Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2022-0046

Vol. **33**, No. 1, pp. 57-76
January 2023

Multiscale Hybrid Modeling of Proteins in Solvent: SARS-CoV2 Spike Protein as Test Case for Lattice Boltzmann – All Atom Molecular Dynamics Coupling

Marco Lauricella[1,*], Letizia Chiodo[2,*], Fabio Bonaccorso[3], Mihir Durve[4], Andrea Montessori[5], Adriano Tiribocchi[1], Alessandro Loppini[2], Simonetta Filippi[2] and Sauro Succi[4]

[1] *Consiglio Nazionale delle Ricerche, Istituto per le Applicazioni del Calcolo IAC-CNR, 00185, Rome Italy.*
[2] *Engineering Department, Campus Bio-Medico University, 00128, Rome, Italy.*
[3] *Physics Department and INFN, University of Rome "Tor Vergata", 00133 Rome, Italy.*
[4] *Istituto Italiano di Tecnologia, Center for Life Nano- & Neuro-Science@Sapienza – IIT, 00161, Rome, Italy.*
[5] *Engineering Department, Universitá degli Studi Roma TRE, via Vito Volterra 62, Rome, 00146, Italy.*

Received 15 February 2022; Accepted (in revised version) 2 June 2022

Abstract. Physiological solvent flows surround biological structures triggering therein collective motions. Notable examples are virus/host-cell interactions and solvent-mediated allosteric regulation. The present work describes a multiscale approach joining the Lattice Boltzmann fluid dynamics (for solvent flows) with the all-atom atomistic molecular dynamics (for proteins) to model functional interactions between flows and molecules. We present, as an applicative scenario, the study of the SARS-CoV-2 virus spike glycoprotein protein interacting with the surrounding solvent, modeled as a mesoscopic fluid. The equilibrium properties of the wild-type spike and of the Alpha variant in implicit solvent are described by suitable observables. The mesoscopic solvent description is critically compared to the all-atom solvent model, to quantify the advantages and limitations of the mesoscopic fluid description.

AMS subject classifications: 82M37, 92C05, 35Q20

Key words: SARS-CoV-2, biophysics, lattice Boltzmann, molecular dynamics.

*Corresponding author. *Email addresses:* marco.lauricella@cnr.it (M. Lauricella), l.chiodo@unicampus.it (L. Chiodo)

1 Introduction

In the last decades, continuous methodological and technological progress paved the successful path of computational tools in fighting contagious diseases by providing in silico simulations of biological molecules and drug design. Further, the graphics processing units (GPUs) granted significant technological progress, boosting the computational power over large system sizes. In this context, we use a multiscale modeling approach that couples a mesoscale solvent representation to the molecular dynamics. The final aim is to deliver an efficient biophysical computational strategy to boost the simulation power and understand the biological mechanisms at the atomistic level, avoiding the computational effort due to the atomistic description of the solvent.

Nowadays, the molecular dynamics (MD) method has shown its massive potential in characterizing the biological mechanisms underlying the activities of several proteins at the atomistic level. Remarkable examples in computational biophysics are the recent simulations of an entire cell organelle, a photosynthetic chromatophore vesicle from a purple bacterium [1] or the study of the N-Methyl-D-Aspartate (NMDA) neuroreceptor by the DE Shaw research group [2]. As of today, an enormous scientific effort has been spent to investigate in-silico the molecular behavior of SARS-CoV-2 proteins, both for drug repurposing and for antibody design [3]. Standard MD simulations have been used, for example, to estimate binding free energies of spike in interaction with the human angiotensin-converting enzyme 2 (ACE2) receptor [4–7] alongside with their interaction scores [8]. The effort devoted to SARS-CoV-2 proteins, with the exceptional focus on its spike protein, somehow put the spotlight on the strengths and limitations of bioinformatics and biophysics computational tools [9, 10] in the field of medicine and drug discovery, bringing these tools also to the attention of the general public.

The widely known and investigated spike protein is here used as a test case to highlight strengths and drawbacks of a mixed multiscale description scheme. Indeed, despite the all-atom molecular dynamics description being the method of election to properly describe the biochemical nature of protein functioning, one of the main issues in its usage is related to the long time scales of biological mechanisms, and on their rare-event nature from a statistical mechanics perspective. Normally, the quaternary movements associated with the allosteric and functional response of biological mechanisms lie in several microseconds, beyond the standard actual high-performance computational limits to obtain a statistically meaningful description [11, 12], even by exploiting optimised codes for GPUs clusters [13, 14]. While a possible solution relies on using enhanced sampling techniques [15], the large simulation size involved in most realistic mechanisms cannot be taken easily into account. Thus, in the last three decades, the development of coarse-grained models has shown great scope in overcoming these limits [16, 17]. The coarse-grained strategy usually aims at reducing the details of the protein structures alongside their aqueous solvent. Such a reduction shall be made with particular care to preserve the detailed description, where necessary to appropriately describe the protein structure and function [18–21].