A long-timescale, ensemble study of SARS-CoV-2 protein-ligand interactions

Art Hoti[1], Agastya Bhati[1], Peter Coveney[1]

[1] Centre for Computational Science, Department of Chemistry, University College London, London, WC1H 0AJ, United Kingdom

Abstract

There is an urgent need for drugs which target SARS-CoV-2, the pathogen responsible for the current coronavirus pandemic. In this regard, a concerted global effort has led to a rapid rise in the number of SARS-CoV-2 protein structures available in the PDB, rendering the virus increasingly susceptible to rational, structure-based drug discovery.

Crucially, this critical situation provides an explicit example which demonstrates the necessity to develop robust, reproducible molecular dynamics (MD) based protocols for the elucidation of ligand binding mechanisms, sites, and interactions. Though molecular dynamics has seen rapid advancements in the ability to shed light into the nature of protein-ligand interactions, trajectories derived from the technique are intrinsically chaotic in nature, meaning that trajectories which are initiated from arbitrarily close regions of phase space will diverge exponentially over time. This causes results derived from individual simulations to be non-reproducible and therefore unreliable. Nevertheless, the current consensus is that individual 'long' (microsecond scale) simulations suffice for the accurate elucidation of biomolecular processes such as protein-ligand binding.

Here, using a series of five protein-ligand systems which contain crucial SARS-CoV-2 targets: 3CLpro, PLpro and ADRP, we demonstrate that this is not the case. By running ensembles of long simulations at an unprecedented scale, we provide clear evidence of the abundance of kinetic traps that are encountered in microsecond long trajectories, showing that even in simulations of this timescale, a system is likely to sample a mere sub-set of the overall set of possible ligand binding sites. As such, we demonstrate that it is only by running ensembles of many independent trajectories that we are able to overcome the inherent stochastic uncertainty of these systems in order to obtain reproducible and precise results.

Furthermore, by subjecting these simulations to extensive analysis, we aim to elucidate binding pathways, bindings sites and allosteric mechanisms of inhibition that are not experimentally determined. As such, providing insights that will be of significant benefit to anti-SARS-CoV-2 drug discovery.