Using Supercomputers for Accelerating Drug Design

Andrew Emerson

The COVID-19 pandemic has highlighted the need to accelerate the rate at which new antiviral therapies are brought to market. Computational methods based on supercomputers are being increasingly used to cut the lead development time, and hence reduce the massive financial investment required to find new drug molecules. A common strategy involves the use of high throughput virtual screening to filter large databases of potential drug molecules (termed ligands) against a target protein associated with the disease. A good example of where this approach has been applied is the EU-funded multi-disciplinary Exscalate4Cov project (April 2020 - September 2021) which was designed with the aim of fighting the SARS-COV-2 virus. The multi-disciplinary team, led by Dompe' pharmaceuticals, have used the Ligen software to screen billions of ligands against selected target proteins from the virus. The software has been highly optimised and shows a high parallel efficiency with the result that in the largest run it was able exploit all the nodes of the CINECA M100 and Eni HPC5 supercomputers simultaneously, at the time the two most powerful supercomputers in Europe. In this experiment, over a trillion molecules were screened in just 60 hours, equivalent to processing over 4.5 million ligands per second. In a related project called LIGATE, the computational pipeline at the heart of Exscalate4Cov is being enhanced with molecular dynamics tools, task schedulers, new AI engines and improved facilities for managing large datasets. An important asset of the project will be program code designed to run on any CPU or GPU architecture.

The utility of high throughput virtual screening has been demonstrated by one of the outcomes of the Exscalate4Cov project – the identification of Raloxifene, a generic osteoporosis drug, as a potential inhibitor of SARS-COV-2 virus replication. Raloxifene is currently in clinical trials for both immunised and non-immunised patients.

It has been predicted that pandemics like COVID-19 may become more frequent in the future, but we believe that the HPC -based tools we have described here will allow us to react more quickly, as well make new progress against other pathogens which are currently endemic in many parts of the world.