boldsymbol{\lambda}}$$

The **thermodynamic length** measures how much the distribution has changes from one value of control parameters to another. (Can also integrate effects of correlation time)

$$\mathcal{L} \equiv \int_0^\tau dt \left(\dot{\boldsymbol{\lambda}}^{\mathrm{T}} \boldsymbol{g} \dot{\boldsymbol{\lambda}} \right)^{1/2}$$

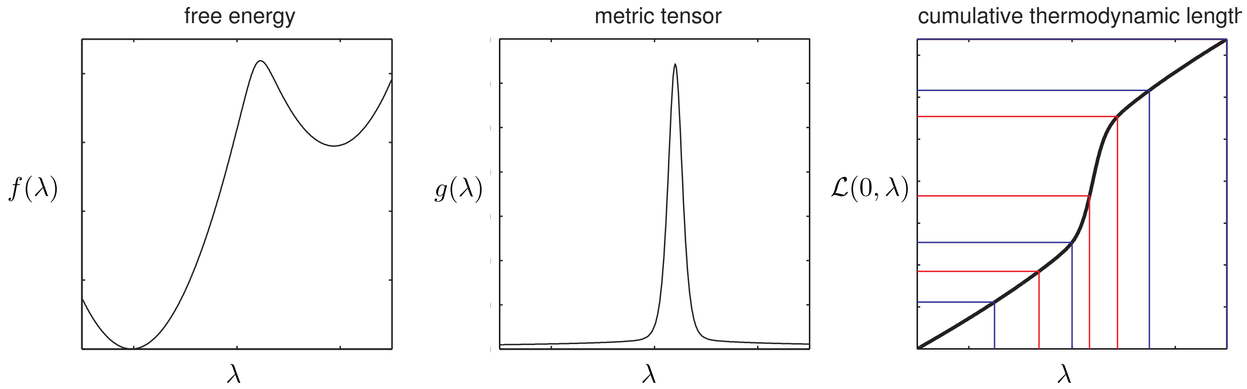
.

Optimal protocols are geodesics in thermodynamic metric space; they equalize thermodynamic length between measurements.

$$\mathrm{var}(\Delta \hat{f}) \geq N^{-1} \mathcal{L}(\lambda_a, \lambda_b)^2$$

The **efficiency of a transformation** is related to how much effort is needed to achieve a given target variance ε. For the same amount of computer effort, we can estimate it via a ratio of variances:

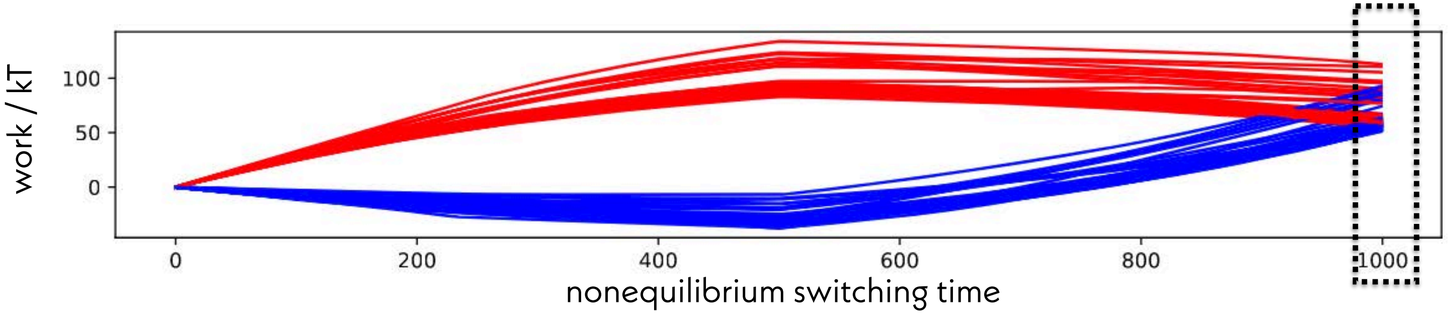
$$E = \frac{\mathrm{var}_1^{-1}(\Delta f)}{\mathrm{var}_2^{-1}(\Delta \hat{f})}$$



optimized paths can yield orders of magnitude reduction in variance!

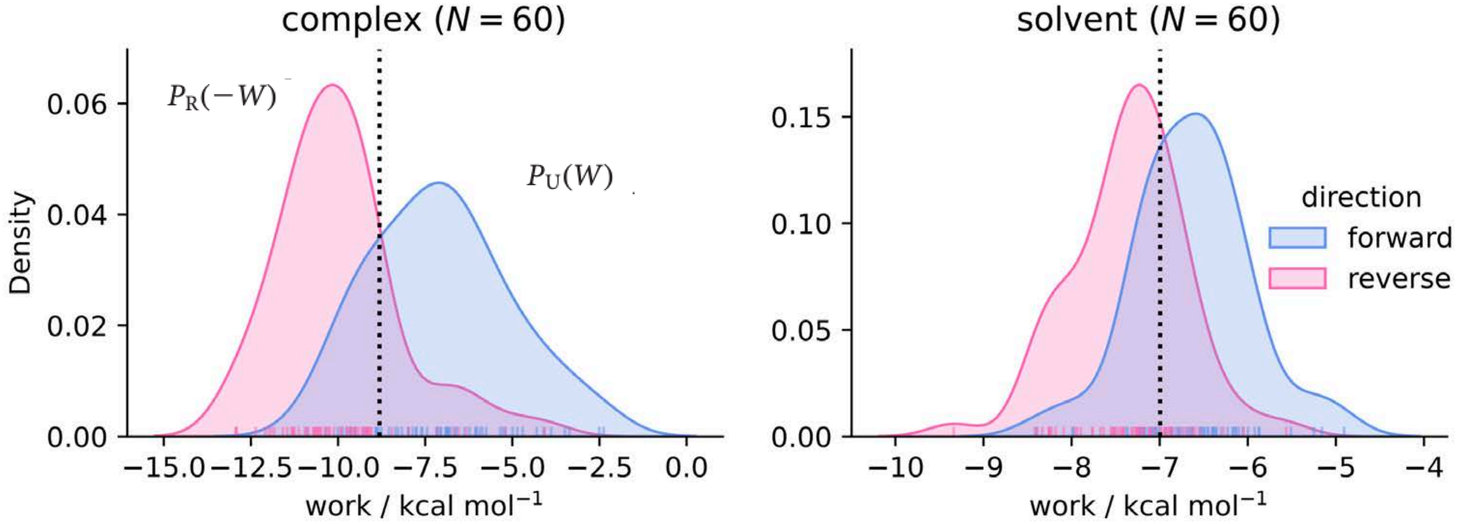
Crooks, PRL 99:100602, 2007; Shenfeld et al. PRE 80:046705, 2009; Sivak and Crooks PRL 108:190602, 2012.
Nilmeier, Crooks, Minh, Chodera PNAS 1008:E1009, 2011

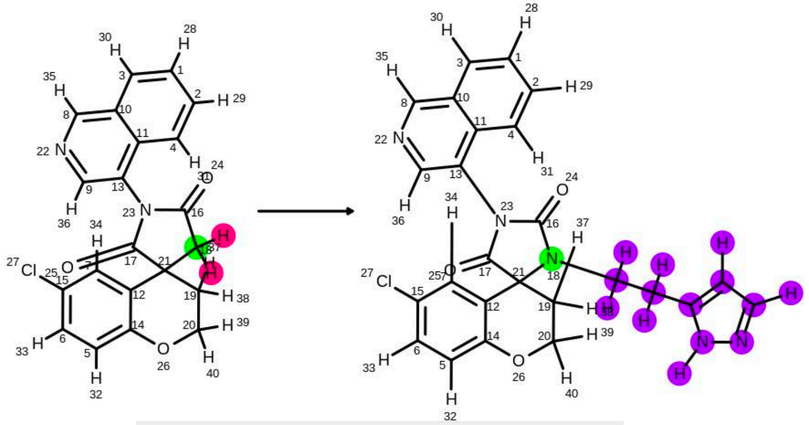
NONEQUILIBRIUM CYCLING CAN EASILY BE RUN IN PARALLEL DISTRIBUTED COMPUTING ENVIRONMENTS



$$\frac{P_{\mathrm{U}}(W)}{P_{\mathrm{R}}(-W)} = \exp\left(\frac{W - \Delta G}{k_{\mathrm{B}}T}\right)$$

Crooks fluctuation theorem





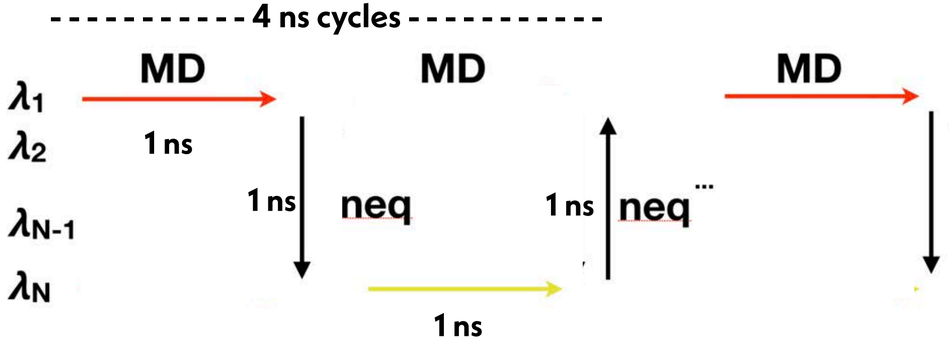
ΔΔG = –1.6 ± 0.1 kcal/mol

(from Bennett acceptance ratio)

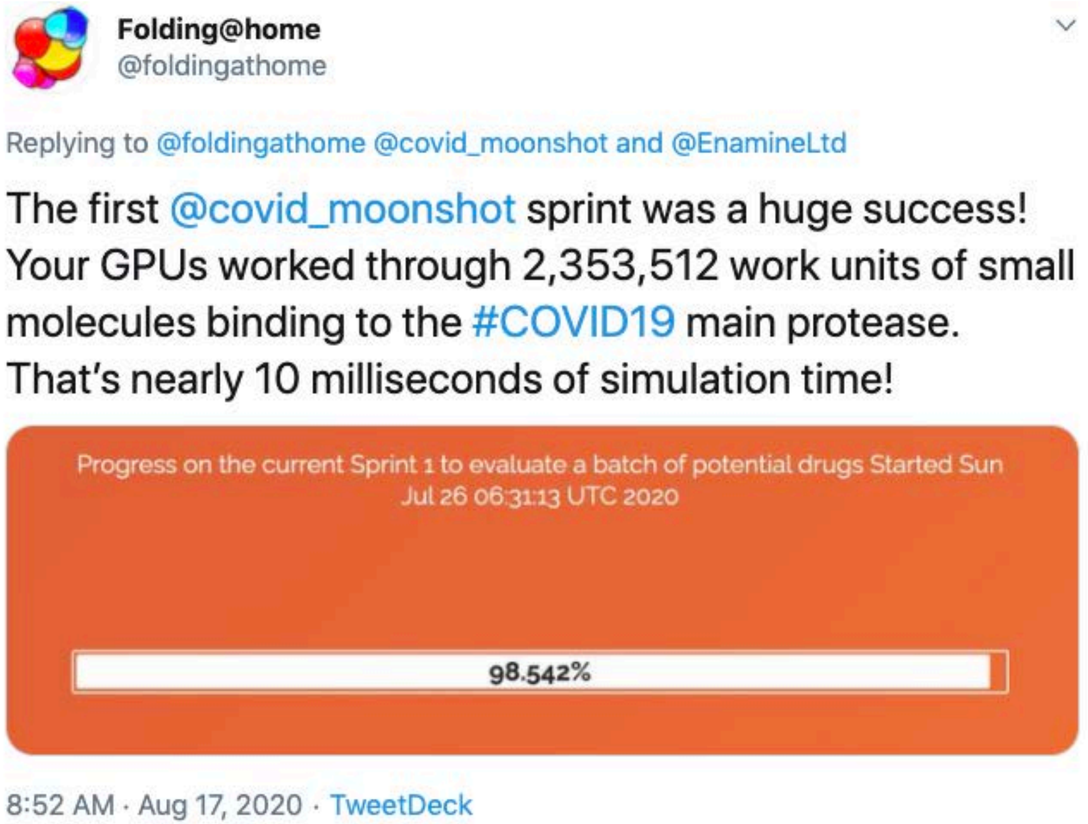
Nonequilibrium cycling

Can approximate nonequilibrium switching if relaxation is fast (or restraints are used to limit motion)

a terrible hack, but it just might work



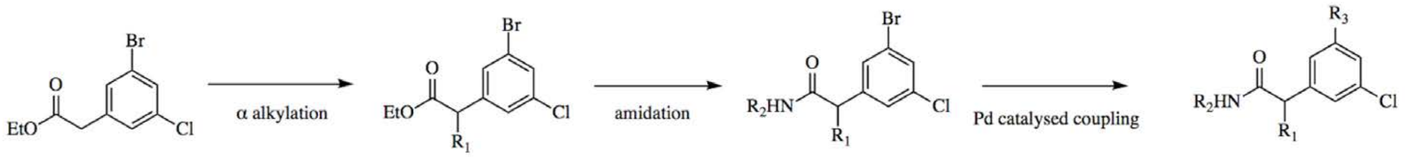
We generated a *lot* of data, which we have shared online via AWS



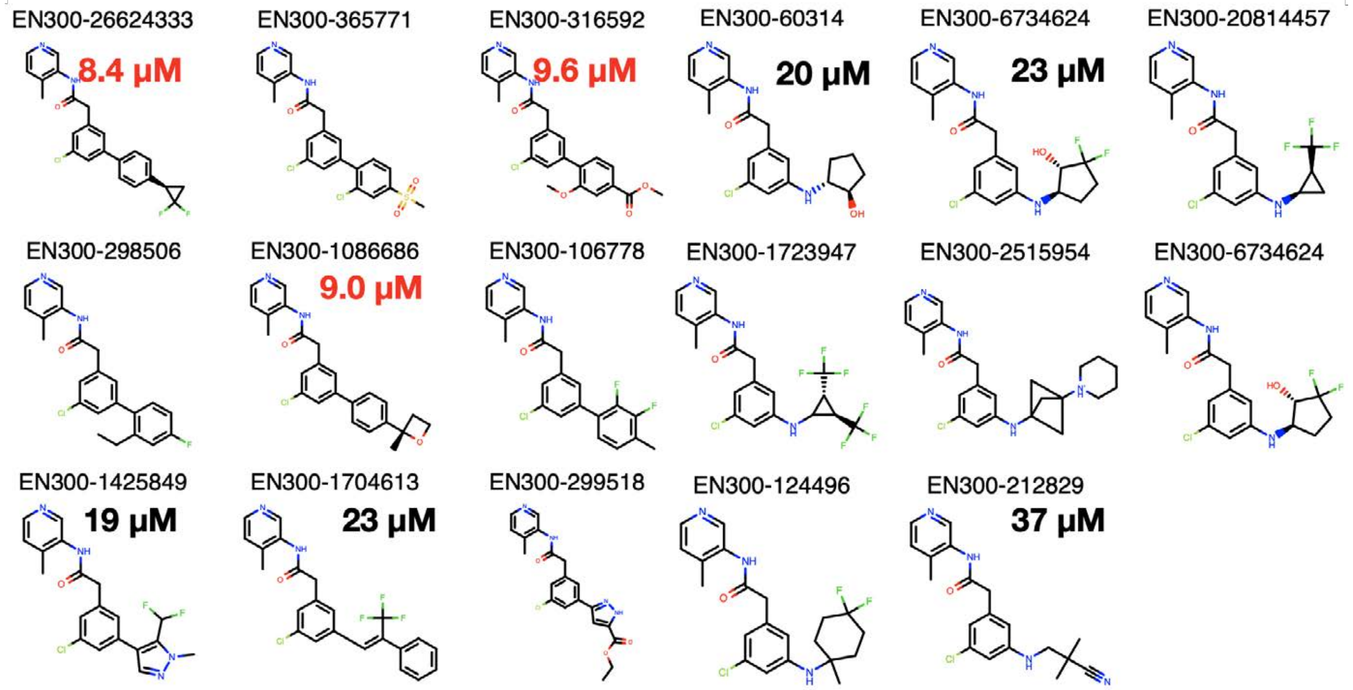
https://registry.opendata.aws/foldingathome-covid19/
https://covid.molssi.org//org-contributions/#folding--home

EVEN LARGE TRANSFORMATIONS WERE SUCCESSFUL IN IDENTIFYING MORE POTENT COMPOUNDS

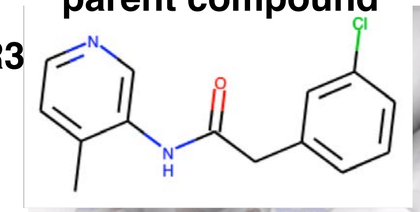
Can we engage S4 from this 5,000-compound virtual synthetic library varying R3



Top free energy calculation compounds and experimental affinity measurements:



parent compound



IC_{50} = 25 μM

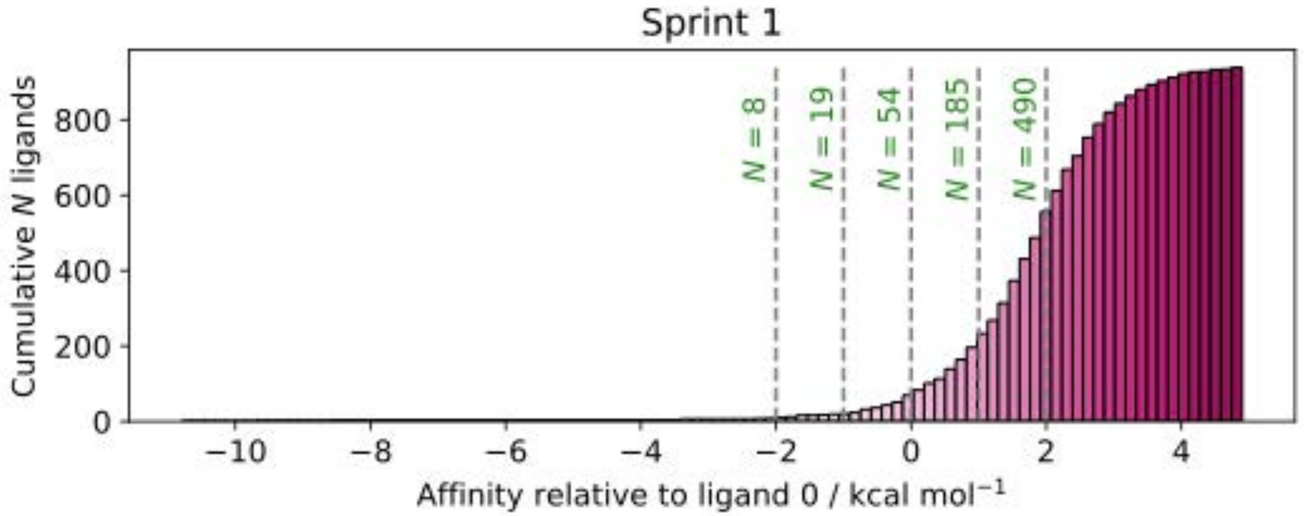
TRY-UNI-714a760b-6

top compounds from free energy calculations

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

MOST VIRTUAL LIBRARY COMPOUNDS WERE BAD

better **worse**

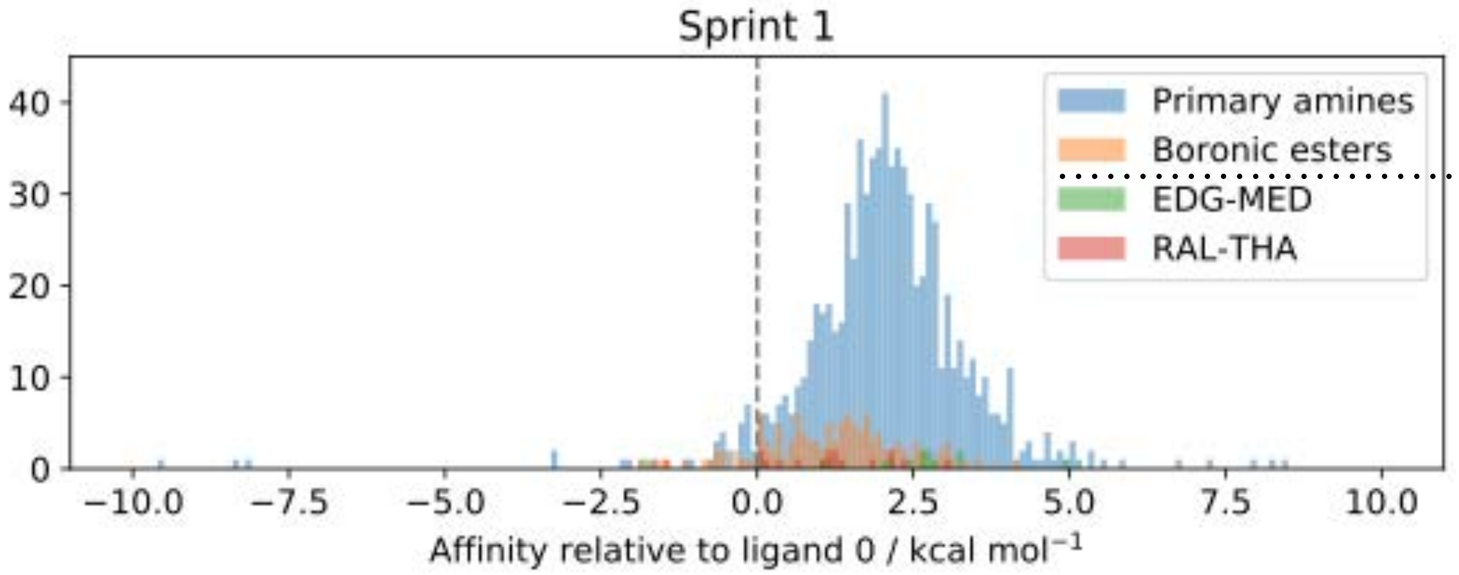


DOMINIC RUFA
TPCB student

HANNAH BRUCE MACDONALD
postdoc

WILLIAM GLASS
postdoc

HUMAN CHEMISTS NOMINATE BETTER COMPOUNDS, BUT ARE LIMITED IN THE NUMBER OF DESIGNS THEY CAN EVALUATE



computer

humans

DOMINIC RUFA
TPCB student

HANNAH BRUCE MACDONALD
postdoc

WILLIAM GLASS
postdoc

WE SET UP A DASHBOARD TO PROVIDE A REAL-TIME LEADERBOARD

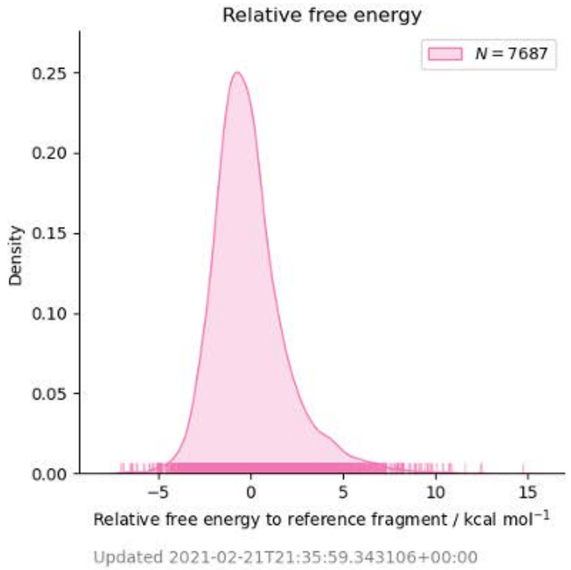
Description

COVID Moonshot Sprint 5 for benzopyran-isoquinoline series retrospective based on x11498 (MAT-POS-b3e365b9-1) to optimize substituents in the P1' pocket with Mpro dimer and neutral Cys145:His41 catalytic dyad

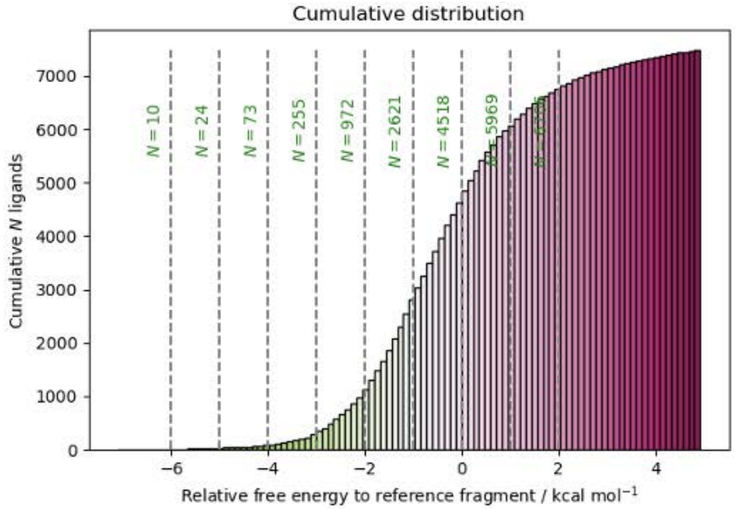
Progress

98.25%

Distributions



Updated 2021-02-21T21:35:59.343106+00:00



Updated 2021-02-21T21:35:59.343106+00:00

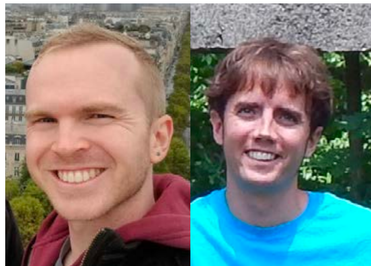
Leaderboard

Rank	Compound		SMILES	ΔG / kcal mol⁻¹	pIC50
1	VLA-UNK-83c3754c-1		c1ccc2c(c1)cncc2N3C(=O)[C@@]4(COc5c4cc(cc5)Cl)NC3=O	−15.9 ± 0.2	11.6 ± 0.2
2	ADA-UCB-dc2b944c-1		c1ccc2c(c1)cncc2N3C(=O)CN([C@@]4(C3=O)CCOc5c4cc(cc5)Cl)CC6CCCCC6	−15.5 ± 0.3	11.3 ± 0.2

[dashboard]

DAVID DOTSON
software scientist

MATT WITTMANN
software scientist

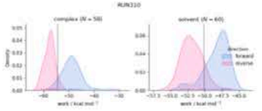
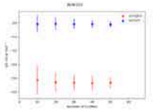
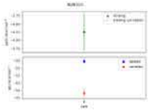
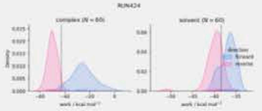
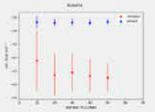
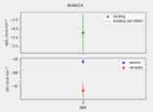
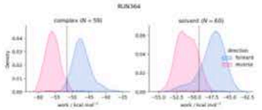
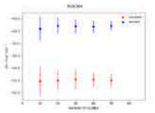
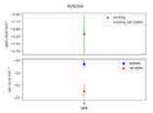
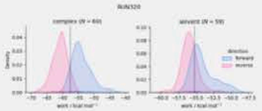
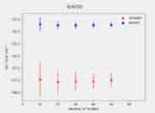
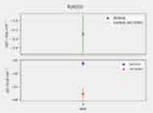
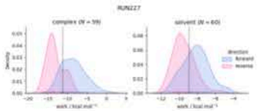
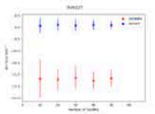
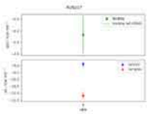
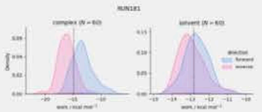
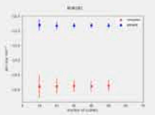
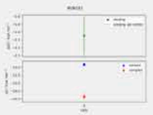


THE DASHBOARD LET CHEMISTS EASILY INSPECT THE RESULTS

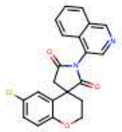
COVID Moonshot Sprint 11 | Summary | Compounds | Microstates | Transformations | **Reliable Transformations** | Retrospective Transformations | Retrospective Compounds

Reliable Transformations

Showing 1 through 100 of 100

RUN	Atom map	Initial microstate		Final microstate		ΔΔG / kcal M^{-1}	Work distribution	Bootstrapping	Convergence
RUN310	map	VLA-UCB-50c39ae8-2_1	pdb	MAT-POS-c2d406ed-1_2	pdb	−4.0 ± 0.3			
RUN424	map	VLA-UCB-50c39ae8-2_1	pdb	LUO-POS-b5068a05-1_2	pdb	−3.4 ± 0.5			
RUN364	map	VLA-UCB-50c39ae8-2_1	pdb	MAT-POS-c2d406ed-2_2	pdb	−2.5 ± 0.3			
RUN320	map	VLA-UCB-50c39ae8-2_1	pdb	MAT-POS-c2d406ed-1_1	pdb	−2.5 ± 0.2			
RUN227	map	VLA-UCB-50c39ae8-2_1	pdb	VLA-UNK-f702bf1c-5_1	pdb	−2.3 ± 0.2			
RUN181	map	VLA-UCB-50c39ae8-2_1	pdb	VLA-UNK-f702bf1c-6_1	pdb	−2.2 ± 0.1			

POTENT HUMAN CHEMIST DESIGNS SOMETIMES UNEXPECTEDLY FLOAT TO THE TOP

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

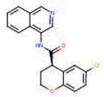
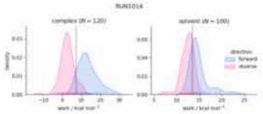
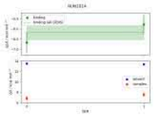
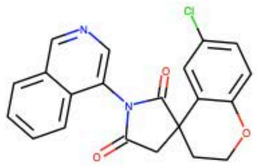
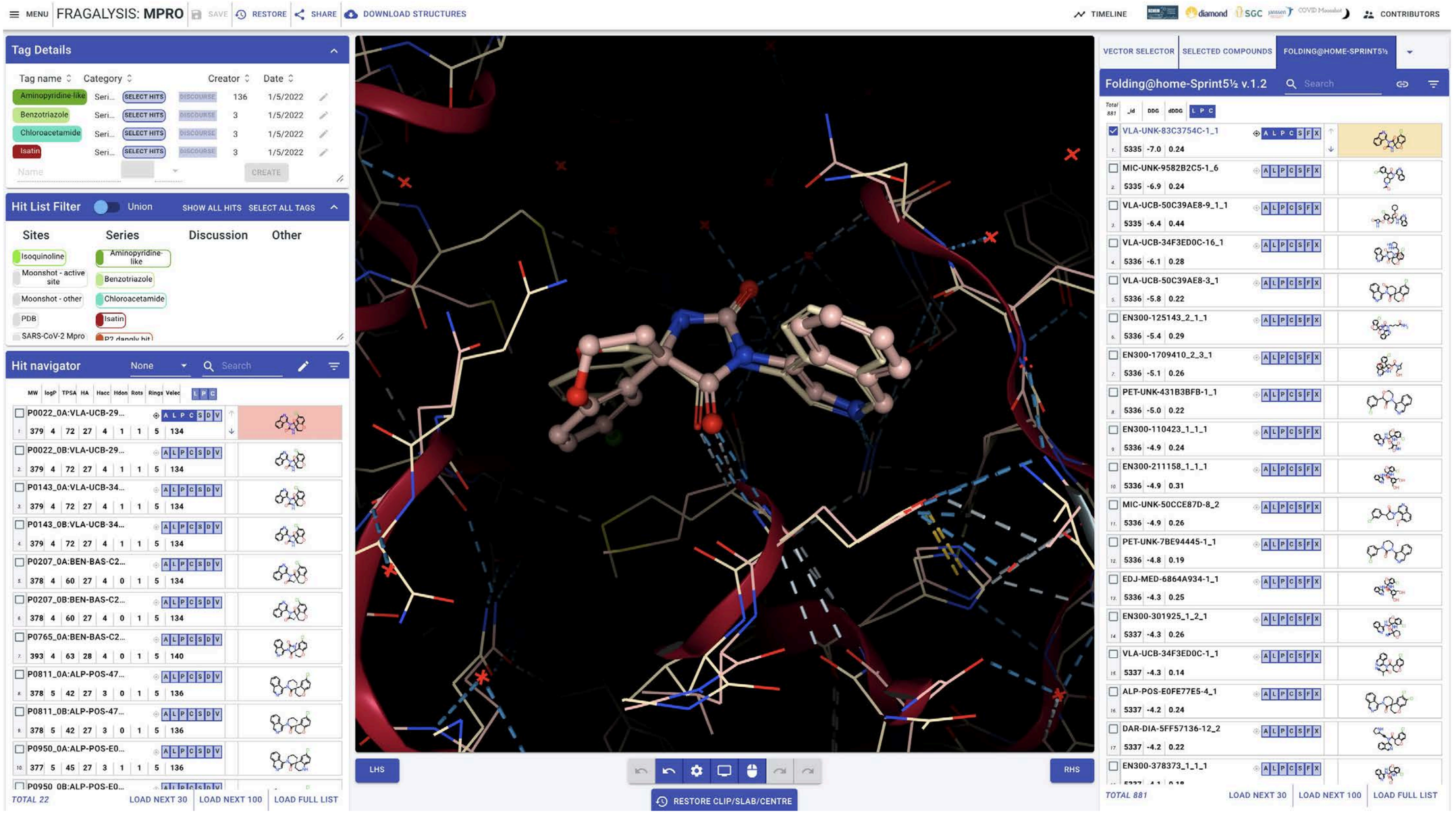
RUN1014	MAT-POS-b3e365b9-1_1		sdf	pdb	BEN-BAS-c2bc0d80-6_1		sdf	pdb	−6.2 ± 0.2	 

Image	Molecule	IC50 Curves	IC50 (µM) - Fluorescence
	BEN-BAS-c2bc0d80-6 O=C1CC2(CC0c3ccc(Cl)cc32)C(=O)N1c1cncc2ccccc12 3-aminopyridine-like Assayed Check Availability on Manifold	Fluorescence RapidFire	0.49

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

IT WAS SURPRISING HOW WELL POSES COULD BE PREDICTED



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

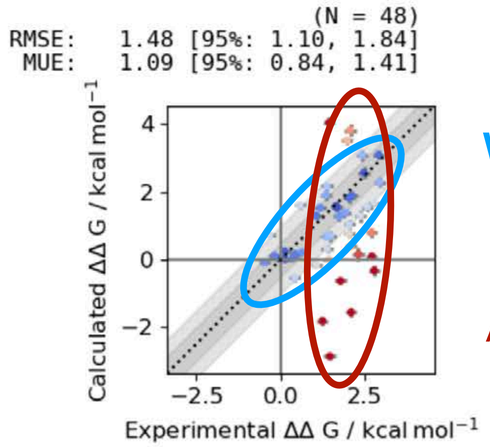
Fragalysis viewer: https://fragalysis.diamond.ac.uk/viewer/react/preview/target/Mpro

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

RAPID CYCLES OF PREDICTION AND POSTMORTEM GENERATES ACTIONABLE INSIGHTS AT AN INCREDIBLE PACE

COVID Moonshot Sprint 10 | Summary | Compounds | Microstates | Transformations | Reliable Transformations | Retrospective Transformations

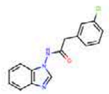
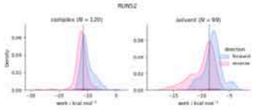
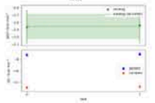
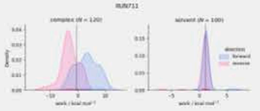
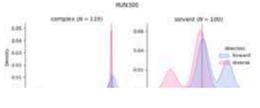
Retrospective Transformations



Well-predicted transformations

All modifications of P1 substituent pKa => **His163 is accepting H-bond, not donating!**

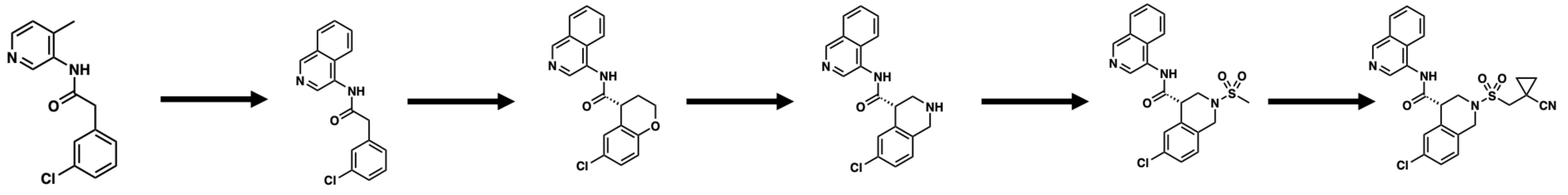
Showing 1 through 48 of 48

RUN	Initial microstate			Final microstate			ΔΔG / kcal M^{-1}	$\Delta\Delta G_{exp}$ / kcal M^{-1}	$\lvert\Delta\Delta G - \Delta\Delta G_{exp}\rvert$ / kcal M^{-1}	Work distribution	Convergence
RUN52	ADA-UCB-6c2cb422-1_1		sdf pdb	JAN-GHE-5a013bed-2_1		sdf pdb	−2.9 ± 0.1	1.5 ± 0.2	4.3 ± 0.2		
RUN711	ADA-UCB-6c2cb422-1_1		sdf pdb	PET-UNK-b1ef24dc-1_1		sdf pdb	−1.6 ± 0.1	2.1 ± 0.2	3.6 ± 0.2		
RUN300	ADA-UCB-6c2cb422-1_1		sdf pdb	EDJ-MED-c8e7a002-4_1		sdf pdb	−0.3 ± 0.2	2.8 ± 0.2	3.1 ± 0.2		

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

P1

P2

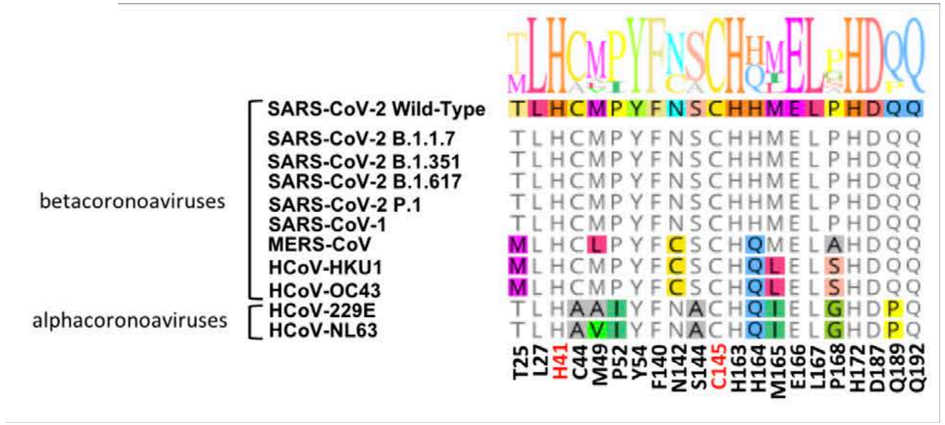


	TRY-UNI-714a760b-6	ADA-UCB-6c2cb422-1	MAT-POS-b3e365b9-1	MAT-POS-3ccb8ef6-1	MAT-POS-e194df51-1	MAT-POS-e194df51-1
IC_{50}(Mpro)/uM	25	0.73	0.21	0.28	0.141	0.037
EC_{50}(SARS-CoV-2, A549)/uM	n.d.	4.5	7.0	1.9	1.65	0.064

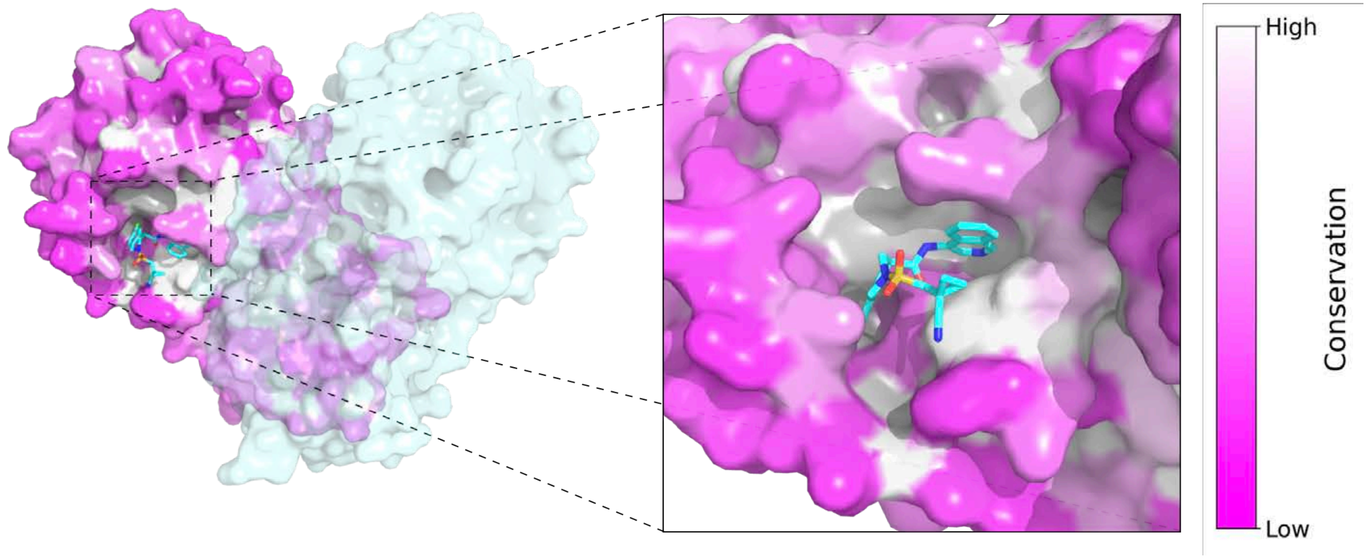
crowdsourced
merged fragment hit

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

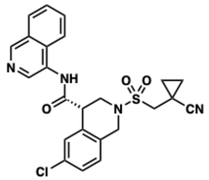
active-site residue conservation
of pathogenic coronaviruses



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.

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

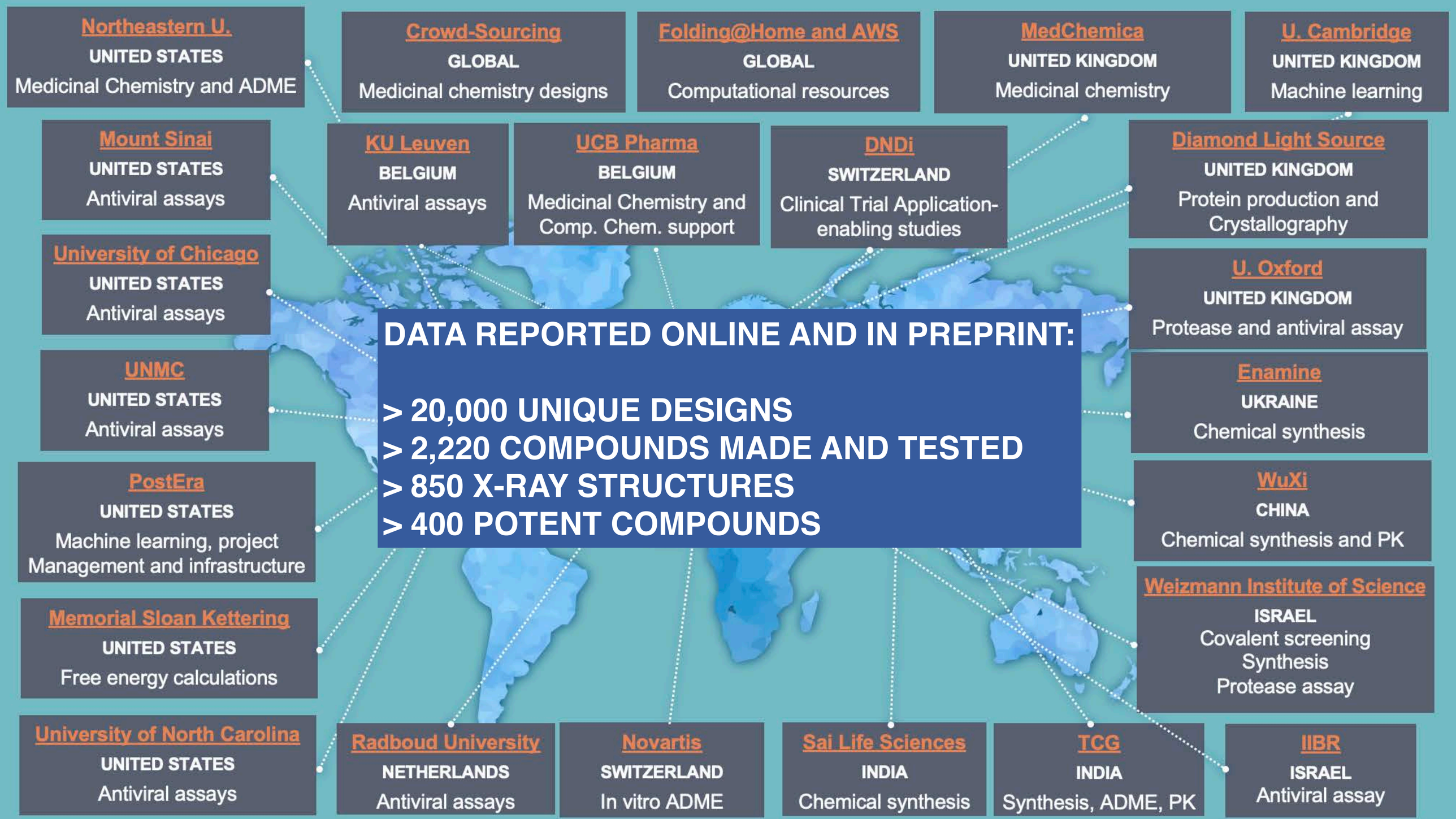
Follow this preprint

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

https://doi.org/10.1101/2020.10.29.339317

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

We're actively pursuing multiple backups in an accelerated preclinical program

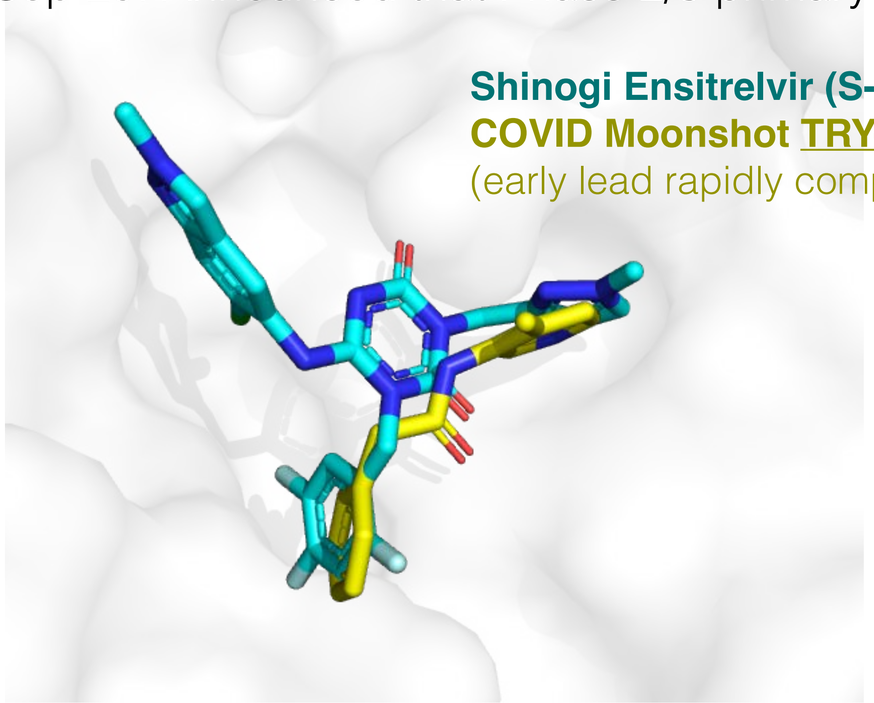

Northeastern U.
UNITED STATES
Medicinal Chemistry and ADME
Crowd-Sourcing
GLOBAL
Medicinal chemistry designs
Folding@Home and AWS
GLOBAL
Computational resources
MedChemica
UNITED KINGDOM
Medicinal chemistry
U. Cambridge
UNITED KINGDOM
Machine learning
Mount Sinai
UNITED STATES
Antiviral assays
KU Leuven
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
UNMC
UNITED STATES
Antiviral assays
Enamine
UKRAINE
Chemical synthesis
DATA REPORTED ONLINE AND IN PREPRINT:
> 20,000 UNIQUE DESIGNS
> 2,220 COMPOUNDS MADE AND TESTED
> 850 X-RAY STRUCTURES
> 400 POTENT COMPOUNDS
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

SHIONOGI RECENTLY REPORTED THE DISCOVERY OF ENSITRELVIR, DISCOVERED WITH THE HELP OF MOONSHOT DATA

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

Shionogi rapidly developed into potent antiviral with extraordinary PK
(1 pill/day; compare with 6 pills/day for Paxlovid with significant DDI risk)

Sep 29: Announced that Phase 2/3 primary endpoint was achieved



Shinogi Ensitrelvir (S-217622)
COVID Moonshot TRY-UNI-714a760b-6
(early lead rapidly compound disclosed online)

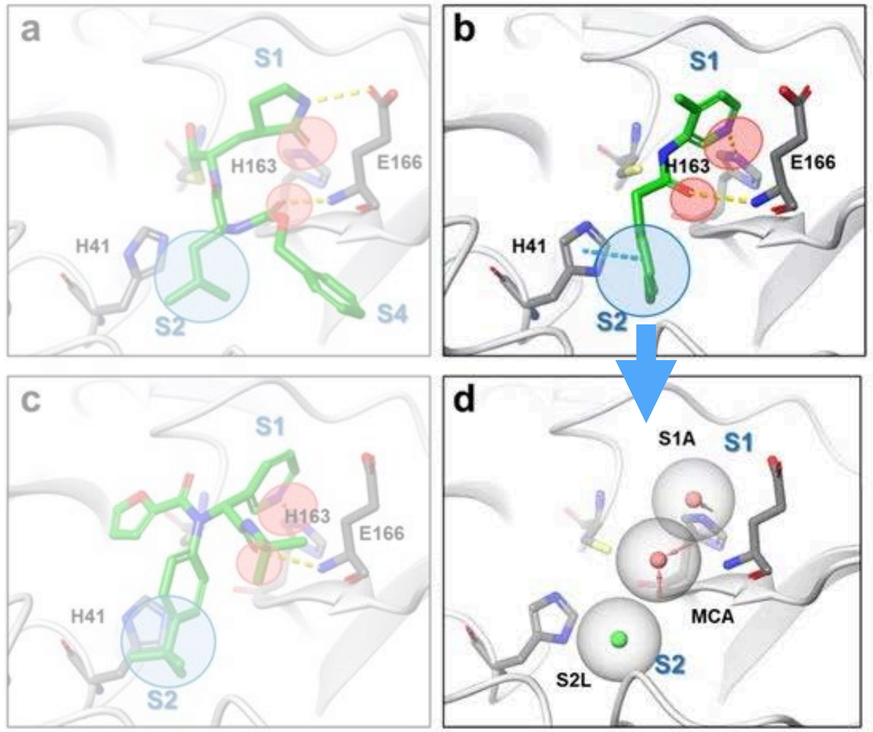
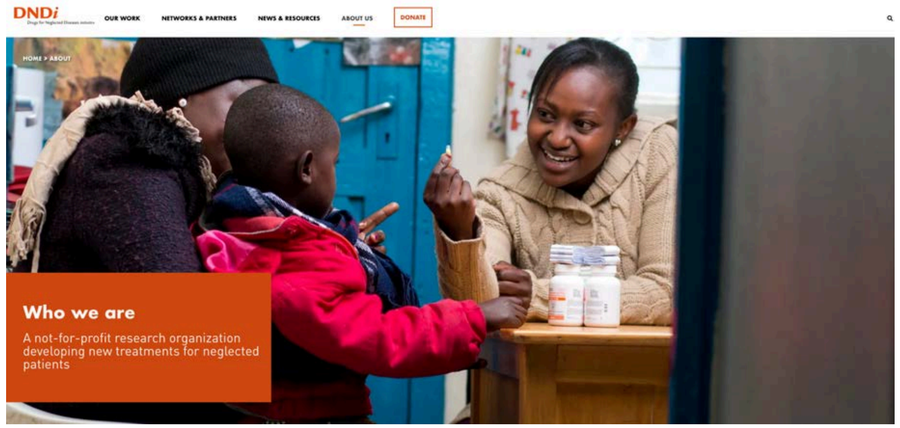


Figure 2. Binding modes of 3CL^{pro} inhibitors, their pharmacophores, and defined pharmacophore filters for virtual screening. (a) Crystal structures of GC376 (PDB: 6WTT), (b) 3-aminopyridine-like compound of the Postera COVID moonshot project (PDB: 5RH2) and (c) ML188 (PDB: 7L0D). The common H-bond 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.

Discovery of S-217622, a Noncovalent Oral SARS-CoV-2 3CL Protease Inhibitor Clinical Candidate for Treating COVID-19
J. Med. Chem. 2022, 65, 9, 6499–6512
https://doi.org/10.1021/acs.jmedchem.2c00117



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.

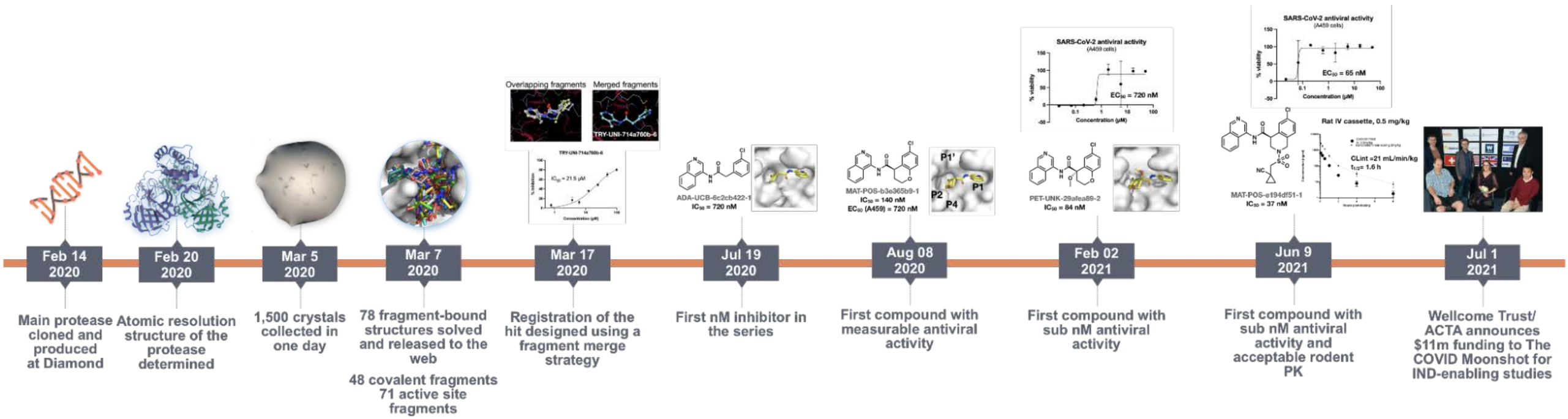
'*Most of the research and funding efforts early in the pandemic focused predominantly on repurposing of existing small molecule drugs and the more rapid development of novel monoclonal antibodies. Now, with the realization that COVID-19 will be a global issue for the foreseeable future we urgently need 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/

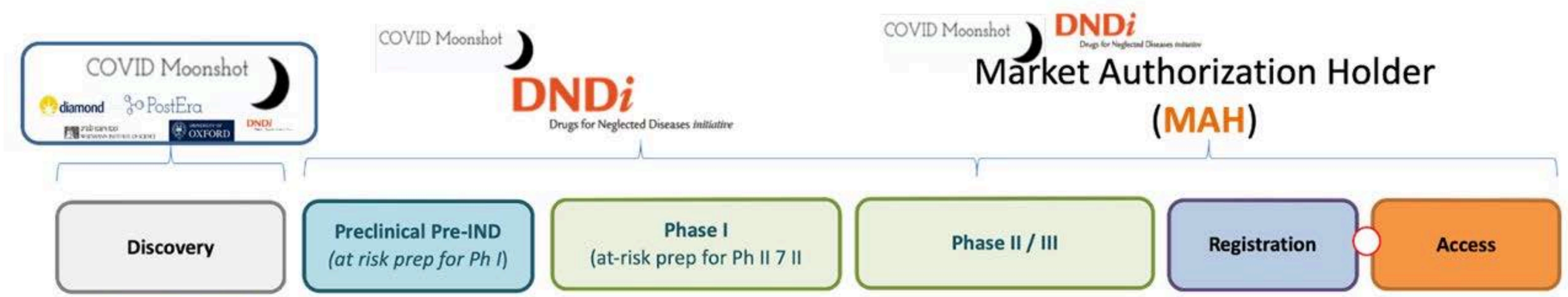
THE COVID MOONSHOT WENT FROM FRAGMENT SCREEN TO PRECLINICAL PHASE IN JUST 18 MONTHS, SPENDING LESS THAN $1M



A white-knuckle ride of open COVID drug discovery
https://www.nature.com/articles/d41586-021-01571-1

Open Science Discovery of Oral Non-Covalent SARS-CoV-2 Main Protease Inhibitor Therapeutics
SARS-CoV-2 Mpro antiviral preclinical candidate in a structure-based drug discovery program.
https://doi.org/10.1101/2020.10.29.339317

WE'RE AIMING TO BRING AN ANTIVIRAL STRAIGHT TO GENERICS MANUFACTURE WITHOUT A PATENT



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

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

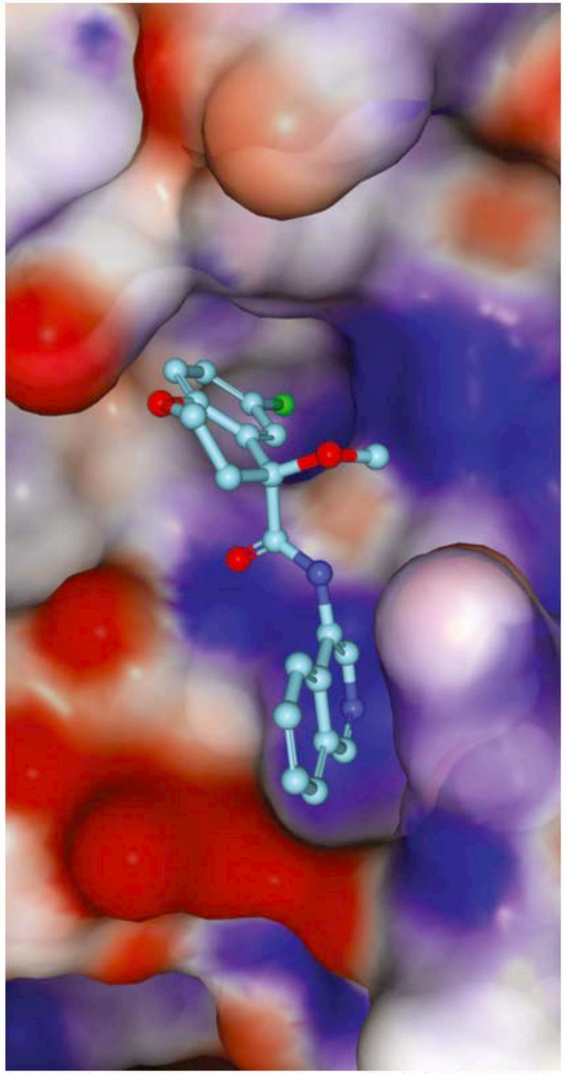
The response surprised us: almost 1,000 retweets in a week, and diverse offers for help. A.L. and M.R., two co-founders of the US-based technology firm PostEra, got in touch to suggest that their machine-learning technology could propose synthetic routes to make new molecules inspired by the fragment 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.

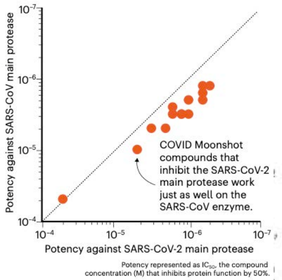
Above all, we can start predicting that these molecules will be straightforward to synthesize and will work as pills that are suitable for vaccine-hesitant or immunocompromised individuals, health-care workers and others in risky situations who could take them prophylactically. Furthermore, we expect 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.



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.

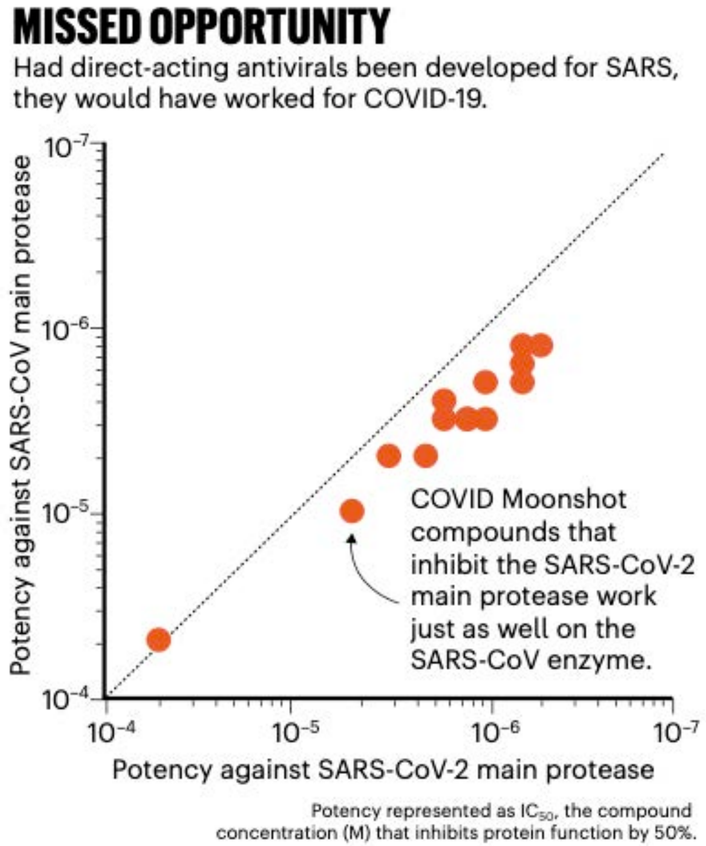
The authors

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 2003 SARS pandemic.

GLOBAL, EQUITABLE ACCESS IS A ENORMOUS 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

THIS PROBLEM WILL NOT GO AWAY.
CLIMATE CHANGE IS CREATING THE "PANDEMICENE"

Climate change

Heating and stirring the global viral soup

Rachel E. Baker & C. Jessica E. Metcalf

Simulations show that rising global temperatures and changes in land use will drive new encounters between mammalian species. This could lead to an increase in virus-sharing events that might threaten both wildlife and humans. **See p.555**

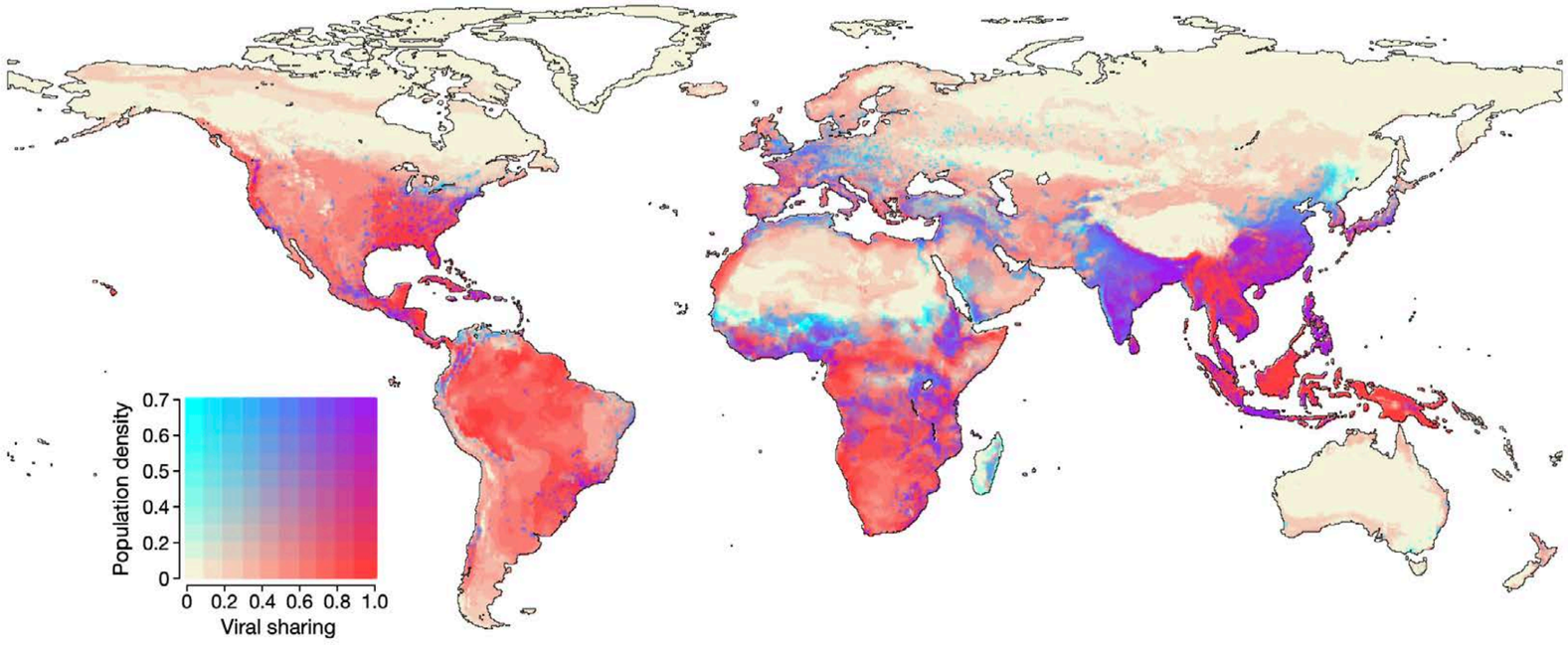
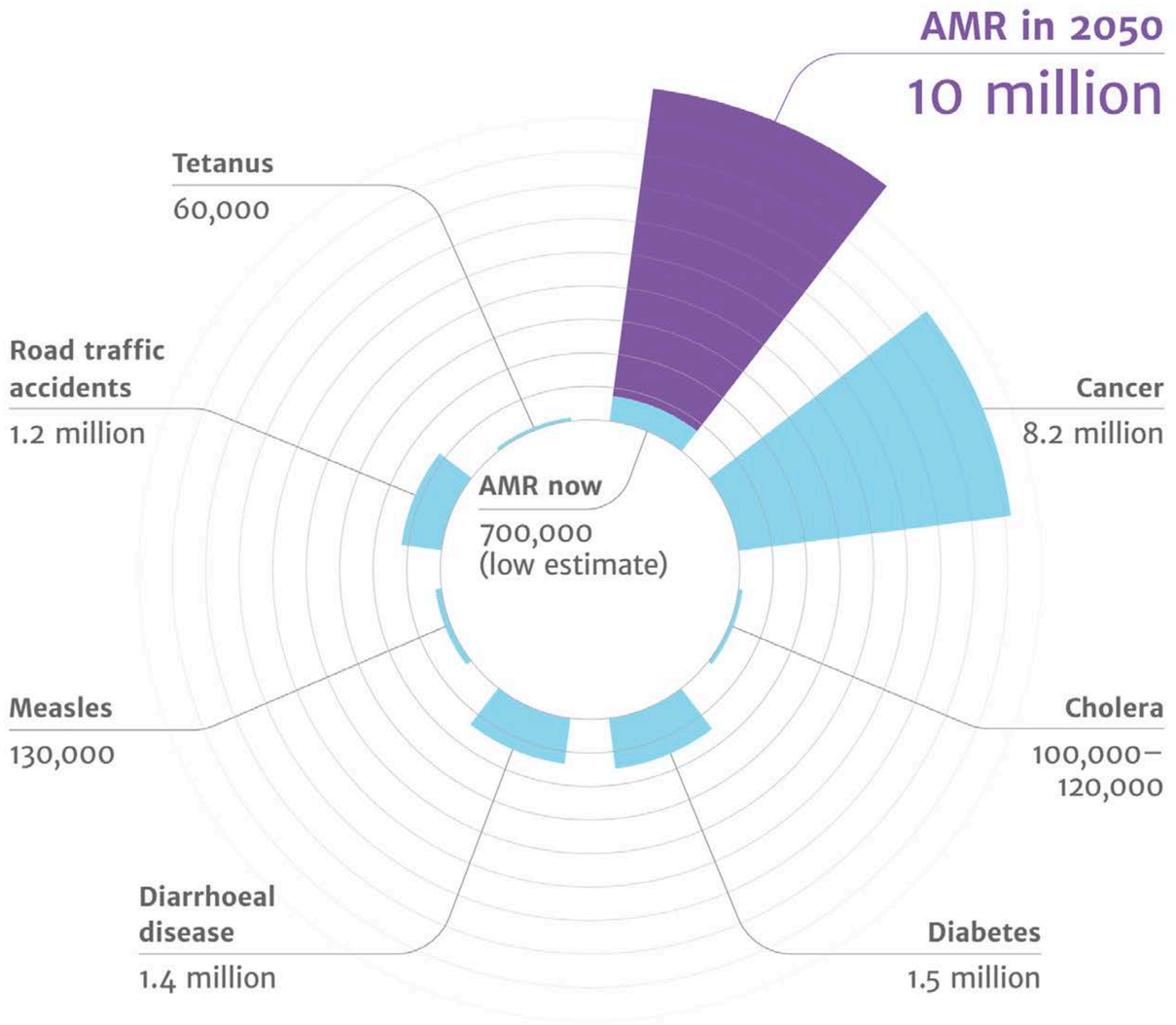


Fig. 4 | Novel viral sharing events coincide with human population centres. In 2070 (SSP 1–RCP 2.6; climate only), human population centres in equatorial Africa, south China, India and southeast Asia will overlap with projected hotspots of cross-species viral transmission in wildlife. Both variables were linearly rescaled to 0 to 1. The results were averaged across nine GCMs.

https://www.nature.com/articles/s41586-022-04788-w
https://www.nature.com/articles/d41586-022-01474-9

BY 2050, ANTIMICROBIAL RESISTANCE WILL KILL 10 MILLION PEOPLE EACH YEAR



Projected deaths

https://amr-review.org/Publications.html

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.

HOW CAN WE PREVENT FUTURE PANDEMICS?

What's the best way to win a race?

1. **Run fast.** Develop a technology platform for accelerated discovery of oral antivirals that can rapidly progress fragments to preclinical candidates leveraging machine learning and physical modeling

2. **Start close to the finish line.** Repeatedly exercise this platform to develop an arsenal of low-cost clinic-ready drug candidates against viruses of pandemic concern

HOW CAN WE PREVENT FUTURE PANDEMICS?

What's the best way to win a race?

1. **Run fast.** Develop a tec al antivirals that can rapidly progress fragments to pre d physical modeling

2. **Start close to the finish** an arsenal of low-cost clinic-ready drug candidates ag

A Pill to Treat Covid-19? The U.S. Is Betting on It.

A new $3.2 billion program will support the development of antiviral pills, which could start arriving by the end of this year.

Dr. Anthony Fauci announced on Thursday that the White House was investing over $3 billion to advance the development of antiviral pills to treat Covid-19 as well as future virus outbreaks.
Agence France-Presse — Getty Images

https://www.nytimes.com/2021/06/17/health/covid-pill-antiviral.html

Consortium formed to discover antivirals for COVID-19 receives NIH funding to develop globally accessible treatments for pandemics

18 May 2022

First $68M award for initial 3 years

DNDi
Drugs for Neglected Diseases *initiative*

A consortium led by international scientists from the non-profit, open-science COVID Moonshot has been awarded an initial $68,662,387 from the US National Institutes of Health (NIH) to discover and develop globally accessible and affordable novel oral antivirals to combat COVID-19 and future pandemics.

'*If we had clinic-ready antivirals suitable for SARS-CoV-2 when the pandemic struck in late 2019, we could have perhaps saved millions of lives*,' said Dr Ben Perry, Discovery Open Innovation Leader at the Drugs for Neglected Diseases *initiative* (DND*i*), and a founder of the COVID Moonshot. '*The world needs a diverse stockpile of novel antiviral compounds that can be quickly advanced for the current pandemic and when the next pandemic strikes, and it is essential that these be affordable and equitably accessible to everyone.*'

The consortium has created the artificial intelligence (AI)-driven Structure-enabled Antiviral Platform (ASAP), which will use cutting-edge technology, encompassing advanced structural biology, AI, machine learning, and computational chemistry on Folding@home, the world's largest distributed computing platform, to build a robust antiviral discovery pipeline.

DND*i* is one of the three institutions leading the consortium, along with AI-driven biotech PostEra and the Memorial Sloan Kettering Cancer Center. ASAP partners include the Diamond Light Source (UK); PostEra (USA); the Memorial Sloan

The ultimate objective of the project is to deliver multiple drug candidates ready for evaluation in humans in the event of an ongoing or emerging pandemic threat. The project will maximize the use of an open science model that prioritizes global, equitable, and affordable access.

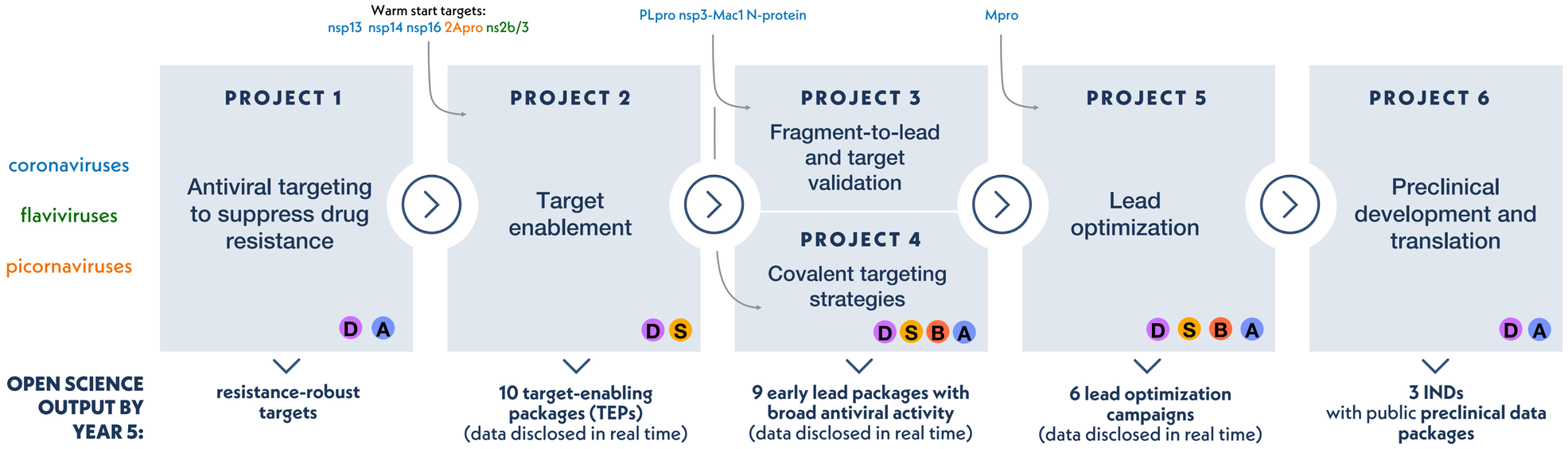
ASAP is built on the successes of the COVID Moonshot, a global, open-science collaboration that began in March 2020 and rapidly identified potent antivirals targeting the main protease of the SARS-CoV-2 virus, which are currently undergoing a preclinical development program funded by the Wellcome/COVID-19 Therapeutics Accelerator. The open science data publicly shared by Moonshot additionally enabled the identification of another promising COVID-19 drug developed by the Japanese pharmaceutical company Shionogi that is now in late-stage clinical trials.

'*The rapid progress of Moonshot demonstrates the power of AI-driven drug design*,' said Dr Alpha Lee, Chief Scientific Officer of PostEra and a founder of the COVID Moonshot. '*Our algorithms generate molecules with optimized properties that can quickly be made and tested in the lab and help us select the most important experiments. ASAP will take this to the next level.*' Dr Lee is one of the leaders of ASAP.

ASAP will target viral families that have been historically neglected by the market,

WE AIM TO TEST AUTONOMOUS DISCOVERY METHODS IN THE AI-DRIVEN STRUCTURE-ENABLED ANTIVIRAL PLATFORM (ASAP)

Open science drug discovery for global equitable and affordable access



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)

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

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)

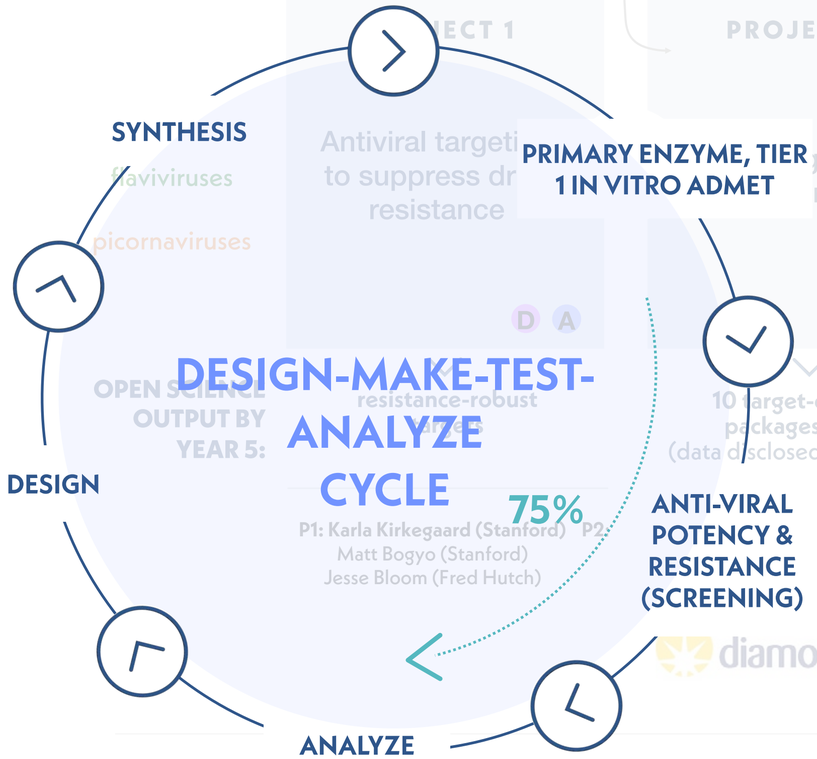
B BIOCHEMICAL ASSAY CORE
Nir London (Weizmann)
Haim Barr (Weizmann)

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

$68M awarded for initial 3 years / up to $110M over 5 years

http://asapdiscovery.org

ASAP WILL ENABLE IS TO REFINE OUR METHODS WITH AN UNPRECEDENTED DEGREE OF PROSPECTIVE EXPERIMENTAL DATA



We will spend **over $12M on chemistry** on compound designs our methods select

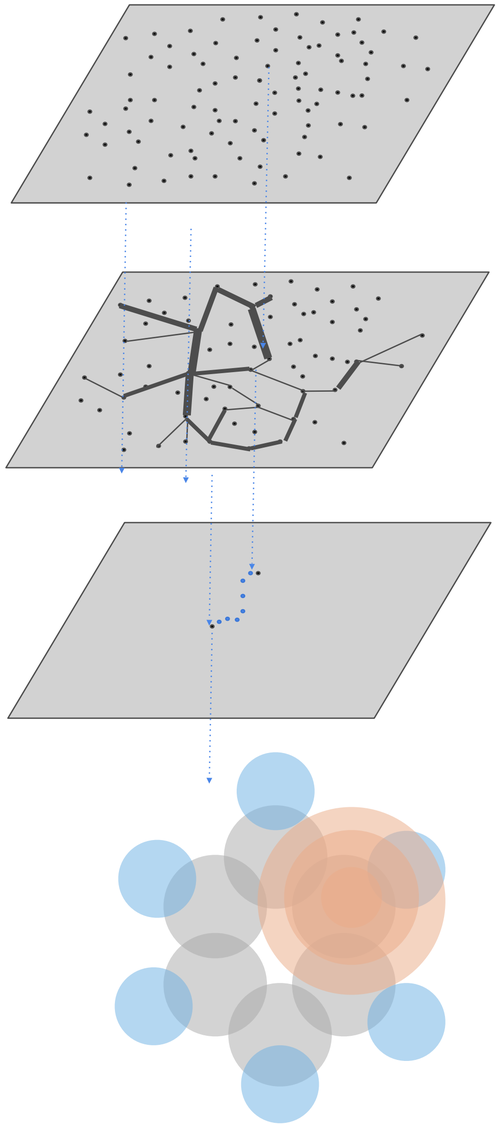
This provides an **unprecedented opportunity** to build better predictive models while producing useful antivirals against future pandemics

The infrastructure we build for accelerated discovery can be directly applied **to cancer**

First $68M awarded for first 3 years / up to $110M over 5 years

http://asapdiscovery.org

WE ARE DRIVING THE DEVELOPMENT OF A NEW GENERATION OF HYBRID PHYSICAL / MACHINE LEARNING MODELS

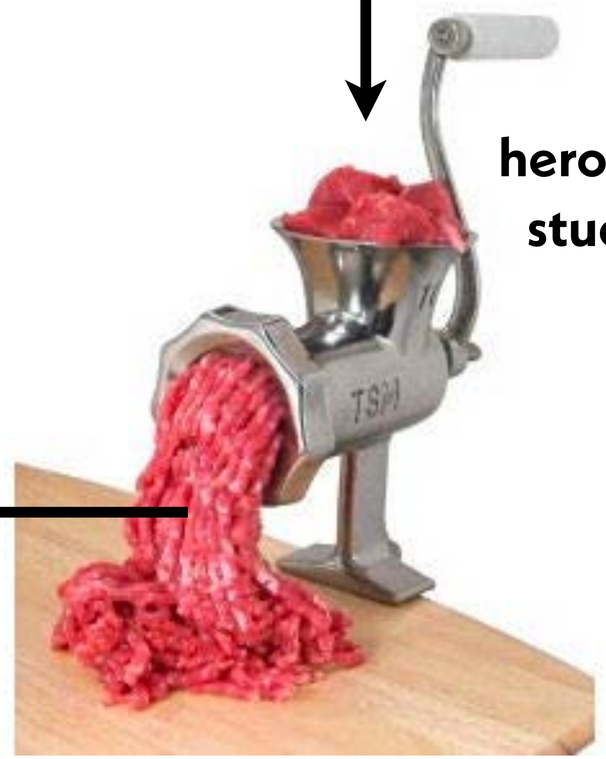


- Fast, structure-based **machine learning surrogates** assess designs over vast synthetic chemical spaces prioritize useful calculations
- Adaptive allocation of effort to alchemical free energy calculations guided by **machine learning cost prediction**s
- **Machine learned optimal alchemical transformations** produce faster estimates of free energy differences more cheaply
- **Learnable machine learning potentials** fit to experimental free energy and quantum chemical data produce higher accuracy predictions

Chemical Science 2022 https://doi.org/10.1039/D2SC02739A
bioRxiv https://www.biorxiv.org/content/10.1101/2021.08.24.457513v2

FORCE FIELDS HAVE TRADITIONALLY BEEN HEROIC PRODUCTS OF HUMAN EFFORT

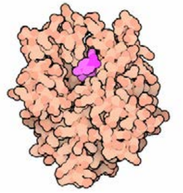
experimental data
quantum chemistry
keen chemical intuition



heroic effort by graduate students and postdocs

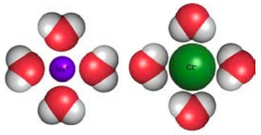
a parameter set we desperately hope someone actually uses

FORCE FIELDS HAVE TRADITIONALLY BEEN HEROIC PRODUCTS OF HUMAN EFFORT



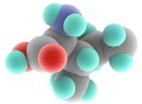
proteins

post-translational modifications

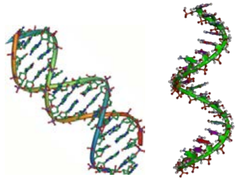


water

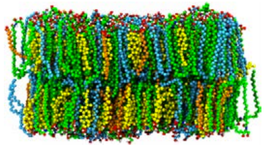
ions



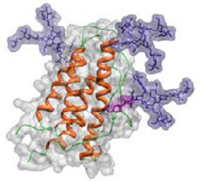
small molecules



nucleic acids



lipids



carbohydrates

Amber20 recommendations

J. A. Maier; C. Martinez; K. Kasavajhala; L. Wickstrom; K. E. Hauser; C. Simmerling. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.*, **2015**, *11*, 3696–3713.

W. D. Cornell; P. Cieplak; C. I. Bayly; I. R. Gould; K. M. Merz, Jr.; D. M. Ferguson; D. C. Spellmeyer; T. Fox; J. W. Caldwell; P. A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, **1995**, *117*, 5179–5197.

N. Homeyer; A. H. C. Horn; H. Lanig; H. Sticht. AMBER force-field parameters for phosphorylated amino acids in different protonation states: phosphoserine, phosphothreonine, phosphotyrosine, and phosphohistidine. *J. Mol. Model.*. **2006**. *12*. 281–289.

H. W. Horn; W. C. Swope; J. W. Pitera; J. D. Madura; T. J. Dick; G. L. Hura; T. Head-Gordon. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.*, **2004**, *120*, 9665–9678.

I. S. Joung; T. E. Cheatham, III. Molecular dynamics simulations of the dynamic and energetic properties of alkali and halide ions using water-model-specific ion parameters. *J. Phys. Chem. B*, **2009**, *113*, 13279–13290.

P. Li; B. P. Roberts; D. K. Chakravorty; K. M. Merz, Jr. Rational Design of Particle Mesh Ewald Compatible Lennard-Jones Parameters for +2 Metal Cations in Explicit Solvent. *J. Chem. Theory Comput.*, **2013**, *9*, 2733–2748.

J. Wang; R. M. Wolf; J. W. Caldwell; P. A. Kollamn; D. A. Case. Development and testing of a general Amber force field. *J. Comput. Chem.*, **2004**, *25*, 1157–1174.

R. Galindo-Murillo; J. C. Robertson; M. Zgarbovic; J. Sponer; M. Otyepka; P. Jureska; T. E. Cheatham. Assessing the Current State of Amber Force Field Modifications for DNA. *J. Chem. Theory Comput.*, **2016**, *12*, 4114–4127.

A. Perez; I. Marchan; D. Svozil; J. Sponer; T. E. Cheatham; C. A. Laughton; M. Orozco. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of alpha/gamma Conformers. *Biophys. J.*, **2007**, *92*, 3817–3829.

M. Zgarbova; M. Otyepka; J. Sponer; A. Mladek; P. Banas; T. E. Cheatham; P. Jurecka. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles. *J. Chem. Theory Comput.*, **2011**, *7*, 2886–2902.

Å. Skjevik; B. D. Madej; R. C. Walker; K. Teigen. Lipid11: A modular framework for lipid simulations using amber. *J. Phys. Chem. B*, **2012**, *116*, 11124–11136.

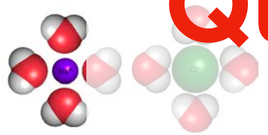
C. J. Dickson; B. D. Madej; A. A. Skjevik; R. M. Betz; K. Teigen; I. R. Gould; R. C. Walker. Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput.*, **2014**, *10*, 865–879.

K. N. Kirschner; A. B. Yongye; S. M. Tschampel; J. González-Outeiriño; C. R. Daniels; B. L. Foley; R. J. Woods. GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.*, **2008**, *29*, 622–655.

FORCE FIELDS HAVE TRADITIONALLY BEEN HEROIC PRODUCTS OF HUMAN EFFORT

proteins

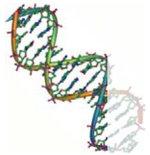
post-translational modifications



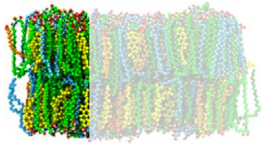
water

ions

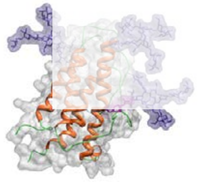
non-standard residues



nucleic acids



lipids



carbohydrates

Quickly adds up to >100 human-years

Intended to be compatible, but not co-parameterized

Significant effort is required to extend to new areas (e.g. covalent inhibitors, bio-inspired polymers, etc.)

Nobody is going to want to refit this based on some new data

How can we bring this problem into the modern era?

Amber20 recommendations

J. A. Maier; C. Martinez; K. Kasavajhala; L. Wickstrom; K. E. Hauser; C. Simmerling. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.*, **2015**, *11*, 3696–3713.

W. D. Cornell; P. Cieplak; C. I. Bayly; I. R. Gould; K. M. Merz, Jr.; D. M. Ferguson; D. C. Spellmeyer; [illegible] Kollman. A second generation force field for the simulation of proteins, nucleic [illegible] *J. Am. Chem. Soc.*, **1995**, *117*, 5179–5197.

N. Homeyer; A. H. C. Horn; H. Lanig; H. Sticht. AMBER force-field parameters for phosphorylated amino acids in different protonation states: phosphoserine, phosphothreonine, phosphotyrosine, and phosphohistidine. *J. Mol. Model.*, **2006**, *12*, 281–289.

H. W. Horn; W. C. Swope; J. W. Pitera; J. D. Madura; T. J. Dick; G. L. Hura; T. Head-Gordon. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.*, **2004**, *120*, 9665–9678.

I. S. Joung; T. E. Cheatham, III. Molecular dynamics simulations of the dynamic and energetic properties [illegible] specific ion parameters. *J. Phys. Chem. B*, **2009**, *113*, 13279–13290.

P. Li; B. P. Roberts; D. K. Chakravorty; K. M. Merz, Jr. Rational Design of Particle Mesh Ewald Compatible [illegible] Simulations in Explicit Solvent. *J. Chem. Theory Comput.*, **2013**, *9*, 2733–2748.

J. Wang; R. M. Wolf; J. W. Caldwell; P. A. Kollamn; D. A. Case. Development and testing of a general [illegible] *J. Comput. Chem.*, **2004**, *25*, 1157–1174.

R. Galindo-Murillo; J. C. Robertson; M. Zgarbovic; J. Sponer; M. Otyepka; P. Jureska; T. E. Cheatham. [illegible] *J. Chem. Theory Comput.*, **2016**, [illegible]

A. Perez; I. Marchan; D. Svozil; J. Sponer; T. E. Cheatham; C. A. Laughton; M. Orozco. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of alpha/gamma Conformers. *Biophys. J.*, **2007**, *92*, 3817–3829.

M. Zgarbova; M. Otyepka; J. Sponer; A. Mladek; P. Banas; T. E. Cheatham; P. Jurecka. Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic [illegible]

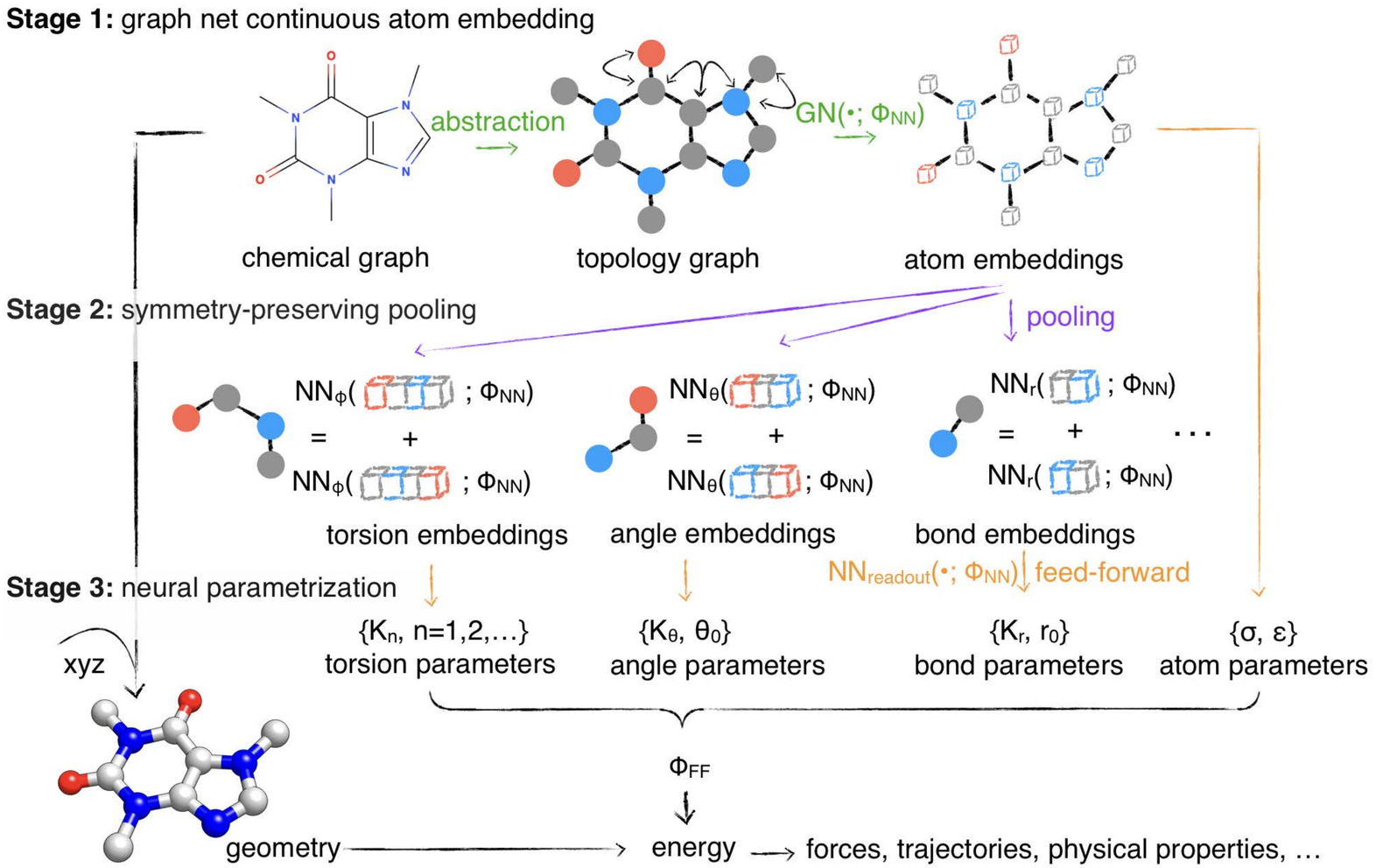
Å. Skjevik; B. D. Madej; R. C. Walker; K. Teigen. Lipid11: A modular framework for lipid simulations using amber. *J. Phys. Chem. B*, **2012**, *116*, 11124–11136.

C. J. Dickson; B. D. Madej; A. A. Skjevik; R. M. Betz; K. Teigen; I. R. Gould; R. C. Walker. Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput.*, **2014**, *10*, 865–879.

K. N. Kirschner; A. B. Yongye; S. M. Tschampel; J. González-Outeiriño; C. R. Daniels; B. L. Foley; R. J. Woods. GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.*, **2008**, 29, 622–655.

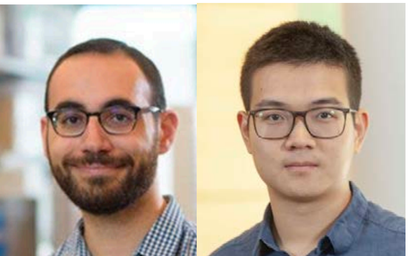
HYBRID PHYSICAL / MACHINE LEARNING MODELS ARE DATA-EFFICIENT AND CAN GENERALIZE BROADLY

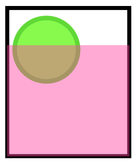
use of only **chemical graph** means that model can generate parameters for small molecules, proteins, nucleic acids, covalent ligands, carbohydrates, etc.



JOSH FASS

YUANQING WANG



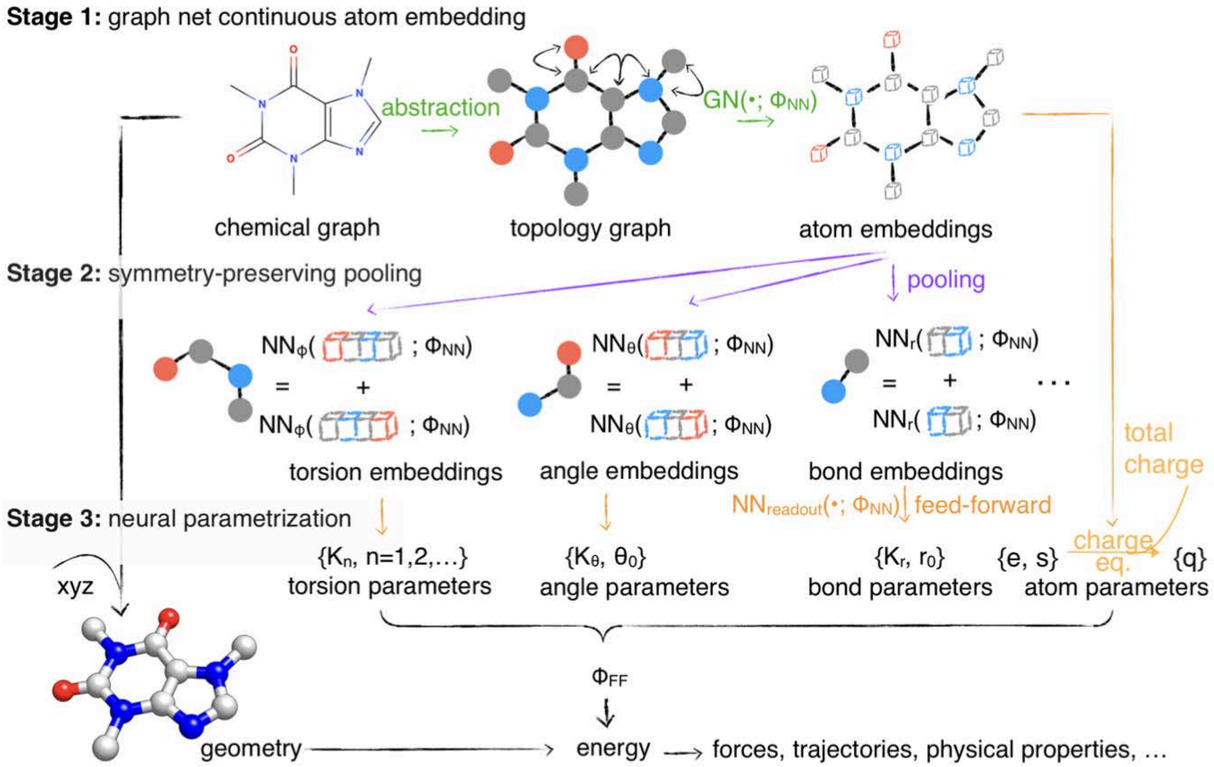


preprint: https://arxiv.org/abs/2010.01196

code: https://github.com/choderalab/espaloma

ESPALOMA MAKES LEARNING NEW PHYSICAL MODELS EASY

espaloma architecture



(implemented in pytorch)

http://github.com/choderalab/espaloma

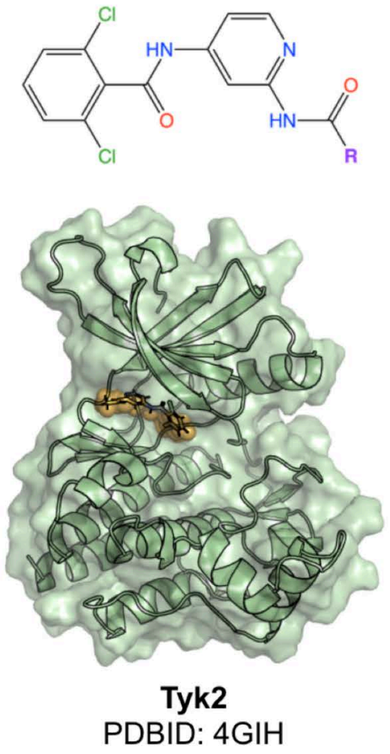
YUANQING WANG

building a new force field

```python
import torch, dgl, espaloma as esp

# retrieve OpenFF Gen2 Optimization Dataset
dataset = esp.data.dataset.GraphDataset.load("gen2").view(batch_size=128)

# define Espaloma stage I: graph -> atom latent representation
representation = esp.nn.Sequential(
    layer=esp.nn.layers.dgl_legacy.gn("SAGEConv"), # use SAGEConv implementation in DGL
    config=[128, "relu", 128, "relu", 128, "relu"], # 3 layers, 128 units, ReLU activation
)

# define Espaloma stage II and III:
# atom latent representation -> bond, angle, and torsion representation and parameters
readout = esp.nn.readout.janossy.JanossyPooling(
    in_features=128, config=[128, "relu", 128, "relu", 128, "relu"],
    out_features={              # define modular MM parameters Espaloma will assign
        1: {"e": 1, "s": 1}, # atom hardness and electronegativity
        2: {"coefficients": 2}, # bond linear combination
        3: {"coefficients": 3}, # angle linear combination
        4: {"k": 6}, # torsion barrier heights (can be positive or negative)
    },
)

# compose all three Espaloma stages into an end-to-end model
espaloma_model = torch.nn.Sequential(
                 representation, readout,
                 esp.mm.geometry.GeometryInGraph(), esp.mm.energy.EnergyInGraph(),
                 esp.nn.readout.charge_equilibrium.ChargeEquilibrium(),
)

# define training metric
metrics = [
    esp.metrics.GraphMetric(
            base_metric=torch.nn.MSELoss(), # use mean-squared error loss
            between=['u', "u_ref"],         # between predicted and QM energies
            level="g", # compare on graph level
    )
    esp.metrics.GraphMetric(
            base_metric=torch.nn.MSELoss(), # use mean-squared error loss
            between=['q', "q_hat"],         # between predicted and reference charges
            level="n1", # compare on node level
    )
]

# fit Espaloma model to training data
results = esp.Train(
    ds_tr=dataset, net=espaloma_model, metrics=metrics,
    device=torch.device('cuda:0'), n_epochs=5000,
    optimizer=lambda net: torch.optim.Adam(net.parameters(), 1e-3), # use Adam optimizer
).run()

torch.save(espaloma_model, "espaloma_model.pt") # save model
```

Listing 1. Defining and training a modular Espaloma model.

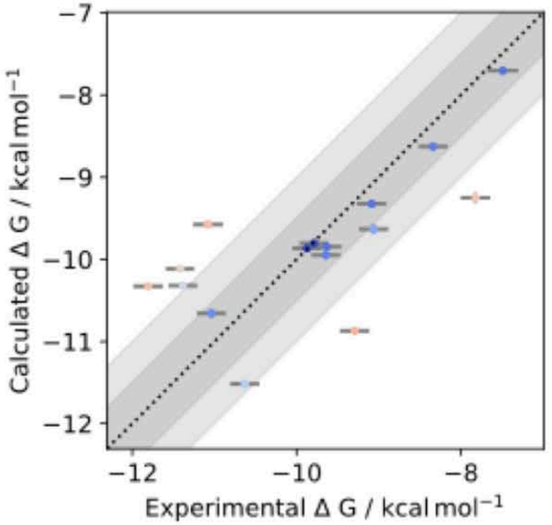
ESPALOMA SMALL MOLECULE PARAMETERS PERFORM AS WELL OR BETTER THAN MODERN BIOMOLECULAR FORCE FIELDS



	ΔG_{exp} / kcal mol^{-1}			ΔG_{exp} / kcal mol^{-1}
1	-9.54		9	-9.56
2	-10.94		10	-7.42
3	-8.98		11	-11.28
4	-11.31		12	-9.00
5	-9.21		13	-9.70
6	-8.26		14	-11.70
7	-10.91		15	-9.78
8	-7.75		16	-10.53

OpenFF 1.2.0 small molecule
Amber ff14SB protein
TIP3P water

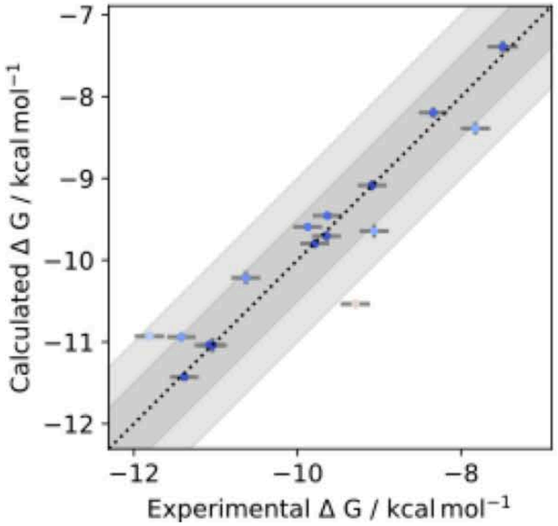
```
Absolute binding energies - tyk2
                          tyk2 (N = 16)
RMSE:    0.91 [95%: 0.66, 1.17]
MUE:     0.72 [95%: 0.47, 1.03]
R2:      0.48 [95%: 0.09, 0.78]
rho:     0.69 [95%: 0.28, 0.89]
```



~1 year of effort

espaloma "joint" 0.2.2 small molecule
Amber ff14SB protein
TIP3P water

```
Absolute binding energies - tyk2
                          tyk2 (N = 16)
RMSE:    0.47 [95%: 0.30, 0.70]
MUE:     0.31 [95%: 0.22, 0.56]
R2:      0.87 [95%: 0.62, 0.96]
rho:     0.93 [95%: 0.80, 0.98]
```



~1 day of effort

MIKE HENRY

IVÁN PULIDO

IVY ZHANG

DOMINIC RUFA

HANNAH BRUCE CDONALD

YUANQING WANG

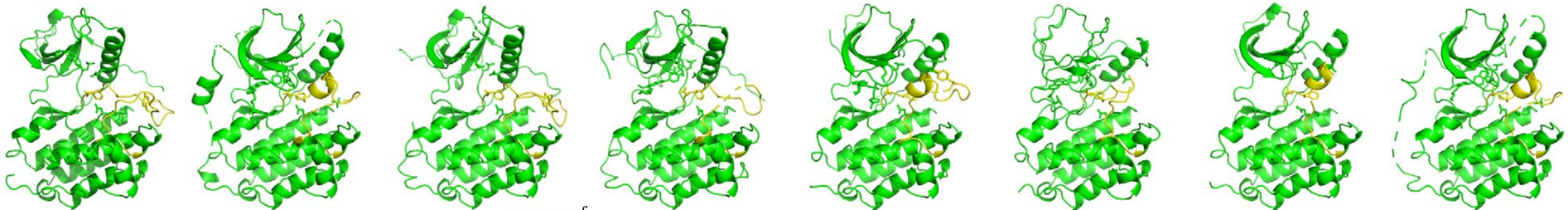
preprint: https://arxiv.org/abs/2010.01196
code: http://github.com/choderalab/espaloma
free energy calculations with http://github.com/choderalab/perses

WE'RE BUILDING NEW HYBRID MACHINE LEARNING / PHYSICAL MODELS TO DRIVE THE DISCOVERY OF MUTANT-SELECTIVE KINASE INHIBITORS FOR CANCER THERAPY

OpenFold-like modeling
of mutant conformations

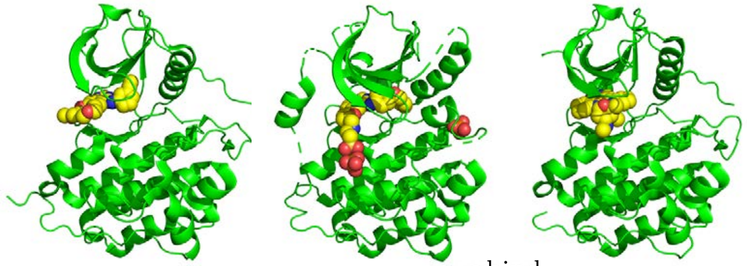
RCSB PDB
PROTEIN DATA BANK

distinct conformations of *apo* kinase



ΔG_i^{conf}

hybrid docking
shape overlay and
physical docking

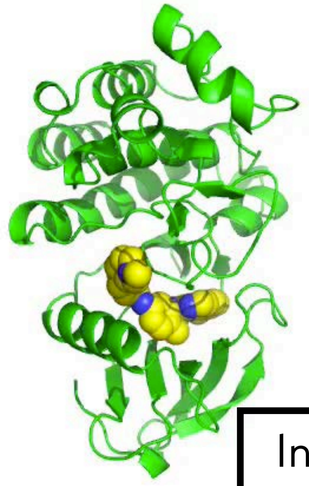


ΔG_i^{bind}

featurize
chemical/structural features

deep learning
to predict conformation/
pose specific affinity

structure-based
ML surrogates



prioritize conformations,
poses for detailed **alchemical**
free energy calculations

Integrated infrastructure can
predict **affinity**, **selectivity**,
and **impact of mutations**

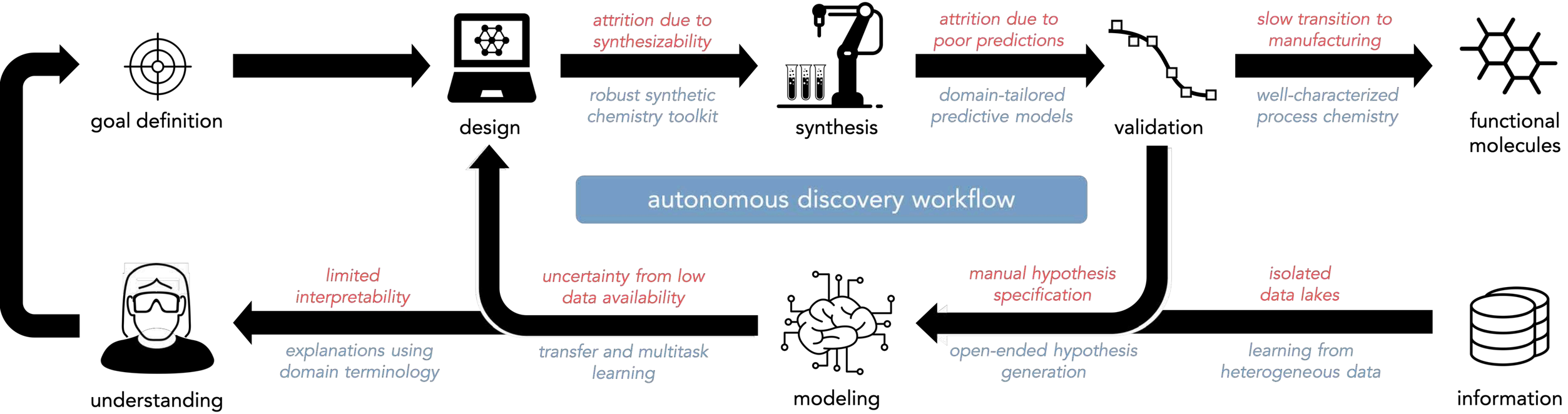
Boltzmann pooling across
conformations/poses
to predict affinities

$$\Delta G = -k_B T \ln \sum_i e^{-\beta(\Delta G_i^{\text{conf}} + \Delta G_i^{\text{bind}})}$$

PROF. DR. ANDREA
VOLKAMER

OUR ULTIMATE GOAL IS TO DEVELOP TOOLS TO ENABLE FULLY AUTONOMOUS DRUG DISCOVERY

MICHAEL RETCHIN
CBM student



By moving humans out of the DMTA loop, humans can focus on **objectives** and **strategies** across many targets, rather than just **which molecules to make**

Diagram from https://coley.mit.edu/research/

WHAT MAKES US THINK WE CAN AUTOMATE DECISIONMAKING IN DRUG DISCOVERY?

Robots fail to complete Grand Challenge

$1 million prize goes unclaimed

By Marsha Walton
CNN
Thursday, May 6, 2004 Posted: 10:44 AM EDT (1444 GMT)

BARSTOW, California (CNN) -- Nobody won. Nobody even came close.

But that didn't stop organizers of the DARPA Grand Challenge from declaring an unusual race across the Mojave Desert a spirited success.

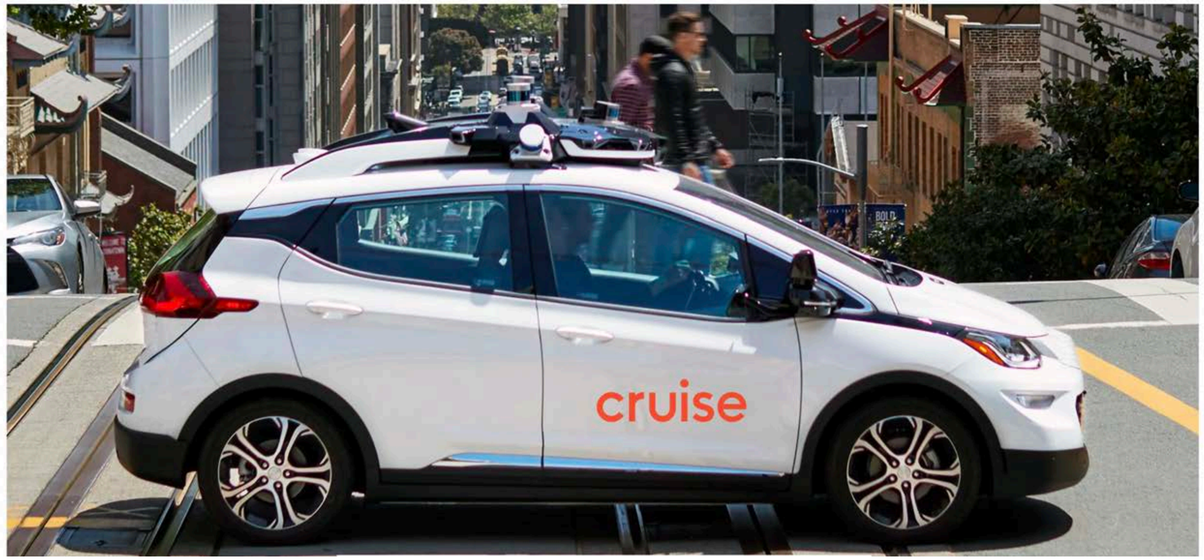
Contextual Link #1
Lorem ipsum dolor sit amet consectateur nonummy lorenzino. Interdum volgus videt, est ubi peccat...
www.contextual_1.com

Contextual Link #2
Lorem ipsum dolor sit amet consectateur

https://www.cnn.com/2004/TECH/ptech/03/14/darpa.race/

GM Cruise takes first fares for paid driverless taxi in San Francisco

Jameson Dow - Jun. 23rd 2022 1:46 pm PT

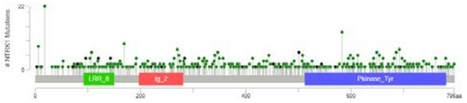


https://electrek.co/2022/06/23/gm-cruise-takes-first-fares-for-paid-driverless-taxi-in-san-francisco/

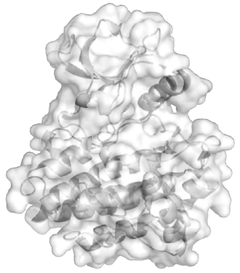
AUTONOMOUS DISCOVERY WILL BE TRANSFORMATIVE IN MUTANT-SELECTIVE KINASE INHIBITOR DISCOVERY

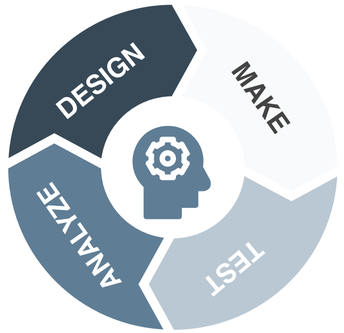
JESSICA WHITE
CBM student

clinical mutant landscape



Target





Autonomous decision-making engine

Preclinical → Phase 1 (safety) → Phase 2 (efficacy) → Phase 3 (effectiveness)
3% patients
$62M
11.1 years

Preclinical → Phase 1 (safety) → Phase 2 (efficacy)

Preclinical → Phase 1 (safety) → Phase 2 (efficacy) → Phase 3 (effectiveness) → NDA (new drug approval)
7% patien
$43M
7.4 y

Preclinical
failure
$3.2M
2.8 years

Preclinical → Phase 1 (safety) → Phase 2 (efficacy) → Phase 3 (effectiveness) → NDA (new drug approval)
13% patien
$112M
9.3 y

Preclinical → Phase 1 (safety)

DECISION POINT WITH OPTIONS

SIMULATE ALL POSSIBLE FUTURES - - - - - - - - - ->

value
total cost
duration

CHODERA LAB

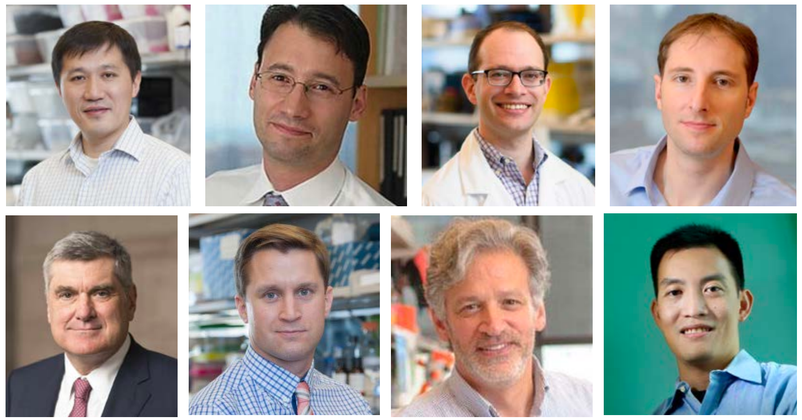
Daniel Parton Bardess Group
Kadeem Ho Sang Gingko Bioworks
Patrick Grinaway Onai
Sarah Boyce Schrödinger
Sonya Hanson Flatiron Institute
Jan-Hendrik Prinz Keylight
Kyle Beauchamp Tempus
Bas Rustenburg SCM
Chaya Stern Odyssey Therapeutics
Rafal Wiewiora Roivant Sciences
Simon Boothroyd Roivant Sciences
Julie Behr Rome Therapeutics
Binisha Karki BioNTech
Corey Taylor Aqemia Therapeutics
William Glass Exscientia
Ivy Zhang
Dom Rufa
Viktor Belay
Michael Retchin
Ben Kaminow
Iván Pulido
David Dotson
Jenke Scheen

Mehtap Isik Moderna
Ashleigh Fowlkes MSKCC
Andrea Rizzi ETH
Josh Fass Relay Therapeutics
Levi Naden Molecular Sciences Software Institute
Steven Albanese Schrödinger
Lucelenie Rodriguez NYU Medical
Ana Silveira Roivant Sciences
Greg Ross Schrodinger
Jiaye Guo Schrödinger
Erin Grundy George Washington University
Hannah Bruce Macdonald Merck
Marcus Wieder Exscientia
David W. H. Swenson Open Free Energy Consortium
Liza Casella Kingdom Supercultures
David Schaller Nuvisan
Yuanqing Wang
Talia Kimber
Alex Payne
Melissa Boby
Ellen Mammen
Mike Henry
Jessica White

Code and data available at http://www.choderalab.org

THANKS!

MSKCC COLLABORATORS



Minkui Luo
Daniel Heller
Michael Kharas
Alex Kentsis
Andrew Koff
Andy Intlekofer
Ingo Mellinghof

Robert Benezra
Sarat Chandarlapaty
Richard Kolesnick
Craig Thompson
Derek Tan
Ouathek Ouerfelli
Hans-Guido Wendel

MSKCC COMPUTING

Monica Chakradeo
Annmarie Pacchia
Richard Knospler
Juan Perin

HPC

Start Folding at http://foldingathome.org

FUNDING

NIH National Institutes of Health
NSF
STIFTUNG CHARITÉ
STARR CANCER CONSORTIUM
CYCLE FOR SURVIVAL
VIR
XtalPi
SILICON Therapeutics
OpenEye SCIENTIFIC
PARKER INSTITUTE for CANCER IMMUNOTHERAPY
SCHRÖDINGER.
BAYER
open forcefield consortium
STIFTUNG CHARITÉ
CZ
wellcome
Gerstner FAMILY FOUNDATION
INTERLINE therapeutics

All funders listed at http://choderalab.org/funding