EXSCALATE4CoV - EXaSCale smArt pLatform Against paThogEns for Corona Virus.

Dompé, EuroCC Italy

25 May 2022

training_eurocc_italy@cineca.it

Industrial HPC Course

Dompé SUCCESS STORY IN DRUG DESIGN INDUSTRIAL DOMAIN

EXSCALATE4CoV – EXaSCale smArt pLatform Against paThogEns for Corona Virus.

The most advanced center of competence aimed at fighting the coronavirus by combining the best supercomputing resources, and AI up to clinical validation.

THE PROBLEM

The E4C project addresses a globally unmet societal and medical need: is there an infrastructure/operating workflow which might help in quickly identifying existing drugs or novel NCE in case of a pandemic outbreak? Virus epidemics are one of the most serious health issues in the world which need to be transacted and the growth in the number of infected cases detected and fatality rapidly alarming the world. Using a solid scientific method was critical for bringing this outbreak under control.

SUCCESS STORY DETAILS

HPC provider: Cineca and ENI
Domain expert: HPC and AI
Country: Italy
Link:
https://www.exscalate4cov.eu/

THE HPC PROBLEM DOMAIN

The exploitation of supercomputer such as Marconi100 (CINECA) and HPC5 (ENI) allowed to apply Computational-Aided Drug Design technologies during Phase 1 of the project, leading to the fast virtual identification of known drugs (repurposing). After the experimental validations, the most promising one, the Raloxifene, is in clinical trials in three EU countries. Given the ongoing health emergency, we have pushed the best hardware and software technologies to the extreme, performing, in the Phase 2, the largest virtual screening experiment ever carried out. More than 70 billion molecules will be simulated on the 15 active interaction sites of the virus for a total of more than a thousand billion interactions evaluated in just 60 hours. This will be possible thanks to the simultaneous availability of the computing power (81 petaflops: millions of billions of operations per second) of Eni's HPC5, the most powerful industrial supercomputer in the world, of CINECA's Marconi100, and the virtual screening software accelerated by the Politecnico di Milano and Cineca, and the Exscalate molecular library from Dompé.

THE SOLUTION

Exscalate4CoV, using a unique combination of high- performance computing power and AI with biological processing, brings together 18 partners and further 15 associated members. This includes supercomputing centers in Italy, Spain and Germany, large research centers, pharmaceutical companies and biological institutes from across Europe.
A private-public consortium aimed to develop a realistic and fast operational workflow to tackle this aspect of scientific answer to pandemic cases was certainly of impact.

THE BENEFITS

The benefits of E4C was twofold:

- to identify molecules capable of targeting the new coronavirus (SARS-CoV2)
- to develop a tool effective for countering future pandemics.

Dompé POLITECNICO MILANO 1863 CINECA EURO

Dompé SUCCESS STORY IN DRUG DESIGN INDUSTRIAL DOMAIN

THE SOLUTION

This result confirms the real possibility of exploiting HPC platforms in cases of pandemics. Also, during the project precious web platforms to support the global research community with bioinformatics and simulation tools were deployed, including MEDIATE – MolEcular DockIng AT home (https://mediate.exscalate4cov.eu/) – will give free access to the largest database available today on the Sars-CoV-2 Virus both from a structural (three-dimensional structures) and functional (proteins interacting with human cells) point of view. In https://viralseq.exscalate4cov.eu/, the viral mutations retrieved using genomic data from public repositories (i.e. GISAID, EMBL COVID-19 data portal) are mapped and analysed in their 3D structural context to investigate their impact in terms of host-immune interaction, ligand/substrate/drug binding sites and SDPs; The other web portal is https://spikemutants.exscalate4cov.eu/, that aims to provide the scientific community with structural information on emerging variants involving the protein sequence of the Sars-CoV-2 Spike protein; The other impressive release will be http://SCoV2-MD.org, a website that include MD trajectories retrieved worldwide and analysis tools to better understand the dynamic behaviour of viral proteins.

THE BENEFITS

More specifically, E4C aims to:

- **Establish a sustainable example for a rapid scientific answer to any future pandemic scenario.** The model leverages a rapid and effective High Performance Computing platform for the generation and analysis of 3D models and experimental 3D X-ray structures of protein targets from pandemic pathogens.
- **Drive a fast virtual identification of known drugs (repurposing) or proprietary/commercial candidate molecules** to be further experimentally characterized.
- **Define a workflow scheme for biochemical and cellular screening test** to validate the candidate molecules in previous points and assure, through phenotypic and genomic assays;
- Prepare, together with EMA, a **development plan for successful candidates for direct "first-in-human" studies** or for further testing in animals for bridging studies.
- **Identify SARS-CoV-2 genomic regions** involved in host adaptation, pathogenicity and mutations.

EURO