# Journée MASIM 2020
# SARS-CoV-2 et Bioinformatique Structurale

Organisateurs : Frédéric Cazals (Inria Sophia)
et Yann Ponty (CNRS/Ecole Polytechnique)

19 Novembre 2020

## Programme

| | |
|---|---|
| 8h45 – 8h50 | Frédéric Cazals & Yann Ponty<br>*Présentation de l'évènement* |
| 8h50 – 9h15 | Stéphane Télétchéa (UFIP, Université de Nantes)<br>*COVID-19 : past, present and future* |
| 9h15 – 9h40 | Edoardo Sarti (LCQB, Sorbonne Université)<br>*Statistical analysis of coevolutionary signals on the SARS-CoV-2 Spike protein* |
| 9h40 – 10h05 | Yasaman Karami & Laura Ortega Varga (Institut Pasteur)<br>*Targeting SARS-CoV-2 Spike Glycoprotein to fight Covid-19* |
| 10h05 – 10h30 | Frédéric Cazals (Inria Sophia)<br>*Spikes of SARS-CoV-2 : mechanisms and visualization* |
| 10h30 – 11h00 | **Pause (café ?)** |
| 11h00 – 11h25 | Jean-Philip Piquemal (LCT, Sorbonne Université)<br>*High-Resolution Modeling of SARS-CoV-2 Proteins Structural Dynamics* |
| 11h25 – 11h50 | Maria Kadukova (Nano-D, Inria Grenoble)<br>*Re-docking of the available MPro co-crystal ligands* |
| 11h50 – 12h15 | Stephane Redon (OneAngstrom)<br>*SAMSON - Integrated Molecular Design* |
| 12h15 – 12h40 | Marc Baaden (CNRS IBPC)<br>*FAIR sharing of molecular visualization experiences : COVID-19 use case* |
| 12h40 – 13h05 | Sebastian Will (LIX, Ecole Polytechnique)<br>*Computational RNA-RNA interaction prediction for elucidating template switching in SARS-CoV-2* |

# Abstracts

### Stéphane Télétchéa (Université de Nantes)

*COVID-19 : past, present and future*

**Abstract:** 2020 saw the emergence of a new pandemic virus spreading so rapidly that we had no other disease-control strategy than to confine humans. The exact virus origin is still under scrutiny, but the knowledge acquired in less than one year on the virus' biology has ramped up at a rate never seen before. In particular, the structural characterization of the COVID-19 genome and proteome has allowed a giant leap into the understanding of the virus organization and function. This enormous deciphering work led to a nearly complete characterization of the atomic protein structures. Yet, after a careful genome analysis, surprises about the number of open reading frames encoded in the RNA still happened. The presentation will start with a rapid introduction about the virus biology and the associated diseases, then will follow with the virus genome and proteome description. Ongoing efforts for fighting against the virus with (new) chemical entities or vaccines will be highlighted the discuss the most promising strategies for 2021 and beyond.

### Edoardo Sarti (Sorbonne Université)

*Statistical analysis of coevolutionary signals on the SARS-CoV-2 Spike protein*

**Abstract:** The viral protein S (Spike) allows the docking of SARS-CoV-2 virion by exploiting the human receptor ACE2, and has a central role in the fusion of the fusion of the viral membrane with that of the host cell. Since it is an outer membrane protein, and therefore in principle capable of being targeted by the host's immune system or drugs, it is at the heart of biomedical approaches against SARS-CoV-2.

The strategy of the Laboratoire de Virologie Humaine at the ENS Lyon is to design peptides with high specificity and binding affinity that prevent the conformational changes that the protein must complete in order to fuse the two membranes. The choice of the binding sites of these peptides on the surface of the S protein is also fundamental : the Analytical Genomics group at Sorbonne Université is developing techniques based on coevolutionary signals for the prediction of residue-residue contacts and allosteric changes.

An analysis which combines signals from different phylogenetic distances as well as several structural properties allows to define interesting candidates for the production of peptides. The predictions will be tested by the Laboratoire de Virologie Humaine using viral pseudo-particles that allow comparing the infectivity of the virion during the fusion process in presence of different solutes.

### Yasaman Karami and Laura Ortega Varga (Institut Pasteur)

*Targeting SARS-CoV-2 Spike Glycoprotein to fight Covid-19*

**Abstract:** Authors : Arnaud Blondel, Max Bonomi, Guillaume Bouvier, Luis Checa Ruano, Michael Nilges, Yasaman Karami, Laura Ortega Varga, Olivier Sperandio

In 2020 the Covid-19 global pandemic has emerged, caused by severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2) outbreak. One promising therapeutic strategy to fight this disease, is to prevent the cellular infection mechanism, which is triggered by the interactions between the receptor binding domain (RBD) of the SARS-COV-2 Spike glycoprotein and the human ACE2 protein. To this end, a set of molecular dynamics (MD) simulations were carried out to explore and decipher the transition path between the open and closed states of the RBD as well as the dynamics of the open state and that of the Spike-hACE2 complex. This allowed us to

detect plausible druggable sites with our in-house program mkgridXf. After analysis of the geometrical evolution of those pockets along MD and of their druggability, using the DeepPPI-Pocket detection tool developed in our lab, we could select promising effector sites.

Two groups of sites were identified : i) one at the interface between RBD and hACE2, and ii) various potential allosteric pockets, positioned in locations where they could block the RBD opening and function. Docking of in-house chemical libraries has been performed on available crystallographic structures and representative conformations from MD, retrieved by self-organizing maps (SOM). For the analysis and selection of molecules, different scoring functions have been used along with buriedness assessment and structural interaction fingerprints. This study will pave the way toward a better understanding of the conformational dynamics behind Spike-hACE2 interactions and will allow us to propose small molecules inhibiting such interactions thanks to virtual screening of the identified pockets.

### Frederic Cazals (Inria)

*Spikes of SARS-cov-2 : mechanisms and visualization*

**Abstract:** Enveloped viruses, including SARS-Cov-2, perform infection by fusing their envelope with the membrane of the target cell. The protein involved, a homotrimeric class I fusion protein called protein S, accomplishes this task via a complex and highly dynamic mechanism with two main steps separated by a cleavage step.

The first part of the talk will review this mechanism, based upon the 10 high resolution structures published to date since March 2020. In the second part, I will present a visualization tool, called MISA for Multiple Interface String Alignment, displaying sequence and structure based pieces of informations, of bound and/or unbound proteins.

As an illustration, we will scrutinize the spike protein, analyzing which regions are stable or not before/after binding, and also studying the binding competition between the cognate partner of protein S in human (ACE2), and antibodies targeting the so-called receptor binding domain of protein S.

MISAs are available in the SBL from https ://sbl.inria.fr/doc/Multiple_interface_string_alignment-user-manual.html

**URL/DOI:** https ://www.biorxiv.org/content/10.1101/2020.09.03.281600v1

### Jean-Philip Piquemal (Sorbonne Université)

*High-Resolution Modeling of SARS-CoV-2 Proteins Structural Dynamics*

**Abstract:** In this talk, I will present some aspects of our COVID-HP PRACE computational project aiming at obtaining a high-resolution structural and dynamical description of the components of the Sars-Cov-2 virus.

To do so, we rely on intensive new generation polarizable molecular dynamics simulations (AMOEBA force field) using the massively parallel/multi-GPUs Tinker-HP software to produce detailled conformational ensembles that could be useful to experimentalists as well as to theoreticians to perform for further ensemble docking/virtual screening studies. I will focus my talk on the description of our modeling efforts dedicated to the Spike (with the Pasteur Institute) and Main protease (Mpro) proteins.

Jaffrelot Inizan, Theo ; Célerse, Frédéric ; Adjoua, Olivier ; El Ahdab, Dina ; Jolly, Luc-Henri ; Liu, Chengwen ; et al. (2020) : High-Resolution Mining of SARS-CoV-2 Main Protease Conformational Space : Supercomputer-Driven Unsupervised Adaptive Sampling. ChemRxiv. Preprint. https ://doi.org/10.26434/chemrxiv.13003166.v5

**URL/DOI:** https ://doi.org/10.26434/chemrxiv.13003166.v5

### Maria Kadukova (Nano-D, Inria Grenoble)

*Re-docking of the available MPro co-crystal ligands*

**Abstract:** Starting from 2013, our team has been developing knowledge-based protein-ligand scoring functions, namely Convex-PL, Convex-PL-R, and KORP-PL. We took the opportunity of assessing our methods on the recent Covid-19 structural data, and ran a series of re-docking experiments with the Mpro structures. The covalent nature of binding of the majority of its known co-crystal ligands made the docking process rather challenging.

We have tried both covalent and non-covalent protocols. Covalent docking done with AutoDock Vina produced poses with rather low RMSDs if the covalent link was known. However, in a realistic scenario with unknown covalent links, docking with the search for the chemically possible links resulted in favoring unnatural binding poses. Non-covalent docking with our in-house version of AutoDock Vina and subsequent re-scoring with KORP-PL provided the 2.73 A median RMSD in cross-docking of 36 ligands with re-generated initial ligand structures and excluded native receptors. This was the most successful protocol among those we assessed and can be further used for the virtual screening or Mpro inhibitors. In this talk we would like to discuss our protocols and the problems we have encountered.

**URL/DOI:** https ://team.inria.fr/nano-d/software/

### Stephane Redon (OneAngstrom)

*SAMSON - Integrated Molecular Design*

**Abstract:** SAMSON is an integrated, open platform for molecular design that makes it possible to connect users and developers thanks to its Software Development Kit and associated marketplace for (free and non-free) extensions. It is available at https ://www.samson-connect.net.

In this presentation, we will present SAMSON and its unique capabilities for building, simulating, analyzing and visualizing large and complex molecular systems. We will also present the recently introduced extensions for cloud computing, as well as how the Software Development Kit makes it possible to integrate and distribute existing or new apps, data and services.

Finally, we will discuss how SAMSON and its SDK may offer an opportunity to help address current global challenges by contributing to federate the molecular modeling community through integration and dissemination of research results.

**URL/DOI:** https ://www.youtube.com/watch ?v=eHgKthSkqiA

### Marc Baaden (CNRS)

*FAIR sharing of molecular visualization experiences : COVID-19 use case*

**Abstract:** Visualization renders structural molecular data accessible to a broad audience. We describe an approach to share molecular visualization experiences based on FAIR principles. Our workflow is exemplified with recent Covid-19 related data.

**URL/DOI:** 10.1101/2020.08.27.270140

**Sebastian Will (Ecole Polytechnique)**

*Structural basis of the template switching mechanisms in SARS-Cov 2*

**Abstract:** TBA