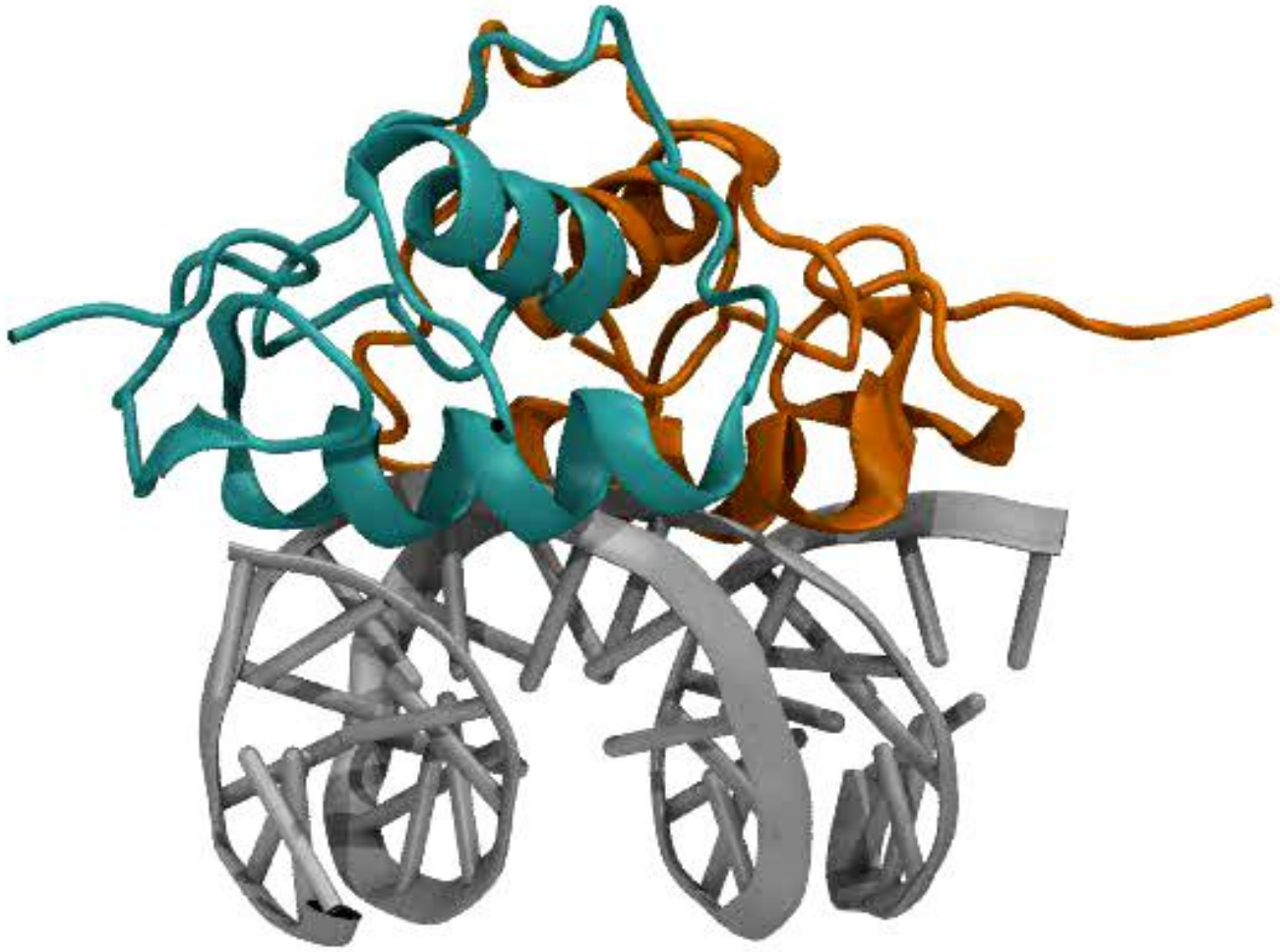
# THE USE OF DEEP DOCKING FOR AUTOMATED, CONSENSUS-BASED HIT IDENTIFICATION IN DRUG DISCOVERY

ARTEM CHERKASOV UBC

# MOLECULAR DOCKING – MAJOR DRUG DISCOVERY TOOL

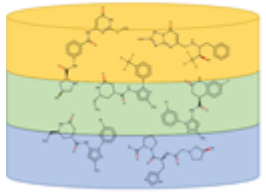


WET-LAB HIGHT-THROUGHPUT SCREENING HIT RATE

IS ~0.03%, DOCKING-SUPPORTED HIT RATE IS 5-20%

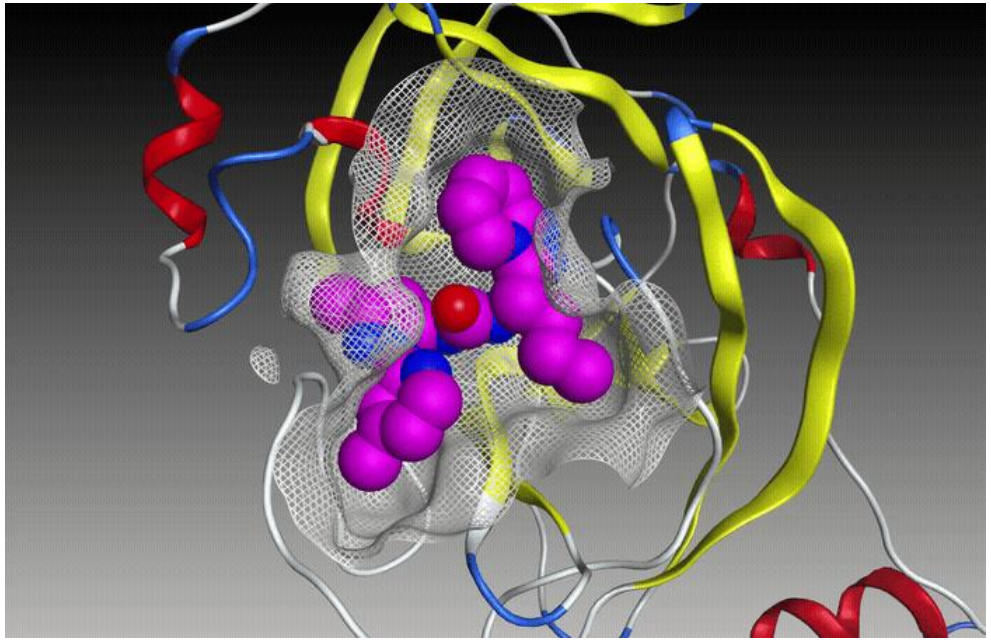
# CONVENTIONAL MOLECULAR DOCKING WORKFLOW

DOCKING DATABASE



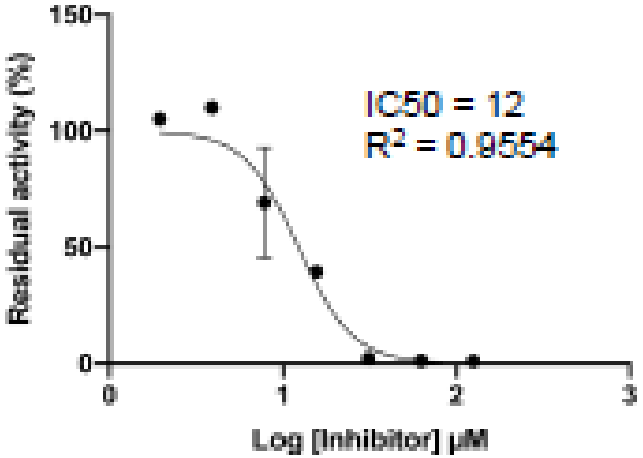
~5-10 MILLION MOLECULES PER TARGET

TARGET PROTEIN/TARGET SITE



DRUG CANDIDATES FOR SCREENING

EXPERIMENTAL ASSAYS



2015
no A.I.
5M Molecules Screened

IN SPORTS

THE VANCOUVER SUN

VRBATA NEEDS A GOOD FEED

Winger struggling without healthy diet of Sedinery » C7

BUSINESSBC C

TUESDAY, DECEMBER 15 | 2015 | BUSINESS COORDINATOR: SCOTT NEUFELD 604-605-2636 | SNEUFELD@VANCOUVERSUN.COM

| S & P/TSX | TSX Venture | Dow Jones | S & P 500 | Dollar | Gold | Oil | Natural Gas |
|---|---|---|---|---|---|---|---|
| 12,695.49 | 896.18 | 17,368.50 | 2,021.94 | 72.79¢ US | 1,064.70 | 38.81 | 1.905 |
| 94.46 | 5.91 | 103.29 | 9.57 | 0.02 | 12.20 | 0.69 | 0.10 |

BRIEFINGS

**Newalta to spend less on growth**

The Newalta environmental services company, which helps recover valuable material from waste produced by the oil and gas industry, says it will spend less on growth projects next year. The Calgary-based company said Monday it has about $50 million of underutilized equipment in inventory. As a result, it will reduce its budget for growth capital to between $20 million and $30 million in 2016, down from $70 million this year. It will also spend $10 million on maintenance capital, the same as 2015.

**Oil, gas weakness curbs Trinidad**

Trinidad Drilling Ltd. has chopped its initial capital spending budget for 2016 to $30 million — 84 per cent less than what it's spending this year — to reflect weak conditions in the oil and gas industry. The Calgary-based company says it's primarily aiming to maintain Trinidad's current operations, although it may be able to spend as much as $45 million if certain growth opportunities arise — still 76 per cent below 2015 levels.

**Pepsi revamps vending tactics**

PepsiCo Inc., facing an anti-soda backlash and health concerns about snack foods, is looking for a resurgence in an especially hard-hit part of the industry: vending machines. The company is rolling out several thousand dual-temperature machines that offer both food and drinks under the new Hello Goodness brand, according to a statement. The units will include healthier products from Pepsico's beverage and Frito-Lay divisions.

**Ferrari designer shares tank**

The Mahindra industrial group based in Mumbai, India, announced Monday it had reached a deal to buy a controlling stake in the Italian design firm Pininfarina, most famous for its designs for Ferrari, Alfa Romeo and Maserati. Shares in the 85-year-old Italian company dropped by nearly 70 per cent on the news, to close at $1.44 US. Under the deal, two of the Mahindra group's units will buy 76.06 per cent of the design firm for $1.21 US a share.

SCIENCE

Researcher Artem Cherkasov displays a computer model simulation used to develop a new treatment for drug-resistant prostate cancer at the Vancouver Prostate Centre. 'Using computer simulations, we sometimes go through 50 million compounds to find a molecule that will seat in a precise and accurate way,' he says.

# Massive cancer-drug deal one of UBC's biggest to date

## University's record agreement worth $140 million — and counting

RANDY SHORE

A promising new treatment for drug-resistant prostate cancer developed by scientists at the University of B.C. has been licensed by the pharmaceutical giant Roche for more than $140 million, the university's richest intellectual property deal in its history.

Researchers Paul Rennie and Artem Cherkasov of the Vancouver Prostate Centre designed the drug to treat prostate cancers

*"As much as that sounds — and it is a lot — the real money is in the royalties, which could exceed $140 million by quite a bit."*

BRAD WHEELER
TECHNOLOGY TRANSFER MANAGER, UBC'S UNIVERSITY-INDUSTRY LIAISON OFFICE

Wheeler, technology transfer manager at the University-Industry Liaison Office and lead negotiator on the Roche licence.

The scientists will share 50 per cent of the net revenue

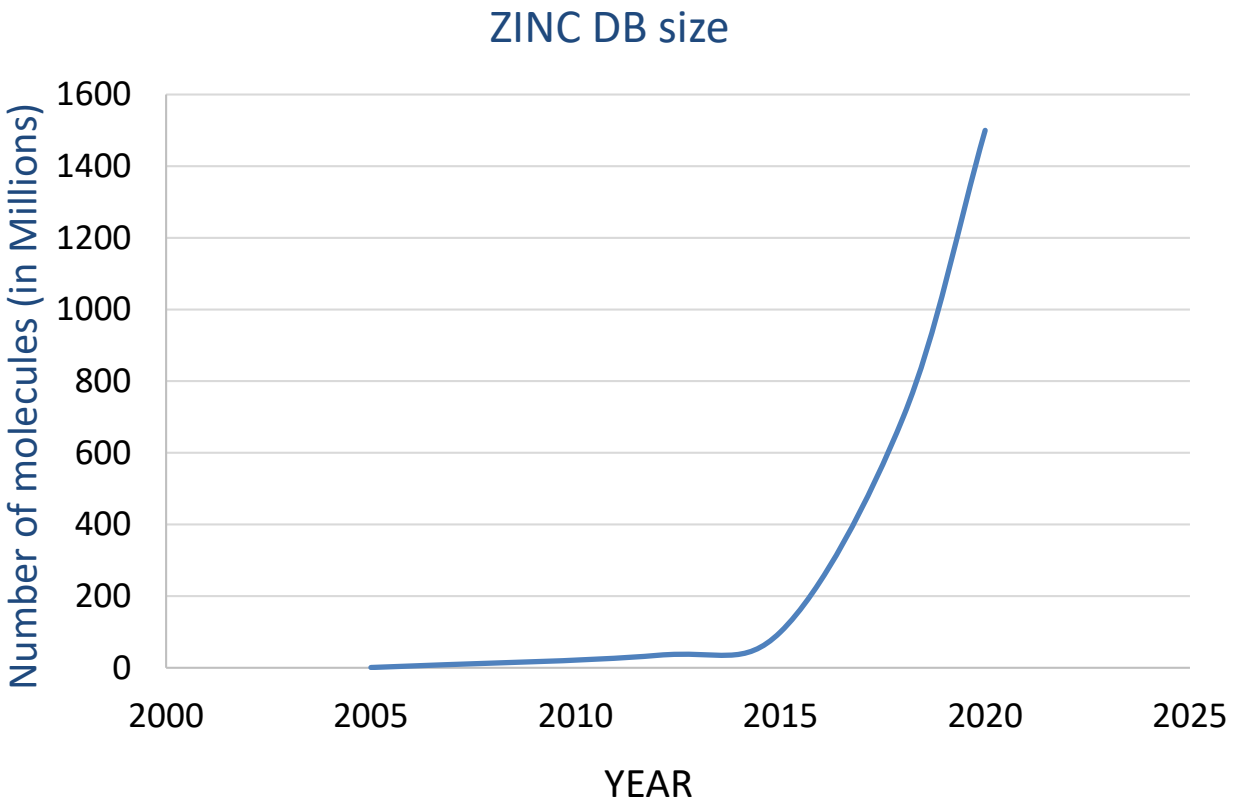
Prostate cancer is the most prevalent and potentially lethal cancer affecting men, with 24,000 new cases leading to more than 4,000 deaths per year, according to the Canadian

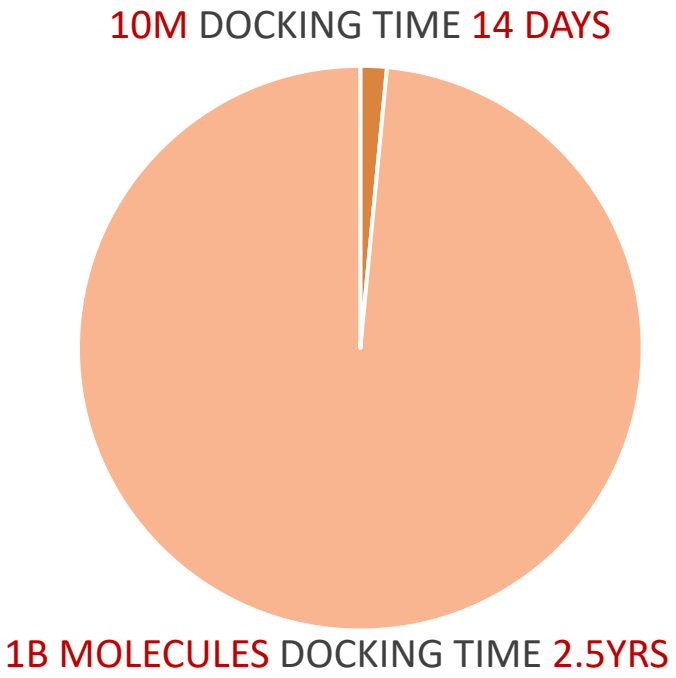
The researchers found the three-dimensional shape of their target location in previous research and set about looking for a molecule that would shut it down.

"Drugs and proteins work like a key in a lock, so we have to find the perfect key for the existing lock," said Cherkasov. "Using computer simulations, we sometimes go through 50 million compounds to find a molecule that will seat in a precise and accurate way."

Cherkasov's powerful search technique produced about 200 candidate molecules, which they

# CHEMICAL SPACE REMAINS INACCESSIBLE TO DRUG DISCOVERY




DOCKING CANNOT KEEP UP WITH EXPLODING CHEMICAL SPACE

CURRENTLY ENAMINE RS DATABASE CONTAINS 38 B MOLECULES

DOCKING MISSES OUT 99.9% OF ALREADY AVAILABLE MOLECULES

TOTAL NUMBER OF POSSIBLE DRUG-LIKE MOLECULES : $10^{60}$ - $10^{100}$

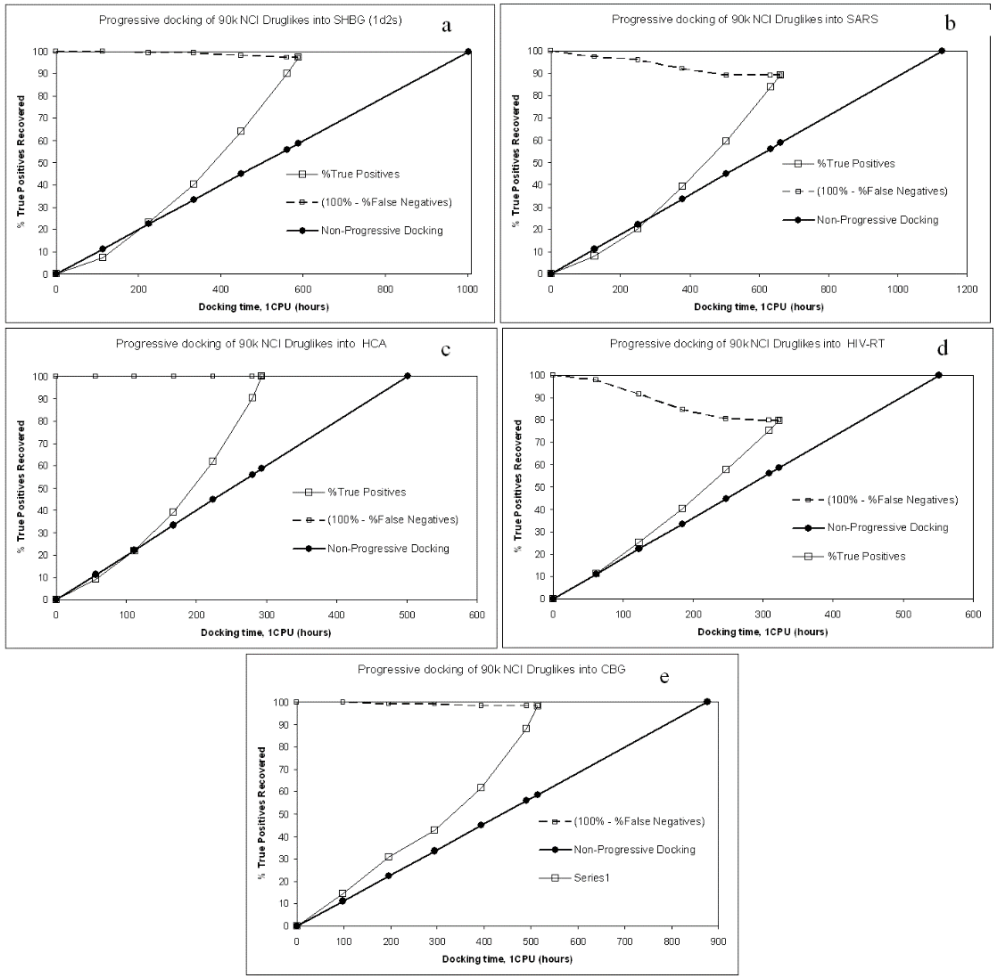
# Progressive Docking 1.0

J Med Chem. 2006 Dec 14;49(25):7466-78.

**Progressive docking: a hybrid QSAR/docking approach for accelerating in silico high throughput screening.**

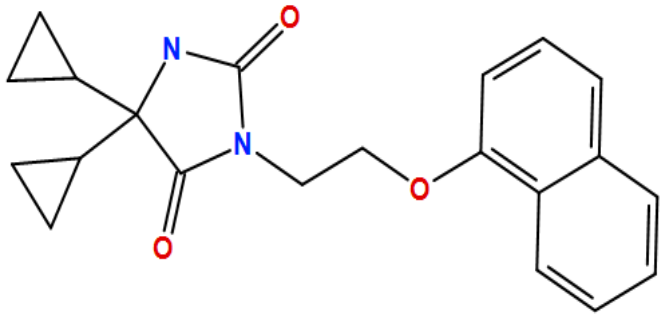
Cherkasov A[1], Ban F, Li Y, Fallahi M, Hammond GL.

Author information

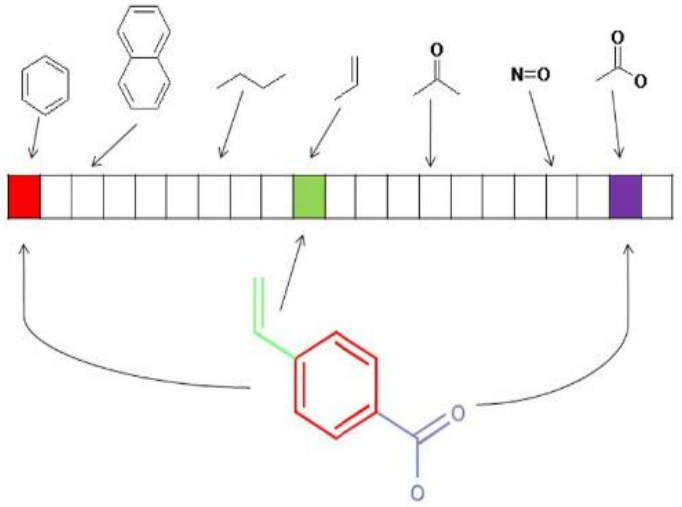


# WHAT IF WE PREDICT DOCKING SCORES WITHOUT DOCKING??

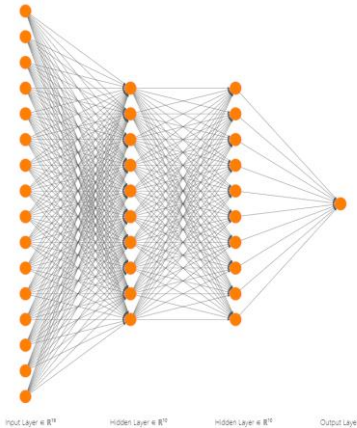
Molecule




Fingerprint



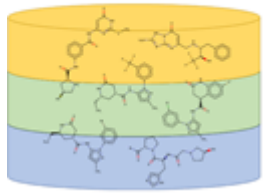

Algorithm



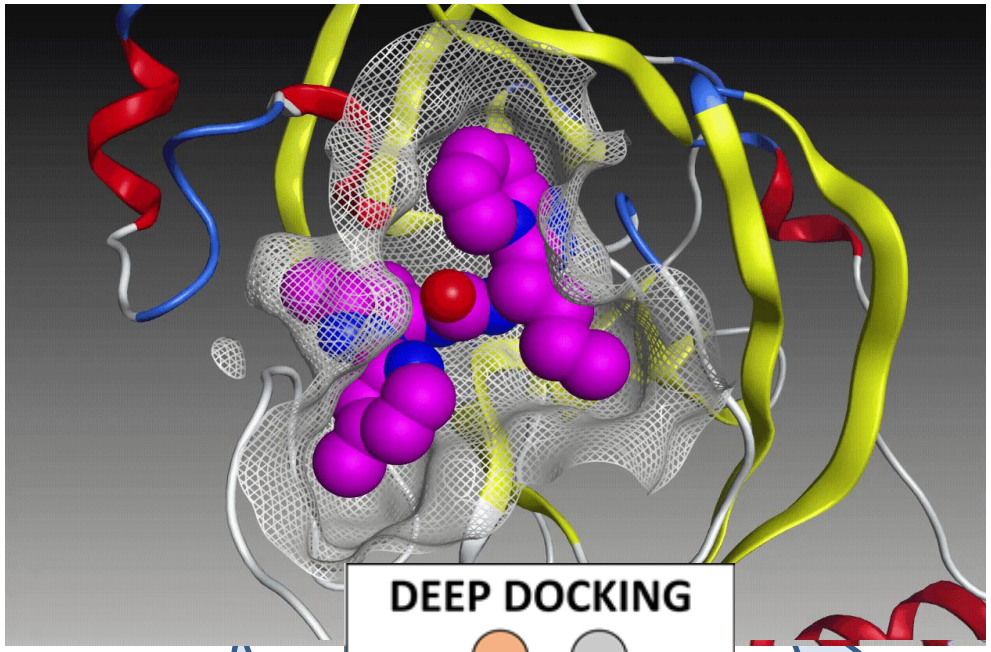
# DEEP DOCKING

TARGET PROTEIN/TARGET SITE

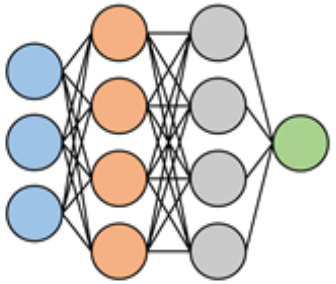
DOCKING DATABASE



1.4 BILLION MOLECULES PER TARGET



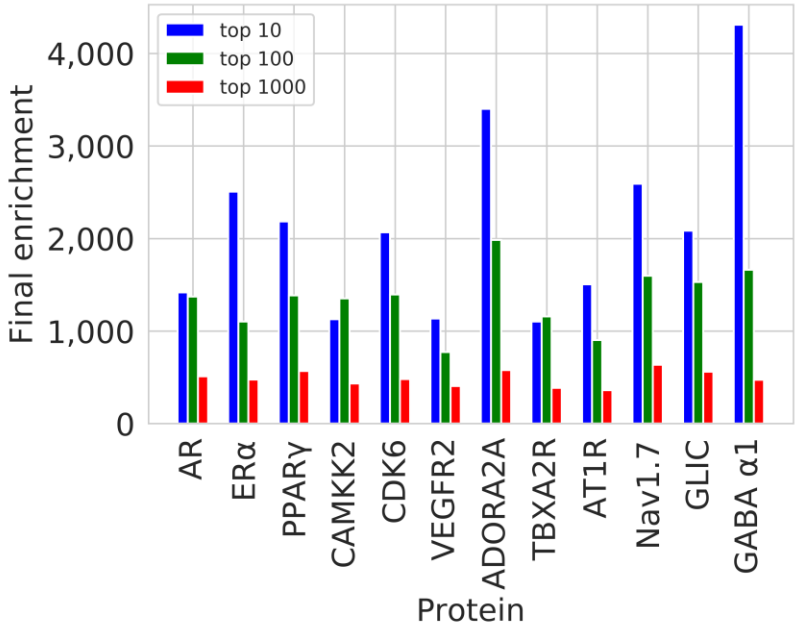
DEEP DOCKING



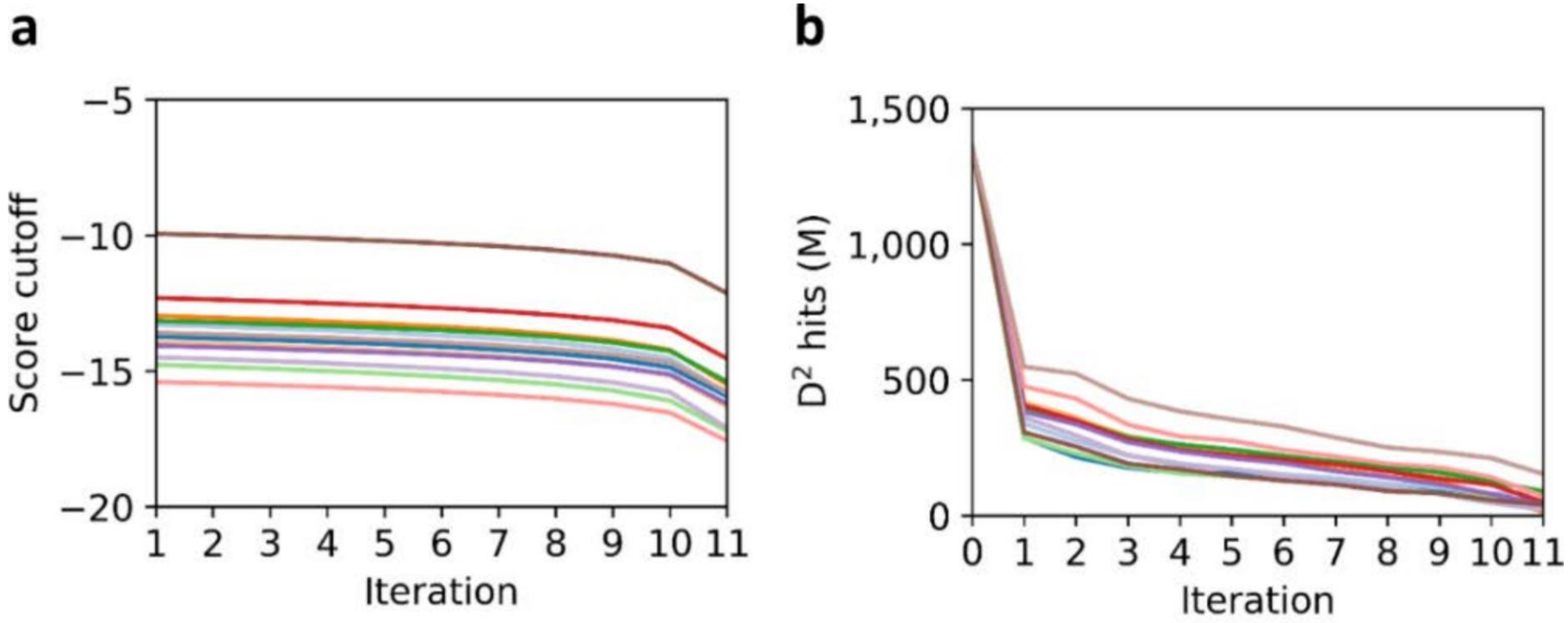
*Predict scores with QSAR models*

100s X FASTER

DRUG CANDIDATES



# DEEP DOCKING PERFORMANCE ON 12 MAJOR DRUG TARGETS

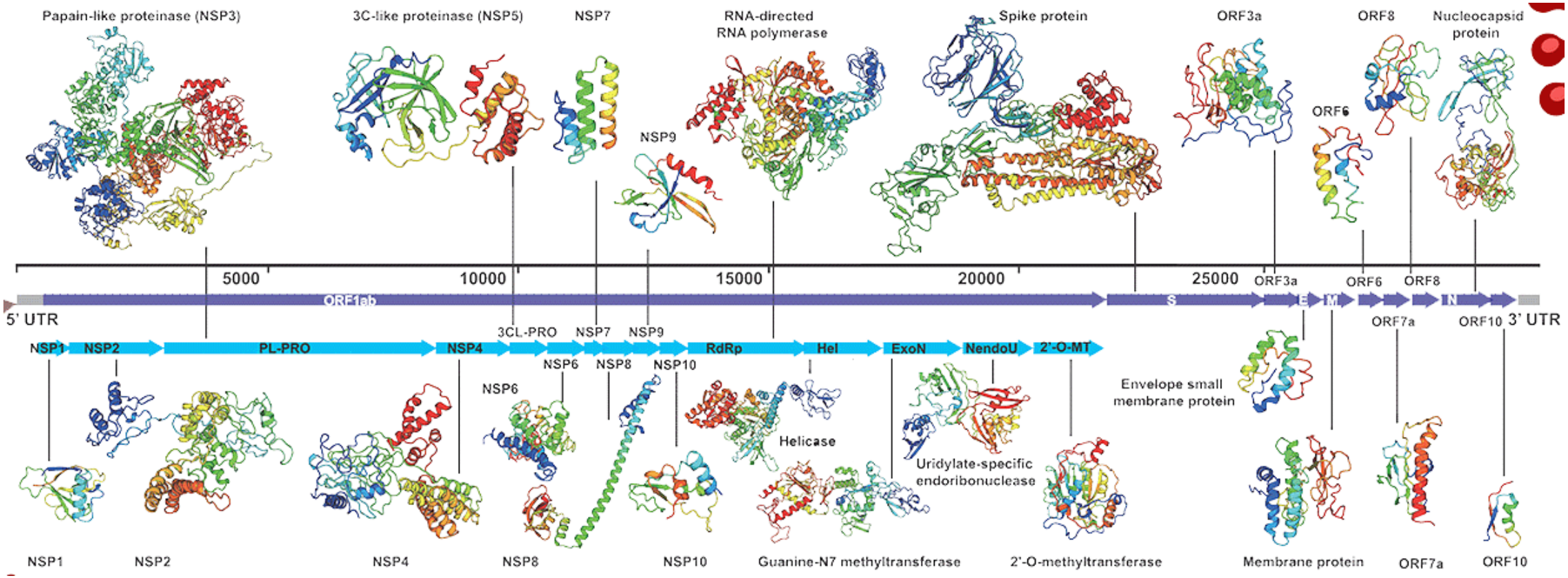



PREDICTED HIGH SCORING MOLECULES AUGMENT THE TRAINING SET OF THE MODEL (1% IN TOTAL)

ACTIVE/INACTIVE CUT-OFF TO IS MADE MORE STRINGENT AT EVERY ITERATION

NR OF MOLECULES PREDICTED AS VIRTUAL HITS AFTER EACH ITERATION IS REDUCED

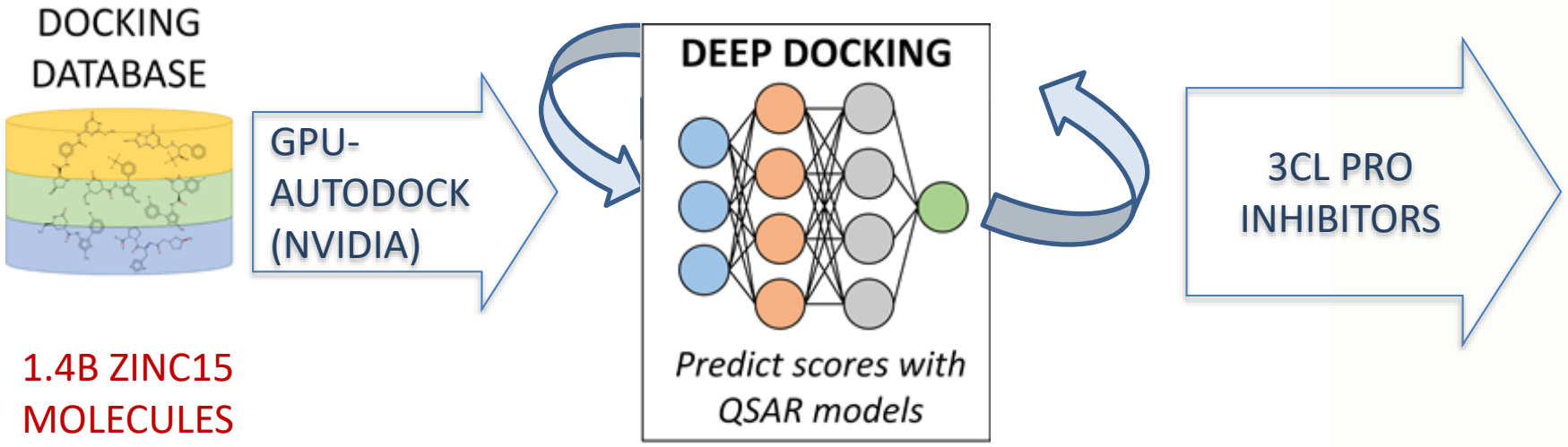
# TARGETTING SARS-CoV-2 GENOME WITH DEEP DOCKING



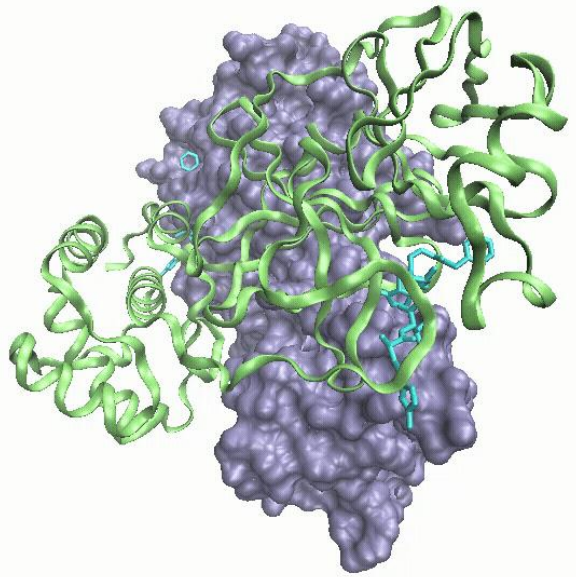

*Zhang et al,, 2020*

WE HAVE IDENTIFIED 25+ POTENTIALLY DRUGGABLE SITES ON VIRAL PROTEINS OF SARS-COV-2. SELECTED 3CL PROTEASE AS MAIN TARGET FOR INITIAL DEEP DOCKING WITH 1.4B ZINC15 COMPOUNDS

# DEEP DOCKING FOR SARS-CoV-2 DRUG DISCOVERY




1.4B ZINC15 MOLECULES



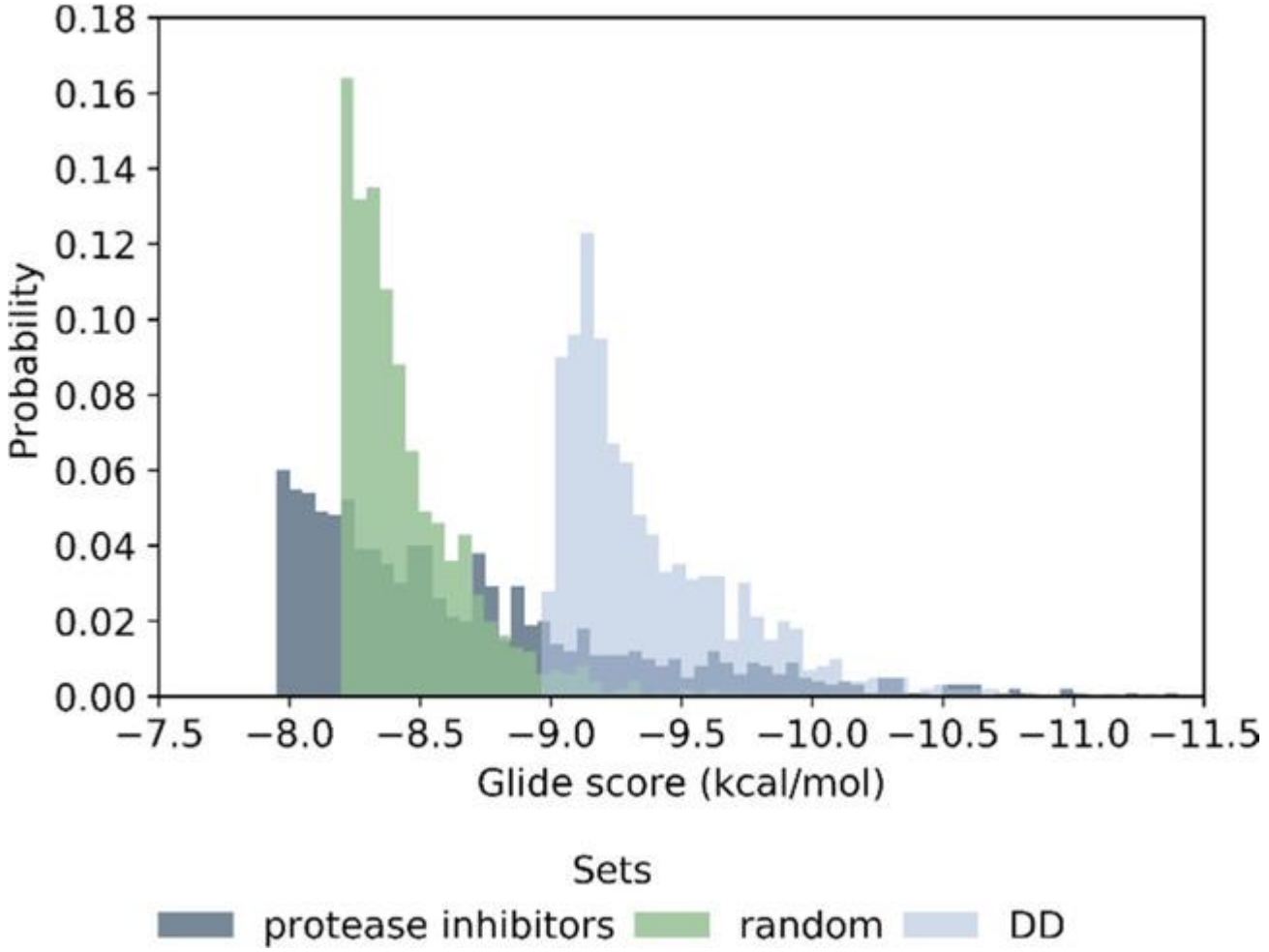
SARS-COV-2 3CL PROTEASE

# DEEP DOCKING IDENTIFIED 585 POTENTIAL 3CL PRO INHIBITORS

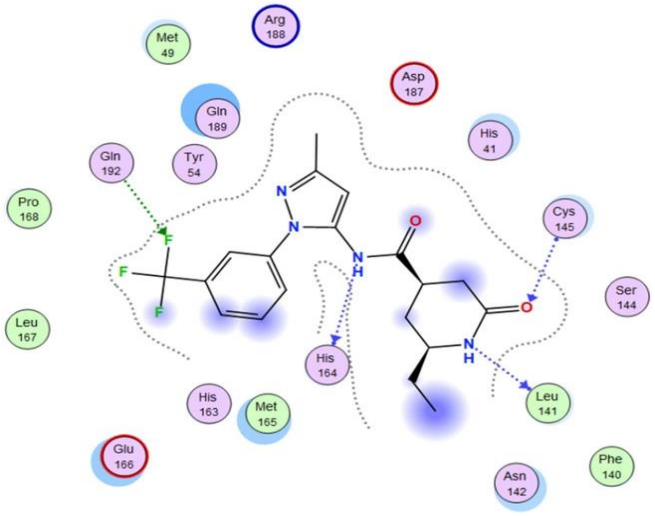
DOCKING SCORES OF TOP 1,000 CANDIDATES SIGNIFICANTLY BETTER THAN OF KNOWN BENCHMARKS

IDENTIFIED 585 UNIQUE SCAFFOLDS FOR SARS-COV-2 3CL PRO, NOT SHARED WITH KNOWN PROTEASE INHIBITORS.

IDENTIFIED 30+ ACTIVES, SOME CONFIRMED BY INDEPENDENT LABS



# 30+ INHIBITORS OF 3CL PRO ENZYME ARE CONFIRMED ACTIVE

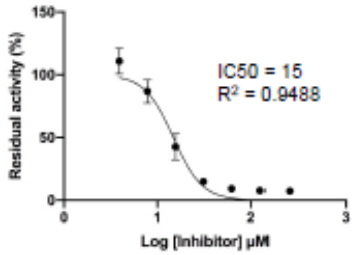


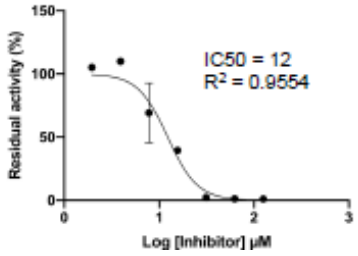

OUR FIRST PUBLICATION WITH INITIAL DRUG CANDIDATES

AGAINST COVID19 APPEARED AS EARLY AS FEB19, 2020

OUT OF 585 PREDICTED COMPOUNDS 30+ ACTIVE (5%)

WET-LAB SCREENING HIT RATE IS USUALLY ~0.03%




molecular informatics
models – molecules – systems

Full Paper | Free Access

## Rapid Identification of Potential Inhibitors of SARS-CoV-2 Main Protease by Deep Docking of 1.3 Billion Compounds

Anh-Tien Ton, Francesco Gentile, Michael Hsing, Fuqiang Ban, Artem Cherkasov

First published: 11 March 2020 | https://doi.org/10.1002/minf.202000028 | Citations: 88

# DEEP DOCKING ENABLES ROUTINE SCREENING >1B LIBRARIES

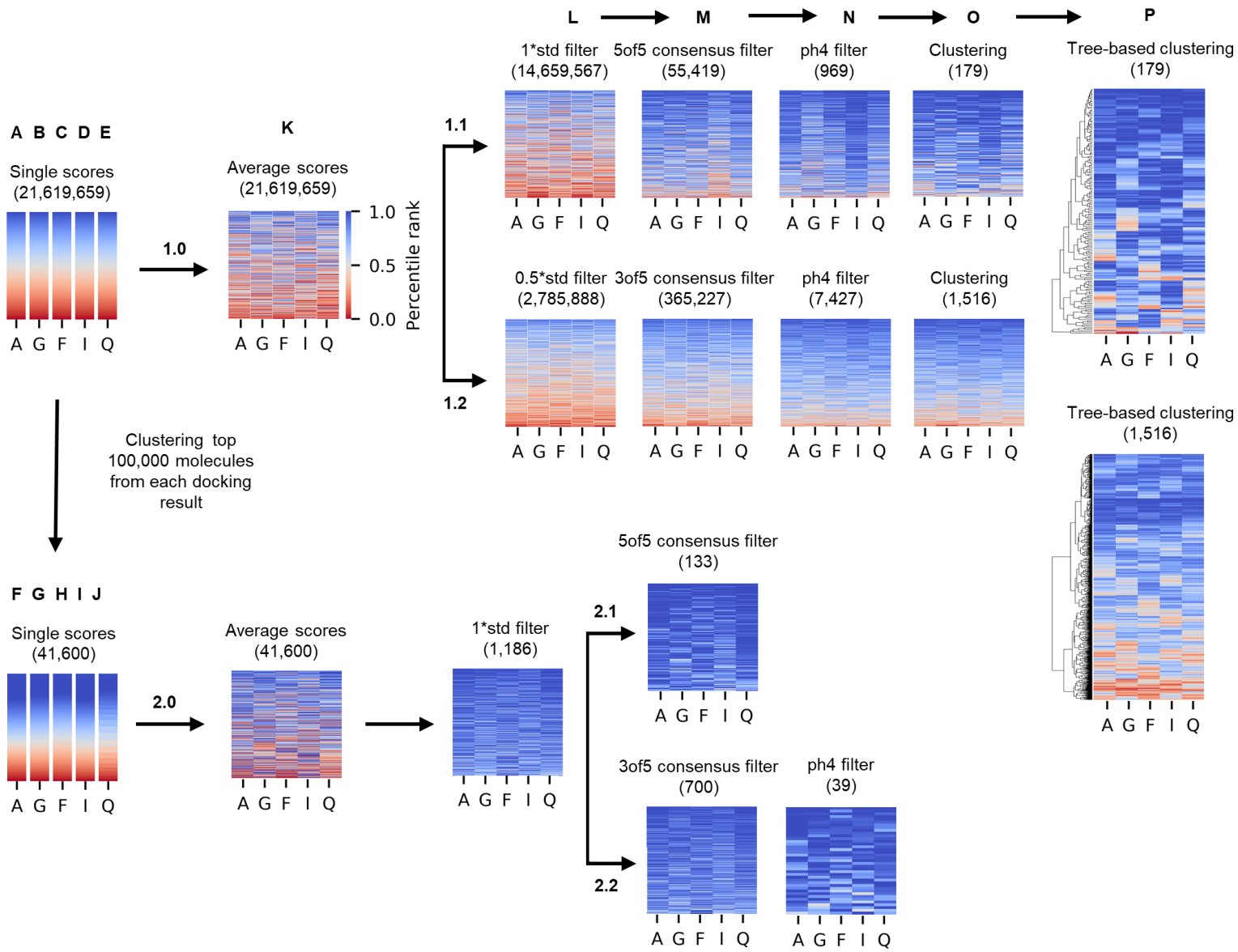
LARGER DOCKING LIBRARIES YIELD BETTER AND MORE HITS (LYU ET AL, NATURE, 2019)

PUBLICLY AVAILABLE CHEMICAL LIBRARIES KEEP EXPLODING: ZINC20 (1.6B), ENAMINE REAL (1.6B)

FEW METHODS PUBLISHED AFTER OUR FEB20 PUBLICATION ON SCREENING 1B+ ULTRA LARGE LIBRARIES

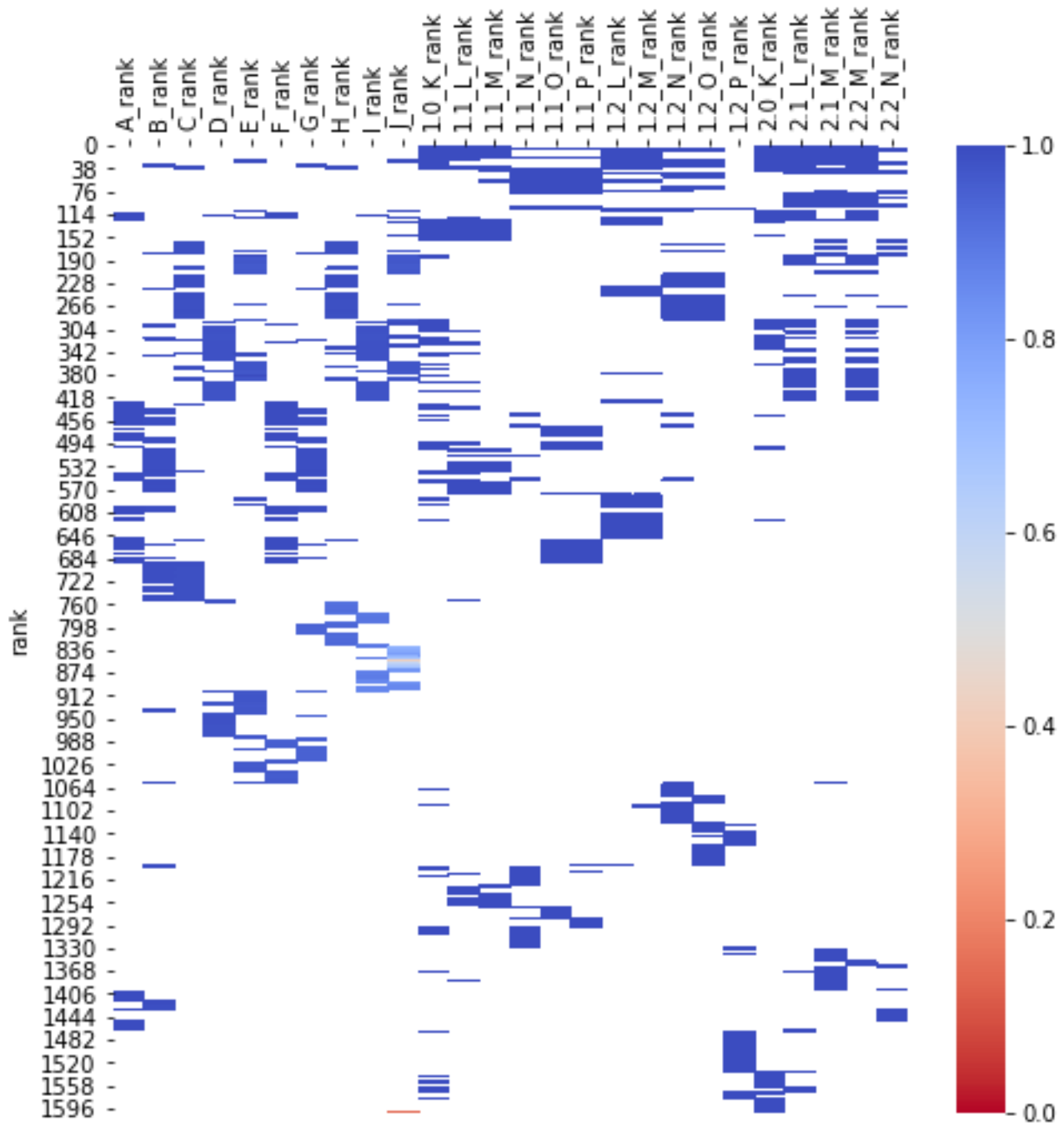
| METHOD | DATABASE | REQUIRED TIME | RESOURCES | DOCKING PROGRAM | TARGET | REFERENCE |
|---|---|---|---|---|---|---|
| OPENEYE ORION | REAL | 1 DAY | 45,000 CORES | FRED | PNP/HSP90 | HTTPS://WWW.EYESOPEN.COM/ORION |
| AUTODOCK-GPU | REAL | 1 DAY | 27,600 GPU | AUTODOCK-GPU | SARS-COV-2 MPRO | ACHARYA ET AL, CHEMRXIV, 2020 |
| VIRTUALFLOW | REAL | 4 WEEKS | 8,000 CORES | QUICKVINA, VINA, … | KEAP1-NRF2 INTERACTION | GORGULLA ET AL, NATURE, 2020 |
| DEEP DOCKING | ZINC15 | 5 WEEKS | 60 CORES, 4 GPU | FRED, GLIDE | MULTIPLE TARGETS | GENTILE ET AL, CENTRAL SCIENCE, 2020 |

# Stringent consensus docking as hitlist filters

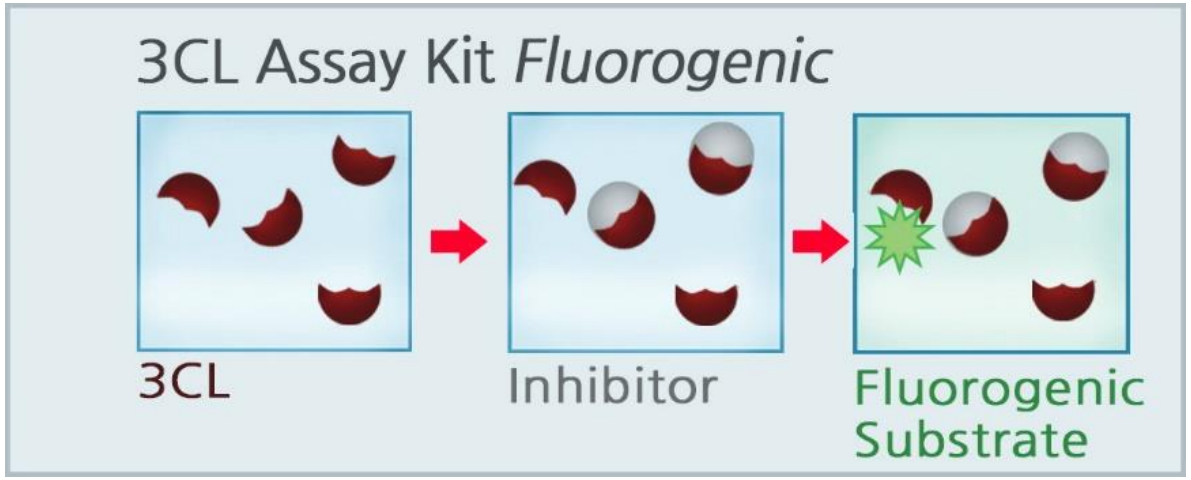


# SELECTED 1700 CANDIDATE MRO INHIBITORS (ROWS) FROM 26 HITLISTS (COLUMNS)

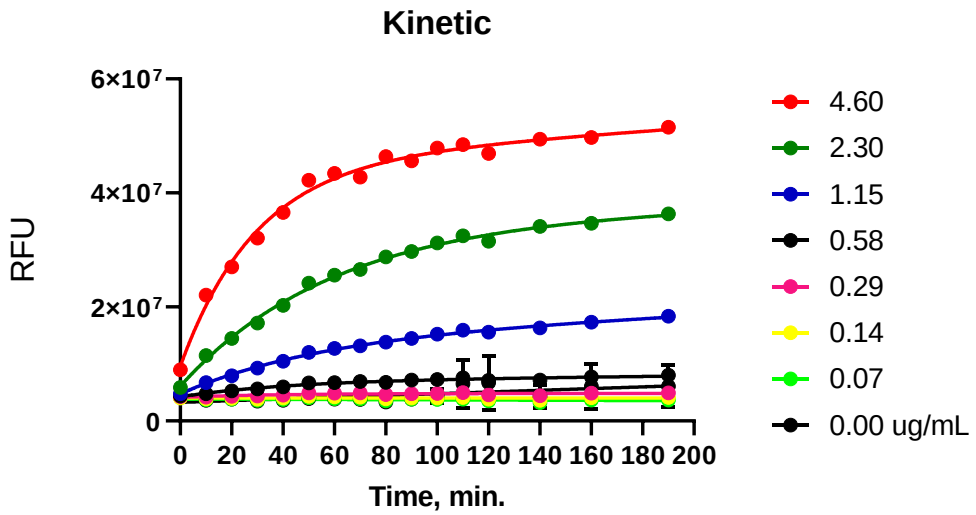
GOAL: COMPLETELY AUTOMATED SELECTION. NO HUMAN INSIGHT

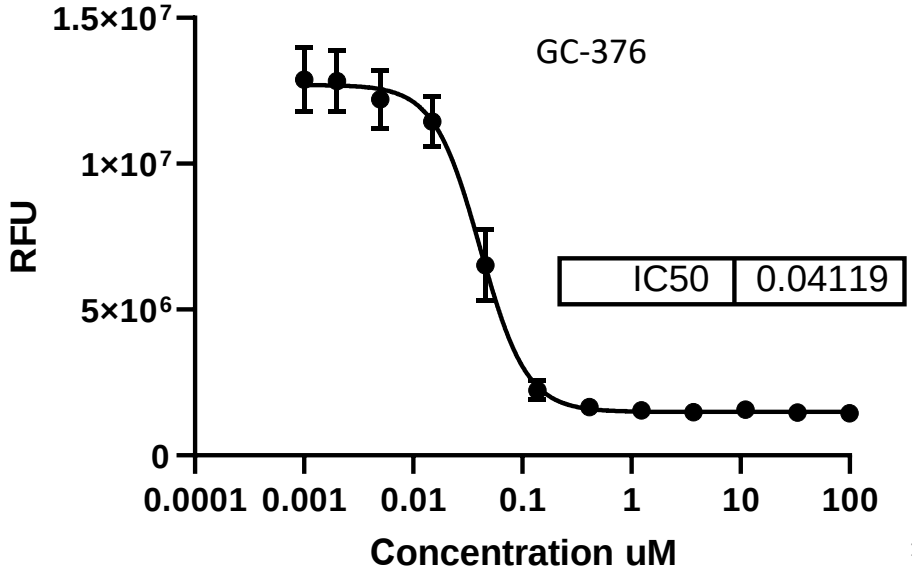


# High throughput screening (HTS) IQFS assay




- Contracted Bienta (Enamine Biological Services) for HTS
- Used similar substrate
- Automated 384-well plate screening
- GC376 IC50 is similar to Jean's lab result

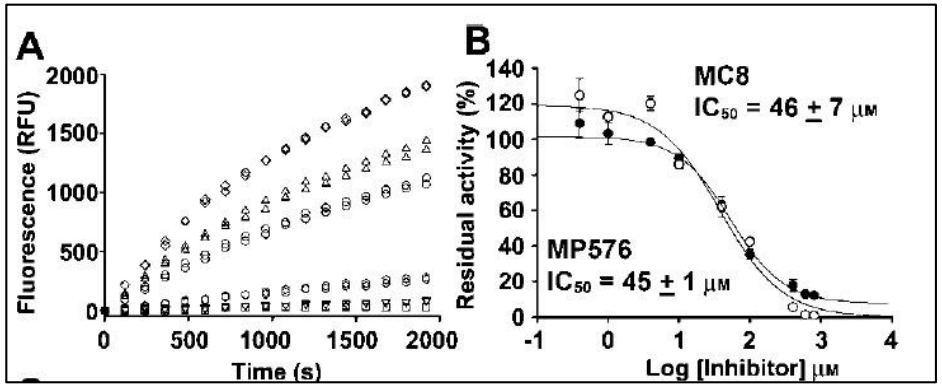


# INTEGRATED EVALUATION PIPELINE FOR SARS-CoV-2 3CL PRO

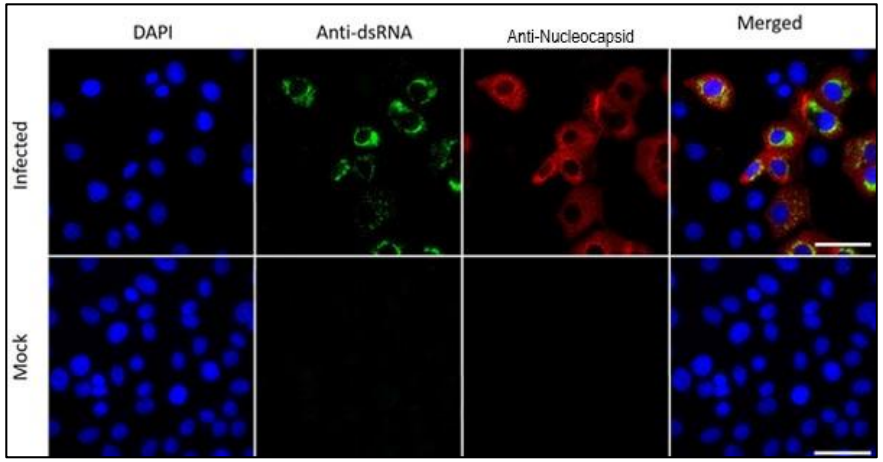
MAIN PROTEASE CATALYTIC ASSAY

CELLULAR ASSAYS CONDUCTED IN CL3/BSL3 UBC FINDER

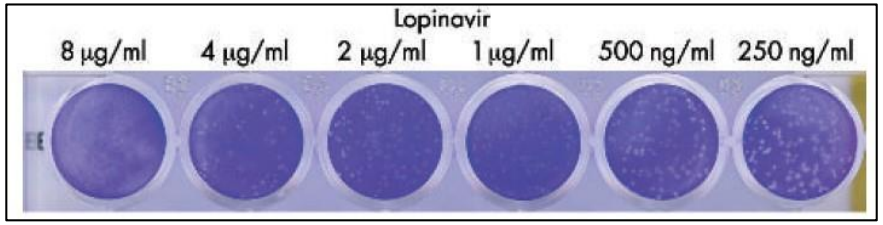



DOUBLE-STRANDED RNA ASSAY: MEASURES SARS-COV-2 VIRAL RNA DURING LIFE CYCLE

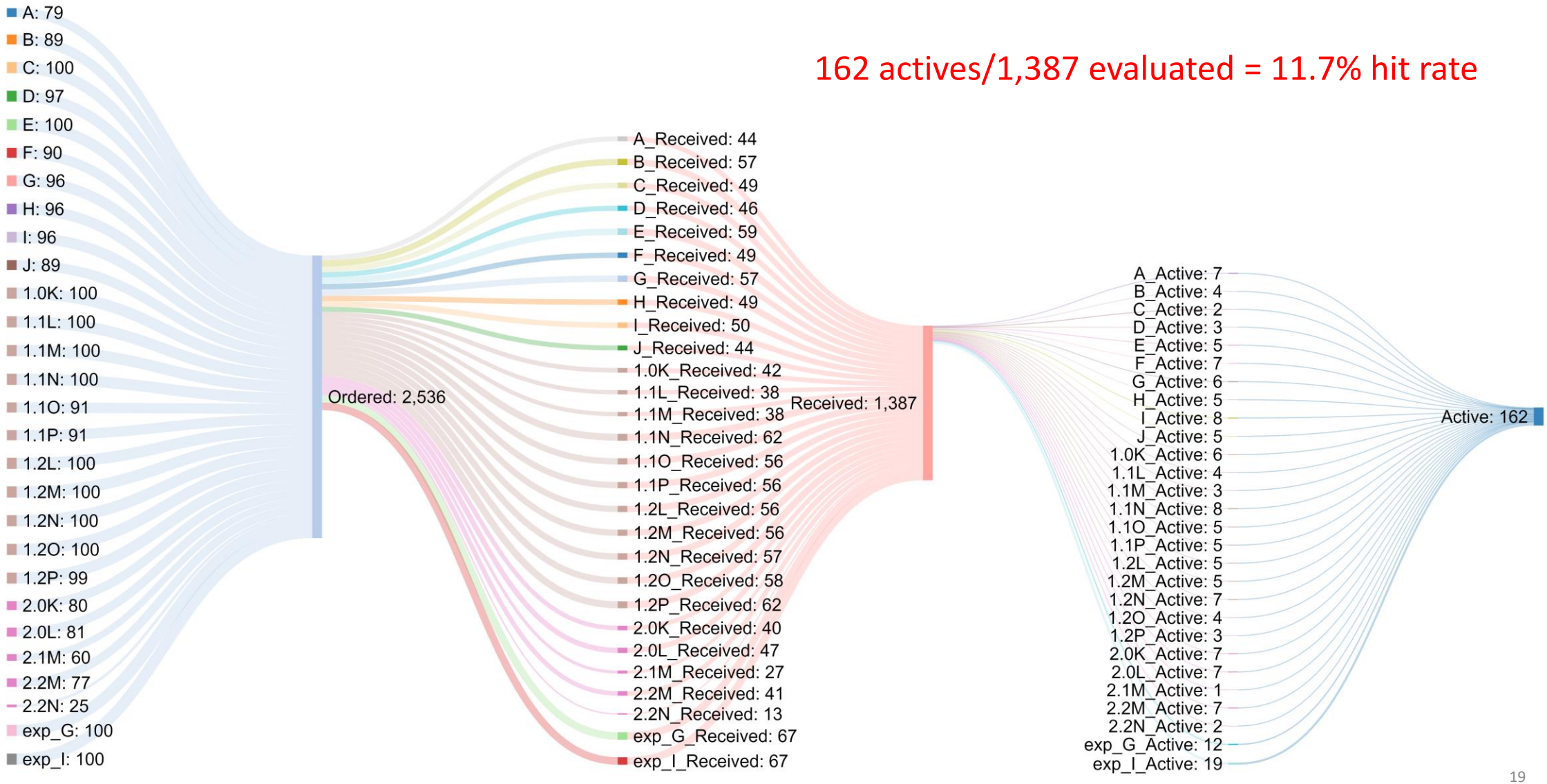
NUCLEOCAPSID ASSAY: MEASURES SARS-COV-2 VIRAL PROTEIN PRODUCTION DURING LIFE CYCLE




PLAQUE ASSAY: MEASURES INFECTION/BUDDING OF SARS-COV-2 VIRUS DURING LIFE CYCLE



# 40B DD hit rate (5 µM; 10% inhibition scored as active)




162 actives/1,387 evaluated = 11.7% hit rate

# **PREPARED - 2021** MOTIVATION

in **2005** on the emergence of SARS-1 outbreak we created an integrated scientific platform PREPARE-2005 aiming to rapidly respond to future SARS-like infectious threats


This project will use one overall approach to uncover the biology of infection of such serious diseases as SARS, influenza, West Nile, BSE, pathogenic E. coli, tuberculosis, malaria and HIV/AIDS. The approach consists of identifying microbial drug targets through the study of protein interaction networks and the application of innovative computational genomics. Protein interaction networks are complex - they are involved in catalytic processes, protein synthesis and gene expression within the cell.

The research team will share experimental approaches to study different pathogens and use whole-genome approaches to investigate common pathogens. This new knowledge base will be particularly valuable in the event that new infectious agents emerge – new strains of existing pathogens, for example, or previously unknown pathogens.

The research project will create new opportunities for the pharmaceutical and biotechnology industries, and will also maintain a rapid response team of highly competent genomics researchers, ready to find scientific solutions for new infectious threats as they arise.

# WE PROPOSE TO ESTABLISH GLOBAL PATHOGEN - DEFENSE SYSTEM

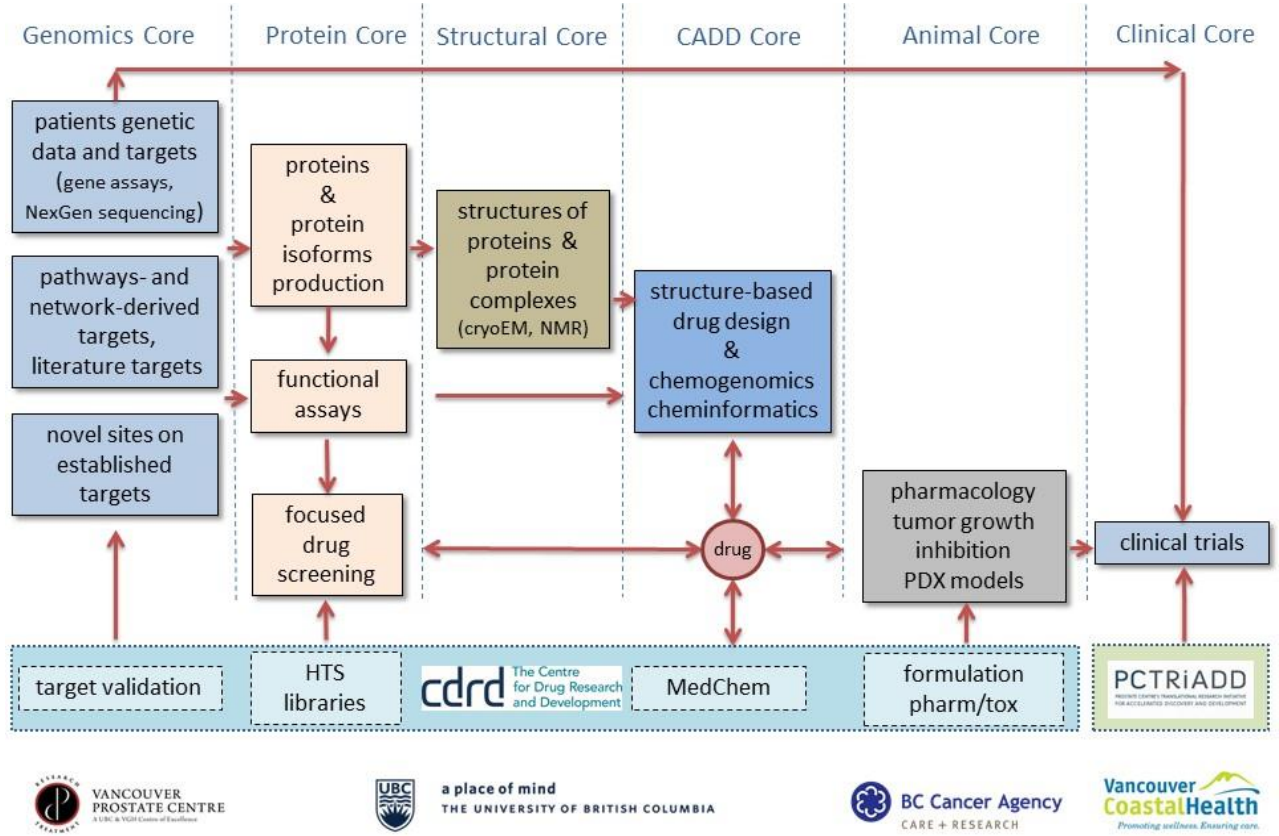
## ON THE BASIS OF

## **PREPARED-2023** PROJECT

INFRASTRUCTURE INCLUDES ALL LATEST TECHNOLOGICAL ADVANCEMENTS

AI CAN SPEED UP CADD CORE WORKFLOW 100-S FOLDS

CAN ALSO SIGNIFICANTLY IMPROVE THE OVERALL WORKFLOW PERFORMANCE



# Acknowledgements


- Cherkasov Lab
  - Dr. Art Cherkasov
  - Dr. Francesco Gentile
  - Dr. Fuqiang Ban
  - Dr. Michael Llamosa
  - Dr. Michael Hsing
  - Dr. Eric Leblanc
  - Dr. Carl Perez
  - Helene Morin
  - Hazem Mslati
  - James Gleave
  - Jean Charle Yaacoub
  - Jane Foo
  - Mohit Pandey
  - Mariia Radaeva
  - Jiaying You
- Young Lab
  - Dr. Robert Young
  - Dr. Michael Bielecki
  - Dr. Jason Smith
- Strynadka Lab
  - Dr. Natalie Strynadka
  - Dr. Liam Worrall
  - Dr. Jaeyong Lee
- Jean Lab
  - Dr. François Jean
  - Dr. Tirosh Shapira
  - Dr. Andrea Omstead
  - Ivan Villanueva
  - Rory Long