# EuroHPC Summit Week 2022

## Using Supercomputers for Accelerating Drug Design

Andrew Emerson, CINECA Supercomputing Centre, Italy.

From **22** to **24 March 2022** | **Paris**, France

#EHPCSW #PRACEdays

# The problem with drug design…

Billion dollar investment in bringing to market

Low success rate – typically only 12% that enter clinical development are approved *

Lead development time ~10 yrs

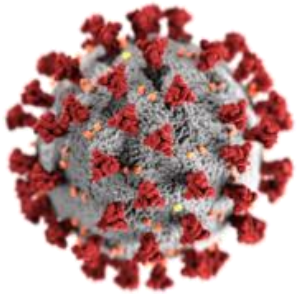
Uneven coverage – few new antivirals and antibiotics

* https://engitix.com/technology/

The EuroHPC Summit Week conference series receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 823767 (PRACE-6IP), 824151 (HPC-GIG).

# COVID-19

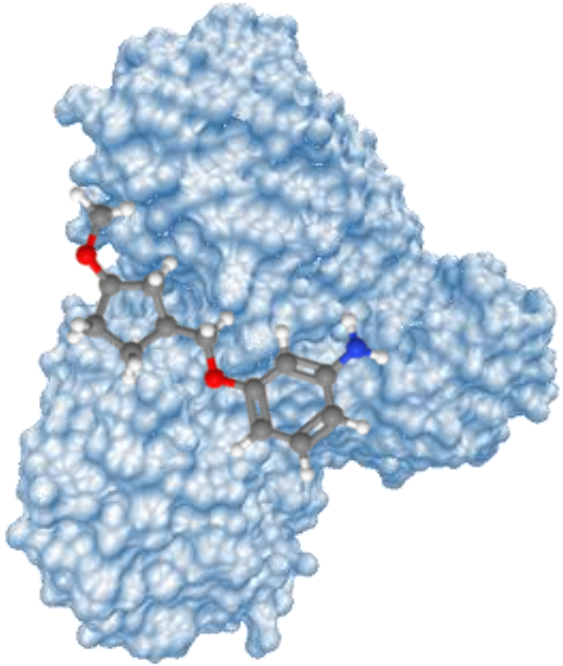
- Rapid vaccine development for COVID-19 is a remarkable success story.
- Much harder to find successful drug therapies.
- Supercomputers around the world were asked to help with drug design



The EuroHPC Summit Week conference series receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 823767 (PRACE-6IP), 824151 (HPC-GIG).

# Computational Drug design - background

- Drugs are often small molecules which bind to specific proteins.
- The effect of binding may inhibit the usual activity of the protein.
- For example, binding a molecule to a viral enzyme may stop viral infection

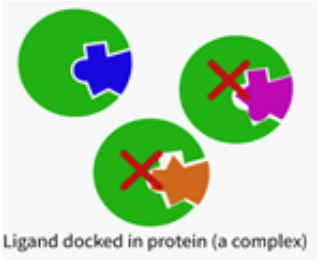


The EuroHPC Summit Week conference series receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 823767 (PRACE-6IP), 824151 (HPC-GIG).

# Virtual Screening

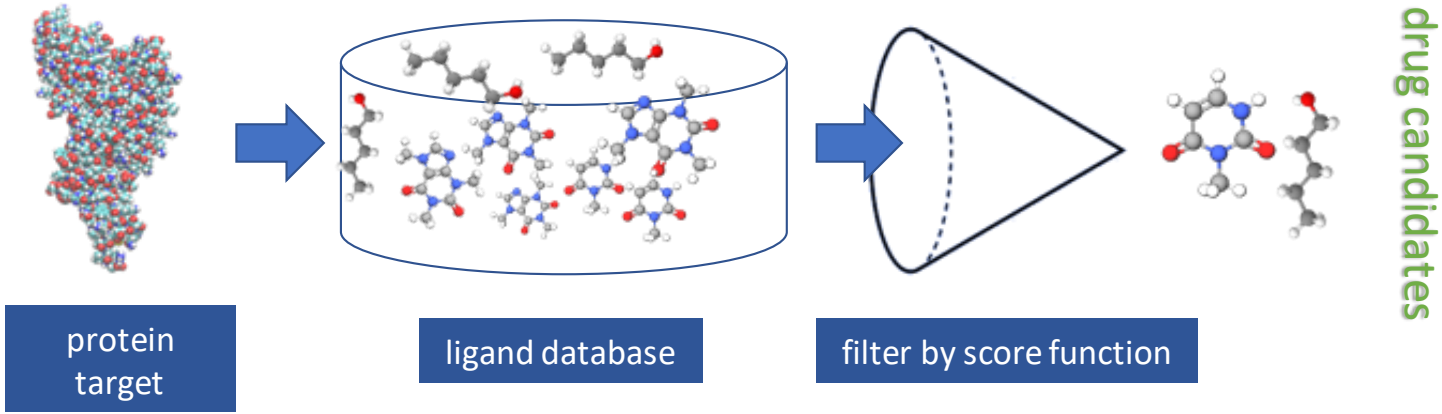
- In virtual screening databases of *ligands* are screened against the target protein.
- At the simplest level, we just look for the right shape.
- Higher levels also consider specific interactions, protein flexibility, etc.
- The docked ligand is assigned a *score*. Highest scoring ligands are studied further.


⇩ *Docking*




The EuroHPC Summit Week conference series receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 823767 (PRACE-6IP), 824151 (HPC-GIG).

# High Throughput Virtual Screening




With massively parallel computers we can rapidly filter extremely large molecule databases (e.g. billions of molecules).

The EuroHPC Summit Week conference series receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 823767 (PRACE-6IP), 824151 (HPC-GIG).

EXSCALATE 4COV

LIGATE

| | |
|---|---|
| Urgent Computing H2020 to find anti-covid19 agents. | EuroHPC JU grant for integrated drug-design solution |
| Jan 2020-June 2021 | Jan 2021-Dec 2023 |
| Multi discpl. approach – bioinformatics, HPC, laboratory, clinical, big data. | Long term focus on integrated platform for drug design. |
| Coordinator: Dompé farmaceutici SpA | |

LIGEN High Throughput Docking system*

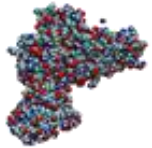
* Dompe' proprietary software, co-developed with Cineca and others.

The EuroHPC Summit Week conference series receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 823767 (PRACE-6IP), 824151 (HPC-GIG).

# Exscalate4Cov computational pipeline


= CINECA M100, ENI HPC5

*Molecule database*


3D protein structures


**LIGEN**
High-Throughput Virtual Screening

In-silico drug molecule candidates

Laboratory in-vitro tests, animal studies

Clinical trials

EXSCALATE 4COV

Raw 3D protein structures from experiments or models


**GROMACS**
10 µs molecular dynamics

The EuroHPC Summit Week conference series receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 823767 (PRACE-6IP), 824151 (HPC-GIG).

# Exscalate4Cov impact

- Initial strategy of *drug repurposing** to rapidly find new drugs.
- Key achievement: **identification of the osteoporosis drug Raloxifene as possibly active against COVID-19.**
- Registered and generic with known safety characteristics (e.g., dosage and side effects).
- In clinical trials (France and Italy), also for vaccinated people.

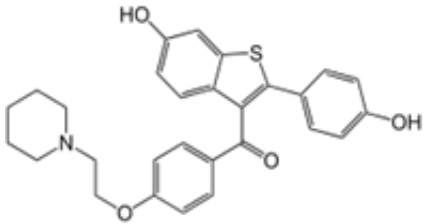
**400k** molecules tested in-silico

**7k** molecule candidates

**100** molecules deemed interesting from in-vitro tests

**40** molecules effective in limiting virus replication

approved drugs or safe in humans




raloxifene

* re-using existing drugs for different conditions.

The EuroHPC Summit Week conference series receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 823767 (PRACE-6IP), 824151 (HPC-GIG).

# The Big Run

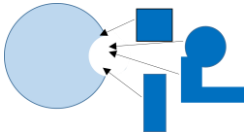
12 viral proteins

70+ billion ligands

60 hours run time
(~ 3M ligands /second per protein)

2 supercomputers (Cineca M100+ Eni HPC5)

15 binding sites for each protein



>1 trillion evaluations

81 Pflops total performance

7.3Tb archive data

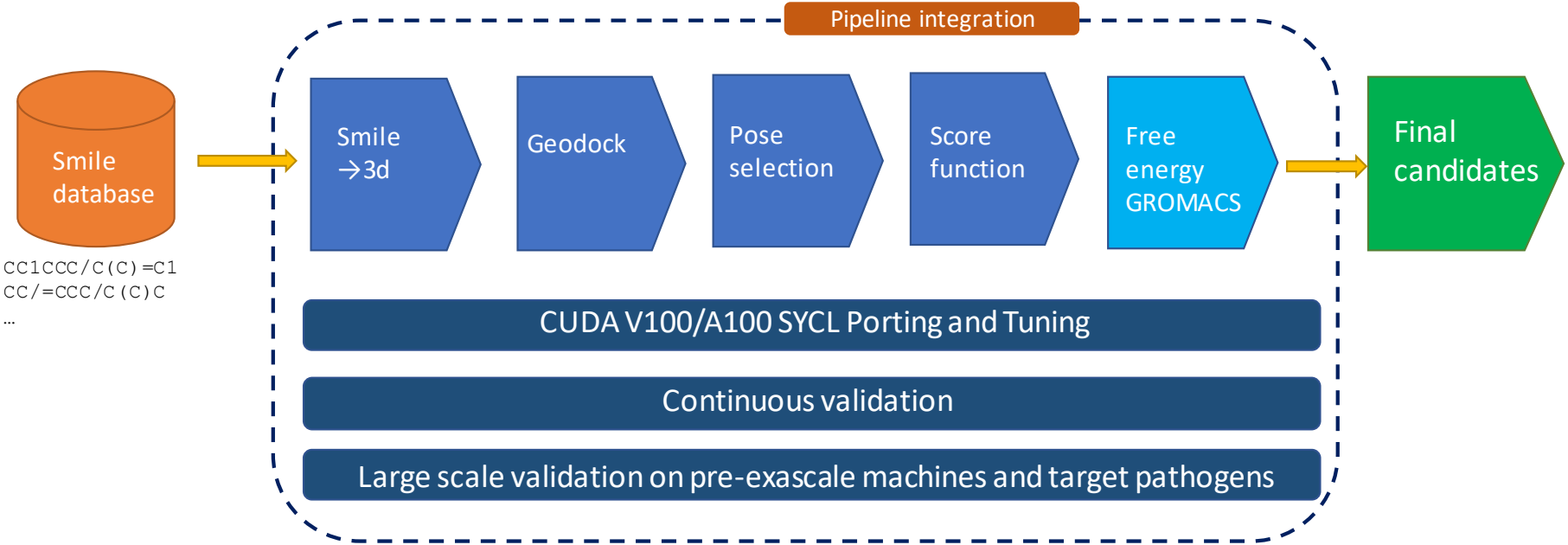
http://arxiv.org/abs/2110.11644

Data still being analysed…

The EuroHPC Summit Week conference series receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 823767 (PRACE-6IP), 824151 (HPC-GIG).

# Ligate Pipeline

LIGATE




The EuroHPC Summit Week conference series receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 823767 (PRACE-6IP), 824151 (HPC-GIG).

# Ligate Components

software

- LIGEN docking system
- GROMACS free energy estimation
- HyperQueue scheduling

libraries

- SYCL
- Celerity

cross validation

- Tof Computer Vision Simulation Use Case

Partners

- 11 partners across Europe

HPC resources

- -Two supercomputers
- -GPU servers

The EuroHPC Summit Week conference series receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 823767 (PRACE-6IP), 824151 (HPC-GIG).

# Acknowledgements

**Exscalate4Cov**

- This research was made possible under the project "***EXaSCale smArt pLatform Against paThogEns for Corona Virus — Exscalate4CoV***" funded by the EU's H2020-SC1-PHE-CORONAVIRUS-2020 call, grant N. 101003551.
- Consortium: Dompe', Politecnico Milano, CINECA, Uni. Milano, KU Leuven, IIMCB, Elettra Sincrotrone Trieste, Fraunhofer, INMI Spallanzani, BSC, FZJ Juelich, Uni. Naples Fed. II, Uni. Cagliari, SIB, KTH, ABD , INFN, Chelonia Applied Science

**Ligate**

- This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 956137. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Italy, Sweden, Austria, Czech Republic and Switzerland.
- Consortium: Dompe', Politecnico Milano, CINECA, KTH, Uni. Salerno, Uni. Innsbruck, E4 Computing, Chelonia Applied Science, tofmotion, IT4I, Uni. Basel.

Big thanks to partners from both projects and in particular CINECA colleagues, F. Ficarelli, C. Latini, G. Frumenzio, A. Piserchia, M. de Pascale, L. Querciagrossa, B. Chandramouli, B. Agarwal, A. Grottesi and N. Besker.

The EuroHPC Summit Week conference series receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 823767 (PRACE-6IP), 824151 (HPC-GIG).

# Thank you for your attention

EuroHPC Summit Week 2022